EM Estimation of Nanostructure Interactions With Incomplete Feature Measurement and Its Tailored Space Filling Designs

Lijuan Xu and Qiang Huang, *Member, IEEE*

***Abstract*—Automatic assessment of nanostructure quality is essential for scale-up nanomanufacturing. In our previous work, we have developed a method to quantify nanostructure growth quality and detect structural defects through interaction analysis. However, because the method builds on complete feature measurement, its direct application to nanomanufacturing systems is severely constrained by nanostructure metrology. For current inspection techniques such as scanning electron microscope (SEM), the major difficulties of measuring nanostructures lie in two aspects: (i) taking and calibrating images for seamless coverage and (ii) extracting and matching feature information from the images. In this paper, we develop a tailored sampling strategy to relax the metrology constraint. It not only explores the growth region with greatly reduced metrology efforts but maintains desired sampling resolution. In addition, we customize Expectation-Maximization algorithm to optimize interaction estimation with corresponding "incomplete" measurement. Our developed approach enables nanostructure characterization within manufacturing relevant time spans and thus provides a supporting tool for nanomanufacturing.**

***Note to Practitioners*—Automatic assessment of nanostructure quality is essential for scale-up nanomanufacturing, but current characterization of nanostructures based on SEM or transmission electron microscopy (TEM) is labor and computation intensive. This paper develops methods to quantify nanostructure local variability and detect defects with minimum metrology efforts.**

***Index Terms*—Defect detection, expectation-maximization (EM), incomplete feature measurement, nanomanufacturing, nanostructure interactions, nanostructure metrology, space filling design.**

I. Introduction

AUTOMATIC assessment of nanostructure quality within manufacturing relevant time spans is essential for scale-up nanomanufacturing. Due to the critical roles played by nanostructure interactions such as competing for source materials during growth process, assessing nanostructures through interaction analysis has proved to be efficient [1]. By extracting interactions, we could not only quantify nanostructure feature variability but also detect structural defects and identify their patterns.

Manuscript received May 30, 2012; revised November 24, 2012; accepted January 29, 2013. Date of publication March 07, 2013; date of current version June 27, 2013. This paper was recommended for publication by Associate Editor M. K. Jeong and Editor K. Bohringer upon evaluation of the reviewers' comments. This work was supported by the National Science Foundation under Grant CMMI-1000972.

The authors are with the Daniel J. Epstein Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, CA 90089 USA (e-mail: qiang.huang@usc.edu).

Digital Object Identifier 10.1109/TASE.2013.2245120

The major constraint for extending our previous approach in [1] to scale-up nanomanufacturing systems is nanostructure metrology. On one hand, because the analysis algorithm was built on complete feature measurement, mapping interaction patterns across the whole substrate would demand complete information of all local regions. On the other hand, characterizing nanostructure features is extremely labor intensive and time consuming under current inspection techniques [1]–[3]. To characterize dense quantities of nano elements under manufacturing-relevant time spans, it is necessary to have revolutionary instead of evolutionary advances. But "no known solutions" are available for 5–10 years down the road [4].

A customized sampling strategy is thus urgently needed to relax the metrology constraint but still maintain the desired characterization resolution. At the same time, methods to extract nanostructure interactions also have to be updated based on the corresponding sampled feature measurement to ensure accurate structure assessment within manufacturing relevant time spans.

In this paper, we will analyze the limitations of current nanostructure inspection techniques in detail and propose in accordance a two-step sampling strategy to maximize the exploration of the growth region and better support subsequent interaction analysis. Moreover, we will customize Expectation-Maximization (EM) algorithm to optimize nanostructure interaction estimation based on corresponding "incomplete" feature measurement. With the developed approaches, we are able to quantify nanostructure feature variation and detect structural defects with affordable metrology efforts and within desirable time spans. Our method provides basis for online quality monitoring and control based on image inspection for scale-up nanomanufacturing systems.

In what follows, Section II presents the overall strategies of nanostructure sampling and interaction analysis. Section III discusses the modeling and estimation of nanostructure interactions with "incomplete" feature measurement. Section IV develops a tailored space filling design that further enhances interaction estimation efficiency and enables further reduction on metrology. Sections V and VI provide simulation and real case studies to validate proposed approaches. In the end, Section VII gives a summary.

II. Overall Strategy

Our objective in this paper is to quantify nanostructure variability and detect defects with minimum metrology efforts and within manufacturing relevant time spans to better support nanomanufacturing scale up. In view of the great potentials

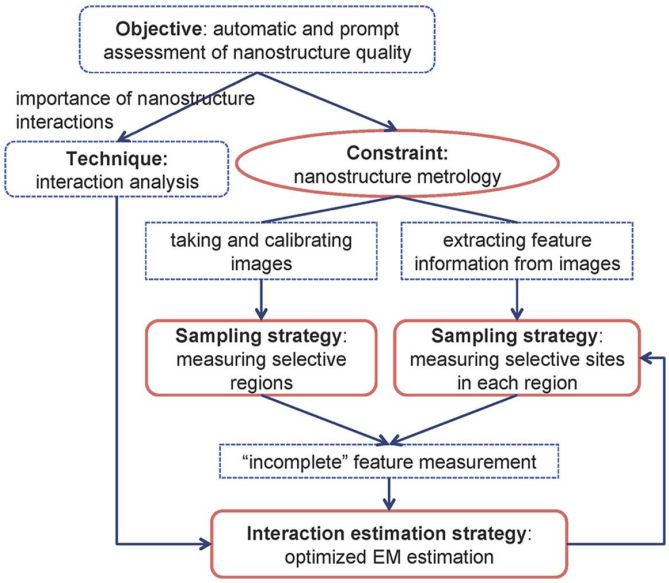

Fig. 1. Overall strategies for automatic and prompt quality assessment of nanostructures under metrology constraint. Modules in solid red are research tasks in this paper.

of interaction analysis in [1] and difficulties of nanostructure metrology, we develop a two-step *nanostructure sampling strategy* as well as an *interaction estimation strategy* to reduce feature measurement and speedup interaction analysis. The overall framework is in Fig. 1.

Sampling Strategy: For current inspection techniques such as scanning electron microscope (SEM), the key difficulties of measuring nanostructures lie in two aspects: 1) taking and calibrating images for seamless coverage and 2) extracting and matching feature information from the images. For instance, due to limited coverage of SEM, we may need to carefully take and calibrate over 200 images of neighboring areas to cover a 0.1 mm $\times$ 0.1 mm local region. Additionally, it is also computation intensive to extract feature information from the two dimensional images. For example, to measure nanowire lengths, we need to match SEM images taken from different angles and find the same nanowire on each image among the nanowire forest [1]–[3].

Accordingly, to reduce metrology efforts but maintain the desired interaction analysis resolution, we take following strategy to sample and measure nanostructures.

Step 1) We selectively measure multiple separated tiny regions spreading on the substrate.

Step 2) For each selected region, we divide it regularly into non-overlapping sites (grids) and only measure nanostructures in selected sites.

The above approach is explained in Fig. 1 and illustrated in Fig. 2.

To better explore the whole growth region, standard space filling designs such as *maximin distance Latin hypercube designs* are applied to select the measurement regions [blue squares in Fig. 2(a)], and for each selected region, we will develop a *tailored space filling design* in Section IV to choose which sites to measure [e.g., white grids in Fig. 2(b)]. By selectively measuring partial of the sites, we can not only lighten the

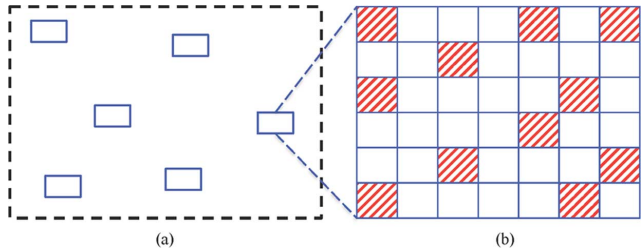

Fig. 2. Nanostructure sampling strategies. (a) Measurement regions (blue squares) across the substrate. (b) Nonmeasurement sites (patterned grids) in each measurement region.

burden of feature extraction but also maintain an appropriate interaction analysis resolution. Moreover, our space filling design minimizes sampling biases and supports subsequent interaction analysis (Fig. 1).

We call those sites to be measured "measurement sites" and those not to measure "nonmeasurement sites." The resulting feature measurement is denoted as "incomplete" feature measurement in this work.

Although we take a two-step procedure to sample nanostructures, our strategy is different from the sequential or nested designs in literature [5]–[10] for two reasons. First, regions and sites are all determined before the real measurement takes place. There is no sequential evaluation. Second, our dataset consists only feature measurement in the sites. There is no multiple levels of sampling or approximation.

In addition, Latin hypercube based space filling designs [11] are not applicable to select nonmeasurement sites for selected regions, because we have both preestablished division of sites driven by desired resolution of interaction analysis and predetermined number of sites to measure due to metrology constraints.

Interaction Estimation Strategy: In accordance with the "incomplete" feature measurement, we also update interaction estimation method to ensure accurate and prompt quality assessment of nanostructures. For the new estimation technique, we would like it to possess following key features. (i) It is capable to deal with "missing values." (ii) It can be easily fitted within our interaction modeling framework. (iii) it accurately estimates nanostructure interactions within desirable time spans. In this work, we optimize EM algorithm to estimate nanostructure interactions. We will also customize the selection of measurement sites to further enhance interaction estimation efficiency.

By integrating our developed sampling strategy and interaction estimation strategy, we are capable of characterizing nanostructures for scale-up nanomanufacturing systems. Details are discussed in following sections.

III. Modeling and Estimation of Nanostructure Interactions

Notation for measurement regions, sites, and nanostructure features is as following.

- R is the total number of measurement regions.
- l is the index of a measurement region. $l \in \{1, 2, \ldots, R\}$.
- n is the total number of sites in a measurement region.
- i or j is the index of a site in a measurement region. $i, j \in \{1, 2, \ldots, n\}$.

- $\boldsymbol{s}_i^{(l)}$ represents the spatial location of site i in measurement region l.
- $\boldsymbol{r}_l = \{\boldsymbol{s}_1^{(l)}, \boldsymbol{s}_2^{(l)}, \ldots, \boldsymbol{s}_n^{(l)}\}$ denotes the complete collection of sites in region l.
- $\boldsymbol{m}_l$ represents the collection of measurement sites in region l.
- $\boldsymbol{o}_l = \boldsymbol{r}_l \backslash \boldsymbol{m}_l$ represents the collection of nonmeasurement (omitted) sites in region l.
- $\boldsymbol{A}_l$ denotes the matrix that maps the complete collection of sites to measured sites in region l, i.e., $\boldsymbol{m}_l = \boldsymbol{A}_l \boldsymbol{r}_l$.
- $X(\boldsymbol{s}_i^{(l)})$ represents nanostructure feature at site $\boldsymbol{s}_i^{(l)}$.
- $\boldsymbol{X}(\boldsymbol{r}_l) = \{X(\boldsymbol{s}_1^{(l)}), X(\boldsymbol{s}_2^{(l)}), \ldots, X(\boldsymbol{s}_n^{(l)})\}$ is the collection of nanostructure features at region l.
- $\boldsymbol{X}(\boldsymbol{m}_l)$ is feature field for measurement sites in region l. It is also called measured feature field.
- $\boldsymbol{X}(\boldsymbol{o}_l) = \boldsymbol{X}(\boldsymbol{r}_l) \backslash \boldsymbol{X}(\boldsymbol{m}_l)$ is feature field for nonmeasurement sites in region l.

Although here we assume different measurement regions are of the same size (i.e., every region contains n sites), the modeling and estimation techniques developed in this paper apply in general cases. In addition, the equal size assumption is not hard to satisfy in practice since we normally determine the region sizes before doing the real measurement.

A. Stationary Modeling of Nanostructure Interactions for Each Measurement Region

In this paper, we adopt the stationary interaction modeling developed in [1] to model the interactions within each measurement region and assume different regions share the same set of spatial parameters. In particular, we assume the feature measurement in each region l can be decomposed into local trend and local variability as following:

$$\boldsymbol{X}(\boldsymbol{r}_l) = \boldsymbol{Z}(\boldsymbol{r}_l)\boldsymbol{\gamma} + \boldsymbol{\Phi}(\boldsymbol{r}_l). \tag{1}$$

The local trend is modeled as a linear combination of explanatory variables $\boldsymbol{Z}$ that can be spatial location functions or concomitant data with each site [12]. For local variability $\boldsymbol{\Phi}(\boldsymbol{r}_l)$, we use stationary Gaussian Markov random field to describe interactions among neighboring sites in each region

$$\Phi\left(\boldsymbol{s}_i^{(l)}\right) \left| \left\{\phi\left(\boldsymbol{s}_j^{(l)}\right)\right\}_{j \neq i} \sim N\left(\sum \beta_{ij} \phi\left(\boldsymbol{s}_j^{(l)}\right),\ \sigma^2\right). \tag{2}$$

Here, $\{\beta_{ij}\}$ are spatial coefficients that depend soly on the relative locations between sites $\boldsymbol{s}_i^{(l)}$ and $\boldsymbol{s}_j^{(l)}$ and are zero unless the two sites are neighbors of each other. σ^2 is the conditional variance of local variability at site $\boldsymbol{s}_i^{(l)}$ given its neighbors.

In this stationary modeling, we assume the local trend of each measurement region has the same functional form and shares the same set of regression parameters $\boldsymbol{\gamma}$. It implies that the global trend for the whole substrate is a smooth surface described by $\boldsymbol{Z}\boldsymbol{\gamma}$. If measurement data objects this assumption, it means we have a more complex trend profile due to large process variations. In this case, one possible extension of the trend modeling here is to take a different set of regression parameters $\boldsymbol{\gamma}_l$ for each region and connect $\boldsymbol{\gamma}_l$ with the process conditions such as temperature T. Interested readers may refer to Huang [13] for the connection.

In addition, our stationary interaction modeling also assumes the interaction parameter β_{ij} between any two sites i and j in region l only depends on their relative locations. That is, we conjecture homogeneous and nonisotropic interaction patterns across the substrate. When the local trend is correctly captured, violation of the stationary interaction assumption here indicates the existence of defects within measured regions.

By following similar derivations in [1], we can obtain the distribution of (measured) nanostructure feature field for each region l. Specifically, we have

$$\boldsymbol{X}(\boldsymbol{r}_l) \sim N(\boldsymbol{Z}(\boldsymbol{r}_l)\boldsymbol{\gamma}, \sigma^2(\boldsymbol{I} - \boldsymbol{B})^{-1}) \tag{3}$$

where $\boldsymbol{B} = (\beta_{ij})_{i,j=1,\ldots,n}$ is a block circulant matrix composed by interaction parameters $\{\beta_{ij}\}$. Additionally, as $\boldsymbol{A}_l$ maps complete collection of sites in region l to measured sites, i.e., $\boldsymbol{m}_l = \boldsymbol{A}_l \boldsymbol{r}_l$, the distribution of measured feature field is easily obtained as following:

$$\boldsymbol{X}(\boldsymbol{m}_l) \sim N\left(\boldsymbol{A}_l \boldsymbol{Z}(\boldsymbol{r}_l)\boldsymbol{\gamma}, \sigma^2 \boldsymbol{A}_l (\boldsymbol{I} - \boldsymbol{B})^{-1} \boldsymbol{A}_l^T\right). \tag{4}$$

Since different regions are purposely selected to be tiny and fairly separated, we may reasonably assume (measured) feature fields in different regions are independent.

Remark: We follow similar framework as in [1] to quantify nanostructure feature variation and detect structural defects. The focus of this paper is thus on interaction analysis under metrology constraints. In particular, we discuss in detail the sampling strategy of nanostructures and the stationary interaction estimation based on the incomplete feature measurement. We will also briefly mention the detection of defects in a real case study but leave nonstationary interaction analysis to interested readers. The extension of stationary interaction analysis to nonstationary interaction analysis will be easy by using local likelihood methods [1], [14].

B. EM Estimation of Nanostructure Interactions With Incomplete Feature Measurement

Denote $\boldsymbol{\eta} = \{\boldsymbol{\gamma}, \boldsymbol{\beta}, \sigma^2\}$ as the collection of unknown parameters in the interaction modeling. Estimating nanostructure interactions is then to estimate values of $\boldsymbol{\eta}$ based on the incomplete feature measurement from R independent measurement regions. That is

$$\hat{\boldsymbol{\eta}} = \arg\max_{\boldsymbol{\eta}} \log f\left(\{\boldsymbol{x}(\boldsymbol{m}_l)\}_{l=1}^R | \boldsymbol{\eta}\right). \tag{5}$$

The estimation difficulty mainly comes from the "incompleteness." The nonsquare mapping matrix $\boldsymbol{A}_l$ in (4) makes Maximum-Likelihood Estimation (MLE) [1], [12], [15]–[17] computationally infeasible.

We use Expectation-Maximization (EM) algorithm to estimate nanostructure interactions. Those omitted sites (i.e., nonmeasurement sites) are treated as "missing at random" observations, though here we omit them purposely. Comparing to other techniques dealing with missing values such as imputation, interpolation and Bayesian data analysis [18], EM algorithm takes advantages of the Markovian property in GMRF modeling of nanostructure interactions in (2) and the analytical property of likelihood function for complete feature field.

The EM algorithm is an iterative method for finding the maximum likelihood estimates of $\boldsymbol{\eta}$. Instead of maximizing directly the log likelihood for measured feature field (what MLE does), the EM algorithm iteratively calculates (**E-Step**) and maximizes (**M-Step**) the conditional expectation of the log likelihood for complete feature field based on the measured feature field and the current estimate for parameters $\boldsymbol{\eta}$ at step p (i.e., $\boldsymbol{\eta}^p$) [19], [20]. In particular, we are to calculate and optimize the following:

$$\begin{aligned} Q(\boldsymbol{\eta}|\boldsymbol{\eta}^p) &= \mathbb{E}\left\{\log f\left(\{\boldsymbol{X}(\boldsymbol{r}_l)\}_{l=1}^R|\boldsymbol{\eta}\right)|\{\boldsymbol{x}(\boldsymbol{m}_l)\}_{l=1}^R, \boldsymbol{\eta}^p\right\} \\ &= \sum_{l=1}^{R} \mathbb{E}\{\log f(\boldsymbol{X}(\boldsymbol{r}_l)|\boldsymbol{\eta})|\boldsymbol{x}(\boldsymbol{m}_l), \boldsymbol{\eta}^p\} \end{aligned} \tag{6}$$

where the decomposition is made possible because of the independence assumption for different measurement regions.

Since the calculation in (6) is similar for each measurement region, we only demonstrate EM approach for *one single* region. To keep the notation simpler, we denote $\boldsymbol{Z}$ for $\boldsymbol{Z}(\boldsymbol{r}_l)$ and omit any subscript l in the notation for following discussions.

In addition, we introduce the following notation for more convenient presentation of the conditional statistics.

- $\mathbb{E}_{\text{cond}}(\cdot) = \mathbb{E}\{\cdot|\boldsymbol{x}(\boldsymbol{m}), \boldsymbol{\eta}^p\}$ represents the conditional expectation given measured feature field $\boldsymbol{x}(\boldsymbol{m})$ and current parameter estimate $\boldsymbol{\eta}^p$.
- $\text{Var}_{\text{cond}}(\cdot) = \text{Var}\{\cdot|\boldsymbol{x}(\boldsymbol{m}), \boldsymbol{\eta}^p\}$ denotes the conditional variance.
- $\text{Cov}_{\text{cond}}(\cdot,\cdot) = \text{Cov}\{\cdot,\cdot|\boldsymbol{x}(\boldsymbol{m}), \boldsymbol{\eta}^p\}$ represents the conditional covariance.

1) Expectation Step of EM Estimation: The complete feature field $\boldsymbol{X}(\boldsymbol{r})$ follows a multivariate normal distribution as given in (3). We can thus easily obtain the conditional expectation of its log-likelihood function as follows:

$$\begin{aligned} Q(\boldsymbol{\eta}|\boldsymbol{\eta}^p) &= \mathbb{E}\{\log f(\boldsymbol{X}(\boldsymbol{r})|\boldsymbol{\eta})|\boldsymbol{x}(\boldsymbol{m}), \boldsymbol{\eta}^p\} \\ &= -\frac{\mathbb{E}_{\text{cond}}\{(\boldsymbol{X}(\boldsymbol{r}) - \boldsymbol{Z}\boldsymbol{\gamma})^T(\boldsymbol{I} - \boldsymbol{B})(\boldsymbol{X}(\boldsymbol{r}) - \boldsymbol{Z}\boldsymbol{\gamma})\}}{2\sigma^2} \\ &\quad - \frac{n}{2}\log(2\pi\sigma^2) + \frac{1}{2}\log(|\boldsymbol{I} - \boldsymbol{B}|) \end{aligned} \tag{7}$$

and the conditional expectation in (7) can be expanded as

$$\begin{aligned} &\mathbb{E}_{\text{cond}}\{(\boldsymbol{X}(\boldsymbol{r}) - \boldsymbol{Z}\boldsymbol{\gamma})^T(\boldsymbol{I} - \boldsymbol{B})(\boldsymbol{X}(\boldsymbol{r}) - \boldsymbol{Z}\boldsymbol{\gamma})\} \\ &\quad = (\mathbb{E}_{\text{cond}}\boldsymbol{X}(\boldsymbol{r}) - \boldsymbol{Z}\boldsymbol{\gamma})^T(\boldsymbol{I} - \boldsymbol{B})(\mathbb{E}_{\text{cond}}\boldsymbol{X}(\boldsymbol{r}) - \boldsymbol{Z}\boldsymbol{\gamma}) \\ &\qquad + \text{Tr}\{(\boldsymbol{I} - \boldsymbol{B})\,\text{Var}_{\text{cond}}\boldsymbol{X}(\boldsymbol{r})\}. \end{aligned} \tag{8}$$

Therefore, to evaluate $Q(\cdot|\cdot)$, keys are to compute the *conditional expectation* for any nonmeasurement site and the *conditional covariance* between any two nonmeasurement sites in (8).

To speedup the calculation, we take advantages of the Markovian property in our interaction modeling in (2). Specifically, Markovian property allows us to classify nonmeasurement sites into conditionally independent groups. In this way, the conditional distribution of nonmeasurement feature field can be decomposed. Mathematically, we have

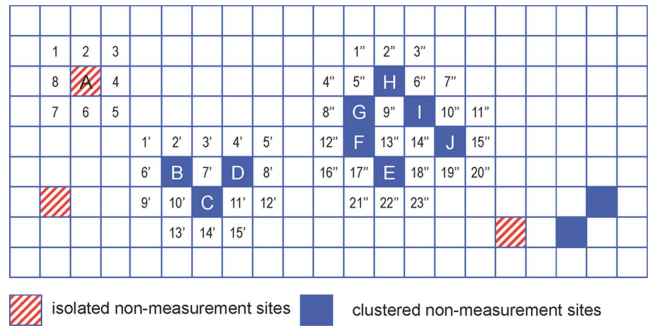

Fig. 3. Schematic illustration of nonmeasurement site classification.

$$\begin{aligned} f(\boldsymbol{X}(\boldsymbol{o})|\boldsymbol{x}(\boldsymbol{m}), \boldsymbol{\eta}^p) &= \prod_k f(\boldsymbol{X}(\boldsymbol{g}_k)|\boldsymbol{x}(\boldsymbol{m}), \boldsymbol{\eta}^p) \\ &= \prod_k f(\boldsymbol{X}(\boldsymbol{g}_k)|\boldsymbol{x}(\boldsymbol{n}_k), \boldsymbol{\eta}^p) \end{aligned} \tag{9}$$

where $\{\boldsymbol{g}_k\}$ represents conditionally independent groups of nonmeasurement sites and $\boldsymbol{n}_k$ denotes the neighboring measurement sites for group k. Consequently, we may obtain the conditional expectations and covariances by evaluating the conditional distribution for each group individually.

The algorithm for classifying nonmeasurement sites is as follows. (i) If a nonmeasurement site has all measured neighbors, e.g., site A in Fig. 3, we call it *isolated nonmeasurement site*. Given measured feature field $\boldsymbol{x}(\boldsymbol{m})$, each "isolated nonmeasurement site" is conditionally independent from any other nonmeasurement sites. (ii) If a nonmeasurement site also has nonmeasurement sites as neighbors, e.g., site B in Fig. 3, we call it *clustered nonmeasurement site*. We further group "clustered nonmeasurement sites" into smallest *clusters* such that a site in one cluster is not the neighbor of any site in other clusters. For instance, in Fig. 3, sites B, C, and D form a cluster and sites E-J form another cluster under second-order neighborhood structure where neighboring sites are the eight nearest sites such as sites 1–8 for site A.

It is easy to prove, groups constructed this way are conditionally independent. In addition, the conditional distributions of each group will only depend on their neighboring measurement sites such as sites $1''$–$23''$ for the cluster of E–J in Fig. 3.

Upon the classification of nonmeasurement sites, the calculation of conditional expectations and covariances are quite routine for each group. In particular, for isolated nonmeasurement sites, we obtain their conditional statistics directly from the interaction modeling in (2). That is

$$\begin{aligned} \mathbb{E}_{\text{cond}}X(\boldsymbol{s}_i) &= (\boldsymbol{Z}\boldsymbol{\gamma}^p)_i + \sum \beta_{ij}^p(x(\boldsymbol{s}_j) - (\boldsymbol{Z}\boldsymbol{\gamma}^p)_j) \\ \text{Var}_{\text{cond}}X(\boldsymbol{s}_i) &= (\sigma^2)^p \end{aligned} \tag{10}$$

and for each cluster of nonmeasurement sites, the calculation follows standard derivations for multivariate normal random variables. Specifically, we have

$$\mathbb{E}_{\text{cond}}\boldsymbol{X}(\boldsymbol{g}_k) = \mathbb{E}(\boldsymbol{X}(\boldsymbol{g}_k)|\boldsymbol{x}(\boldsymbol{n}_k), \boldsymbol{\eta}^p)$$
$$\text{Var}_{\text{cond}}\boldsymbol{X}(\boldsymbol{g}_k) = \text{Var}(\boldsymbol{X}(\boldsymbol{g}_k)|\boldsymbol{x}(\boldsymbol{n}_k), \boldsymbol{\eta}^p) \quad (11)$$

Equation (11) clearly shows that the interaction estimation is mostly affected by those *large clusters of nonmeasurement sites*. To efficiently estimate nanostructure interactions, it is thus desirable to have nonmeasurement sites in smaller groups. This is exactly the motivation to develop tailored space filling designs in Section IV.

2) Maximization Step of EM Estimation: Maximizing $Q(\boldsymbol{\eta}|\boldsymbol{\eta}^p)$ based on its evaluation in Section III-B1 is a quite standard procedure. Denote $\boldsymbol{\eta}^{p+1}$ to be the EM estimates at step $p+1$. Then, we have

$$\boldsymbol{\eta}^{p+1} = \arg\max_{\boldsymbol{\eta}} Q(\boldsymbol{\eta}|\boldsymbol{\eta}^p) \quad (12)$$

which is easily obtained by setting $\partial Q(\boldsymbol{\eta}|\boldsymbol{\eta}^p)/\partial\boldsymbol{\eta} = \mathbf{0}$.

We iteratively perform the Expectation and Maximization steps described above until EM estimates $\boldsymbol{\eta}^p$ converge with certain tolerance. The converged values $\tilde{\boldsymbol{\eta}}$ will be our final parameter estimation.

IV. Tailored Space Filling Design for Site Selection in Each Region

This section aims to develop a sampling approach for selecting nonmeasurement sites in each measurement region with given nonmeasurement ratio (number of nonmeasurement sites to total number of sites). Our objective is to optimize both the accuracy and the efficiency of nanostructure interaction estimation.

A problem with simple random sampling is that with moderately large nonmeasurement ratios it often produces clustering and poorly covered area. The associated consequences include: 1) nanostructures measured may not be representative over the whole region, as a result of which the sampling brings extra biases to interaction estimation; and 2) EM estimation of interactions are not computation efficient due to those large clusters of dependent nonmeasurement sites. While classical space-filling designs including distance-based designs [21], [22] and uniform designs [23] may well solve problem (1), they have no control in clustering of dependent nonmeasurement sites. Fig. 4 shows a sample maximin distance design with nine points on a 5×5 griding. All nonmeasurement sites (patterned sites) are conditionally dependent under diamond neighborhood structure [Fig. 4(b)] which is usually evaluated for interaction estimation. Consequently, all nonmeasurement sites form a single cluster making EM algorithm inefficient.

To optimize interaction estimation efficiency, we develop a tailored space filling design for selecting nonmeasurement sites.

Design criteria are as follows with descending priorities.

- C.1 Size of the largest cluster of dependent nonmeasurement sites, *smaller the better*.
- C.2 Number of isolated nonmeasurement sites, *larger the better*.

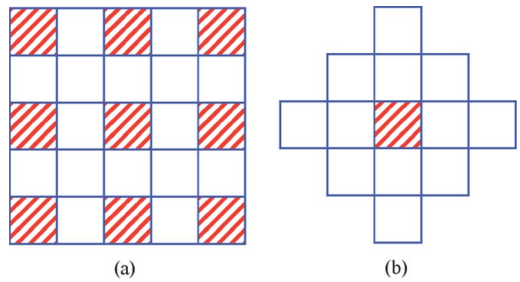

Fig. 4. Sample maximin distance design that produces *single* conditionally dependent group of nonmeasurement sites under diamond neighborhood structure.

- C.3 Number of conditionally independent groups, *larger the better*.

It is intuitive that with smaller clusters, more isolated nonmeasurement sites and more conditionally independent groups, we have more "space filling" nonmeasurement sites.

We use simulated annealing [22], [24] which is an iterative optimization method to search the best space filling design with given nonmeasurement ratio. Four components are contained in the algorithm which are: 1) generation of initial designs; 2) perturbation scheme; 3) updating criteria; and 4) termination conditions.

Initial Design Generation: When nonmeasurement ratio is low (e.g., $< 15\%$), we use *simple random sampling* to generate the initial design. When nonmeasurement ratio is high, we either inherit *optimized designs of lower nonmeasurement ratios* or use *stratified sampling* to select sites from conditionally independent strata.

Perturbation Scheme: At each step, we perturb the current design $\mathbb{D}_{\text{current}}$ by randomly replacing a nonmeasurement site in the largest cluster with an existing measurement site. Denote the new design as $\mathbb{D}_{\text{try}}$, the newly added nonmeasurement site as $moveToSite$ and the replaced design point as $moveFromSite$. Instead of summarizing the new design $\mathbb{D}_{\text{try}}$ from scratch, we reclassify only those nonmeasurement sites affected by the movement. The reclassification algorithm is in Algorithm 1.

Updating Criteria: For each perturbation, we will either keep the current design unchanged or update it with the newly generated one. The probability of updating is

$$p(\mathbb{D}_{\text{try}}, \mathbb{D}_{\text{current}}) = \min\left\{1, \exp\left(\frac{g(\mathbb{D}_{\text{try}}) - g(\mathbb{D}_{\text{current}})}{T}\right)\right\} \quad (13)$$

where $g(\cdot) = -a \times \text{largestClusterSize} + b \times \text{numIsoSites} + c \times \text{numClusters}$ with $a > b > c$ being positive constants. T is the temperature of annealing that decreases along iterations. It is clear a better design based on criteria **C**. 1–3 has larger chances to be accepted. Besides, we put most efforts to control the largest cluster size.

Termination Conditions: The optimization process stops when either of the following two is met: (i) the largest cluster size in current design is smaller than a predefined value or (ii) there is no design update in certain number of consequent iterations.

When nonmeasurement ratio increases, nonmeasurement sites form large clusters more easily. As a consequence, more efforts are needed to obtain the space filling design. However, as a design can be used for any dataset with compatible dimensions, time to generate the sampling strategy is not a big concern.

V. Simulation Case Studies

In this section, we will demonstrate and validate the effectiveness of our proposed approaches by simulation case studies.

Algorithm 1: Reclassification of nonmeasurement sites

if moveToSite is an isolated nonmeasurement site **then**
  isoSites in $\mathbb{D}_{\text{try}}$ = isoSites in $\mathbb{D}_{\text{current}}$ + moveToSite
  # The origCluster may decompose after the replacement
  classiSites= otherSitesInOrigCluster in $\mathbb{D}_{\text{current}}$
else
  # Find nonmeasurement neighbors of moveToSite
  nonMNeighs=nonmeasurement neighbors of moveToSite in $\mathbb{D}_{\text{try}}$
  if ALL nonMNeighs belong to origCluster in $\mathbb{D}_{\text{current}}$ **then**
    classiSites= otherSitesInOrigCluster+ moveToSite
  else if NONE of nonMNeighs belong to origCluster **then**
    nonNeigGroups= groups nonMNeighs belong to in $\mathbb{D}_{\text{current}}$
    # moveToSite joins nonNeigGroups together
    aNewCluster= moveToSite + sitesInNonNeigGroups
    classiSites= otherSitesInOrigCluster
  else
    # Some belong to origCluster, some don't
    nonNeigGroups= groups nonMNeighs belong to in $\mathbb{D}_{\text{current}}$
    otherSites= sites in nonNeigGroups but not in origCluster
    classiSites= otherSitesInOrigCluster+ moveToSite + otherSites
  end if
end if
Classify classiSites
Update $\mathbb{D}_{\text{try}}$ accordingly

A. Multiple Tiny Regions Versus One Large Region

For the same amount of data collection, if we distribute it in multiple separated regions instead of a single large region we can reduce SEM calibration since calibration is only needed for neighboring images. In this subsection, we will validate the interaction estimation accuracy is maintained by this multiple region approach as well.

We simulated 100 i.i.d. datasets of size 50×100 to mimic replicates of feature measurement from large continuous regions. At the same time, we simulated 100 i.i.d. data collections each of which contains eight independent 25×25 datasets to represent replicates of feature measurement obtained in eight separated small regions. Structure of simulated datasets is depicted in Fig. 5. All datasets are GMRF with linear trend and second-order interaction patterns. Specifically, for any site $\boldsymbol{s}_i = (s,t)$ where s and t are row and column indexes in the datasets, we assume

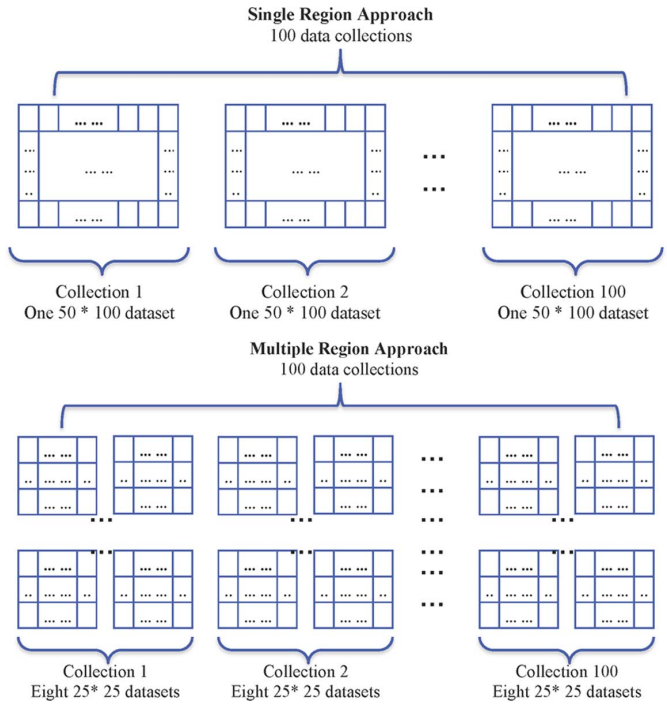

Fig. 5. Schematic illustration of simulated data structures.

$$\mathbb{E}\{X(s,t)\} = \gamma_0 + \gamma_1 * s + \gamma_2 * t \tag{14}$$

$$\begin{aligned}\mathbb{E}\Phi(s,t)|\{\phi(s',t')\}_{(s',t')\neq(s,t)} &= \beta_1\{\phi(s-1,t-1)+\phi(s+1,t+1)\} \\ &\quad + \beta_2\{\phi(s-1,t)+\phi(s+1,t)\} \\ &\quad + \beta_3\{\phi(s-1,t+1)+\phi(s+1,t-1)\} \\ &\quad + \beta_4\{\phi(s,t-1)+\phi(s,t+1)\}\end{aligned} \tag{15}$$

$$\text{Var}\ \Phi(s,t)|\{\phi(s',t')\}_{(s',t')\neq(s,t)} = \sigma^2. \tag{16}$$

Values set for simulating the datasets are summarized in Table I under the column True Val.

Parameter estimation results for multiple region approach and single large region approach are summarized and compared in Table I. Results show both approaches give excellent quantification of feature variability. Sampling multiple regions indeed reduces SEM calibration and achieves as good interaction estimation accuracy.

B. EM Estimation of Nanostructure Interactions With Incomplete Measurement

We use the 100 i.i.d 50×100 datasets simulated in Section V-A as baseline and study the EM estimation of interactions with different nonmeasurement ratios in this sub-

TABLE I
MLE ESTIMATION OF NANOSTRUCTURE INTERACTIONS BASED ON SIMULATED DATASETS

		Multiple Region		Single Region	
	True Val.	Est. Mean	95% C.I.	Est. Mean	95% C.I.
γ_0	6.00	6.010	[5.888, 6.158]	6.001	[5.838, 6.142]
γ_1	-0.03	-0.031	[-0.037, -0.026]	-0.030	[-0.033, -0.027]
γ_2	-0.01	-0.010	[-0.018, -0.003]	-0.010	[-0.012, -0.008]
β_1	-0.10	-0.099	[-0.129, -0.078]	-0.101	[-0.127, -0.080]
β_2	0.25	0.250	[0.223, 0.269]	0.251	[0.230, 0.270]
β_3	0.05	0.049	[0.027, 0.074]	0.048	[0.030, 0.069]
β_4	0.15	0.150	[0.125, 0.175]	0.150	[0.129, 0.182]
σ^2	1.00	0.997	[0.956, 1.037]	0.996	[0.952, 1.038]

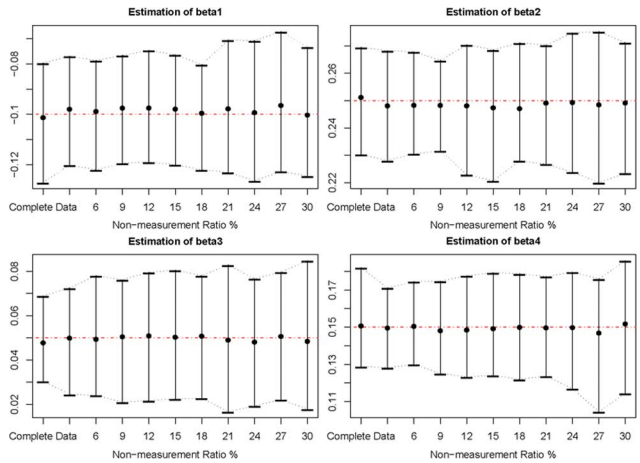

Fig. 6. EM estimation of interactions for simulated datasets under different nonmeasurement ratios. ▬ ▬: True Val., ●: Est. Mean, |: 95% empirical C.I. based on 100 datasets.

section. In particular, we simulate ten extra data collections by adding 3%–30% missing values in the complete datasets. Nonmeasurement sites in these data collections are generated by the *simple random sampling* approach.

We summarize EM estimated means and 95% empirical confidence intervals for interaction parameters $\beta_1, \ldots, \beta_4$ in Fig. 6. Results for each nonmeasurement ratio are calculated based on 100 i.i.d. datasets with corresponding amount of "missed" measurement. MLEs of $\{\beta_i\}$ for complete datasets are also depicted for comparison. It shows that we can reduce more than 30% nanostructure feature measurement but still achieve comparable quantification of nanostructure variations.

Besides estimation accuracy, we also investigate the dependence of EM estimation speed on nonmeasurement ratios. From Fig. 7, we can see both averages and standard deviations of estimation time increase monotonically with nonmeasurement ratios. This rapid increase roots in those large clusters of dependent nonmeasurement sites generated by *simple random sampling*. Fig. 8 validates our claim by depicting EM estimation time with the largest cluster size under its highest evaluated interaction structure. Empirically, estimation time increases at least quadratically with maximum cluster sizes since we have to evaluate the covariance matrix for each cluster.

With *simple random sampling*, there are even cases where all nonmeasurement sites are clustered together, e.g., the case with maximum cluster size $50 \times 100 \times 30\% = 1500$. Therefore, to enhance interaction estimation efficiency so as to further reduce metrology efforts, the key is to develop a new sampling

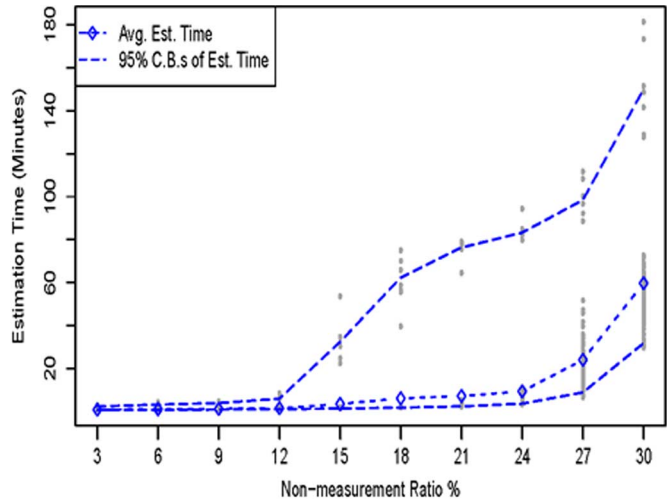

Fig. 7. EM estimation time with nonmeasurement ratios (simple random sampling).

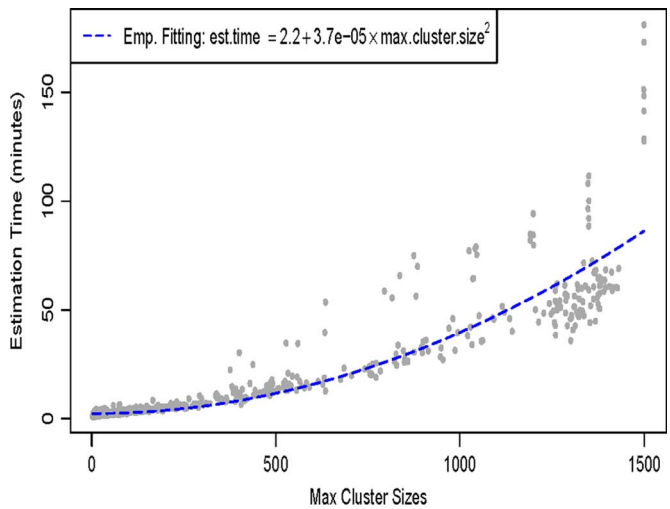

Fig. 8. EM estimation time with maximum cluster sizes (simple random sampling).

approach that controls the largest cluster size. This is exactly what we do in Section IV. We will examine the effectiveness of our tailored space filling design in Section V-C.

C. Tailored Space Filling Design to Select Nonmeasurement Sites in Each Region

We simulate ten data collections with 3%–30% nonmeasurement sites similar to those in Section V-B. The only difference is that nonmeasurement sites here are selected by our *tailored space filling design* developed in Section IV instead of *simple random sampling*. With our new approach, we obtain the largest cluster sizes in Fig. 9(a) for each dataset under their highest evaluated interaction structures. Comparing to simple random sampling, our tailored space filling design successfully precludes large clusters. As a result, both averages and standard deviations of EM estimation time are greatly reduced [Fig. 9(b) and (c)].

Summarizing simulation case studies, we conclude our approaches of sampling nanostructures and analyzing nanostructure interactions indeed greatly reduce the metrology

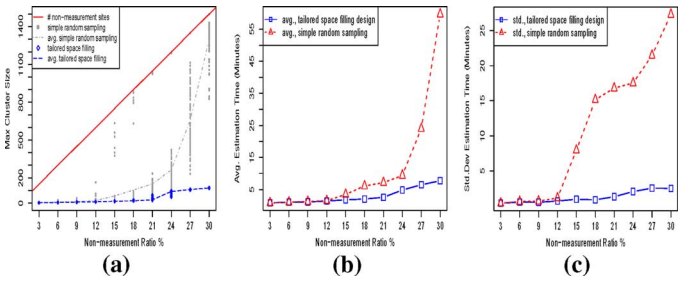

Fig. 9. Comparison of *tailored space filling sampling* and simple random sampling of nonmeasurement sites: (a) max cluster sizes, (b) avg. est. time, and (c) std. est. time.

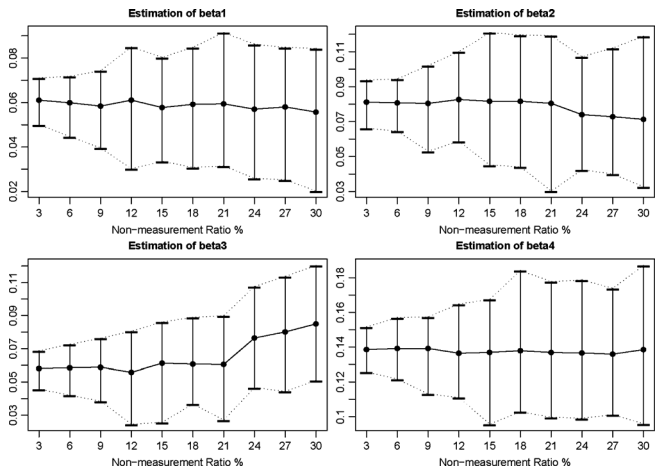

Fig. 10. EM estimation of interactions for datasets generated from real nanowire length data. •: Est. Mean, |: 95% empirical C.I. based on 100 datasets.

efforts and achieve excellent interaction estimation accuracy and efficiency.

VI. Real Case Studies

In this section, we will study a 45×94 real collected ZnO nanowire length dataset that includes measurement for both normally grown nanowires and nanowire bundles (one type of structural defects). Interested readers may find more details in [1] about the dataset.

Our primary interest is to investigate whether and how much incomplete length measurement affects interaction analysis and nanowire bundle detection. To this end, we design ten levels of incompleteness ranging from 3% to 30% by adding "missing observations" in the original length data. These "missing observations" are used to mimic the sites we neglect during real measurement. In addition, under each level of incompleteness, we generate 100 replicates of datasets each of which has the same amount of "missing observations" but different missing locations based on distinctive space filling designs.

We use optimized EM algorithm (Section III-B) to extract interaction patterns from each dataset. Since most of the datasets identify second-order interaction structure [(15)], we summarize estimation results for interaction parameters $\beta_1, \ldots, \beta_4$ in Fig. 10.

Intuitively, we do not see dramatic change in mean estimation of interaction parameters with different levels of incompleteness except for β_3 when nonmeasurement ratio exceeds 21%. However, empirical confidence intervals are generally wider with larger nonmeasurement ratios for all β_i, $i = 1, \ldots, 4$. Comparing to simulated stationary datasets (Fig. 6), unstable confidence intervals here indicate larger variations among the 100 datasets with different missing observations, which may *qualitatively* implies the existence of defects.

To *quantitatively* detect defects, we do normality tests on transformed residuals of measured length field similarly as for complete data case [1]. Specifically, under the null hypothesis that there are no defects within the field, interaction patterns are stationary across the field and thus measured feature field follows a normal distribution as stated in (4). The mapping matrix $\boldsymbol{A}_l$ is known for each measurement region l and the parameters are estimated as $\tilde{\boldsymbol{\eta}}$ by the EM algorithm. Based on the consistency property of EM estimates, we then have the transformed residuals follow i.i.d. normal distribution asymptotically. That is

$$\left\{\tilde{\boldsymbol{e}}_l = \left[\boldsymbol{A}_l(\boldsymbol{I} - \tilde{\boldsymbol{B}})\boldsymbol{A}_l^T\right]^{-\frac{1}{2}} \left[\boldsymbol{X}(\boldsymbol{m}_l) - \boldsymbol{A}_l \boldsymbol{Z}(\boldsymbol{r}_l)\tilde{\boldsymbol{\gamma}}\right]\right\}_{l=1}^{R} \sim \text{i.i.d. } N(0, \sigma^2) \text{ asymptotically.} \quad (17)$$

Therefore, we may perform Anderson–Darling tests on transformed residuals $\{\tilde{\boldsymbol{e}}_l\}$ to determine whether there are structural defects [1], [25].

With the normality test, all datasets here containing up to 30% missing observations are detected to have defects. The defect detection result is 100% correct.

Real case studies in this section show that our new method of interaction analysis guarantees to reduce metrology efforts more than 30% and at the same time provides comparable interaction estimation accuracy and competitive defect detection precision. It thus provides a supporting tool for scale-up nanomanufacturing.

VII. Conclusion

In this paper, we focus on the relaxation of metrology constraints for nanostructure interaction analysis to quantify nanostructure growth quality and detect structural defects in scale-up nanomanufacturing systems.

Targeting on the limitations of current nanostructure inspection techniques, we develop a two-step sampling strategy to reduce the metrology efforts. Specifically, we sample multiple tiny regions over the substrate and develop tailored space filling designs to selectively measure partial of the sites in each measurement region. This sampling strategy not only maximizes the exploration of the growth region but also supports subsequent interaction analysis. In addition, we optimize the EM algorithm to estimate nanostructure interactions based on corresponding "incomplete" feature measurement. By classifying nonmeasurement sites into conditionally independent groups, we further enhance the efficiency of interaction estimation.

Both simulation and real case studies show that our methods provide accurate and prompt nanostructure assessment with significantly reduced metrology efforts. The developed sampling and interaction estimation strategies facilitate online quality

monitoring and control in scale-up nanomanufacturing systems. They serve as a supporting tool for future nanomanufacturing.

ACKNOWLEDGMENT

The authors would like to thank Dr. C. Zhou's group at the University of Southern California (USC) for providing ZnO nanowire samples.

REFERENCES

[1] L. Xu and Q. Huang, "Modeling the interactions among neighboring nanostructures for local feature characterization and defect detection," *IEEE Trans. Autom. Sci. Eng.*, vol. 9, no. 4, pp. 745–754, Oct. 2012.
[2] F. Wang, Y. Hwang, P. Qian, and X. Wang, "A statistics-guided approach to precise characterization of nanowire morphology," *ACS Nano.*, vol. 4, no. 2, pp. 855–862, 2010.
[3] F. Bao and J. Xuan, "Recovery of nanowire morphology and distribution by a computer vision-assisted approach," in *Proc. IEEE Int. Conf. Wavelet Analysis and Pattern Recognition (ICWAPR'12)*, 2012, pp. 38–44.
[4] M. Postek and R. Hocken, Instrumentation and metrology for nanotechnology: Report of the National Nanotechnology Initiative Workshop, Jan. 27–29 2004, National Nanotechnology Coordination Office, Monograph (Tech. Rep.), 2006.
[5] B. Husslage, E. Van Dam, and D. Den Hertog, "Nested maximin Latin hypercube designs in two dimensions," Center Discussion Paper No. 2005–79, 2005.
[6] P. Qian, B. Tang, and C. J. Wu, "Nested space-filling designs for computer experiments with two levels of accuracy," *Statistica Sinica*, vol. 19, no. 1, pp. 287–300, 2009.
[7] P. Qian, "Nested Latin hypercube designs," *Biometrika*, vol. 96, no. 4, pp. 957–970, 2009.
[8] P. Qian, M. Ai, and C. Wu, "Construction of nested space-filling designs," *The Ann. Statist.*, vol. 37, no. 6A, pp. 3616–3643, 2009.
[9] E. van Dam, B. Husslage, and D. den Hertog, "One-dimensional nested maximin designs," *J. Global Optim.*, vol. 46, no. 2, pp. 287–306, 2010.
[10] G. Rennen, B. Husslage, E. Van Dam, and D. Den Hertog, "Nested maximin latin hypercube designs," *Structural and Multidisciplinary Optim.*, vol. 41, no. 3, pp. 371–395, 2010.
[11] M. McKay, R. Beckman, and W. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, pp. 239–245, 1979.
[12] N. Cressie, *Statistics for Spatial Data*, 2nd ed. New York, NU, USA: Wiley, 1993.
[13] Q. Huang, "Physics-driven Bayesian hierarchical modeling of the nanowire growth process at each scale," *IIE Trans.*, vol. 43, no. 1, pp. 1–11, 2011.
[14] G. Kauermann and G. Tutz, "Local likelihood estimation in varying-coefficient models including additive bias correction," *J. Nonparametric Statist.*, vol. 12, no. 3, pp. 343–371, 2000.
[15] P. Whittle, "On stationary processes in the plane," *Biometrika*, vol. 41, no. 3, pp. 434–449, 1954.
[16] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Royal Statist. Soc.*, ser. B (Methodological), vol. 36, no. 2, pp. 192–236, 1974.
[17] X. Guyon, "Parameter estimation for a stationary process on a d-dimensional lattice," *Biometrika*, vol. 69, no. 1, p. 95, 1982.
[18] R. Little and D. Rubin, *Statistical Analysis With Missing Data*. New York, NY, USA: Wiley, 1987.
[19] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *J. Royal Statist. Soc.*, ser. Series B (Methodological), vol. 39, no. 1, pp. 1–38, 1977.
[20] C. Wu, "On the convergence properties of the em algorithm," *Ann. Statist.*, pp. 95–103, 1983.
[21] M. Johnson, L. Moore, and D. Ylvisaker, "Minimax and maximin distance designs," *J. Statist. Planning Inference*, vol. 26, no. 2, pp. 131–148, 1990.
[22] M. Morris and T. Mitchell, "Exploratory designs for computational experiments," *J. Statist. Planning Inference*, vol. 43, no. 3, pp. 381–402, 1995.
[23] K. Fang, D. Lin, P. Winker, and Y. Zhang, "Uniform design: Theory and application," *Technometrics*, pp. 237–248, 2000.
[24] P. Laarhoven and E. Aarts, *Simulated Annealing: Theory and Applications*. New York, NY, USA: Springer, 1987, vol. 37.
[25] H. Thode, *Testing for Normality*. New York, NY, USA: CRC Press, 2002.

Lijuan Xu received the B.S. degree from the Department of Industrial Engineering, Tsinghua University, Beijing, China, in 2009 and the M.S. degree in statistics from the Department of Mathematics, University of Southern California, Los Angeles, CA, USA, in 2011. She is currently a Ph.D. candidate at the Daniel J. Epstein Department of Industrial and Systems Engineering, University of Southern California.

Her general research interests include statistical modeling, diagnosis and control of dynamic systems/networks, applied statistics, and data mining.

Ms. Xu is a member of the Institute for Operations Research and the Management Sciences (INFORMS).

Qiang Huang (M'10) received the Ph.D. degree in industrial and operations engineering from the University of Michigan, Ann Arbor, MI, USA, in 2003 .

He is currently an Associate Professor and the Gordon S. Marshall Early Career Chair in Engineering in the Daniel J. Epstein Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, CA, USA. Funded by the National Science Foundation (including the CAREER Award) and ONR, his research focuses on modeling and analysis of complex systems for quality and productivity improvement, with special interest in integrated nanomanufacturing and nanoinformatics, and additive manufacturing.

Prof. Huang is a member of the Institute of Industrial Engineers (IIE), the Institute for Operations Research and the Management Sciences (INFORMS), and the American Society of Mechanical Engineers (ASME). He has been an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING since 2012 and has been a member of the scientific committee (Editorial Board) for the North American Manufacturing Research Institution (NAMRI) of SME, 2009–2011 and 2013–2015. He served as Editor (Quality, Micro and Nanomanufacturing Systems) for the *Journal of Manufacturing Systems* from 2008 to 2011. He was one of the editors for the Special Issue of the *IIE Transactions:* Quality, Sensing and Prognostics Issues in Nanomanufacturing.