Partially Observable Markov Decision Process for Monitoring Multilayer Wafer Fabrication

Article *in* IEEE Transactions on Automation Science and Engineering · August 2020

DOI: 10.1109/TASE.2020.3017481

3 authors, including:

Marzieh Khakifirooz
Tecnológico de Monterrey
59 PUBLICATIONS **555** CITATIONS
SEE PROFILE

Michel Fathi
University of North Texas
113 PUBLICATIONS **762** CITATIONS
SEE PROFILE

Partially Observable Markov Decision Process for Monitoring Multilayer Wafer Fabrication

Marzieh Khakifirooz, *Member, IEEE*, Mahdi Fathi, *Member, IEEE*, and Chen-Fu Chien, *Member, IEEE*

***Abstract*— The properties of a learning-based system are particularly relevant to the process study of the unknown behavior of a system or environment. In the semiconductor industry, there is regularly a partially observable system in which the entire state of the process is not directly or fully visible due to uncertainties or disturbances. The model for studying such a system that permits uncertainties regarding the stochastic (Markov) process for state information acquisition is called a partially observable Markov decision process (POMDP). This study aims to deal with the optimization issue of compensation control bias of a dynamic multilayer lithography process in wafer fabrication with prior information, the existence of high-dimensionality, and unmeasurable uncertainties. We show how the POMDP on a linear state-space model with uncertainties can encode the information from past runs and layers, and deal with accumulated overlay error at the current run and layer. The Gibbs sampling is applied to optimize the belief function of POMDP optimization approach.**

***Note to Practitioners*—The multilayer overlay error of the photolithography process is one of the remarkable and challenging issues in wafer fabrication. In a multilevel manufacturing process, errors occur at each level, which would be accumulated in the upstream operations. The optimization objective will be even more critical in a high-mixed fabrication process. In this study, the learning-based control system emerged with the state-space model compensates the multilayer overlay error. The Gibbs sampling as a Bayesian approach as a core structure of optimization algorithm is utilized, which can be updated with information from engineering's domain knowledge or estimated information about previous runs. The robustness of the proposed optimization algorithm is shown by comparing the distribution of overlay error with conventional methods and with a fast convergence rate of the learning algorithm.**

***Index Terms*— Bayesian optimization, Gibbs sampler, multilayer overlay, partially observable Markov decision process (POMDP), semiconductor photolithography process.**

Manuscript received February 23, 2020; revised June 8, 2020; accepted August 14, 2020. This article was recommended for publication by Associate Editor J. R. Morrison and Editor F.-T. Cheng upon evaluation of the reviewers' comments. This work was supported in part by the Ministry of Science and Technology, Taiwan, under Grant MOST 109-2634-F-007-019 and in part by the Micron Foundation. *(Corresponding author: Marzieh Khakifirooz.)*

Marzieh Khakifirooz is with the Department of Industrial Engineering, Tecnológico de Monterrey, Monterrey 64849, Mexico (e-mail: mkhakifirooz@tec.mx).

Mahdi Fathi is with the Department of Information Technology and Decision Sciences, G. Brint Ryan College of Business, University of North Texas, Denton, TX 76203 USA (e-mail: mahdi.fathi@unt.edu).

Chen-Fu Chien is with the Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu 30013, Taiwan, and also with the Artificial Intelligence for Intelligent Manufacturing Systems (AIMS) Research Center, National Tsing Hua University, Hsinchu 30013, Taiwan (e-mail: cfchien@mx.nthu.edu.tw).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASE.2020.3017481

I. Introduction and Motivation

THE misalignment in the photolithography process is a demanding area which has received increasing attention by the semiconductor manufacturing strategists [1], [2]. Misalignment also known as overlay error has several significant effects on other manufacturing process such as dry-etching process [3] and yield enhancement [4]. The major dilemmas associated with overlay error in the literature are stochastic metrology delay [5], unmeasurable noise and disturbance [6], [7], online monitoring [8], high-mixed manufacturing structure [9], [10], control system structure (i.e., multi-input multioutput, single-input single-output) [11], accurate and stable convex optimization algorithm (i.e., exponentially weighted moving average (EWMA) and proportional integral derivative) [12], and multilayer overlay [13].

The specific goal of this study is to propose an optimal control system for monitoring a multilayer overlay error. In the multilayer photolithography process, if the pattern of each layer cannot align well with the previous layer, the whole process will fail in a way that the variation in the previous reference layer influences misalignment in the subsequent layer. The consequences of multilayer overlay errors are more noticeable for the 3-D integrated circuit (3D IC) [14]. In 3D IC technology, the process requires building circuits with multiple layers of active devices; therefore, every single device (wafer/chip) enforced to coordinate well for the conventional photolithography process. The misalignment in wafer-level 3D IC is crucial since if any layers of a single chip are defective, then the entire 3D IC will be defective.

Fig. 1 shows the overlay error in the multilayer production process for two different cases of yaw misalignment (case A) and axial misalignment (case B). The errors in adjacent layers are similar in both cases; however, the total superposed error is better in case B than case A.

An accurate error model for presenting the accumulation and aggregation of misalignment is essential to control the overlay error efficiently. To the best of our knowledge, only limited researchers have investigated the monitoring system of multilayer overlay errors in the literature (see Table I). However, they did not consider uncertainty and other complexity, such as high-mixed scheduling and delay in the system.

On the other hand, the run-to-run (R2R) control is the most general monitoring technique in semiconductor manufacturing [20]. The main objective of R2R control is variability reduction of the process through the shrinking process output error. The EWMA controller is the most preferred design in R2R control [21]. However, the EWMA controller is known

TABLE I
LITERATURE REVIEW OF PROCESS CONTROL OF MULTILAYER OVERLAY ERROR

Reference	Problem	Core principals	Core contribution	Optimization technique	Compare with current study
Conway et al. (2003) [15]	How can the perturbations of the reference level, impact overlay controllability at cascading levels?	Correlation analysis	Reference layer feedforward can improve control objective at the target layer	-	Delay is happened in layer to layer level by missing the information of perturbations of reference layers.
Yu and Qin (2009) [16]	How significant are the mean/variance of each component in different layers to detect the error source?	Multistage state-space model	Root-cause detection of multilayer overlay error	Minimum norm quadratic unbiased estimation	Considered only layers with significant deviation from the target value.
Jiao and Djurdjanovic (2011) [17]	How is the stack-up effect of the multilayer overlay errors?	Minimizing the expectation of the weighted sum of stack-up overlay errors	Stack-up effects of multilayer overlay can be significantly reduced compared to the traditional run-to-run control	Stochastic sequential optimization	Considered stochastic behavior via the noise variable by Gaussian distribution with mean zero and unknown variance.
He and Zhang (2015) [18], and He and Zhang (2017) [19]	How is the cascading effect of multilayer overlay error?	State-space model for overlay variation	Categorizing the overlay error on intrafiled and interfiled coordination	Universal multivariate least-squares method	nonlinear noise for process uncertainties due to missing flatness in layer-stacking process.

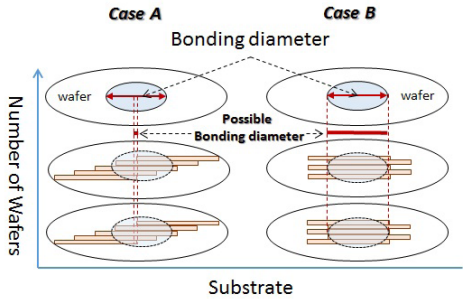

Fig. 1. Schematic of multilayer overlay error with bonding pad for "via" design.

to have several limitations. Some of the underlying limitations of EWMA include the following [22]:

1) dependence on maximum likelihood estimator (MLE);
2) dependence on limited control action by fixed filtering parameters;
3) dependence on multiple filtering steps;
4) inefficiency to deal with the large-scale disturbance of the real-world system.

Therefore, the EWMA controller is unfeasible for applying to a wide range of applications. In this study, motivated by the concept shown in Fig. 1, for compensating the multilayer overlay error, cumulative process information from previous layers are used as the constraints to update the process control law in the manufacturing layer. The proposed method in this study is followed by the high-mixed scheduling strategy in [1] considering different sources of uncertainties such as stochastic metrology delay and process disturbance.

For dealing with difficulties engaged in compensating multilayer overlay errors, such as stack-up error and dynamic behavior of the production process, the environment is formulated as a Markov decision process (MDP). Besides, for the system with some degrees of uncertainty where information is partially observable, MDP is accelerated to the partially observable MDP (POMDP) method [23]. Optimizing the throughput of an adaptive control system in the Markovian process is a demanding task. Therefore, the proposed optimization algorithm in this study is derived from a Markov chain Monte Carlo (MCMC) algorithm inspired by the Gibbs optimization technique in [4], [24], and [25].

The remainder of this study is organized as follows. Section II introduces the fundamental assumptions and definitions. Section III proposes the optimization solution for compensating the multistage (layer) overlay error. Section IV exposes the numerical illustration of the optimization-based POMDP approach and details the validation of the proposed method with Gibbs sampling techniques for simulation. Section V summarizes the main results and recommendations for further research.

II. FUNDAMENTALS

The notation and terminologies used in this study are listed as follows.

k	The layer index.
j	The overlay factor index.
t	The process run index, $t \geq 1$.
m	Number of layers.
N	Number of overlay factors.
l, l'	The length of process delay, and metrology delay, respectively.
i, i'	The iteration indexes for Gibbs sampler.
n	Number of iterations for Gibbs sampler.
$\mathbf{u}_t(k)$	$N \times 1$ input vector for layer k at run t.
$\mathbf{Q}_t(k)$	$N \times 1$ output vector for layer k at run t.
$\mathbf{x}_t(k)$	$N \times 1$ state vector for layer k at run t.
$d_t(k)$	The process disturbance for layer k at run t.
$e_t(k)$	The measurement noise for layer k at run t.
$\mathbf{s}_t(k)$	$N \times 1$ stack-up overlay vector for layer k at run t.
$\gamma_t(k)$	The weighting parameter for stack-up overlay error.
$\mathbf{E}_t$	$N \times 1$ deviation vector for layer k at run t.
T	The target of overlay factors.
S_t	A finite set of states at run t.
A_t	A finite set of actions at run t.
O_t	A finite set of observations at run t.
B_t	The distribution over state S_t at run t.
R	The reward function.

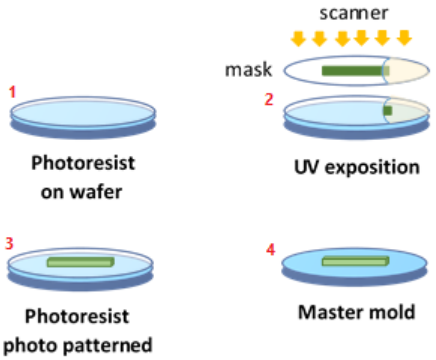

Fig. 2. Wafer fabrication in photolithography process.

R^*	The average reward function.
$G(\mathbf{s}_t(k))$	The learning based reward function for stack-up overlay for layer k at run t.
r_t	The discount factor at run t.
π_t	The policy function of state at run t.
$P(\cdot)$	The probability function.
$E(\cdot)$	The expectation function.
K	A large number for balancing learning rates.
$\boldsymbol{\theta} = [\mathbf{a}(k), \mathbf{b}(k), \mathbf{c}(k)]$	Coefficient matrices in state-space model for layer k.
λ, η	Parameters of zero-inflated Poisson distribution.
α_t, α_t'	Learning parameters.
ε	Maximum tolerance for learning algorithm.
$\boldsymbol{\beta}, \boldsymbol{\beta}^*$	Parameters of EWMA controller.
ω	The discount factor of EWMA controller.

A. Multilayer Overlay Model

Lithography is the most frequently used processes in wafer fabrication. Currently, the step-and-scan (shortened scanner) method is one of the most commercially used systems, in the lithography process. The purpose of the scanner is to superimpose a masking pattern on top of the existing wafer pattern. Fig. 2 illustrates the initial steps of wafer fabrication in the photolithography process when using the scanner. The gap between the actual position of the mask and substrate is known as overlay error [26], [27].

Overlay errors are measured from the misalignment between the current and previous exposure layers, through the box-in-box design. When the inside box is accurately patterned in the center of the outside box, no overlay error is apparent (Fig. 3). Consider how the direction of movement in the box-in-box design, the error can be described in x- and y-axes separately. The major source of misalignment in the lithography process can classify into two categories, the interfield overlay errors,

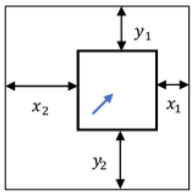

Fig. 3. Overlay error measurement (blue arrow in the small box shows the movement direction).

which is the result of misalignment between the mask and the wafer, and the intrafield overlay errors as a result of the mismatch between the light source filter lens and the mask. The overlay error is accumulated by several overlay factors such as translation, rotation, magnification, expansion, and rotation, which are solicited to be as small as possible to the target point (T) or zero [27].

The overlay error in the multilayer photolithography process is impressed by the following phenomena.

1) *Layer to Layer Misalignment:* In each layer, there is a cumulative error from the first to the current. The control parameters at each level are estimated to minimize the total overlay error, and the initial parameter setting is the cumulative error from the previous layers [16].
2) *Recipe Confliction:* The photolithography process with a single scanner device is a bottleneck in semiconductor manufacturing, which generates a high-mixed setting of recipe adjustment for different products at different levels of production. Each recipe has its alignment and control laws, and therefore, the process parameters are unique for each recipe [28].
3) *Process and Metrology Delays:* Delays hinder information from reaching the controller at the right time and thus affect the control performance. Process delay is inherent in the photolithography process due to being a bottle. On the other hand, metrology delay appears because of the time and tools' limitations in the metrology station [29].

Considering this fact that, typically, different recipes apply in different layers, for simplicity, this study considers only the effect of multilayer phenomena along with the effects of process and metrology delays.

Fig. 4 illustrates the block diagram of a control system for compensating overlay errors in the state-space model. According to Fig. 4, the input ($\mathbf{u}_t$) is the variation in each overlay variable from the target (T), which is affected by process delay (for l runs), and process disturbance (d_t). The control plant continually and effectively measures the value of all overlay variables ($\mathbf{x}_t$). However, this measurement is accompanied by another source of uncertainties, namely measurement noise (e_t) and measurement delay (for l' runs). The variation from the target value, which is an output of the manufacturing system ($\mathbf{Q}_t$) compared with the input ($\mathbf{u}_t$), is calculated by the control plant and the feedback command for corrective action forward to the controller.

In controlling overlay error the assumption is that all overlay variables are independent in each layer. Therefore, the general

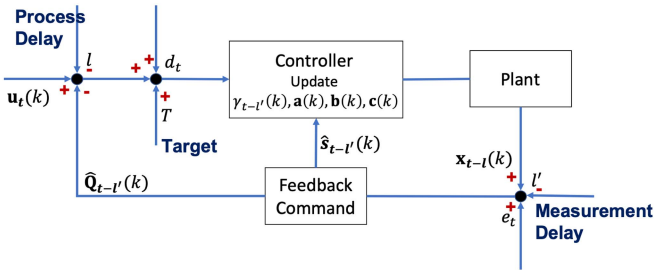

Fig. 4. Block diagram of POMPD controller for single layer overlay error.

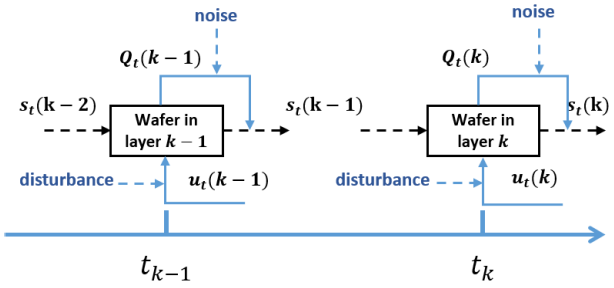

Fig. 5. Schematic of multilayer overlay model with stack-up error.

linear state-space model for describing Fig. 4 can be modeled as follows:

$$\begin{aligned}\mathbf{x}_{t-l}(k) &= \mathbf{a}(k)\mathbf{x}_{t-l-1}(k) + \mathbf{b}(k)\mathbf{u}_{t-l}(k) + d_t(k)\\ \mathbf{Q}_{t-l'}(k) &= \mathbf{c}(k)\mathbf{x}_{t-l}(k) + e_t(k)\end{aligned} \tag{1}$$

where the initial state of the system at time t is $\mathbf{x}_t(0) = 0$.

In practice, as depicted in Fig. 4, the system uncertainties including noise e_t and disturbance d_t are included into the random effects of $\mathbf{Q}_t$ and $\mathbf{u}_t$, respectively. In addition, there is a metric for describing the integral of error from the first layer to the current layer called stack-up overlay error

$$\mathbf{s}_{t-l'}(k) = \mathbf{s}_{t-l'}(k-1) + \mathbf{Q}_{t-l'}(k). \tag{2}$$

The schematic of the stack-up overlay error is shown in Fig. 5. Therefore, the process error is

$$\mathbf{E}_t(k) = \mathbf{Q}_t(k) - T \tag{3}$$

where $T = 0$ can be combined with aggregated stack-up error in (2) to define the control laws for minimizing the total overlay error. Therefore, the stochastic optimization problem which minimize the multilayer overlay model at kth layer by finding a sequence of vectors of controllable process parameters $\{\mathbf{a}(k), \mathbf{b}(k), \mathbf{c}(k)\}$ can be characterized as follows:

$$\begin{aligned}\arg \min_{\mathbf{a}(k),\mathbf{b}(k),\mathbf{c}(k)} & \sum_{k=1}^{m} \gamma_{t-l'}(k)(\|\mathbf{s}_{t-l'}(k))\|)^2\\ \text{s.t. } & \mathbf{s}_{t-l'}(k) = \mathbf{s}_{t-l'}(k-1) + \mathbf{Q}_{t-l'}(k)\\ & \mathbf{Q}_{t-l'}(k) = \mathbf{c}(k)\mathbf{x}_t(k)\\ & \mathbf{x}_t(k) = \mathbf{a}(k)\mathbf{x}_{t-l}(k) + \mathbf{b}(k)\mathbf{u}_{t-l}(k)\\ & \mathbf{s}_{t-l'}(0) = 0\\ & \forall l' < t, k : 0 < \gamma_t(k) < 1\\ & \sum_{k=1}^{m} \gamma_t(k) = 1\end{aligned} \tag{4}$$

where $\gamma(k)$ can be selected by tuning algorithm and expresses the relative importance of overlay error corresponding to each layer. In practice, $\gamma(k)$ can be selected based on expert opinion.

B. POMDP

The POMDP [24] is a generalization of MDP when only part of the information is unavailable about the current state (for instance, due to delay), and this leads to the uncertainties (i.e., noise).

Consider a class of algorithms for finding good approximations to a class of learning problems in which agents interact in a dynamic, noisy, and stochastic environment; this interaction is conventionally modeled as a POMDP with the following properties.

1) S_t: Finite set of states.
2) A_t: Finite set of actions.
3) $R(S_t, A_t)$: Reward function.
4) $P(S_{t+1}|S_t, A_t)$: State transition probability function.
5) O_t: Set of observations.
6) $P(O_{t+1}|S_{t+1}, A_t)$: Observation probability.
7) $r_t \in [0, 1]$: Discount factor.
8) B_t: Distribution over state S_t called "Belief State."

POMDP can be identified as an optimal or near-optimal behavior for an uncertain system [30]. The MDP problem seeks to find a mapping from states to actions; however, due to partially available data, the challenge in the POMDP problem is to find a mapping from probability distributions over states to actions. For dealing with this phenomenon, the key step is to calculate the value of a given policy function (π). The policy function (π) is the mapping function from the state to the action, for maximizing the expected sum of the discounted factors. The block diagram of POMDP is demonstrated in Fig. 6 and the optimization procedure is described as the following steps.

1) Set up the unobserved state S_t of the system at each time/step t.
2) Select an action A_t.
3) Maintain the distribution over S_t as B_t.
4) Receive the reward function $R(S_t, A_t)$.
5) Transit to the unobserved state S_{t+1} with probability $P(S_{t+1}|S_t, A_t)$.
6) Receive the observation O_{t+1} with probability $P(O_{t+1}|S_{t+1}, A_t)$.
7) Estimate the distribution of state S_{t+1} as $B_{t+1}(S_{t+1}) = P(S_{t+1}|O_{t+1}, A_t, B_t(S_t))$.
8) Update the reward function by $R(S_{t+1}, A_{t+1}) = B_{t+1}(S_{t+1}) \times R(S_t, A_t)$.
9) Optimize the return function by policy $\pi(S_{t+1}) = \max_{A_{t+1}} \sum r_{t+1} R(S_{t+1}, A_{t+1}))$ and select the best action A_{t+1}.
10) Update and repeat the process.

In the deterministic setting, techniques such as dynamic programming can be used to tackle the optimal control problem [31]. However, the Bellman equation is often the most convenient method of solving stochastic optimal control problems [32]. Bellman's optimality equation [33] says that

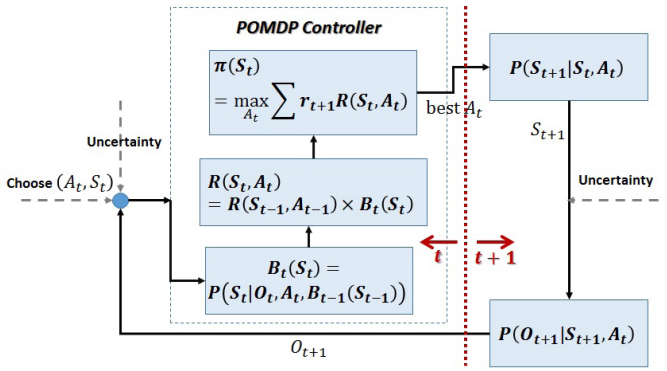

Fig. 6. Block diagram of a POMDP.

under principal of stochastic approximation the average reward $R^*(S_t, A_t)$ from t periods simulation-based solution is

$$R^*(S_t) = \min_{A_t}\Bigg[R(S_t, A_t) - \frac{1}{t}E\left\{\sum_t R(S_t, A_t)\right\} + \sum_{S_t} P(S_t|S_{t-1}, A_{t-1}) \times E\left\{\sum_t\left(R(S_t, A_t) - \frac{1}{t}E\left\{\sum_t R(S_t, A_t)\right\}\right)\right\}\Bigg]. \tag{5}$$

In practice, the average reward function $[R^*(S_t)]$ in (5) is the best (in this study minimum) reward function that is learned from the previous results and can be led to the best action.

The optimization problem in (5) is a simulation-based method for solving POMDP. Note that the optimization problem in (5) is expensive when the size of observation or time-steps grows and leads to the exponential growth of the size of the policy space [34]. In this article, the approximated Gibbs sampling solution method based on the MCMC technique is applied to optimize the system policy and minimize the computation cost/time. Additionally, for compensating the overlay error of a multilayer system, we study the multistage POMDP setting [35], where the objective is to optimize m stage with the same state–action space, but different dynamics and rewards.

C. Bayesian Optimization

In this study, the Bayesian optimization technique is used to minimize the expected value of reward functions in (5). According to Bayesian probability theory, the likelihood of each observation is relative to the likelihood of other observations that already happened. Bayesian models can induce the noise at each level of the explanatory variables, and represent the dependence among variables.

To demonstrate the optimization process-based Bayesian inferences, let $P(X|x)$ be the probability for estimating an unobserved population parameter X on the basis of given structure x. Assume prior distribution $P(x)$ for the likelihood of each structure, then the posterior distribution on x given by the Bayes rule is

$$P(x|X) = \frac{P(X|x)P(x)}{P(X)}. \tag{6}$$

Algorithm 1 Gibbs Sampling

Initiate $\gamma_1^{(0)}, \gamma_2^{(0)}, \ldots, \gamma_m^{(0)}$
for $i \leftarrow 1 \rightarrow n$ **do**
$\gamma_1^{(i)} \sim P(\boldsymbol{\gamma}_1 = \gamma_1|\boldsymbol{\gamma}_2 = \gamma_2^{(i-1)}, \ldots, \boldsymbol{\gamma}_m = \gamma_m^{(i-1)})$
$\gamma_2^{(i)} \sim P(\boldsymbol{\gamma}_2 = \gamma_2|\boldsymbol{\gamma}_1 = \gamma_1^{(i)}, \ldots, \boldsymbol{\gamma}_m = \gamma_m^{(i-1)})$
$\ldots$
$\gamma_m^{(i)} \sim P(\boldsymbol{\gamma}_m = \gamma_m|\boldsymbol{\gamma}_1 = \gamma_1^{(i)}, \ldots, \boldsymbol{\gamma}_{m-1} = \gamma_{m-1}^{(i-1)})$
end for

The general aim of Bayes rule is to find the most probable value (mode) of x given the observation X

$$\hat{x} = \arg\max_X P(x|X). \tag{7}$$

However, quantifying the idea of a Bayesian model is difficult when the distribution of observations is stochastic or unknown. Nevertheless, one approach to facilitating this difficulty is to sample from the distribution before computing the sample statistics. The MCMC method has facilitated Bayesian statistics for this purpose [36]. One basis of Markov chain theory posits that, if the probability associated with different events is constructed correctly, and the chain has a sufficient length, then the event distribution can be made equal to any arbitrary distribution, including a posterior distribution.

Gibbs sampling is an MCMC technique suitable for sampling from the distribution and estimating sample statistics [37]. Gibbs sampling conditionally generates the posterior sample by eliminating one variable at each iteration, while the remaining variables fixed to their latest estimated value (see Algorithm 1)

Gibbs sampling has been extensively adapted to analyze a variety of predictive analytics challenges. For more details, one can refer to comprehensive reviews of studies on application of Gibbs sampling on state-space model [38], graphical modeling [39], adaptive sampling [40], and variable selection [41].

III. Proposed Framework

This section describes the structure of the proposed optimization solution for controlling a multistage (multilayer) system with partially available information. In the first step, after defining the stack-up overlay error for multilayer wafer fabrication, due to the high dependence of the stack-up error to previous information and delay, POMDP is applied to optimize the error. Since updating the belief function of POMDP by growing the process run becomes infeasible, Gibbs sampling is applied to estimate the belief function (see Fig. 7). Following this section, the details of the optimization process are explained.

A. Multistage Control System Using POMDP

The POMDP relies on defining a set of states, the expected observations from those states, the action transition matrix, and the reward function. Following this section, each component's applications for computing the optimal action given a specific belief about the current state of the system is discussed. The detailed descriptions of each component for modeling a

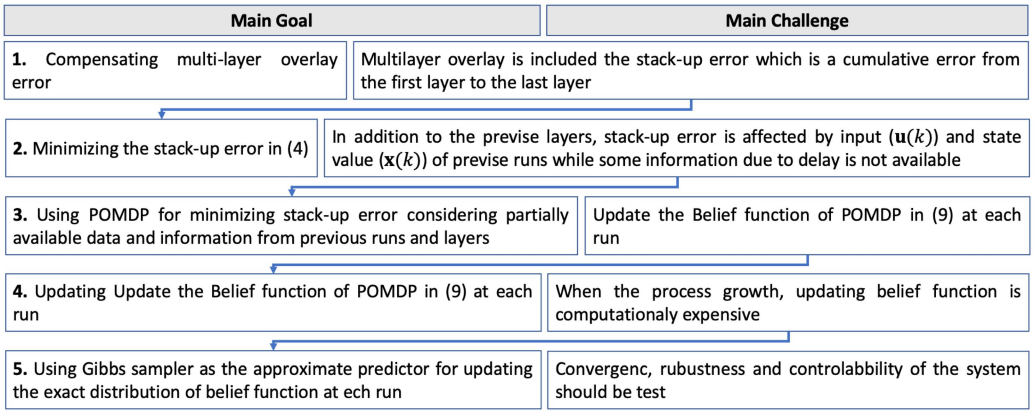

Fig. 7. Structure of proposed Bayesian optimization algorithm for compensating multilayer overlay error.

control system of multilayer overlay error are summarized as follows.

1) *State-Space:* Consider (1), the output of controller to the plant ($\mathbf{x}_t$), given the action ($\mathbf{u}_t$), including disturbances and process delay. There are $m \times N$ different states at each run in the system. Note that, in practice, the actual value of $\mathbf{x}_t$ is not observable and should be predicted.
2) *Observation-Space:* The actual output of the plant ($\mathbf{Q}_t$) for N overlay factors and m layers including metrology delay and measurement noise. Without loss of generality, regards to the definition of stack-up overlay in (2), the observation $\mathbf{Q}_t$ can be replaced by $\mathbf{s}_t$.
3) *Action Transition Matrix:* The probability matrix for each state ($\mathbf{x}_t$) of a specific layer can be appeared in the sequence of the photolithography process. The transition matrix can be computed based on the historical data and updated after each run. The elements of transition matrix can be derived by

$$P(\mathbf{x}_t(k)|\mathbf{x}_{t-1}(k), \mathbf{u}_{t-1}(k)). \tag{8}$$

4) *Belief Updating:* The probability distribution of $\mathbf{x}_t$ given the state of previous belief, and observation and action at the current run.
5) *Reward Function:* The actual error $\mathbf{E}_t(k)$ in (3) or stack-up error $\mathbf{s}_t(k)$ in (2) which results from action $\mathbf{u}_t(k)$.
6) *Average Reward Function:* The optimal value of actual error after t runs with regard to the transition matrix at each run and learning from the previous runs. Consider $\mathbf{s}_t(k)$ as the reward function in tth run. Then the simulation-based method for solving the average optimality reward function based on [22] is

$$\begin{aligned} &R^*(\mathbf{x}_t(k), \mathbf{u}_t(k)) \\ &= (1-\alpha_t)R^*(\mathbf{x}_{t-1}(k), \mathbf{u}_{t-1}(k)) + \alpha_t \\ &\quad \times \left[\mathbf{s}_t - \frac{1}{t} E\left\{ \sum_t G(\mathbf{s}_t(k)) \right\} \right. \\ &\quad + \min_{\mathbf{u}_t} \sum_{x_t} P(\mathbf{x}_t(k)|\mathbf{x}_{t-1}(k), \mathbf{u}_{t-1}(k)) \\ &\quad \left. \times E\left\{ \sum_t \mathbf{s}_t - \frac{1}{t} E\left\{ \sum_t G(\mathbf{s}_t(k)) \right\} \right\} \right] \end{aligned} \tag{10}$$

where

$$\begin{aligned} G(\mathbf{s}_t(k)) &= (1-\alpha_t')G(\mathbf{s}_{t-1}(k)) \\ &\quad + \alpha_t' \left[\frac{(t-1)G(\mathbf{s}_{t-1}(k)) + \mathbf{s}_t(k)}{t} \right] \end{aligned} \tag{11}$$

where at $t = 1$, $G(\mathbf{s}_1(k)) = \alpha_1' \mathbf{s}_1(k)$.
The learning parameters α and α' are both decayed by the following rule:

$$\alpha_t, \alpha_t' = \frac{\alpha_0, \alpha_0'}{1 + \frac{t^2}{K+t}} \tag{12}$$

where K is a very large number.
7) *Optimal Action:* The objective function of a control system in (4) can be minimized by the optimal solution of stationary policy given by the observation space (stack-up overlay error)

$$\pi(B(\mathbf{x}_t)) = \arg\min_{\gamma_t} [\boldsymbol{\gamma}_t B(\mathbf{x}_t) \times R^*(\mathbf{x}_t(k), \mathbf{u}_t(k)) \boldsymbol{\gamma}_t']. \tag{13}$$

8) *Controllability and Observability:* For having a controllable and observable system, the following assumptions should be satisfied.
   a) The model applies over an infinite number of run, implying that the control system is stationary.
   b) Conditioned on the true $\mathbf{u}_t(k)$ and control setting at run $t-1$, the $P(\mathbf{s}_{t-1}|\mathbf{x}_t)$ is independent of information related to the run $t-1$.
   c) The measurement noise and process disturbance are accumulated to the $\mathbf{Q}_t(k)$ and $\mathbf{u}_t(k)$, respectively.
   d) Regards to policy function in (9), as shown at the bottom of the next page, optimizing objective function in (4) is updated by

$$\begin{aligned} \arg\min_{\gamma(k)} \sum_{k=1}^{m} &(\gamma_{t-l'}(k) B(\mathbf{x}_t(k)) \\ &\times R^*(\mathbf{x}_t(k), \mathbf{u}_t(k)) \gamma_{t-l'}(k)). \end{aligned} \tag{14}$$

9) *Convergence Assumptions:* The POMDP is converged under a set of assumptions as follows.

a) The values of α_t and α'_t should be chosen very small in order to allow slow learning and corresponding convergence.
b) $|R^*(\mathbf{x}_t(k), \mathbf{u}_t(k)) - R^*(\mathbf{x}_{t-1}(k), \mathbf{u}_{t-1}(k))| < \varepsilon$.
c) $\lim_{t\to\infty} \sup(\alpha_t/\alpha'_t) = 0$.
d) $\lim_{t\to\infty} var(E\{\sum_t \mathbf{s}_t - (1/t)E\{\sum_t G(\mathbf{s}_t)\}\}) < \infty$.
e) $\lim_{t\to\infty} var((1/t)E\{\sum_t G(\mathbf{s}_t)\}) < \infty$.

B. Gibbs Sampling Optimization

In practice, POMDP is often computationally expensive to be solved precisely, and several studies have been developed approximate solutions for POMDP [42]. In this study, we use sampling techniques through the Gibbs sampler to update the belief probability.

For generating the state vector using Gibbs sampler as optimization tool, consider the state-space model in (1). Let $P(\mathbf{x}_{1:t}(k), \boldsymbol{\theta}, \mathbf{s}_{1:t-1}(k))$ be the joint posterior density of $\mathbf{x}_{1:t}(k)$, $\boldsymbol{\theta}$ and $\mathbf{s}_{1:t-1}(k)$. Therefore, Gibbs sampler can generate $\mathbf{x}_{1:t}(k)$, and $\boldsymbol{\theta}$ from the conditional densities $P(\mathbf{x}_{1:t}(k)|\boldsymbol{\theta}, \mathbf{s}_{1:t-1}(k))$, and $P(\boldsymbol{\theta}|\mathbf{x}_{1:t}(k), \mathbf{s}_{1:t-1}(k))$, respectively, until eventually $(\mathbf{x}_{1:t}(k), \boldsymbol{\theta})$ is generated from the joint posterior distribution $P(\mathbf{x}_{1:t}(k), \boldsymbol{\theta}, \mathbf{s}_{1:t-1}(k))$. Therefore, the only essential assumption is θ_1 to be known.

Assume that $\mathbf{x}_1(k)$ has a known distribution with condition on $\mathbf{u}(k)$, $\mathbf{s}(k)$ and parameter vector $[\mathbf{a}(k), \mathbf{b}(k), \mathbf{c}(k)]$. For simplicity, we name the parameter vector $[\mathbf{a}(k), \mathbf{b}(k), \mathbf{c}(k)]$ as $\boldsymbol{\theta}$. Lemma 1 shows how to generate $\mathbf{x}(k)$ given $\mathbf{s}(k)$ and $\boldsymbol{\theta}$ [37].

Lemma 1: We have

$$P(\mathbf{x}_{1:t}(k)|\boldsymbol{\theta}, \mathbf{s}_{1:t-1}(k)) \propto P(\mathbf{x}_t(k)|\boldsymbol{\theta}, \mathbf{s}_{t-1}(k)) \prod_{i=1}^{t-1} P(\mathbf{x}_i(k)|\boldsymbol{\theta}, \mathbf{s}_{i-1}(k), \mathbf{x}_{i+1}(k)).$$

Thus to generate $\mathbf{x}(k)$, the simplest approach would be to simulate $\mathbf{x}_{1:t-1}(k)$ from $P(\mathbf{x}_{1:t-1}(k)|\mathbf{s}_{1:t-2}(k))$, $\boldsymbol{\theta}$ from $P(\boldsymbol{\theta}|\mathbf{s}_{t-1}(k), \mathbf{x}_{t-1}(k))$ then $\mathbf{x}_t(k)$ from $P(\mathbf{x}_t(k)|\mathbf{x}_{1:t-1}(k))$.

IV. Illustration and Results

For validating the performance of the proposed POMDP controller using Gibbs sampler, a simulation study is conducted, followed by sensitivity analysis and discussion on how to implement the result in the real setting.

To run the proposed control system, we consider a continuous-time POMDP system with a finite-state space $\mathbf{x}_t$ for each layer ($k = 1, \ldots, m$). Let $\mathbf{u}_t$ be a finite action space (input of controller), when the system is in state $\mathbf{x}_t(k)$ for the layer k and action $\mathbf{u}_t(k)$ is taken, the system will transit to state $\mathbf{x}_{t+1}(k)$ at the next run with probability $P(\mathbf{x}_{t+1}(k)|\mathbf{x}_t(k), \mathbf{u}_t(k))$, and a reward $\mathbf{Q}_t(k)$ is received which is accumulated as $\mathbf{s}_t(k)$ at current layer k. In practice, $\mathbf{x}_t$ is unknown and the measured $\mathbf{s}_t(k)$ is used as the action $\mathbf{u}_{t+1}$ for the next run. However, $\mathbf{s}_t(k)$ could involve delay in measurement for l' runs which is denoted by $\mathbf{s}_{t-l'}(k)$. Therefore, the observation (output) of controller is the action (input) of controller for the next run. Consider this phenomenon, only the first action ($u_1(k)$) is necessary to be defined for the controller, which ideally is close to the target value (T). Following these facts, to develop the simulation study, $u_1(k = 1)$ and T set to 0. Three scenarios for describing uncertainties are designed as follows.

1) *Structural Change Model:* The data-generating process is given by

$$\mathbf{u}_t(k) = \hat{\mathbf{s}}_{t-1} + d_t(k) \tag{15}$$

where d_t periodically changes to 1 and -1 when k changes. This model corresponds to the case B (axial misalignment) in Fig. 1 and represents the sudden shifts due to changing the recipe for each layer.
2) *Crooked Change Model:* The data-generation process is the same as the structural change model in (15), where $d_t = d_{t-1} + 1$ when k changes. This model corresponds to case A (yaw misalignment) in Fig. 1.
3) *Stochastic Delay:* The metrology delay (l') and process delay (l) randomly generated from the zero-inflated Poisson (ZIP) distribution [43], with $\lambda = 2$ and $\eta = 0.6$ and with maximum delay length 6 for each lot.

To evaluate the performance of the proposed POMDP controller, we simulate a control system for the photolithography process with ten overlay factors and each factor with five layers. Two experiments are designed to consider the structural change with metrology delay and crooked change with metrology delay. The characteristics of uncontrollable outputs $\mathbf{Q}_t(k)$ with Gaussian distribution for each factor and layer are described in Table II. Different standard deviations are selected for each layer with an increasing pattern such that the last layer has the most significant standard deviation. The zero mean and small standard deviation for the Gaussian distribution of the first layer are considered due to a minimal but important value of nuisance factors in the wafer production process. The purpose of this pattern for data selection is because $\mathbf{s}_t(k)$ at layer k is influenced by all other previous layers, in which many sources of uncertainty (besides delay and structural/crook model of d_t) are involved in the system, including measurement noise and optimization uncertainty.

For performance evaluation of the proposed POMDP process control, as mentioned earlier $\mathbf{u}_1(k) = 0$ for $k = 1, \ldots, 5$, $\mathbf{Q}_t(k)$ is considered as the value of all sources of noise and disturbances in the system before using a control model, and the scheduling system is set based on the uniform random value for order of layers. Following the details of simulation scenario is described.

$$B(\mathbf{x}_t(k)) = \frac{P(\mathbf{s}_{t-1}(k)|\mathbf{x}_t(k), \mathbf{u}_{t-1}(k)) \sum_{\mathbf{x}_t} P(\mathbf{x}_t(k)|\mathbf{x}_{t-1}(k), \mathbf{u}_{t-1}(k)) B(\mathbf{x}_{t-1}(k))}{\sum_{\mathbf{x}_t} P(\mathbf{s}_{t-1}(k)|\mathbf{x}_t(k), \mathbf{u}_{t-1}(k)) P(\mathbf{x}_t(k)|\mathbf{x}_{t-1}(k), \mathbf{u}_{t-1}(k)) B(\mathbf{x}_{t-1}(k))}. \tag{9}$$

TABLE II
GAUSSIAN DISTRIBUTION OF $\mathbf{Q}_t(k)$ FOR 10×5 MULTILAYER OVERLAY FACTORS

	Overlay Factors									
Layers	*V.1*	*V.2*	*V.3*	*V.4*	*V.5*	*V.6*	*V.7*	*V.8*	*V.9*	*V.10*
L.5	$N(0, 10)$	$N(0, 9)$	$N(0, 10)$	$N(0, 9)$	$N(0, 8)$	$N(0, 10)$	$N(0, 9)$	$N(0, 8)$	$N(0, 7)$	$N(0, 10)$
L.4	$N(0, 5)$	$N(0, 2)$	$N(0, 3)$	$N(0, 4)$	$N(0, 5)$	$N(0, 3)$	$N(0, 4)$	$N(0, 5)$	$N(0, 4)$	$N(0, 5)$
L.3	$N(0, 1)$	$N(0, 0.9)$	$N(0, 0.7)$	$N(0, 0.8)$	$N(0, 0.9)$	$N(0, 1)$	$N(0, 0.8)$	$N(0, 0.9)$	$N(0, 01)$	$N(0, 0.9)$
L.2	$N(0, 0.5)$	$N(0, 0.4)$	$N(0, 0.3)$	$N(0, 0.2)$	$N(0, 0.5)$	$N(0, 0.4)$	$N(0, 0.3)$	$N(0, 0.2)$	$N(0, 0.5)$	$N(0, 0.4)$
L.1	$N(0, 0.1)$	$N(0, 0.09)$	$N(0, 0.08)$	$N(0, 0.07)$	$N(0, 0.08)$	$N(0, 0.09)$	$N(0, 0.1)$	$N(0, 0.09)$	$N(0, 0.08)$	$N(0, 0.07)$

1) Generating 200 runs of uncontrollable outputs $\mathbf{Q}_t(k)$ based on the model parameters in Table II, and different process uncertainties.
2) Generating 200 random numbers from $U(1, 5)$ to represent the order of layers at each run.
3) Using the distribution of $\mathbf{Q}_t(k)$ in Table II and considering the rows (k) indicate the index of layer at $t-1$th run and the columns (k') to the index of layer at tth run, the (k, k')th element of transition matrix is defined using Gibss sampler as follows:

$$P(\mathbf{x}_t(k')|\mathbf{x}_{t-1}(k), \mathbf{u}_{t-1}(k)) \tag{16}$$

where $\mathbf{x}_t(k)$ is estimated using (1) based on the initial value of $\mathbf{u}_1(k)$. Also, $\mathbf{u}_t(k)$ is calculated based on (15) for structural change and crooked changed scenarios separately, such that

$$\sum_{k,k'=1}^{5} P(\mathbf{x}_t(k')|\mathbf{x}_{t-1}(k), \mathbf{u}_{t-1}(k)) = 1.$$

These considerations and conditions remain the dependence situation of information at each layer, and each runs to its previous layers and runs, while overlay factors are considered independent.
A simple R-code to describe the Gibbs sampling process for the first overlay variable, and the first layer is presented as follows:

```
gibbs<-function (a,b,c)
{ Q <- rnorm(1, 0, 0.1)
s[i-1] <- Q[i-1]
s[i] <- s[i-1] + c*x[i-1]
u[i] <- s[i] + d[i]
x[i] <- a*x[i-1] + b*u[i]
}
```

4) Calculate belief and average reward functions using (9) and (10), respectively, where $\mathbf{s}_t(k)$ is estimated based on (2) and $P(\mathbf{s}_{t-1}(k)|\mathbf{x}_t(k), \mathbf{u}_{t-1}(k))$ is determined in a similar way as $P(\mathbf{x}_t(k')|\mathbf{x}_{t-1}(k), \mathbf{u}_{t-1}(k))$ is described in the last step. According to the first and third convergence rules, the value of α_t and α'_t are selected close to zero and initiated as 0.001 and 0.01, respectively (due to convergence assumptions in Section III-A, see step 9 for more details). In addition, for the fifth layer (the layer with the highest variation) $K = 1000$, and for simplification the expected value in (10) is considered as the weighted average value.
5) Optimizing problem in (4) with objective function (14) based on Lemma 1, using the estimated outputs, belief, and average reward functions, by utilizing Gibbs sampler. At the first step, $\boldsymbol{\theta}_1$ is initiated then Gibbs function in step 3 is used for simulation process while parameters a, b, c and γ in (14) are estimated using the tuning procedure. According to [4], the iteration time for the burn-in process of Gibbs sampler is selected equal to the sample size, which is 200 in this study.
6) Defining the sample mean and standard deviation of $\mathbf{s}_t(k)$ and $\mathbf{s}_t(k)$ for each overlay factor and layer over 200 runs for performance comparison.
7) Evaluating the performance of the proposed controller with 200 runs in comparison with EWMA controller in (17), as the most popular control system for batch processing, with $\omega = 0.3$, and $\mathbf{E}_t(k)$ as defined in (3). For a comprehensive review on EWMA controller for the semiconductor industry one can refer to [21], [28], [44]

$$\begin{aligned} \mathbf{Q}_t(k) &= \boldsymbol{\beta}_t(k) + \boldsymbol{\beta}_t^*(k)\mathbf{u}_t(k) + d_t(k) \\ \mathbf{u}_t(k) &= \mathbf{u}_{t-1}(k) - \frac{\omega}{\boldsymbol{\beta}_t^*(k)}\mathbf{E}_t(k). \end{aligned} \tag{17}$$

The comparison results for the cumulative stack-up layer $(\mathbf{s}_t(k))$, and predicted input $(\mathbf{u}_t(k))$ for each overlay factor and layer are summarized in Tables III and IV.
According to Tables III and IV, it is clear that the POMDP has smoother variations and improved compensation performance (e.g., closer to the target, $T = 0$) in comparison with the EWMA controller.
The result shows that the proposed POMDP controller tightens the error variation and eventually achieves a lower cost (overlay error) and disturbance compared with the EWMA control system. When the variation increases, the differences are more tangible, such that when the unmeasurable disturbance makes a significant shift in overlay factors, POMDP can competently deal with process shift, while EWMA is disabled to deal with this phenomenon. It is apparent that when disturbance and noise are increasing, the performance of the proposed POMDP controller is significantly better than EWMA.
For proofing the efficiency of the proposed POMDP controller on structural and crooked changed, the third layer with the medium range of noise is selected for comparison. Fig. 8 illustrates this comparison. Although for case B, both POMDP and EWMA controller are performing slightly better than case A, yet POMDP can compensate both types of disturbances supremely.

TABLE III

MEAN AND STANDARD DEVIATION OF $s_t(k)(\mathbf{u}_t(k))$ FOR EACH OVERLAY FACTOR AND LAYER WITH STRUCTURAL CHANGE AND METROLOGY DELAY

Layers		Overlay Factors									
		V.1	V.2	V.3	V.4	V.5	V.6	V.7	V.8	V.9	V.10
L.5	POMDP	$1, 1.1$	$1, 1.6$	$1, 2.1$	$1, 1.4$	$1, 1.55$	$1, 1.63$	$1, 1.9$	$1, 2.3$	$1, 2.4$	$1, 1.7$
		$(0.08, 0.4)$	$(0.08, 0.5)$	$(0.08, 0.5)$	$(0.08, 0.4)$	$(0.08, 0.4)$	$(0.09, 0.4)$	$(0.1, 0.4)$	$(0.09, 0.4)$	$(0.1, 0.4)$	$(0.1, 0.5)$
	EWMA	$111, 66$	$-53, 40$	$-12, 38$	$9, 47$	$0.4, 34$	$77, 79$	$-79, 38$	$-11, 28$	$11, 16$	$-42, 50$
		$(-0.73, 10)$	$(-0.52, 9)$	$(-0.46, 10)$	$(-0.34, 9)$	$(0.25, 7)$	$(0.58, 11)$	$(-0.56, 10)$	$(0.47, 9)$	$(-0.32, 7)$	$(0.11, 9)$
L.4	POMDP	$1, 0.3$	$1, 0.3$	$1, 0.3$	$1, 0.4$	$1, 0.3$	$1, 0.3$	$1, 0.3$	$1, 0.34$	$1, 0.41$	$1, 0.26$
		$(0.08, 0.5)$	$(0.05, 0.5)$	$(0.04, 0.5)$	$(0.1, 0.5)$	$(0.07, 0.4)$	$(0.07, 0.5)$	$(0.08, 0.5)$	$(0.05, 0.5)$	$(0.02, 0.5)$	$(0.03, 0.5)$
	EWMA	$-53, 34$	$-17, 9$	$5, 14$	$-86, 47$	$10, 15$	$5, 8$	$44, 27$	$-1.5, 11$	$-16, 12$	$-67, 38$
		$(-0.13, 5)$	$(-0.22, 2)$	$(-0.15, 3)$	$(-0.59, 4)$	$(0.31, 5)$	$(-0.11, 3)$	$(0.19, 4)$	$(0.32, 5)$	$(-0.43, 4)$	$(-0.35, 5)$
L.3	POMDP	$1, 0.17$	$1, 0.17$	$1, 0.16$	$1, 0.17$	$1, 0.18$	$1, 0.19$	$1, 0.18$	$1, 0.19$	$1, 0.16$	$1, 0.16$
		$(0.08, 0.5)$	$(0.04, 0.5)$	$(0, 0.5)$	$(-0.1, 0.6)$	$(0.04, 0.5)$	$(0.05, 0.6)$	$(0, 0.6)$	$(-0.02, 0.5)$	$(-0.03, 0.5)$	$(-0.04, 0.5)$
	EWMA	$4, 4$	$9, 9$	$-1.7, 2$	$6, 3$	$-1.2, 3$	$2.67, 3$	$0, 5$	$7, 4$	$18, 10$	$-9, 4$
		$(-0.11, 1.4)$	$(0.1, 1.3)$	$(-0.08, 1.2)$	$(0.05, 1.3)$	$(-0.04, 1.3)$	$(-0.11, 1.4)$	$(0, 1.3)$	$(-0.04, 1.3)$	$(0.17, 1.4)$	$(-0.11, 1.4)$
L.2	POMDP	$1, 0.13$	$1, 0.14$	$1, 0.14$	$1, 0.14$	$1, 0.15$	$1, 0.14$	$1, 0.13$	$1, 0.15$	$1, 0.13$	$1, 0.16$
		$(-0.03, 0.5)$	$(0.07, 0.5)$	$(0.1, 0.5)$	$(0.08, 0.4)$	$(0.06, 0.5)$	$(0.04, 0.4)$	$(0.04, 0.4)$	$(0.03, 0.4)$	$(0.03, 0.4)$	$(0.07, 0.4)$
	EWMA	$-0.7, 1.1$	$-3.9, 2.9$	$3.4, 2.1$	$-2.1, 0.98$	$-1.08, 1.7$	$0.5, 2.1$	$-1, 1.5$	$1.6, 0.94$	$-3.9, 2.5$	$-1.4, 1.4$
		$(0.02, 1.1)$	$(0.03, 1.1)$	$(0.06, 1)$	$(0.03, 1)$	$(0.06, 1.1)$	$(0.03, 1.1)$	$(0.04, 1.1)$	$(0.04, 1)$	$(0.04, 1.1)$	$(0.02, 1.1)$
L.1	POMDP	$1, 0.14$	$0.97, 0.14$	$1, 0.13$	$0.98, 0.13$	$1, 0.13$	$1, 0.13$	$0.97, 0.14$	$1, 0.14$	$0.97, 0.13$	$0.97, 0.13$
		$(-0.4, 0.5)$	$(-0.01, 0.5)$	$(-0.02, 0.5)$	$(0.02, 0.5)$	$(-0.01, 0.5)$	$(-0.03, 0.5)$	$(0, 0.5)$	$(0.01, 0.5)$	$(-0.01, 0.5)$	$(-0.01, 0.5)$
	EWMA	$-0.11, 0.6$	$0.84, 0.5$	$0.27, 0.3$	$-0.82, 0.6$	$-0.3, 0.4$	$-0.15, 0.25$	$-0.01, 0.6$	$0.75, 0.35$	$1.03, 0.44$	$0.35, 0.42$
		$(0.11, 1)$	$(0.1, 1)$	$(0.11, 1)$	$(0.1, 1)$	$(0.1, 1)$	$(0.1, 1)$	$(0.1, 1)$	$(0.1, 1)$	$(0.1, 1)$	$(0.1, 1)$

TABLE IV

MEAN AND STANDARD DEVIATION OF $s_t(k)(\mathbf{u}_t(k))$ FOR EACH OVERLAY FACTOR AND LAYER WITH CROOKED CHANGE WITH METROLOGY DELAY

Layers		Overlay Factors									
		V.1	V.2	V.3	V.4	V.5	V.6	V.7	V.8	V.9	V.10
L.5	POMDP	$1, 0.97$	$1, 0.77$	$1, 0.86$	$1, 0.92$	$1, 0.73$	$1, 0.8$	$1, 1.08$	$1, 0.82$	$1, 1.02$	$1, 0.85$
		$(0.02, 0.16)$	$(0.03, 0.15)$	$(0.03, 0.16)$	$(0.03, 0.16)$	$(0.03, 0.16)$	$(0.03, 0.17)$	$(0.03, 0.16)$	$(0.02, 0.16)$	$(0.03, 0.16)$	$(0.03, 0.17)$
	EWMA	$(689, 81)$	$(689, 81)$	$(689, 81)$	$(689, 81)$	$(689, 81)$	$(689, 81)$	$(689, 81)$	$(689, 81)$	$(689, 81)$	$(689, 81)$
		$-7, 28$	$22, 42$	$-24, 38$	$93, 43$	$4, 44$	$-120, 112$	$148, 106$	$66, 29$	$-29, 44$	$20, 18$
L.4	POMDP	$1, 0.34$	$1, 0.42$	$1, 0.41$	$1, 0.41$	$1, 0.36$	$1, 0.31$	$1, 0.31$	$1, 0.34$	$1, 0.32$	$1, 0.37$
		$(0, 0.007)$	$(0, 0.007)$	$(0, 0.007)$	$(0, 0.007)$	$(0, 0.007)$	$(0, 0.007)$	$(0, 0.007)$	$(0, 0.007)$	$(0, 0.007)$	$(0, 0.007)$
	EWMA	$(535, 57)$	$(535, 52)$	$(535, 53)$	$(535, 54)$	$(535, 55)$	$(535, 53)$	$(535, 54)$	$(535, 55)$	$(535, 54)$	$(535, 55)$
		$-20, 23$	$-4, 8$	$-24, 22$	$-41, 20$	$-17, 35$	$8, 8$	$19, 14$	$-9, 21$	$22, 15$	$-16, 20$
L.3	POMDP	$1, 0.18$	$1, 0.19$	$1, 0.19$	$1, 0.17$	$1, 0.16$	$1, 0.16$	$1, 0.18$	$1, 0.17$	$1, 0.16$	$1, 0.15$
		$(0, 0.001)$	$(0, 0.001)$	$(0, 0.001)$	$(0, 0.001)$	$(0, 0.001)$	$(0, 0.001)$	$(0, 0.001)$	$(0, 0.001)$	$(0, 0.001)$	$(0, 0.001)$
	EWMA	$(382, 68)$	$(382, 69)$	$(382, 69)$	$(382, 69)$	$(382, 69)$	$(382, 69)$	$(382, 69)$	$(382, 69)$	$(382, 69)$	$(382, 69)$
		$3, 6$	$8, 3$	$-1.7, 1$	$-13, 4$	$7, 6$	$2.7, 4$	$-0.9, 3$	$-5.9, 4$	$-3.4, 5$	$6.9, 3$
L.2	POMDP	$1, 0.15$	$1, 0.17$	$1, 0.18$	$1, 0.23$	$1, 0.18$	$1, 0.17$	$1, 0.14$	$1, 0.14$	$1, 0.12$	$1, 0.16$
		$(0, 0.001)$	$(0, 0.0008)$	$(0, 0.0009)$	$(0, 0.0009)$	$(0, 0.001)$	$(0, 0.0007)$	$(0, 0.0006)$	$(0, 0.001)$	$(0, 0.0008)$	$(0, 0.0009)$
	EWMA	$(228, 48)$	$(228, 48)$	$(228, 48)$	$(228, 48)$	$(228, 48)$	$(228, 48)$	$(228, 48)$	$(228, 48)$	$(228, 48)$	$(228, 49)$
		$-0.14, 2$	$0.75, 1$	$4, 3$	$-0.13, 1$	$-2.6, 3$	$2.3, 2$	$-0.07, 1$	$1.19, 1$	$-2.28, 2$	$3.5, 3$
L.1	POMDP	$1, 0.12$	$0.97, 0.12$	$1.02, 0.13$	$1, 0.13$	$1.04, 0.13$	$1, 0.13$	$0.99, 0.13$	$1.02, 0.12$	$1.01, 0.12$	$0.98, 0.17$
		$(0, 0.0004)$	$(0, 0.0004)$	$(0, 0.0004)$	$(0, 0.0004)$	$(0, 0.0004)$	$(0, 0.0004)$	$(0, 0.0004)$	$(0, 0.0004)$	$(0, 0.0004)$	$(0, 0.0004)$
	EWMA	$(75, 44)$	$(75, 39)$	$(75, 45)$	$(75, 43)$	$(76, 45)$	$(73, 46)$	$(77, 43)$	$(74, 44)$	$(74, 39)$	$(75, 45)$
		$-0.18, 0.37$	$-0.36, 0.3$	$1, 0.9$	$0.2, 0.4$	$1.35, 0.7$	$-0.8, 1.04$	$-0.11, 0.4$	$0.8, 0.62$	$0.9, 0.32$	$0.12, 0.19$

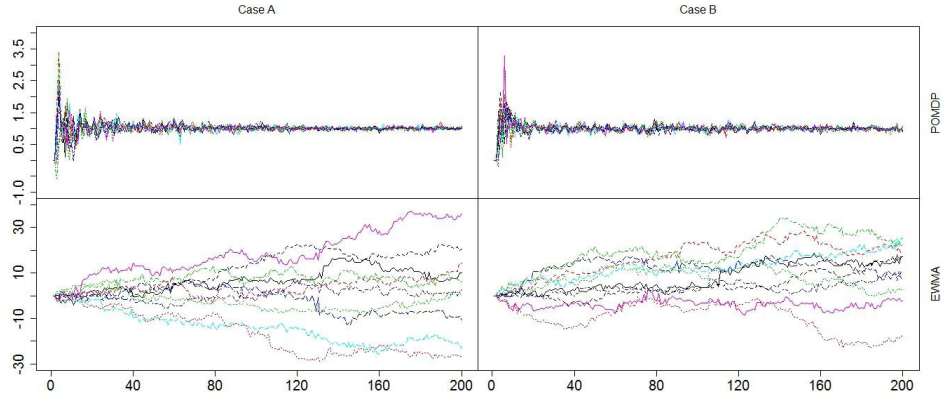

Fig. 8. Effect of disturbance (structural/crooked) comparison on $\mathbf{s}_t(k)$ between POMDP and EWMA controller for third layer, and ten overlay factor (colors are specifying the overlay factors).

8) *Evaluating the performance of γ_t:* In order to test the effect of the weighting parameter, γ_t, on the performance of the POMDP controller and optimization process, $\gamma_t \in (0, 1)$ is tested separately in a simulation experiment. The simulation scenario for testing the effects of γ_t is demonstrated based on the fact that γ_t can

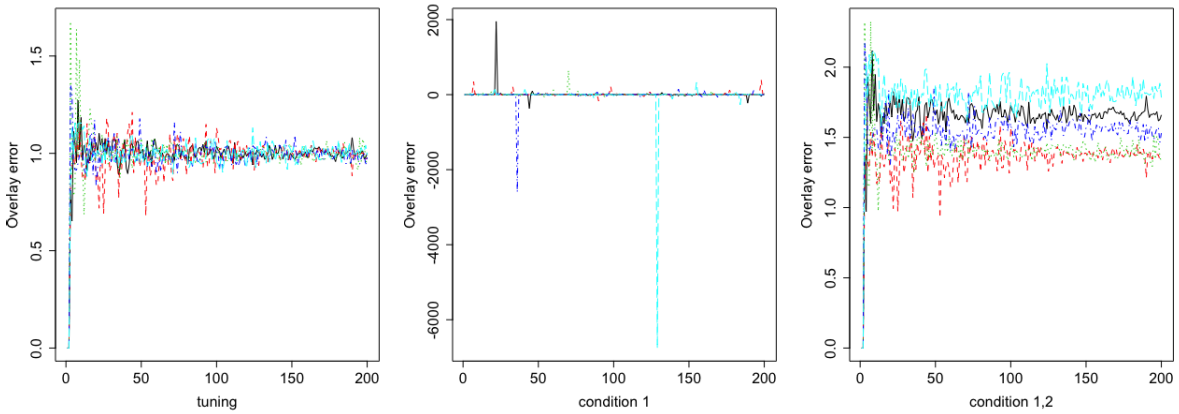

Fig. 9. Effects of tuning algorithm for weighting the γ_t parameter for the crooked change (case A), different colors show different layers.

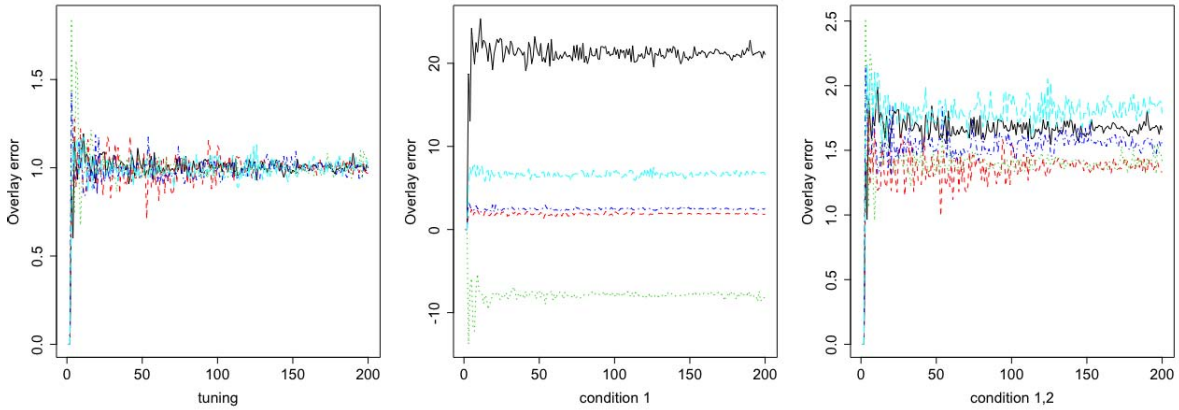

Fig. 10. Effects of tuning algorithm for weighting the γ_t parameter for the structural change (case B), different colors show different layers.

be assigned/predicted for each layer based on expert opinions regards the importance of that layer. Therefore, two additional conditions are considered in previous simulation scenario.

a) $\sum_{k=1}^{m} \gamma_t(k) = 1$.
b) $\gamma_t(k) \propto var(\mathbf{Q}_t(k))$.

The simulation is conducted when the error has either model A or B patterns in Fig. 1, concerning both conditions. Finally, the results are compared with the initial cases in the previous simulation without any conditions on γ_t. Figs. 9 and 10 illustrate the comparison results. The simulation results reveal that the tuning algorithm is the best solution for estimating γ_t. However, if there is an interest in prioritizing layers with higher variation in the real setting, the weighting parameter γ_t should be selected in a direct relationship with the size of overlay error for each specific layer.

9) *Texting controllability and convergence:* In order to evaluate the convergence assumptions in Section III-A, learning rates α_t and α'_t are used to ensure the convergence rate of the proposed POMDP controllers according to the POMDP convergence checking rules in Section III-A. Consider the first and third convergence rules, the value of α_t and α'_t are close to zero and initiated as 0.001 and 0.01, respectively.

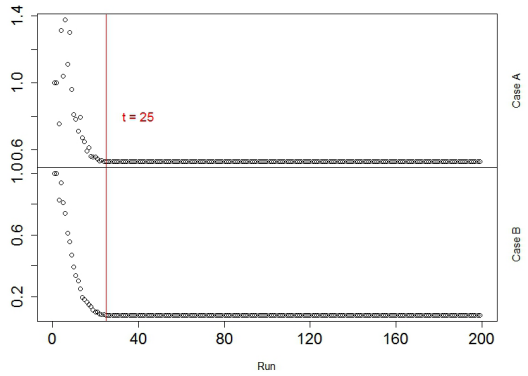

Fig. 11. Learning rate comparison for cases A and B in Fig. 1, between 200 runs and fifth layer.

The result of the learning rate comparison between 200 runs is calculated using

$$|R^*(\mathbf{x}_t(k), \mathbf{u}_t(k)) - R^*(\mathbf{x}_{t-1}(k), \mathbf{u}_{t-1}(k))| < 0.01.$$

The result proves that for both cases, on average, after 25 runs (one lot) with $\varepsilon = 0.01$, the POMDP system reaches the steady-state condition and can reduce the cumulative error (see Fig. 11).

Moreover, as the values of α and α' are close to 0, then the relaxed convergence condition based on (10), $|R^*(\mathbf{x}_t(k), \mathbf{u}_t(k)) - (1-\alpha_t)R^*(\mathbf{x}_{t-1}(k), \mathbf{u}_{t-1}(k))| < \epsilon$, could be approximated as

$$|\alpha_t \times \left[\mathbf{s}_t - \frac{1}{t} E\left\{\sum_t G(\mathbf{s}_t(k))\right\}\right]| < \epsilon. \tag{18}$$

V. Conclusion

The summary of our contributions in this article are as follows.

1) The problem of controlling the state-space model of multilayer overlay error in the photolithography process is drafted as a stochastic optimization problem to minimize the integrated stack-up error and the policy of selecting the state vector.
2) The state-space model is modulated in the Bayesian framework, where the uncertainty parameters are assumed to be stochastic and unknown. For assessing the best parameter setting of the control system without relying on the tuning algorithm and compensating the effect of uncertainties, the Gibbs sampler is used to estimate the probability of the state vector.
3) The POMDP control system is engaged with the optimization goal to handle the drawbacks of the traditional EWMA controller, which provides only a performance measure for response outputs rather than the desired output. The proposed strategy is tested for the MIMO system, and the result is outperformed better than the EWMA controller.
4) POMDP control system is designed to deal with unobservable information due to process delay and measurement delay.
5) The Gibbs sampler algorithm is applied to deal with the difficulties of optimizing the belief function of the POMDP technique.

For future research, applying the same technique for investigating the model for 3D IC is planned out. This approach would be applied to different stacking levels, including wafer-to-wafer, chip-to-wafer, and chip-to-chip.

As POMDPs may require a large state and/or action space, a limitation of our study is related to "curse of dimensionality" and "curse of ambiguity" [45] which would be a fruit future research direction.

The learning-based systems can be applied to many smart manufacturing and Industry 4.0 applications [46], [47] based on partially observable processes. Our proposed POMDP the method can be useful for coping with many practical problems such as dynamic decision-making for both microlevel and macrolevel problems in industry 4.0 and society 5.0 [48], [49], often, decision-makers have incomplete information about the state of the system, and consequently, this situation leads them to apply POMDPs.

References

[1] M. Khakifirooz, C.-F. Chien, and M. Fathi, "Compensating misalignment using dynamic random-effect control system: A case of high-mixed wafer fabrication," *IEEE Trans. Autom. Sci. Eng.*, vol. 16, no. 4, pp. 1788–1799, Oct. 2019.

[2] M. Khakifirooz, C.-F. Chien, and Y.-J. Chen, "Dynamic support vector regression control system for overlay error compensation with stochastic metrology delay," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 1, pp. 502–512, Jan. 2020.

[3] M. Khakifirooz, M. Fathi, and C.-F. Chien, "Modelling and decision support system for intelligent manufacturing: An empirical study for feedforward-feedback learning-based run-to-run controller for semiconductor dry-etching process," *Int. J. Ind. Eng. Theory, Appl. Pract.*, vol. 25, no. 6, pp. 828–842, 2018.

[4] M. Khakifirooz, C. F. Chien, and Y.-J. Chen, "Bayesian inference for mining semiconductor manufacturing big data for yield enhancement and smart production to empower industry 4.0," *Appl. Soft Comput.*, vol. 68, pp. 990–999, Jul. 2018.

[5] C. Garvin, X. Chen, and M. Hankinson, "Advanced process control of overlay with optimal sampling," *Proc. SPIE*, vol. 4689, pp. 817–826, Jul. 2002.

[6] C. A. Bode, B. S. Ko, and T. F. Edgar, "Run-to-run control and performance monitoring of overlay in semiconductor manufacturing," *Control Eng. Pract.*, vol. 12, no. 7, pp. 893–900, Jul. 2004.

[7] M. Khakifirooz, M. Fathi, and P. M. Pardalos, "Disturbance rejection run-to-run controller for semiconductor manufacturing," in *Computational Intelligence and Optimization Methods for Control Engineering*. Cham, Switzerland: Springer, 2019, pp. 301–319.

[8] D. Gkorou *et al.*, "Towards big data visualization for monitoring and diagnostics of high volume semiconductor manufacturing," in *Proc. Comput. Frontiers Conf.*, 2017, pp. 338–342.

[9] M.-D. Ma, X.-L. Zeng, and G.-R. Duan, "A survey of run-to-run control algorithms for high-mix semiconductor manufacturing processes," in *Proc. 30th Chin. Control Conf.*, 2011, pp. 5474–5479.

[10] M. Khakifirooz, C.-F. Chien, M. Fathi, and P. M. Pardalos, "Minimax optimization for recipe management in high-mixed semiconductor lithography process," *IEEE Trans. Ind. Informat.*, vol. 16, no. 8, pp. 4975–4985, Aug. 2020.

[11] M. Janakiram and S. Goernitz, "Real-time lithography registration, exposure, and focus control—A framework for success," *IEEE Trans. Semicond. Manuf.*, vol. 18, no. 4, pp. 534–538, Nov. 2005.

[12] C.-F. Chien, Y.-J. Chen, C.-Y. Hsu, and H.-K. Wang, "Overlay error compensation using advanced process control with dynamically adjusted proportional-integral R2R controller," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 2, pp. 473–484, Apr. 2014.

[13] C. Ausschnitt *et al.*, "Multilayer overlay metrology," *Proc. SPIE*, vol. 6152, Mar. 2006, Art. no. 615210.

[14] A. W. Topol *et al.*, "Three-dimensional integrated circuits," *IBM J. Res. Develop.*, vol. 50, nos. 4–5, pp. 491–506, Jul./Sep. 2006.

[15] T. H. Conway, M. Misra, A. P. Carlson, and D. A. Crow, "Effect of overlay APC control on cascading levels: Perturbations of the reference level," *Proc. SPIE*, vol. 5038, pp. 1002–1011, Jun. 2003.

[16] J. Yu and S. J. Qin, "Variance component analysis based fault diagnosis of multi-layer overlay lithography processes," *IIE Trans.*, vol. 41, no. 9, pp. 764–775, Jul. 2009.

[17] Y. Jiao and D. Djurdjanovic, "Stochastic control of multilayer overlay in lithography processes," *IEEE Trans. Semicond. Manuf.*, vol. 24, no. 3, pp. 404–417, Aug. 2011.

[18] F. He and Z. Zhang, "An empirical study-based state space model for multilayer overlay errors in the step-scan lithography process," *RSC Adv.*, vol. 5, no. 126, pp. 103901–103906, 2015.

[19] F. He and Z. Zhang, "State space model and numerical simulation of overlay error for multilayer overlay lithography processes," in *Proc. 2nd Int. Conf. Image, Vis. Comput. (ICIVC)*, Jun. 2017, pp. 1123–1127.

[20] J. Moyne, E. Del Castillo, and A. M. Hurwitz, *Run-to-Run Control in Semiconductor Manufacturing*. Boca Raton, FL, USA: CRC Press, 2000.

[21] A. Ingolfsson and E. Sachs, "Stability and sensitivity of an EWMA controller," *J. Qual. Technol.*, vol. 25, no. 4, pp. 271–287, Oct. 1993.

[22] R. Ganesan, T. K. Das, and K. M. Ramachandran, "A multiresolution analysis-assisted reinforcement learning approach to run-by-run control," *IEEE Trans. Autom. Sci. Eng.*, vol. 4, no. 2, pp. 182–193, Apr. 2007.
P. R. Thie, *Markov Decision Processes*. Comap, 1983.

[23] M. T. Spaan, "Partially observable Markov decision processes," in *Reinforcement Learning*. Berlin, Germany: Springer, 2012, pp. 387–414.

[24] G. E. Monahan, "State of the art—A survey of partially observable Markov decision processes: Theory, models, and algorithms," *Manage. Sci.*, vol. 28, no. 1, pp. 1–16, Jan. 1982.

[25] A. Wilson, A. Fern, S. Ray, and P. Tadepalli, "Multi-task reinforcement learning: A hierarchical Bayesian approach," in *Proc. 24th Int. Conf. Mach. Learn. (ICML)*, 2007, pp. 1015–1022.
[26] C.-F. Chien, K.-H. Chang, and C.-P. Chen, "Design of a sampling strategy for measuring and compensating for overlay errors in semiconductor manufacturing," *Int. J. Prod. Res.*, vol. 41, no. 11, pp. 2547–2561, Jan. 2003.
[27] C.-F. Chien and C.-Y. Hsu, "UNISON analysis to model and reduce step-and-scan overlay errors for semiconductor manufacturing," *J. Intell. Manuf.*, vol. 22, no. 3, pp. 399–412, Jun. 2011.
[28] S.-P. Lee, "EWMA controller with concurrent adjustment for a high-mixed production process," *Adv. Technol. Innov.*, vol. 2, no. 1, pp. 25–28, 2017.
[29] N. N. Patwardhan, "Evaluation and extension of threaded control for high-mix semiconductor manufacturing," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Texas Austin, Austin, TX, USA, 2010.
[30] A. R. Cassandra, "Exact and approximate algorithms for partially observable Markov decision processes," Ph.D. dissertation, Dept. Comput. Sci., Brown Univ., Providence, RI, USA, 1998.
[31] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1, no. 2. Belmont, MA, USA: Athena Scientific, 1995.
[32] H. Kushner and P. G. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, vol. 24. New York, NY, USA: Springer, 2013.
[33] R. Bellman, "The theory of dynamic programming," *Bull. Amer. Math. Soc.*, vol. 60, no. 6, pp. 503–515, 1954.
[34] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artif. Intell.*, vol. 101, nos. 1–2, pp. 99–134, May 1998.
[35] D. Calandriello, A. Lazaric, and M. Restelli, "Sparse multi-task reinforcement learning," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 819–827.
[36] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Mach. Learn.*, vol. 50, no. 1, pp. 5–43, Jan. 2003.
[37] C. K. Carter and R. Kohn, "On Gibbs sampling for state space models," *Biometrika*, vol. 81, no. 3, pp. 541–553, 1994.
[38] C.-J. Kim and C. R. Nelson, *State-Space Models With Regime Switching: Classical and Gibbs-Sampling Approaches With Applications*, vol. 1. Cambridge, MA, USA: MIT Press, 1999.
[39] M. Plummer, "JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling," in *Proc. 3rd Int. Workshop Distrib. Stat. Comput.*, Vienna, Austria, vol. 124, 2003, pp. 1–10.
[40] W. R. Gilks and P. Wild, "Adaptive rejection sampling for Gibbs sampling," *Appl. Statist.*, vol. 41, no. 2, pp. 337–348, 1992.
[41] G. García-Donato and M. A. Martínez-Beneito, "On sampling strategies in Bayesian variable selection problems with large model spaces," *J. Amer. Stat. Assoc.*, vol. 108, no. 501, pp. 340–352, Mar. 2013.
[42] M. Hauskrecht, "Value-function approximations for partially observable Markov decision processes," *J. Artif. Intell. Res.*, vol. 13, pp. 33–94, Aug. 2000.
[43] D. Lambert, "Zero-inflated Poisson regression, with an application to defects in manufacturing," *Technometrics*, vol. 34, no. 1, pp. 1–14, 1992.
[44] N. S. Patel and S. T. Jenkins, "Adaptive optimization of run-to-run controllers: The EWMA example," *IEEE Trans. Semicond. Manuf.*, vol. 13, no. 1, pp. 97–107, Feb. 2000.
[45] S. Saghafian, "Ambiguous partially observable Markov decision processes: Structural results and applications," *J. Econ. Theory*, vol. 178, pp. 1–35, Nov. 2018.
[46] M. Khakifirooz, M. Fathi, C. F. Chien, and P. M. Pardalos, "Management suggestions for process control of semiconductor manufacturing: An operations research and data science perspective," in *Computational Intelligence and Optimization Methods for Control Engineering*. Cham, Switzerland: Springer, 2019, pp. 245–274.
[47] M. Khakifirooz, M. Fathi, and K. Wu, "Development of smart semiconductor manufacturing: Operations research and data science perspectives," *IEEE Access*, vol. 7, pp. 108419–108430, 2019.
[48] J. M. Velásquez-Bermúdez, M. Khakifirooz, and M. Fathi, *Large Scale Optimization in Supply Chains and Smart Manufacturing: Theory and Applications*, vol. 149. Cham, Switzerland: Springer, 2019.
[49] M. Fathi, M. Khakifirooz, and P. M. Pardalos, *Optimization in Large Scale Problems: Industry 4.0 and Society 5.0 Applications*. Cham, Switzerland: Springer, 2019.