

# Statistical Modeling and Analysis for Robust Synthesis of Nanostructures

Tirthankar Dasgupta<sup>†</sup>, Christopher Ma<sup>‡</sup>, V. Roshan Joseph<sup>†</sup>,  
Z. L. Wang<sup>‡</sup>, C. F. Jeff Wu<sup>†\*</sup>

<sup>†</sup>*School of Industrial and Systems Engineering,*

<sup>‡</sup>*School of Materials Science and Engineering ,*

*Georgia Institute of Technology, Atlanta, GA*

*\*Corresponding author*

## Abstract

An effort is made to systematically investigate the best process conditions that ensures synthesis of different types of one dimensional cadmium selenide nanostructures with high yield and reproducibility. Through a designed experiment and rigorous statistical analysis of experimental data, models linking the probabilities of obtaining specific morphologies to the process variables are developed. A new iterative algorithm for fitting a multinomial GLM is proposed and used. The optimum process conditions, which maximize the above probabilities and make the synthesis process robust (i.e., less sensitive) to variations of process variables around set values, are derived from the fitted models using Monte-Carlo simulations.

Cadmium Selenide (CdSe) has been found to exhibit one-dimensional morphologies of nanowires, nanobelts and nanosaws, often with the three morphologies being

intimately intermingled within the as-deposited material. A slight change in growth condition can result in a totally different morphology. In order to identify the optimal process conditions that maximize the yield of each type of nanostructure and, at the same time, make the synthesis process robust (i.e., less sensitive) to variations of process variables around set values, a large number of trials were conducted with varying process conditions. Here, the response is a vector whose elements correspond to the numbers of appearance of different types of nanostructures. The fitted statistical models would enable nano-manufacturers to identify the probability of transition from one nanostructure to another when changes, even tiny ones, are made in one or more process variables. Inferential methods associated with the modeling procedure help in judging the relative impact of the process variables and their interactions on the growth of different nanostructures. Owing to the presence of internal noise, i.e., variation around the set value, each predictor variable is a random variable. Using Monte-Carlo simulations, the mean and variance of transformed probabilities are expressed as functions of the set points of the predictor variables. The mean is then maximized to find the optimum nominal values of the process variables, with the constraint that the variance is under control.

KEY WORDS: Nanotechnology, Statistical modeling, Robust Design, Cadmium Selenide nanostructures, Multinomial, Generalized Liner Model.

## 1 Introduction

Nanotechnology is the construction and use of functional structures designed from atomic or molecular scale with at least one characteristic dimension measured in nanometers (one nanometer =  $10^{-9}$  meter, which is about 1/50,000 of the width of human hair). The size of these nanostructures allows them to exhibit novel and significantly improved physical, chemical, and biological properties, phenomena, and processes. Nanotechnology can provide unprecedented understanding about materials and devices and is likely to impact many fields. By using structure at nanoscale as a tunable physical variable, scientists can greatly expand the range of performance of existing chemicals and materials. Alignment of linear molecules

in an ordered array on a substrate surface (self-assembled monolayers) can function as a new generation of chemical and biological sensors. Switching devices and functional units at nanoscale can improve computer storage and operation capacity by a factor of a million. Entirely new biological sensors facilitate early diagnostics and disease prevention of cancers. Nanostructured ceramics and metals have greatly improved mechanical properties, both in ductility and strength.

Current research by nanoscientists typically focuses on novelty, discovering new growth phenomena and new morphologies. However, within the next five years there will likely be a shift in the nanotechnology community towards controlled and large-scale synthesis with high yield and reproducibility. This transition from laboratory-level synthesis to large scale, controlled and designed synthesis of nanostructures necessarily demands systematic investigation of the manufacturing conditions under which the desired nanostructures are synthesized *reproducibly*, in *large* quantity and with *controlled* or isolated morphology. Application of statistical techniques can play a key role in achieving these objectives. This article reports a systematic study on the growth of 1D CdSe nanostructures through statistical modeling and optimization of the experimental parameters required for synthesizing desired nanostructures. This work is based on the experimental data presented in this paper and research published in Ma and Wang (2005). Some general statistical issues and research opportunities related to the synthesis of nanostructures are discussed in the concluding section.

Cadmium selenide (CdSe) has been investigated over the past decade for applications in optoelectronics (Hodes, Albu-Yaron, Decker and Motisuke 1987), luminescent materials (Bawendi, Kortan, Steigerwald and Brus 1989), lasing materials (Ma, Ding, Moore, Wang and Wang 2004) and biomedical imaging. It is the most extensively studied quantum-dot material and is therefore regarded as the model system for investigating a wide range of nanoscale processes. CdSe is found to exhibit one-dimensional morphologies of nanowires, nanobelts and nanosaws (Ma and Wang 2005), often with the three morphologies being intimately intermingled within the as-deposited material. Images of these three nanostructures obtained using scanning electron microscope are shown in Figure 1.

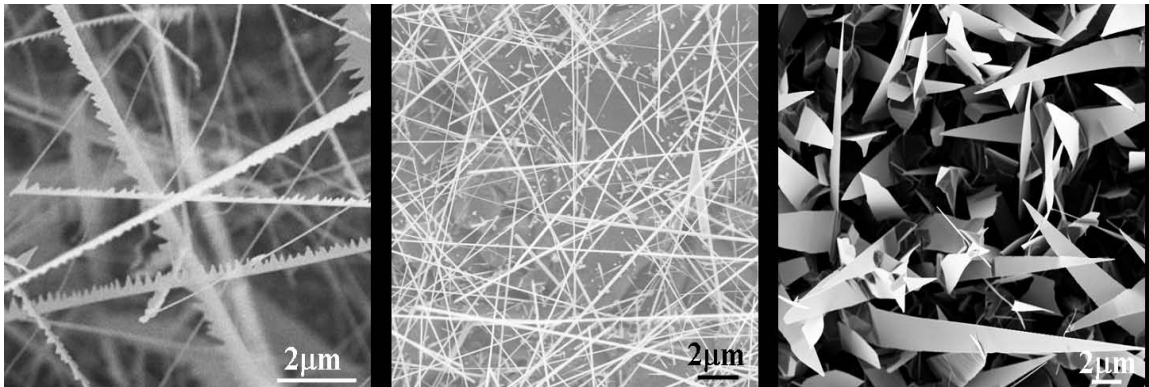


Figure 1: SEM images of nanostructures (from the left: nanosaws, nanowires, nanobelts)

In this experiment, the response is a vector whose elements correspond to the numbers of appearance of different types of nanostructures and hence is a multinomial random variable. Thus a multinomial generalized linear model (GLM) is the appropriate tool for analyzing the experimental data and expressing the multinomial logits as functions of the predictor variables (McCullagh and Nelder 1989; Faraway 2006). A new iterative algorithm for fitting multinomial GLM that has certain advantages over the existing methods is proposed and implemented. The probability of obtaining each nanostructure is expressed as a function of the predictor variables. Owing to the presence of inner noise, i.e., variation around the set value, each predictor variable is a random variable. Using Monte-Carlo simulations, the expectation and variance of transformed probabilities are expressed as functions of the set points of the predictor variables. The expectation is then maximized to find the optimum set values of the process variables, ensuring at the same time that the variance is under control. The idea is thus similar to the two-step robust parameter design for larger-the-better responses (Wu and Hamada 2000, chap. 10).

The article is organized as follows. In Section 2, we give a brief account of the synthesis process, the experimental design and collection of data. Section 3 is devoted to fitting of appropriate statistical models to the experimental data. This section consists of two subsections. In Section 3.1 a preliminary analysis using a binomial GLM is shown. Estimates of the parameters obtained here are used as initial estimates in the iterative algorithm for multinomial GLM, which is developed and described in Section 3.2. In Section 4, we study the

optimization of the process variables to maximize the expected yield of each nanostructure. Some general statistical issues and challenges in nanostructure synthesis and opportunities for future research are discussed in Section 5.

## 2 The synthesis process, design of experiment and data collection

The CdSe nanostructures were synthesized (see Figure 2) through a thermal evaporation process in a single zone horizontal tube furnace (Thermolyne 79300). A 30-inch polycrystalline  $Al_2O_3$  tube (99.9% purity) with an inner diameter of 1.5 inches was placed inside the furnace. Commercial grade CdSe (Alfa Aesar, 99.995% purity, metal basis) was placed at the center of the tube as use for a source material. Single-crystal silicon substrates with a 2-nanometer thermally evaporated non-continuous layer of gold were placed downstream of the source in order to collect the deposition of the CdSe nanostructures. The system was held at the set temperature and pressure for a period of 60 minutes and cooled to room temperature afterwards. The as-deposited products were characterized and analyzed by scanning electron microscopy (SEM) (LEO 1530 FEG), transmission electron microscopy (TEM) (Hitachi HF-2000 FEG at 200 kV). As many as 180 individual nanostructures were counted from the deposition on each substrate.

The two key process variables affecting morphology of CdSe nanostructures are temperature and pressure. A  $5 \times 9$  full factorial experiment was conducted with five levels of source temperature (630, 700, 750, 800, 850<sup>0</sup> C) and nine levels of pressure (4, 100, 200, 300, 400, 500, 600, 700, 800 mbar). For a specific combination of source temperature and pressure, 4-6 substrates were placed downstream of the source to collect the deposition of nanostructures. The distance of the mid-point of the substrate from the source was measured and treated as a covariate.

Three experimental runs were conducted with each of the 45 combinations of temperature and pressure. However, these three runs cannot be considered to be replicates, since the number and location of substrates were not the same in the three runs. Consider, for

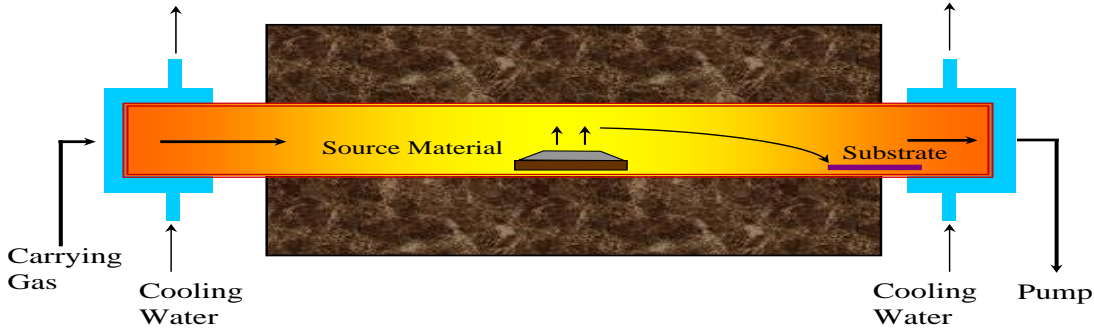


Figure 2: Synthesis Process

example, the three runs performed with a temperature of  $630^{\circ}\text{C}$  and pressure of 4 mb. In the first run, six substrates were placed at distances of 1.9, 4.2, 4.9, 6.4, 8.1, 10.2 cm from the source. In the second run, four substrates were placed at distances of 1.7, 4.6, 7.1, 8.9 cm from the source. Seven substrates were placed at distances of 2.0, 4.3, 4.9, 6.4, 8.5, 10.6, 13.0 cm from the source in the third run. Therefore 17 ( $=6+4+7$ ) individual substrates were obtained with the temperature and pressure combination of ( $630^{\circ}\text{C}$ , 4 mb). Each of these 17 substrates constitute a row in Table 1. The total number of substrates obtained from the 135 ( $=45 \times 3$ ) runs was 415. Note that this is not a multiple of 45 owing to an unequal number of substrates corresponding to each run.

Considering each of the 415 substrates as an experimental unit, the design matrix can thus be considered to be a  $415 \times 3$  matrix, where the three columns correspond to source temperature ( $TEMP$ ), pressure ( $PRES$ ) and distance from the source ( $DIST$ ). Each row corresponds to a substrate, on which a deposition is formed with a specific combination of  $TEMP$ ,  $PRES$  and  $DIST$  (see Table 1).

Recall that from the deposition on each substrate, 180 individual nanostructures were counted using SEM images. The response was thus a vector  $\mathbf{Y} = (Y_1, Y_2, Y_3, Y_4)$ , where  $Y_1, Y_2, Y_3$ , and  $Y_4$  denote respectively the number of nanosaws, nanowires, nanobelts and no morphology, with  $\sum_{j=1}^4 Y_j = 180$ . For demonstration purposes, the first 17 rows of the com-

plete dataset are shown in Table 1. These rows correspond to the temperature-pressure combination (630,4). The complete data can be downloaded from [www.isye.gatech.edu/~roshan](http://www.isye.gatech.edu/~roshan).

Table 1: Partial data (first 17 rows out of 415) obtained from the nano-experiment

Temperature	Pressure	Distance	Nanosaws	Nanowires	Nanobelts	No growth
630	4	12.4	0	0	0	180
630	4	14.7	74	106	0	0
630	4	15.4	59	121	0	0
630	4	16.9	92	38	50	0
630	4	18.6	0	99	81	0
630	4	20.7	0	180	0	0
630	4	12.2	50	94	36	0
630	4	15.1	90	90	0	0
630	4	17.6	41	81	58	0
630	4	19.4	0	121	59	0
630	4	12.5	49	86	45	0
630	4	14.8	108	72	0	0
630	4	15.4	180	0	0	0
630	4	16.9	140	40	0	0
630	4	19.0	77	47	56	0
630	4	21.1	0	88	92	0
630	4	23.5	0	0	0	180

It was observed that, at a source temperature of 850<sup>0</sup> C, almost no morphology was observed. Therefore, results obtained from the 67 experimental units involving this level of temperature were excluded and the data for the remaining 348 units were considered for analysis.

Henceforth, we shall use the suffixes 1,2,3 and 4 to represent quantities associated with nanosaws, nanowires, nanobelts and no growth respectively.

### 3 Model fitting

#### 3.1 Individual modeling of the probability of obtaining each nanostructure using binomial GLM

Here, the response is considered binary, depending on whether we get a specific nanostructure or not. Let  $p_1, p_2$  and  $p_3$  denote respectively the probabilities of getting a nanosaw/nanocomb, nanowire and nanobelt. Then, for  $j = 1, 2, 3$ , the marginal distribution of  $Y_j$  is binomial with  $n = 180$  and probability of success  $p_j$ . The log-odds ratio of obtaining the  $j^{th}$  type of morphology is given by

$$\zeta_j = \log \frac{p_j}{1 - p_j}.$$

Our objective is to fit a model that expresses the above log-odds ratios in terms *TEMP*, *PRES* and *DIST*.

From the main effects plot of *TEMP*, *PRES* and *DIST* against observed proportions of nanosaws, nanowires and nanobelts (Figure 3), we observe that a quadratic model should be able to express the effect of each variable on  $p_j$  adequately. The interaction plots (not shown here) give a preliminary impression that all the three two-factor interactions are likely to be important. We therefore decide to fit a quadratic response model to the data.

Each of three process variables are scaled to  $[-1, 1]$  by appropriate transformations. Let  $T, P$  and  $D$  denote the scaled variables obtained by transforming *TEMP*, *PRES* and *DIST* respectively.

Using a binomial GLM with a logit link (McCullagh and Nelder, 1989), we obtain the following models that express the log-odds ratios of getting a nanosaw/nanocomb, nanowire and nanobelt as functions of  $T, P, D$  :



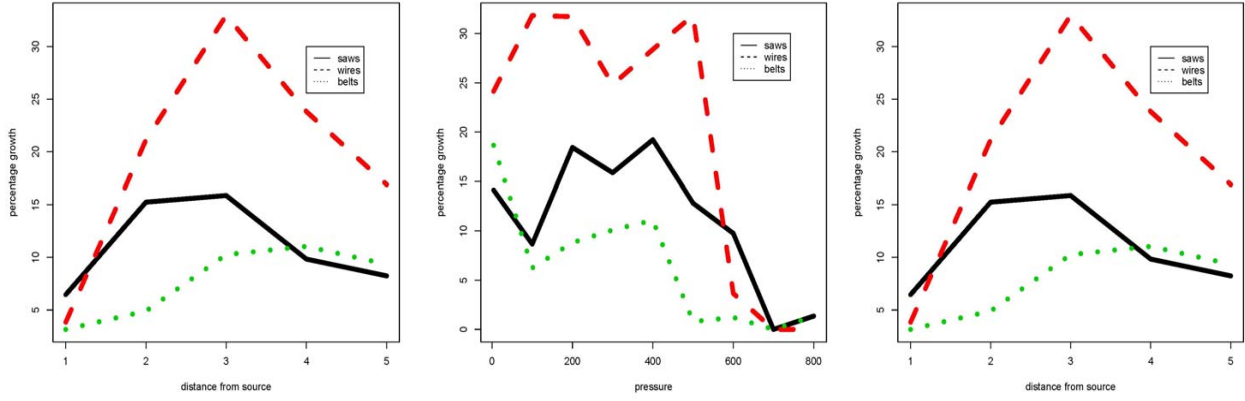


Figure 3: From left - growth vs temperature, growth vs pressure, growth vs distance

$$\begin{aligned}\hat{\zeta}_1 = & - 0.99 - 0.29 T - 1.52 P - 2.11 D - 0.95 T^2 - 1.30 P^2 - 5.64 D^2 \\ & - 0.18 TP - 1.03 PD + 4.29 TD,\end{aligned}\tag{1}$$

$$\begin{aligned}\hat{\zeta}_2 = & - 0.56 + 0.82 T - 2.53 P - 1.59 D - 0.58 T^2 - 2.04 P^2 - 2.62 D^2 \\ & + 1.17 TP - 1.44 PD + 0.87 DT,\end{aligned}\tag{2}$$

$$\begin{aligned}\hat{\zeta}_3 = & - 1.68 + 0.19 T - 1.88 P - 0.58 D - 1.69 T^2 - 0.34 P^2 - 3.20 D^2 \\ & + 0.87 TP - 0.94 PD - 2.58 TD.\end{aligned}\tag{3}$$

All the terms are seen to be highly significant. The residual plots for all the three models do not exhibit any unusual pattern.

### 3.2 Simultaneous modeling of the probability vector using multinomial GLM

Denoting the probability of not obtaining any nanostructure by  $p_4$ , we must have  $\sum_{j=1}^4 p_j = 1$ . Although the results obtained by using the binomial GLM are easily interpretable and useful, the method suffers from the inherent drawback that, for specific values of  $T, P$  and  $D$ , the fitted values of the probabilities may be such that  $\sum_{j=1}^3 p_j > 1$ . This is due to the fact that the *correlation* structure of  $\mathbf{Y}$  is completely ignored in this approach.

A more appropriate modeling strategy is to utilize the fact that the response vector  $\mathbf{Y}$  follows a multinomial distribution with  $n = 180$  and probability vector  $\mathbf{p} = (p_1, p_2, p_3, p_4)$ . In this case, one can express the multinomial logits  $\eta_j = \log(\frac{p_j}{p_4}), j = 1, 2, 3$  as functions of  $T, P$  and  $D$ . Note that  $\eta_j$  can be easily interpreted as the log-odds ratio of obtaining the  $j^{th}$  morphology as compared to no nanostructure, with  $\eta_4 = 0$ .

Methods for fitting multinomial logistic models by maximizing the multinomial likelihood have been discussed by several authors (McCullagh and Nelder 1989, Aitkin, Anderson, Francis and Hinde 1989, Agresti 2002, Faraway 2006, Long and Freese 2006). These methods have been implemented in several software packages like R/S-plus (multinom function), STATA (.mlogit function), LIMDEP (Mlogit\$ function), SAS (CATMOD function) and SPSS (Nomreg function). All of these functions use some algorithm for maximization of the multinomial likelihood (e.g., the multinom function in R/S-plus uses the neural network based optimizer provided by Venebles and Ripley (2002)). They produce more or less similar outputs, the default output generally consisting of the model coefficients, their standard errors and  $z$ -values, and model deviance.

Another popular algorithm to indirectly maximize the multinomial likelihood is to create a pseudo factor with a level for each data point, and use a Poisson GLM with log link. This method, although appropriate for small data sets, becomes cumbersome when the number of data points is large. In the presence of a large number of levels of the pseudo factor, a large part of the output generated by standard statistical softwares like R becomes redundant, because only the terms involving interaction between the categories and the predictor variables are of interest. Faraway (2006) points out some practical inconveniences of

using this method. Its application to the current problem clearly becomes very cumbersome owing to the large number (348) of data points.

We propose a new iterative method of fitting multinomial logit models. The method is based on an iterative application of binomial GLMs. Besides the intuitive extension of binomial GLMs to a multinomial GLM, the method has certain advantages over the existing methods which are described towards the end of the section.

Let  $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{i4})$  denote the response vector corresponding to the  $i^{th}$  data point,  $i = 1$  to  $N$ . Let  $n_i = \sum_{j=1}^4 Y_{ij}$ . Here,  $N = 348$  and  $n_i = n = 180$  for all  $i$ . We have,

$$P(Y_{i1} = y_{i1}, \dots, Y_{i4} = y_{i4}) = \frac{n_i!}{y_{i1}! \dots y_{i4}!} p_{i1}^{y_{i1}} \dots p_{i4}^{y_{i4}}.$$

Thus the likelihood function is given by

$$\begin{aligned} L(\mathbf{Y}_1, \dots, \mathbf{Y}_N) &= \prod_{i=1}^N \frac{n_i!}{y_{i1}! \dots y_{i4}!} p_{i1}^{y_{i1}} \dots p_{i4}^{y_{i4}} \\ &= \prod_{i=1}^N \frac{n_i!}{y_{i1}! \dots y_{i4}!} \prod_{j=1}^3 \left( \frac{p_{ij}}{p_{i4}} \right)^{y_{ij}} p_{i4}^{\sum_{j=1}^4 y_{ij}}. \end{aligned}$$

Defining  $\eta_{ij} = \log \frac{p_{ij}}{p_{i4}}$ , we have

$$p_{ij} = \frac{\eta_{ij}}{1 + \sum_{j=1}^3 \exp(\eta_{ij})} \quad j = 1, 2, 3, \quad (4)$$

and

$$p_{i4} = \frac{1}{1 + \sum_{j=1}^3 \exp(\eta_{ij})}. \quad (5)$$

Therefore the log-likelihood can be written as

$$\begin{aligned} \log(L) &= \sum_{i=1}^N \left( \log n_i! - \sum_{j=1}^4 \log y_{ij}! + \sum_{j=1}^3 y_{ij} \log \frac{p_{ij}}{p_{i4}} + n_i \log p_{i4} \right) \\ &= \sum_{i=1}^N \left( \log n_i! - \sum_{j=1}^4 \log y_{ij}! + \sum_{j=1}^3 y_{ij} \eta_{ij} - n_i \log \left( 1 + \sum_{j=1}^3 \exp(\eta_{ij}) \right) \right). \quad (6) \end{aligned}$$

Let  $\mathbf{x}_i = (1, T_i, P_i, D_i, T_i^2, P_i^2, D_i^2, T_i P_i, P_i D_i, T_i D_i)'$ ,  $i = 1, \dots, N$ . The objective is to express

the  $\eta$ 's as functions of  $\mathbf{x}$ . Substituting  $\eta_{ij} = \mathbf{x}_i' \boldsymbol{\beta}_j$  in (6) and successively differentiating with respect to each  $\boldsymbol{\beta}_j$ , we get the maximum likelihood (ML) equations as

$$\sum_{i=1}^N \mathbf{x}_i \left( y_{ij} - n_i \frac{\exp(\eta_{ij})}{1 + \sum_{j=1}^3 \exp(\eta_{ij})} \right) = \mathbf{0}, \quad j = 1, 2, 3, \quad (7)$$

$$\sum_{i=1}^N \mathbf{x}_i \left( y_{i4} - n_i \frac{1}{1 + \sum_{j=1}^3 \exp(\eta_{ij})} \right) = \mathbf{0}, \quad (8)$$

where  $\mathbf{0}$  denotes a vector of zeros having length 10. Writing  $\exp(\gamma_{il}) = \left( 1 + \sum_{l \neq j} \exp(\eta_{il}) \right)^{-1}$ , we obtain from (7)

$$\sum_{i=1}^N \mathbf{x}_i \left( y_{ij} - n_i \frac{\exp(\eta_{ij} + \gamma_{ij})}{1 + \exp(\eta_{ij} + \gamma_{ij})} \right) = \mathbf{0}, \quad j = 1, 2, 3. \quad (9)$$

Note that each equation in (9) is the maximum likelihood (ML) equation of a binomial GLM with logit link. Thus, if some initial estimates of  $\boldsymbol{\beta}_2, \boldsymbol{\beta}_3$  are available, and consequently  $\gamma_{i1}$  can be computed, then  $\boldsymbol{\beta}_1$  can be estimated by fitting a binomial GLM of  $Y_1$  on  $\mathbf{x}$ . Similarly,  $\boldsymbol{\beta}_2$  and  $\boldsymbol{\beta}_3$  can be estimated. The following algorithm is thus proposed.

### **Binomial GLM-based iterative algorithm for fitting a multinomial GLM :**

Let  $\boldsymbol{\beta}_j^{(k)}$  be the estimate of  $\boldsymbol{\beta}_j, j = 1, 2, 3$ , at the end of the  $k^{th}$  iteration.

Step 1. Using  $\boldsymbol{\beta}_2^{(k)}$  and  $\boldsymbol{\beta}_3^{(k)}$ , compute  $\eta_{i2}^{(k)} = \mathbf{x}_i' \boldsymbol{\beta}_2^{(k)}$  and  $\eta_{i3}^{(k)} = \mathbf{x}_i' \boldsymbol{\beta}_3^{(k)}$  for  $i = 1, \dots, n$ .

Step 2. Compute  $\gamma_{i1}^{(k)} = \log \frac{1}{1 + \exp(\eta_{i2}^{(k)}) + \exp(\eta_{i3}^{(k)})}$ ,  $i = 1, \dots, n$ .

Step 3. Treating  $Y_1$  as the response and using the same design matrix, fit a binomial GLM with logit link. The vector of coefficients thus obtained is  $\boldsymbol{\beta}_1^{(k+1)}$ .

Step 4. Repeat steps 1-3 by successively updating  $\gamma_{i2}$  and  $\gamma_{i3}$  and estimating  $\boldsymbol{\beta}_2^{(k+1)}$  and  $\boldsymbol{\beta}_3^{(k+1)}$ .

Repeat steps 1-4 until convergence. A proof of convergence is given in Appendix 1. Note that we use the 'offset' command in statistical software R to separate the coefficients associated with  $\eta_1$  from those with  $\gamma_1$ .

To obtain the initial estimates  $\hat{\eta}_{i2}^{(0)}$  and  $\hat{\eta}_{i3}^{(0)}$ , we use the results obtained from the binomial GLM as described in Section 4.1. Let

$$\log \frac{\hat{p}_{ij}}{1 - \hat{p}_{ij}} = \mathbf{x}_i' \hat{\boldsymbol{\delta}}_j, \quad (10)$$

where  $\hat{\boldsymbol{\delta}}_j$  is obtained by using binomial GLM. Recalling the definition of  $\eta_{ij}$ , the initial estimates are obtained as

$$\hat{\eta}_{ij}^{(0)} = \log \frac{\hat{p}_{ij}}{1 - \sum_{l=1}^3 \hat{p}_{il}}, \quad j = 2, 3, \quad (11)$$

where  $\hat{p}_{il}, l = 1, 2, 3$  are estimated from (10). It is possible, however, that for some  $i$ ,  $\sum_{l=1}^3 \hat{p}_{il} = \pi_i \geq 1$ . For those data points, we provide a small correction as follows:

$$\hat{p}_{il}^c = \begin{cases} \frac{\hat{p}_{il}}{\pi_i} (1 - \frac{1}{2n_i}) & , l = 1, 2, 3 \\ \frac{1}{2n_i} & , l = 4 \end{cases}.$$

where  $\hat{p}_{il}^c$  denotes the corrected estimated probability. To justify the correction, we note that it is a common practice to give a correction of  $\frac{1}{2n_i}$  (Cox 1970, chap. 3) in estimation of probabilities from binary data. The correction given to category 4 is adjusted among the other three categories in the same proportion as the estimated probabilities. This ensures that  $\hat{p}_{il}^c > 0$  for all  $i$  and  $\sum_{l=1}^4 \hat{p}_{il}^c = 1$

In this example, there were 18 data points (out of 348) corresponding to which we had  $\sum_{l=1}^3 \hat{p}_{il} \geq 1$ . Following the procedure described above to obtain the initial estimates, the following models were obtained after convergence:

$$\begin{aligned} \hat{\eta}_1 = & \quad 0.42 - 0.12 \ T - 3.08 \ P - 3.68 \ D - 1.84 \ T^2 - 1.52 \ P^2 - 9.09 \ D^2 \\ & + \ 0.60 \ TP - 2.31 \ PD + 5.75 \ TD, \end{aligned} \quad (12)$$

$$\begin{aligned} \hat{\eta}_2 = & \quad 0.54 + 0.88 \ T - 3.85 \ P - 3.13 \ D - 1.21 \ T^2 - 2.28 \ P^2 - 5.26 \ D^2 \\ & + \ 1.83 \ TP - 2.62 \ PD + 2.07 \ TD, \end{aligned} \quad (13)$$

$$\begin{aligned} \hat{\eta}_3 = & \quad - \ 0.10 + 0.39 \ T - 3.67 \ P - 2.51 \ D - 2.51 \ T^2 - 1.12 \ P^2 - 7.07 \ D^2 \\ & + \ 1.72 \ TP - 2.38 \ PD + 4.47 \ TD. \end{aligned} \quad (14)$$

### Inference for the proposed method:

To test the significance of the terms in the model, one can use the asymptotic normality of the maximum likelihood estimates. Let  $\mathbf{H}_\beta$  denote the  $30 \times 30$  matrix consisting of the negative expectations of second-order partial derivatives of the log-likelihood function in (6), the derivatives being taken with respect to the components of  $\beta_1, \beta_2$  and  $\beta_3$ . Denoting the final estimator of  $\beta$  as  $\beta^*$ , the estimated asymptotic variance-covariance matrix of the estimated model coefficients is given by  $\Sigma_{\beta^*} = \mathbf{H}_{\beta^*}^{-1}$ . For a specific coefficient  $\beta_l$ , the null hypothesis  $H_0 : \beta_l = 0$  can be tested using the test statistic  $z = \hat{\beta}_l / s(\hat{\beta}_l)$ , where  $s^2(\hat{\beta}_l)$  is the  $l^{th}$  diagonal element of  $\Sigma_{\beta^*}$ .

Let  $\beta_l^{(k)}$  denote the estimate of  $\beta_l$  obtained after the  $k^{th}$  iteration of the proposed algorithm. Let  $s^2(\beta_l^{(k)})$  denote the estimated asymptotic variance of  $\beta_l^{(k)}$ . It can easily be seen (Appendix 2) that  $s^2(\beta_l^{(k)})$  converges to  $s^2(\beta_l^*)$ . Thus, as the parameter estimates converge to the maximum likelihood estimates, their standard errors also converge to the standard error of the MLE. More generally, if  $\Sigma_{\beta^{(k)}}$  denotes the asymptotic covariance matrix of the parameter estimates at the end of the  $k^{th}$  iteration, then  $\Sigma_{\beta^{(k)}} \longrightarrow \Sigma_{\beta^*}$ .

The above property of the proposed algorithm ensures that one does not have to spend any extra computational effort in judging the significance of the model terms. The binomial GLM function in R used in every iteration automatically tests the significance of the model terms, and the  $p$ -values associated with the estimated coefficients after convergence can be used for inference. Thus, the inferential procedures and diagnostic tools of the binomial GLM can easily be used in the multinomial GLM model. This is clearly an advantage of the proposed algorithm over existing methods. Further, the three models for nanosaws, nanobelts and nanowires can be compared using these diagnostic tools. Such facilities are not available in the current implementation of other software packages.

In the fitted models given by (12)-(14), all the 30 coefficients are seen to be highly significant with  $p$  values of the order  $10^{-6}$  or less. To check the model adequacy, we use the generalized  $R^2$  statistic derived by Naglekerke (1991) defined as  $R^2 = (1 - \exp((D - D_{null})/n)) / (1 - \exp(-D_{null}/n))$ , where  $D$  and  $D_{null}$  denote the residual deviance and the null deviance respectively. The  $R^2$  associated with the models for nanosaws, nanowires and nanobelts are

Table 2: Computed values of the test statistic for each estimated coefficient

Term	Nanosaws ( $\hat{\eta}_1$ )			Nanowires ( $\hat{\eta}_2$ )			Nanobelts( $\hat{\eta}_3$ )		
	$\hat{\beta}$	S.E.	$p$ -val.	$\hat{\beta}$	S.E.	$p$ -val.	$\hat{\beta}$	S.E.	$p$ -val.
Intercept	0.42	0.24	0.0807	0.54	0.25	0.0343	-0.10	0.24	0.6763
$T$	-0.12	0.30	0.6855	0.88	0.28	0.0020	0.39	0.38	0.3125
$P$	-3.08	0.41	0.0000	-3.85	0.50	0.0000	-3.67	0.51	0.0000
$D$	-3.69	0.67	0.0000	-3.13	0.57	0.0000	-2.51	0.66	0.0001
$T^2$	-1.84	0.34	0.0000	-1.21	0.27	0.0000	-2.51	0.34	0.0000
$P^2$	-1.52	0.44	0.0006	-2.28	0.52	0.0000	-1.12	0.54	0.0381
$D^2$	-9.09	0.99	0.0000	-5.26	0.66	0.0000	-7.08	0.77	0.0000
$TP$	0.60	0.42	0.1515	1.83	0.41	0.0000	1.72	0.53	0.0011
$PD$	-2.31	0.83	0.0053	-2.62	0.70	0.0000	-2.38	0.84	0.0043
$TD$	5.75	0.80	0.0000	2.07	0.45	0.0000	4.47	0.69	0.0000

obtained as 61%, 50% and 76% respectively. This shows that the prediction error associated with the model for nanowires is the largest. This finding is consistent with the observation made by Ma and Wang (2005) that growth of nanowires is less restrictive compared to nanosaws and nanowires, and can be carried out over wide ranges of temperature and pressure.

However, the small  $p$  values may also arise from the fact that some improper variance is used in testing. This overdispersion may be attributed either to some correlation among the outcomes from a given run in the experiment or some unexplained heterogeneity within a group owing to the effect of some unobserved variable. A multitude of external noise factors in the system make the second reason a more plausible one. We re-perform the testing by introducing three dispersion parameters  $\sigma_1^2, \sigma_2^2, \sigma_3^2$ , which are estimated from the initial binomial fits using  $\hat{\sigma}_j^2 = \chi_j^2 / (N - 10)$ , where  $\chi_j^2$  denotes Pearson's  $\chi^2$  statistic for the  $j$ th nanostructure,  $j = 1, 2, 3$  and  $N - 10 = 338$  is the residual degrees of freedom. The estimated standard errors of the coefficients and the corresponding  $p$  values are shown in Table 2.

From Table 2 we find that the linear effect of temperature is not significant for nanosaws and nanobelts. However, since the quadratic term  $T^2$  and the interactions involving  $T$  are significant, we prefer to retain  $T$  in the models.

Note that although techniques for analyzing overdispersed binomial data are well known (e.g., Faraway 2006, Ch. 2), methods for handling overdispersion in multinomial logit models

are not readily available. The proposed algorithm provides us with a very simple heuristic way to do this and thereby has an additional advantage over the existing methods.

## 4 Optimization of the synthesis process

In the previous subsections, the three process variables have been treated as non-stochastic. However, in reality, none of these variables can be controlled precisely and each of them exhibits certain fluctuations around the set (nominal) value. Such fluctuation is a form of noise, called internal noise (Wu and Hamada 2000, chap. 10) associated with the synthesis process and needs to be considered in performing optimization.

It is therefore reasonable to consider  $TEMP$ ,  $PRES$  and  $DIST$  as random variables. Let  $\mu_{TEMP}$ ,  $\mu_{PRES}$ ,  $\mu_{DIST}$  denote the set values of  $TEMP$ ,  $PRES$  and  $DIST$  respectively. Then we assume

$$TEMP \sim N(\mu_{TEMP}, \sigma_{TEMP}^2), PRES \sim N(\mu_{PRES}, \sigma_{PRES}^2), DIST \sim N(\mu_{DIST}, \sigma_{DIST}^2).$$

where  $\sigma_{TEMP}^2$ ,  $\sigma_{PRES}^2$ ,  $\sigma_{DIST}^2$  are the respective variances of  $TEMP$ ,  $PRES$  and  $DIST$  around their set values and are estimated from process data (Section 5.1). The task now is to determine the optimal nominal values  $\mu_{TEMP}$ ,  $\mu_{PRES}$  and  $\mu_{DIST}$  so that the expected yield of each nanostructure is maximized subject to the condition that the variance in yield is acceptable.

### 4.1 Measurement of internal noise in the synthesis process

Some surrogate process data collected from the furnace were used for estimation of the above variance components. Temperature and pressure were set at specific levels (those used in the experiment), and their actual values were measured repeatedly over a certain period of time. The range of temperature and pressure corresponding to each set value was noted. The variation in distance, which is due to repeatability and reproducibility errors associated with the measurement system, was assessed separately. The summarized data in Table 3 show the observed ranges of  $TEMP$ ,  $PRES$  and  $DIST$  against different nominal values.



Under the assumption of normality, the range can be assumed to be approximately equal to six times the standard deviation.

Table 3: Fluctuation of process parameters around set values

Temperature		Pressure		Distance	
Nominal value ( $\mu_T$ )	Observed range ( $\approx \pm 3\sigma_T$ )	Nominal value ( $\mu_P$ )	Observed range ( $\approx \pm 3\sigma_P$ )	Nominal value ( $\mu_D$ )	Observed range ( $\approx \pm 3\sigma_D$ )
630	$\pm 7$	4	$\pm 10$	11	$\pm 0.02$
700	$\pm 7$	100	$\pm 10$	13	$\pm 0.02$
750	$\pm 6$	200	$\pm 20$	15	$\pm 0.02$
800	$\pm 6$	300	$\pm 20$	17	$\pm 0.02$
850	$\pm 6$	400	$\pm 20$	19	$\pm 0.02$
		500	$\pm 40$	21	$\pm 0.02$
		600	$\pm 40$		

We observe from Table 3 that, for the process variable  $DIST$ , the range of values around the nominal  $\mu_{DIST}$  is a constant ( $2 \times 0.02 = 0.04$  mm) and independent of  $\mu_{DIST}$ . Equating this range to  $6\sigma_{DIST}$ , we obtain an estimate of  $\sigma_{DIST}$  as  $0.04/6 = 0.067$  mm.

Similarly, for  $TEMP$ , the range can be taken to be almost a constant. Equating the mean range of 12.8 ( $= 2 \times (2 \times 7 + 3 \times 6)/5$ ) degrees to  $6\sigma_{TEMP}$ , an estimate of  $\sigma_{TEMP}$  is obtained as  $12.8/6 = 2.13$  degree C.

The case of  $PRES$  is, however, different. The range, and hence  $\sigma_{PRES}$  is seen to be an increasing function of  $\mu_{PRES}$ . Corresponding to each value of  $\mu_{PRES}$ , an estimate of  $\sigma_{PRES}$  is obtained by dividing the range by 6. Using these values of  $\sigma_{PRES}$ , the following regression line is fitted through the origin to express the relationship between  $\sigma_{PRES}$  and  $\mu_{PRES}$

$$\sigma_{PRES} = 0.025\mu_{PRES}. \quad (15)$$

Recall that all the models are fitted with the transformed variables  $T, P, D$ . The means  $\mu_T, \mu_P, \mu_D$  and the variances  $\sigma_T^2, \sigma_P^2, \sigma_D^2$  can easily be expressed in terms of the respective means and variances of the original variables.

## 4.2 Obtaining the mean and variance functions of $p_1, p_2, p_3$

From (4), we have the estimated probability functions as  $\hat{p}_j = \exp(\hat{\eta}_j) / (1 + \sum_{j=1}^3 \exp(\hat{\eta}_j))$ , where  $\hat{\eta}_j$  are given by (12)-(14).

Expressing  $E(p_j)$  and  $Var(p_j)$  in terms of  $\mu_T, \mu_P, \mu_D$  is not a straightforward task. To do this, we use Monte Carlo simulations. For each of the 180 combinations of  $\mu_{TEMP}, \mu_{PRES}, \mu_{DIST}$  ( $\mu_{TEMP} = 630, 700, 750, 800$ ;  $\mu_{PRES} = 4, 100, 200, \dots, 800$ ;  $\mu_{DIST} = 12, 14, 16, 18, 20$ ) the following are done:

1.  $\mu_T, \mu_P$  and  $\mu_D$  are obtained by appropriate transformation.
2. 5000 observations on  $T, P$  and  $D$  are generated from the respective normal distributions and  $\eta_j$  is obtained using equation (12), (13) or (14).
3. From the  $\eta_j$  values thus obtained,  $p_j$ 's are computed using (4) and transformed to  $\zeta_j = \log \frac{p_j}{1-p_j}$ .
4. The mean and variance of those 5000  $\zeta_j$  values (denoted by  $\bar{\zeta}_j$  and  $s^2(\zeta_j)$  respectively) are computed.
5. Using linear regression,  $\bar{\zeta}_j$  and  $\log s^2(\zeta_j)$  are expressed in terms of  $\mu_T, \mu_P$  and  $\mu_D$ .

## 4.3 Maximizing the average yield

Since the response here is of larger-the-better type, maximizing the mean is more important than minimizing the variance in the two-step optimization procedure (Wu and Hamada 2000, chap. 10) associated with robust parameter design.

The problem can thus be formulated as :

$$\begin{aligned} & \text{maximize } \bar{\zeta}_j \quad \text{subject to} \\ & -1 \leq \mu_T \leq 1, \quad -1 \leq \mu_P \leq 1, \quad -1 \leq \mu_D \leq 1 \quad \text{for } j = 1, 2, 3. \end{aligned}$$

Physically, this would mean maximizing the average log-odds ratio of getting a specific morphology.

The following models are obtained from the simulated data:

$$\begin{aligned}\overline{\zeta_1} = & - 0.75 + 0.20\mu_T - 1.02\mu_P - 1.39\mu_D - 1.50\mu_T^2 - 3.54\mu_P^2 - 11.02\mu_D^2 \\ & + 1.58\mu_T\mu_P - 2.22\mu_P\mu_D + 8.41\mu_T\mu_D,\end{aligned}\tag{16}$$

$$\begin{aligned}\overline{\zeta_2} = & - 0.40 + 0.80\mu_T - 2.96\mu_P - 1.43\mu_D - 0.98\mu_T^2 - 2.45\mu_P^2 - 6.05\mu_D^2 \\ & + 1.87\mu_T\mu_P - 3.41\mu_P\mu_D + 2.13\mu_T\mu_D,\end{aligned}\tag{17}$$

$$\begin{aligned}\overline{\zeta_3} = & - 1.25 + 0.26\mu_T - 2.6\mu_P - 0.42\mu_D - 2.36\mu_T^2 - 1.24\mu_P^2 - 8.03\mu_D^2 \\ & + 1.74\mu_T\mu_P - 3.32\mu_P\mu_D + 4.57\mu_T\mu_D.\end{aligned}\tag{18}$$

Maximizing these three functions using the *optim* command in R, we get the optimal conditions for maximizing the expected yield of nanosaws, nanowires and nanobelts in terms of  $\mu_T, \mu_P$  and  $\mu_D$ . These optimal values are transformed to the original units (i.e., in terms of  $\mu_{TEMP}, \mu_{PRES}$ , and  $\mu_{DIST}$ ) and are summarized in Table 4.

Table 4: Optimal process conditions for maximizing expected yield of nanostructures

Nanostructure	Temperature	Pressure	Distance
Nanosaws	630	307	15.1
Nanowires	695	113	19.0
Nanobelts	683	4	17.0

Contour plots of average and variance of the yield probabilities of nanosaws, nanowires and nanobelts against temperature and pressure (at optimal distances) are shown in Figure 4. The white regions on the top (average) panels and the black regions on the bottom (variance) panels are robust regions that promote high yield with minimal variation.

On the basis of these contour plots and the optimization output summarized in Table 4, the following conclusions can be drawn:

1. For nanosaws, the process is seen to be fairly robust at a pressure below 400 mb, irrespective of source temperature.

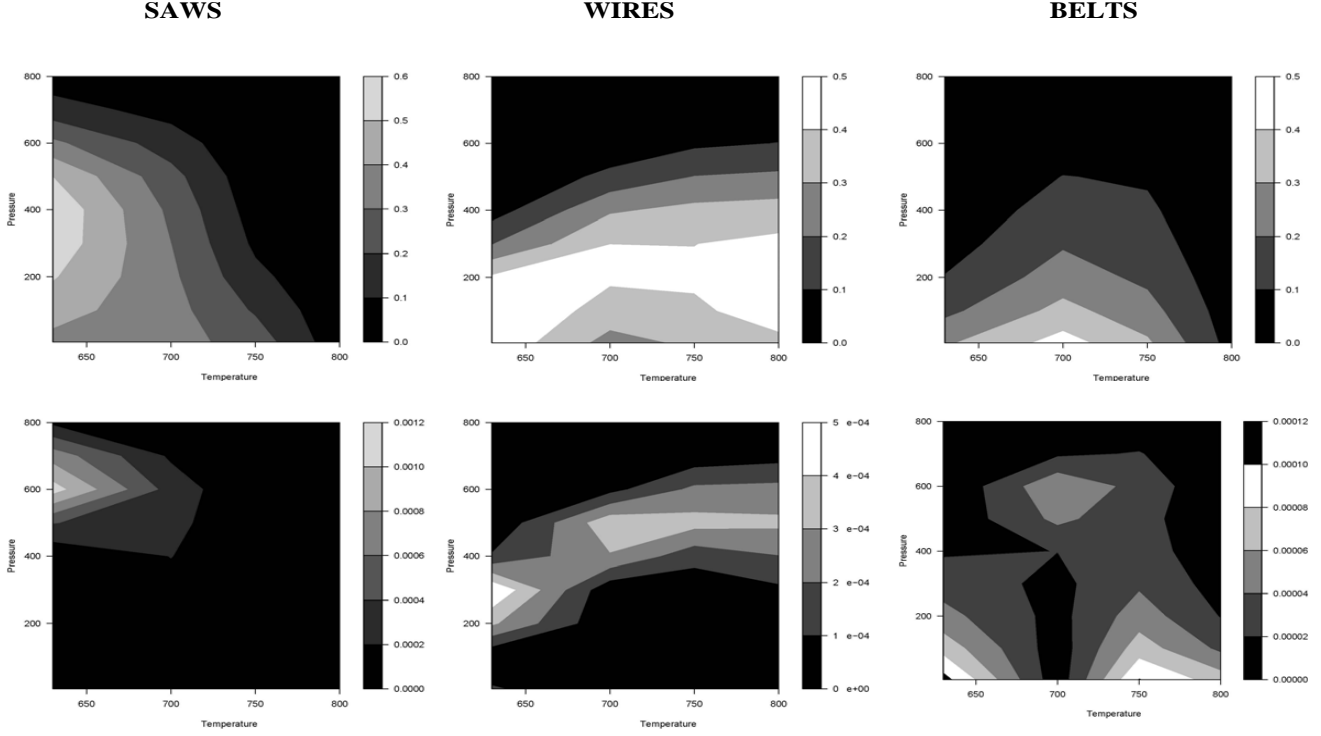


Figure 4: Contour Plots

2. For nanobelts, temperature affects robustness strongly, and for a pressure of less than 400 mb, the process is very robust only when the temperature is in close proximity of 700° C.
3. A temperature of around 630 degrees and pressure of 310 mb simultaneously maximizes the average and minimize the variance of probability of obtaining nanosaws.
4. A temperature of around 700 degrees and pressure of around 120 degrees results in highest average yield for nanowires. Low variance is also observed in this region.
5. Highest yield of nanobelts is achieved at a temperature of 680 degrees and pressure of 4 mb for nanobelts. This is also a low-variance region.
6. There is a large temperature-pressure region (white region in the top-middle panel of Figure 4) that promotes high and consistent yield of nanowires.
7. Highest yields of nanobelts and nanowires are achieved at higher distance (i.e., lower local temperature) as compared to nanosaws.

Except for the robustness-related conclusions, most of the above findings are summarized and discussed by Ma and Wang (2005). They also provide plausible and in-depth physical interpretations of some of the above phenomena.

## **5 Some general statistical issues in nanomaterial synthesis and scope for future research**

In this article, we report an early application of statistical techniques in nanomaterial research. In terms of reporting results of experiments to synthesize nanostructures, this methodology can be considered a significant advancement over the rudimentary data analysis methods using simple graphs, charts and summary statistics (e.g., Song, Wang, Riedo and Wang 2005; Ma and Wang 2005) that have been reported in the nanomaterial literature so far. Here we discuss features of the data arising from a specific experiment and use a multinomial model to express the probabilities of three different morphologies as functions of the process variables. A new iterative algorithm which is more appropriate than conventional methods for the present problem, is proposed for fitting the multinomial model. Inner noise is incorporated into the fitted models and robust settings of process variables that maximize the expected yield of each type of nanostructure are determined.

Apart from the advantages discussed earlier in this paper and mentioned by Ma and Wang (2005), this study demonstrates how statistical techniques can help in identifying important higher-order effects (like quadratic effects or complex interactions among the process variables) and utilize such knowledge in fine-tuning the optimal synthesis conditions. This work is also an important step towards large-scale controlled synthesis of CdSe nanostructures, since in addition to determining conditions for high yield, it also identifies robust settings of the process variables that are likely to guarantee consistent output.

Although statistical design of experiments (planning, analysis and optimization) have been applied very successfully to various other branches of scientific and engineering research to determine high-yield and reproducible process conditions, its application in nanotechnology has been limited till date. Some unique aspects associated with the synthesis

of nanostructures that make the application of the above techniques in this area challenging are: (i) complete disappearance of nanostructure morphology with slight changes in process conditions; (ii) complex response surface with multiple optima, making exploration of optimal experimental settings very difficult (although in the current case study, a quadratic response surface was found more or less adequate, such is not the case in general); (iii) different types of nanostructures (saws, wires, belts) intermingled; (iv) categorical response variables in most cases; (v) functional inputs (control factors that are functions of time); (vi) a multitude of internal and external noise factors heavily affecting reproducibility of experimental results; and (vii) expensive and time consuming experimentation. In view of the above phenomena, the following are likely to be some of the major statistical challenges in the area of nanostructure synthesis :

1. *Developing a sequential space-filling design for maximization of yield* : Fractional factorial designs and orthogonal arrays are the most commonly used (Wu and Hamada 2000) designs, but are not suitable for nanomaterial synthesis, because the number of runs becomes prohibitively large as the number of levels increases. Moreover, they do not facilitate sequential experimentation, which is necessary to keep the run size to a minimum. Response surface methodology (Myers and Montgomery 2002) may not be useful because the underlying response surface encountered in nano-research can be very complex with multiple local optima, and the binary nature of data adds to the complexity. Space-filling designs like Latin hypercube designs, uniform designs, and scrambled nets are highly suitable for exploring complex response surfaces with minimum number of runs. They are now widely used in computer experimentation (Santner, Williams and Notz 2003). However, they are used in the literature for only one-time experimentation. We need designs that are model independent, quickly “carve out” regions with no observable nanostructure morphology, allow for the exploration of complex response surfaces, and can be used for sequential experimentation.
2. *Developing experimental strategies where one or more of the control variables is a function of time* : In experiments for nanostructure synthesis, there are factors whose profiles or curves with respect to time are often crucial with respect to the output.

For example, although the peak temperature is a critical factor, how this temperature is attained over time is very important. There is an ideal curve that is expected to result in the best performance. Planning and analysis of experiments with such factors (which are called *functional factors*) are not much discussed in the literature and may be an important topic for future research.

3. *Scale up* : One of the important future tasks of the nanomaterial community is to develop industrial-scale manufacture of the nanoparticles and devices that are rapidly being developed. This transition from laboratory-level synthesis to large scale, controlled and designed synthesis of nanostructures poses plenty of challenges. The key issues to be addressed are rate of production, process capability, robustness, yield, efficiency and cost. The following specific tasks may be necessary: (1) Deriving specifications for key quality characteristics of nanostructures based on intended usage. Quality loss functions (Joseph 2004) may be used for this purpose. (2) Expanding the laboratory set-up to simulate additional conditions that are likely to be present in an industrial process. (3) Conducting experiments and identifying robust settings of the process variables that will ensure manufacturing of nanostructures of specified quality with high yield. (4) Statistical analysis of experimental data to compute the capability of the production process with respect to each quality characteristic.

## Acknowledgements

We are thankful to the Associate Editor and the referees, whose comments helped us to improve the contents and presentation of the paper. The research of VRJ was supported by NSF grant DMI-0448774. The research of ZLW was supported by grants from NSF, DARPA and NASA.

# Appendices

## Appendix-1: Proof of Convergence of the proposed algorithm:

For simplicity, consider a single predictor variable and assume that  $\eta_{ij} = \beta_j x_i$ , where  $\beta_j$  is a scalar ( $i = 1, 2, \dots, N, j = 1, 2, 3$ ). Let  $Q(\beta_1, \beta_2, \beta_3) = \sum_{i=1}^N \left( \sum_{j=1}^3 y_{ij} \eta_{ij} - n_i \log \left( 1 + \sum_{j=1}^3 \exp(\eta_{ij}) \right) \right)$ . Recall that  $\beta_j^{(k)}$  denotes the estimate of  $\beta_j$  obtained after the  $k^{th}$  iteration. Then, it suffices to show that

- (i)  $Q(\beta_1, \beta_2, \beta_3)$  is a concave function of  $\beta_j$ ,  $j = 1, 2, 3$ , and
- (ii)  $Q(\beta_1^{(k+1)}, \beta_2^{(k)}, \beta_3^{(k)}) \geq Q(\beta_1^{(k)}, \beta_2^{(k)}, \beta_3^{(k)})$ .

It is easy to see that for  $l = 1, 2, 3$ ,

$$\frac{\partial^2 Q}{\partial \beta_l^2} = - \sum_{i=1}^N \frac{n_i x_i^2 e^{\beta_l x_i} (1 + \sum_{j \neq l} e^{\beta_j x_i})}{(1 + \sum_{j=1}^3 e^{\beta_j x_i})^2} \leq 0,$$

which proves the concavity of  $Q$ .

To prove (ii), we note that for given  $\beta_2^{(k)}, \beta_3^{(k)}$  the solution for  $\beta_1$  to the equation

$$\sum_{i=1}^N \left( y_{i1} - n_i \frac{e^{\beta_1 x_i}}{1 + e^{\beta_1 x_i} + \sum_{j=2}^3 e^{\beta_j^{(k)} x_i}} \right) x_i = 0$$

maximizes  $Q(\beta_1, \beta_2^{(k)}, \beta_3^{(k)})$ .

From the first equation of (9) and steps 1-3 of the algorithm, we have

$$\sum_{i=1}^N \left( y_{i1} - n_i \frac{e^{\beta_1^{(k+1)} x_i}}{1 + e^{\beta_1^{(k+1)} x_i} + \sum_{j=2}^3 e^{\beta_j^{(k)} x_i}} \right) x_i = 0,$$

which means  $\beta_1^{(k+1)} = \arg \max Q(\beta_1, \beta_2^{(k)}, \beta_3^{(k)})$ . Therefore (ii) holds.

## Appendix-2: Proof of convergence of the estimated covariance matrix

Again, for simplicity, consider a single predictor variable and assume that  $\eta_{ij} = \beta_j x_i$  where  $\beta_j$  is a scalar ( $i = 1, 2, \dots, N, j = 1, 2, 3$ ). Let  $\beta_j^{(k)}$  denote the estimate of  $\beta_j$  obtained



after steps 1-3 of the  $k^{th}$  iteration and  $\beta_j^*$  denote the final estimate of  $\beta_j$  obtained by the proposed algorithm.

The estimated asymptotic variance of  $\beta_1^{(k)}$ , denoted by  $s^2(\beta_1^{(k)})$ , is given by the negative expectation of  $\frac{\partial^2 \log L_{b1}}{\partial \beta_1^2} |_{\beta_1^{(k)}, \beta_2^{(k-1)}, \beta_3^{(k-1)}}$ , where  $\log L_{b1}$  denotes the binomial log-likelihood function of  $y_{i1} (i = 1, \dots, N)$  that corresponds to the first of the three equations in (9) and is given by

$$\log L_{b1} = \sum_{i=1}^N \log \binom{n}{y_{i1}} + \sum_{i=1}^N y_{i1}(\eta_{i1} + \gamma_{i1}) - n_i \sum_{i=1}^N \log \left( 1 + \exp(\eta_{i1} + \gamma_{i1}) \right).$$

Now,  $s^2(\beta_1^*)$ , the estimated asymptotic variance of  $\beta_1^*$ , is given by the negative expectation of  $\frac{\partial^2 \log L}{\partial \beta_1^2} |_{\beta_j = \beta_j^*, j=1,2,3}$ , where  $\log L$  is the multinomial likelihood given by (6).

It can easily be seen that

$$\frac{\partial^2 \log L_{b1}}{\partial \beta_1^2} = - \sum_{i=1}^N n_i x_i^2 \frac{1 + \exp(\eta_{i2}) + \exp(\eta_{i3})}{(1 + \exp(\eta_{i1}) + \exp(\eta_{i2}) + \exp(\eta_{i3}))^2} = \frac{\partial^2 \log L}{\partial \beta_1^2}.$$

By convergence of  $\beta_j^{(k)}$  to  $\beta_j^*$  for  $j = 1, 2, 3$ , it follows that  $s^2(\beta_1^{(k)}) \longrightarrow s^2(\beta_1^*)$ .

Similarly, each component in the covariance matrix  $\Sigma_{\beta^{(k)}}$  can be proven to converge to each component of  $\Sigma_{\beta^*}$ .

## References

- Aitkin, M., Anderson, D.A., Francis, B.J. and Hinde, J.P. (1989), *Statistical Modelling in GLIM*, Clarendon Press, Oxford.
- Agresti, A. (2002), *Categorical Data Analysis*, New York: Wiley.
- Bawendi, M.G., Kortan, A.R., Steigerwald, M.L., and Brus, L.E. (1989), "X-ray Structural Characterization of Larger Cadmium Selenide (CdSe) Semiconductor Clusters," *Journal of Chemical Physics*, 91, 7282-7290.
- Cox, D.R. (1970), *Analysis of Binary Data*, London : Chapman & Hall.

- Faraway, J.J. (2006), *Extending the Linear Model with R : Generalized Linear, Mixed Effects and Nonparametric Regression Models*, Boca Raton : Chapman & Hall/CRC.
- Hodes, G., Albu-Yaren, A., Decker, F., Motisuke, P. (1987), "Three-Dimensional Quantum-Size Effect in Chemically Deposited Cadmium Selenide Films," *Physics Review B*, 36, 4215-4221.
- Joseph (2004), "Quality Loss Functions for Nonnegative Variables and Their Applications," *Journal of Quality Technology* 36, 129-138.
- Long, J.S., and Freese, J. (2006), *Regression Models for Categorical Dependent Variables Using Stata*, College Station, TX: Stata Press.
- Ma, C., Ding, Y., Moore, D.F., Wang, X. and Wang, Z.L. (2004), "Single-Crystal CdSe Nanosaws," *Journal of American Chemical Society*, 126, 708-709.
- Ma, C. and Wang, Z.L. (2005), "Roadmap for Controlled Synthesis of CdSe Nanowires, Nanobelts and Nanosaws," *Advanced Materials*, 17, 1-6.
- McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*, Chapman and Hall, London.
- Myers, H. M. and Montgomery, D.C. (2002), *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, New York: Wiley.
- Naglekerke, N. (1991), "A Note on a General Definition of the Coefficient of Determination," *Biometrika*, 78, 691-692.
- Song, J., Wang, X., Riedo, E. and Wang, Z.L. (2005), "Systematic Study on Experimental Conditions for Large-Scale Growth of Aligned ZnO Nanowires on Nitrides," *Journal of Physical Chemistry B*, 109, 9869-9872.
- Santer, T. J., Williams, B. J. and Notz, W. I. (2003), *The Design and Analysis of Computer Experiments*, New York: Springer.
- Wu, C.F.J., and Hamada, M. (2000), *Experiments: Planning, Analysis, and Parameter Design Optimization*, New York: Wiley.
- Venables, W.N. and Ripley, B.D. (2002), *Modern Applied Statistics with S-PLUS*, New York: Springer.