Robust Control Settings for Repeated Manufacturing Processes

Debraj Chakraborty
Department of Electrical Engineering
Indian Institute of Technology Bombay
Mumbai 400076, India

Siuli Mukhopadhyay
Department of Mathematics
Indian Institute of Technology Bombay
Mumbai 400076, India

***Abstract*—[1]The product quality in manufacturing processes is dependent on various process parameters (called control settings) usually set by an human overseer at the beginning of each process run. However quality is often adversely influenced by several additional "noise" parameters which cannot be controlled and hence can take random values for each process run. The problem of computing the optimal sequence of control settings for an infinite series of such manufacturing runs, such that the product quality is maintained for all runs and for all possible values of the noise parameter, is addressed in this article. The dependence of product quality on control and noise parameters is modeled using the response surface methodology. Subsequently, the concept of mixed strategy equilibrium is borrowed from the theory of games to compute a probability distribution which optimally determines the control settings for each process run. These choices of control settings maximally improves product quality over all possible values of the noise parameters for all runs of the process.**

I. Introduction

Genichi Taguchi [1] discovered a statistical method called robust parameter design for product quality improvement. Taguchi identified two types of inputs in a manufacturing process: some easy-to-manipulate control factors and also some difficult-to-control noise factors. The noise factors are the sources of uncontrollable variations in the product quality. In a particular run of a manufacturing process, the best settings of the control factors are those for which the product quality is robust to the effects of the noise variables. However, most manufacturing processes are repetitive and the product quality must be maintained at the best possible value for each process run. In this article we propose a min-max formulation of the robust parameter design problem where we guarantee the best average product quality for manufacturing processes which are assumed to be infinitely repeated. A novel game theoretic approach is introduced to address this problem.

As an example, consider a drug manufacturing processes, where batches or lots of drugs are manufactured in each process run and such runs are repeated many times. Evidently, *every* instance (batch, lot or package) of the manufactured drug, irrespective of the uncontrollable noise factors, should be near the target quality. Suppose, the quality of a manufactured drug is quantified by the percentage of impurities it may safely contain. Then the production process should ensure that this percentage is never exceeded no matter what particular value the noise factors takes during any of the runs of the manufacturing process. Such a situation is also prevalent in manufacturing of military and safety devices (e.g. in the automobile industry), where certain safety standards must be guaranteed for each and every product. For instance, life saving devices used by the military must have a guaranteed product quality for each and every device being used.

Traditional approaches (e.g. see [2], [3] and the references therein) in the statistics literature have developed algorithms to choose the best control factors for a particular run of the process. Various approaches aimed at minimizing the process variance due to noise has been studied. However, none of these approaches deal with the scenario of repeated production runs, which however is the practical situation in most manufacturing processes. Moreover, we show that on an average (taken over various process runs), it is possible to have better product quality than for a single process run.

To the best of our knowledge, such a game theoretic approach to the robust parameter design of repeated process runs has not been addressed in the literature. Two conventional approaches to solve the robust parameter design problem for a single process run are the Taguchi method and the response surface approach. Several articles and books by Taguchi (e.g. [4], [1] and [5]), as well as papers by other authors (e.g. [6], [7] and [8]) discuss Taguchi's methodology in details. While several authors in the statistical design area (such as [9], [3], [10], [11] and references therein) have investigated the response surface approach.

The method of finding optimal strategy sequences in repeated game like situations was introduced by [12] and later studied by several authors in game theory. Algorithms to compute the optimal strategies in games with polynomial payoff were introduced by [13] and have been recently improved by [14]. In this article we use an algorithm proposed by [15].

The remainder of the article is organized as follows. In Section 2 we formulate the problem mathematically and introduce the idea of probability distribution based control setting selection for a repeated set of production runs. The response surface approach used in this article to model production processes is also briefly reviewed here. In Section 3, a result in [13] is extended to the multivariate situation needed for this application. While, Section 4 describes the numerical algorithm named IER, which is used to calculate the optimal control settings for the problem described above. A numerical example illustrating the theory introduced, is also included in Section 4.

[1]This work is partially supported by Department of Science and Technology grant No SR/S3/EECE/0017/2009.

II. Problem Formulation

A. Single Process Run

First consider the case of a single process run. We model the dependence of the process response (i.e. product quality) on the control and the noise parameters by a response surface based on a single experimental design [9]. As mentioned above, this process is repeated infinitely many times and in each run, the noise parameters can assume any arbitrary value within a given range. Let $\hat{y} := \hat{y}(\mathbf{x}, \mathbf{z})$ be the estimated process response where $\mathbf{x} \in \mathbb{R}^k$ are the k control factors and $\mathbf{z} \in \mathbb{R}^l$ are the l noise factors (details about the model $\hat{y}$ is given in subsection II-C). We assume that, $\mathbf{x}$ and $\mathbf{z}$ can take values only in the sets $R_x \subset \mathbb{R}^k$ and $R_z \subset \mathbb{R}^l$, respectively. If we were concerned about only a single process run, we might want to choose the control factor vector $\mathbf{x}$ for each run of the process so as to keep the estimated response near the target for all possible values of $\mathbf{z} \in R_z$. In other words we would like to minimize the worst deviation of the estimated response from the target: let the target response be T; Then for any particular run, and for any fixed choice $\mathbf{x}$ of the control factor, the worst possible deviation of the estimated response from the target is

$$\max_{\mathbf{z} \in R_z} \|T - \hat{y}(\mathbf{x}, \mathbf{z})\| \tag{1}$$

where $\|\cdot\|$ is some appropriately defined error criterion. In this article we will use the squared deviation of $\hat{y}$ about the target as the error criterion:

$$\|T - \hat{y}(\mathbf{x}, \mathbf{z})\| := (T - \hat{y}(\mathbf{x}, \mathbf{z}))^2 \tag{2}$$

However this worst deviation is still a function of the control variable $\mathbf{x}$. Hence, it would seem, that, in each run we would like to choose a $\mathbf{x} \in R_x$ which minimizes $\sup_{\mathbf{z} \in R_z}(T - \hat{y}(\mathbf{x}, \mathbf{z}))^2$. Consequently, we are looking for a $\mathbf{x}^* \in R_x$ which satisfies the following:

$$\max_{\mathbf{z} \in R_z}(T - \hat{y}(\mathbf{x}^*, \mathbf{z}))^2 = \min_{\mathbf{x} \in R_x} \max_{\mathbf{z} \in R_z}(T - \hat{y}(\mathbf{x}, \mathbf{z}))^2 \tag{3}$$

However, recall that in reality, the process is run for infinitely many (at least for large number of) times. Surprisingly, it turns out that it is sub-optimal to use the solution $\mathbf{x}^*$ of (3) for each run of an infinitely sequence of repeated production runs. For such situations, the theory of mathematical games [12], [13], predicts that the control setting for each run should be selected randomly from R_x, according to a carefully chosen optimal probability distribution over R_x. We next elaborate on this concept:

B. Repeated Process Runs:

First note that the optimal value described in (3) can be bounded below in the following way [13]:

$$\min_{\mathbf{x} \in R_x} \max_{\mathbf{z} \in R_z} (T - \hat{y}(\mathbf{x}, \mathbf{z}))^2 \geq \max_{\mathbf{z} \in R_z} \min_{\mathbf{x} \in R_x} (T - \hat{y}(\mathbf{x}, \mathbf{z}))^2. \tag{4}$$

In general, the $minmax$ and the $maxmin$ quantities are not equal. This in turn seems to indicate that there is still room for improvement by reducing the $minmax$, till we arrive at an equilibrium solution where the two expressions are equal. Situations where such an equilibrium solution exists have been widely studied in the mathematical theory of games (e.g. see [12], [16], [13] and references therein). It has been shown that equality always holds in (4) if we re-interpret the above inequality (4) in the following way.

Assume that the production process is ran repeatedly for several times, where each run is independent of the previous runs. We already know that the noise variable $\mathbf{z}$ is random in nature taking values from the permissible region R_z. Let us now assume that $\mathbf{z}$ takes values from R_z according to some *unknown* probability distribution with the cumulative distribution function (cdf) $G(\mathbf{z})$, (where $G(\mathbf{z})$ is defined over R_z) for each run of the process. Similarly, the human overseer, instead of using the same $\mathbf{x}$ for all runs of the process, chooses the control factor $\mathbf{x} \in R_x$ for each run of the process according to a probability distribution with a cumulative distribution function $F(\mathbf{x})$ defined over R_x. [Once the value of the control factor is chosen for a run it is then kept fixed at that value for the particular run].

Define the set of all cumulative distribution functions over R_z and R_x as $\mathcal{G}$ and $\mathcal{F}$, respectively. Then the estimated error $(T - \hat{y}(\mathbf{x}, \mathbf{z}))^2$ can be redefined in terms of the expected estimated error, over the product space $R_x \times R_z$ as follows:

$$\mathcal{M}(F, G) = \int_{R_z} \int_{R_x} (T - \hat{y}(\mathbf{x}, \mathbf{z}))^2 dF(\mathbf{x}) dG(\mathbf{z}), \tag{5}$$

where the integral is considered in the Stieltjes sense [13]. Thus, the problem of choosing the minmax control setting $(\mathbf{x}^*)$ as posed in 3 can now be translated to the problem of choosing the optimal cdf $F^* \in \mathcal{F}$ while assuming that the noise factors are chosen from the worst case cdf $G^* \in \mathcal{G}$ for repeated runs of the process. In other words, our objective in this setting of repeated process runs, is to find the optimal distribution function $F^* \in \mathcal{F}$ such that

$$\max_{G \in \mathcal{G}} \mathcal{M}(F^*, G) = \min_{F \in \mathcal{F}} \max_{G \in \mathcal{G}} \mathcal{M}(F, G). \tag{6}$$

Moreover according to the well-known result in game theory derived by [16] (also see [13]), the following always holds:

$$\min_{F \in \mathcal{F}} \max_{G \in \mathcal{G}} \mathcal{M}(F, G) = \max_{G \in \mathcal{G}} \min_{F \in \mathcal{F}} \mathcal{M}(F, G), \tag{7}$$

and at least one pair of distribution functions (F^*, G^*) achieving the above equilibrium, is guaranteed to exist. Such distribution function pairs (F, G) are called mixed strategies in game theory and the equilibrium defined by (7) is called a mixed strategy equilibrium. This equilibrium solution has the following properties and advantages over the solution to (3).

1) The min-max value of $\mathcal{M}(F, G)$ always lower bounds the min-max value of $M(\mathbf{x}, \mathbf{z})$. In other words: $\mathcal{M}(F^*, G^*) \leq \min_{\mathbf{x} \in R_x} \max_{\mathbf{z} \in R_z}(T - \hat{y}(\mathbf{x}, \mathbf{z}))^2$. This property fulfills the promise of improving the min-max solution of (3), thus providing the best possible method of choosing the sequence of control settings to get guaranteed average performance.
2) In reality, the noise variables are randomly created, and hence, $\mathbf{z}$ may not follow the worst case distribution

$G^*(\mathbf{z})$. But the optimal distribution function $F^*(\mathbf{x})$ guards against this possibility, by guaranteeing that for any incidence of the noise variable (say $\mathbf{z} = \mathbf{z}_1$),

$$\mathcal{M}(F^*, G^*) \leq \int_{R_x} (T - \hat{y}(\mathbf{x}, \mathbf{z}_1))^2 dF^*(\mathbf{x})$$

The optimal distribution function thus provides a method for selecting control settings for repeated production runs in manufacturing processes. However the computation of such a distribution function seems to be an issue. We show first this infinite dimensional optimization can be reduced to finite dimension using a property of moments (Section III). In Section IV, we outline a method for the computation of the optimal probability distribution for the control setting $\mathbf{x}^*$, based on the iterative entropic regularization algorithm proposed in [15]. It is shown that this algorithm solves (3) satisfactorily for the numerical example considered.

C. Statistical Model

A single experimental design known as the combined array (see [9] and [3]) is used for both the control and noise variables. Combined arrays are useful in estimating main effects and interaction effects among control variables, main effects and interaction effects among noise variables and interaction effects between control and noise variables. We make some standard assumptions: Due to the easy to control/manipulate nature of the control variables, they are considered as fixed effects, while noise variables are considered to be random. However, for the sake of experimentation and estimating the response, the noise factors are considered to be fixed at a certain level in the experiment; whereby we obtain measurements on the response and then use these data points to fit a model to the response. The fitted model relates the process response y as a function of the control ($\mathbf{x}$) and the noise variables ($\mathbf{z}$). We assume that the control and noise variables are continuous in nature. Using the full single response model proposed by [3] we write,

$$y(\mathbf{x}, \mathbf{z}) = \beta_0 + \mathbf{x}'\beta + \mathbf{x}'B\mathbf{x} + \mathbf{z}'\gamma + \mathbf{x}'\Delta\mathbf{z} + \epsilon, \quad (8)$$

where $\mathbf{x} = (x_1, \ldots, x_k)^T$ and $\mathbf{z} = (z_1, \ldots, z_l)^T$ are the vectors of control and noise variables, respectively; $\beta_{\{k\times 1\}}$ and $\gamma_{\{l\times 1\}}$ are the coefficients of the control and noise variables, respectively; $B_{\{k\times k\}}$ is a symmetric matrix containing the coefficients of the interaction effects between control variables; $\Delta_{\{k\times l\}}$ is a matrix containing the coefficients of the interaction effects between the control and the noise variables and $\epsilon \sim N(0, \sigma^2)$ is the error term.

We assume that both the control and noise variables take values in some closed bounded region: Let $\mathbf{x} \in R_x$ and $\mathbf{z} \in R_z$, where both R_x and R_z are closed and bounded (compact) sets; then the experimental region, R, is the Cartesian product of the sets R_x and R_z, i.e., $R = R_x \times R_z$. When planning the experiment the researcher selects a number of levels of both the control and the noise variables over the experimental region R to obtain measurements on the response variable. Using these measurements and considering both $\mathbf{x}$ and $\mathbf{z}$ to be fixed in model (8) we estimate $y(\mathbf{x}, \mathbf{z})$ by least squares regression. The fitted model is,

$$\hat{y}(\mathbf{x}, \mathbf{z}) = \hat{\beta}_0 + \mathbf{x}'\hat{\beta} + \mathbf{x}'\hat{B}\mathbf{x} + \mathbf{z}'\hat{\gamma} + \mathbf{x}'\hat{\Delta}\mathbf{z}. \quad (9)$$

where $\hat{\theta} = (\hat{\beta}_0, \hat{\beta}, \hat{B}, \hat{\Delta})'$ is, $\hat{\theta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$,$\mathbf{X}$ is the $n \times p$ design matrix, $p = 1 + k + l + \frac{k(k-1)}{2} + kl$, and $\mathbf{y}$ is the $n \times 1$ vector of responses.

We can summarize the problem described above as follows:

Problem 1. Given the estimated response model as in (9), a specified target response T and the design region $R_x \times R_z$, find $F^* \in \mathcal{F}$ such that $\max_{G\in\mathcal{G}} \mathcal{M}(F^*, G) = \min_{F\in\mathcal{F}} \max_{G\in\mathcal{G}} \mathcal{M}(F, G)$

III. The Optimal Distribution Function

We have seen that the optimal sequence of control setting for a repeated production process can be computed from an optimal distribution function. However, in order to compute the optimal distribution, we need to search over the spaces ($\mathcal{F}(\mathbf{x})$ and $\mathcal{G}(\mathbf{z})$) of all possible distribution functions, defined over R_x and R_z. It is evident that searching for distribution functions is, in general, numerically impossible. In this section, we show that for $\hat{y}(\mathbf{x}, \mathbf{z})$ described by (9), the optimal distribution functions (F^*, G^*) turn out to be step functions with a finite number of steps. It is further shown that the step function nature of the optimal distribution functions, reduces the optimization over all possible distribution functions implied by (6) into a much simpler finite dimensional optimization over Euclidean spaces. First let us define a step cumulative distribution function:

Definition 2. Let $\mathbf{P} \subset \mathcal{R}^k$, $\mathbf{p} = [p_1, ..., p_k]^T \in \mathbf{P}$ and $\bar{\boldsymbol{\alpha}} = [\alpha_1, ..., \alpha_k]^T \in \mathbf{P}$. Then a cdf $I(\mathbf{p};\bar{\boldsymbol{\alpha}})$ over $\mathbf{P}$ is said to be a step function with one step, if it is of the form:

$$I(\mathbf{p}; \bar{\boldsymbol{\alpha}}) = \begin{cases} 1 \text{ if } p_1 \geq \alpha_1 \text{ and } p_2 \geq \alpha_2 \text{ and } ....p_k \geq \alpha_k \\ 0 \qquad\qquad\qquad \text{otherwise} \end{cases}$$

It follows that for $\bar{\boldsymbol{\alpha}}^i \in \mathbf{P}$ $(i = 1, ..., m)$, the distribution function

$$F(\mathbf{p}) = \gamma_1 I(\mathbf{p}; \bar{\boldsymbol{\alpha}}^1) + \gamma_2 I(\mathbf{p}; \bar{\boldsymbol{\alpha}}^{\mathbf{2}}) + ... + \gamma_m I(\mathbf{p}; \bar{\boldsymbol{\alpha}}^{\mathbf{m}})$$

where $\gamma_i \geq 0$ and $\sum_i \gamma_i = 1$, is a step function with m steps.

Now recall the single model in both control and noise variables fitted to the response (9) and the min-max optimizations posed in the equation (3). Here the estimated response $\hat{y}(\mathbf{x}, \mathbf{z})$ is a function (see (9)) of the control factors $\mathbf{x} \in R_x \subset \mathcal{R}^k$ and the noise factors $\mathbf{z} \in R_z \subset \mathcal{R}^l$. Define the function $M(\mathbf{x}, \mathbf{z}) = (T - \hat{y}(\mathbf{x}, \mathbf{z}))^2$, which is a multivariate polynomial in the $(k + l)$ components of $\mathbf{x} = (x_1, \ldots, x_k)^T$ and $\mathbf{z} = (z_1, \ldots, z_l)^T$. It is easy to see that each term of the polynomial $M(\mathbf{x}, \mathbf{z})$ is of the form $a\{(x_1)^{t_1}(x_2)^{t_2}...(x_k)^{t_k}\}\{(z_1)^{u_1}(z_2)^{u_2}...(z_l)^{u_l}\}$ where $a \in \mathcal{R}$ is a constant, each of $t_1, t_2...t_k$ take one of the values in the set $\{0, 1, 2, 3, 4\}$ and each of $u_1, u_2...u_l$ take one of

the values in the set $\{0,1,2\}$. Hence, the function $M(\mathbf{x},\mathbf{z})$ can be written as the following sum [13]:

$$M(\mathbf{x},\mathbf{z}) = \sum_{j=1}^{m}\sum_{i=1}^{n} a_{ij} s_i(x_1,...,x_k) v_j(z_1,...,z_l) \qquad (10)$$

where $a_{ij} \in \mathcal{R}$ are constants, the functions $s_i(x_1,...,x_k) = \prod_{r=1}^{k}(x_r)^{t_r^i}$ and $t_r^i \in \{0,1,2,3,4\}$ for $r = 1,2,..,k$. Similarly the functions $v_j(z_1,...,z_l) = \prod_{q=1}^{l}(z_q)^{u_q^j}$ and $u_q^j \in \{0,1,2\}$ for $q = 1,2,..,l$. Here m and n are positive finite integers whose values depend on the number of non-zero elements in the parameter matrix $\hat{\theta}$ (see (9)). The expected estimated squared error can then be expressed as:

$$\mathcal{M}(F(\mathbf{x}),\, G(\mathbf{z})) = \mathcal{M}(\mathbf{S},\mathbf{V}) = \sum_{j=1}^{m}\sum_{i=1}^{n} a_{ij} S_i V_j \qquad (11)$$

where

$$\begin{aligned} S_i &= \int_{R_x} s_i(x_1,...,x_k) dF(\mathbf{x}) \text{ and} \\ V_j &= \int_{R_z} v_j(z_1,...,z_l) dG(\mathbf{z}). \end{aligned} \qquad (12)$$

In the rest of this article, we denote in shorthand $s_i(\mathbf{x}) := s_i(x_1,...,x_k)$, $v_j(\mathbf{z}) := v_j(z_1,...,z_l)$, $\mathbf{S} := (S_1,S_2,...,S_n)$ and $\mathbf{V} := (V_1,V_2,...,V_m)$.

We define the following notation. Let

$$\begin{aligned} C &= \{(s_1(\mathbf{x}),...,s_n(\mathbf{x})) \in \mathcal{R}^n : \mathbf{x} \in R_x\} \text{ and} \\ J &= \{(v_1(\mathbf{z}),...,v_m(\mathbf{z})) \in \mathcal{R}^m : \mathbf{z} \in R_z\} \end{aligned} \qquad (13)$$

While the sets C and J are not convex in general, we denote the convex closures of C and J by

$$D = conv(C) \text{ and } H = conv(J). \qquad (14)$$

In addition, using (12), we define the following sets created by all possible distribution functions over $\mathcal{F}$ and $\mathcal{G}$:

$$\begin{aligned} Q &= \{(S_1,S_2,...,S_n) \in \mathcal{R}^n : F \in \mathcal{F}\} \text{ and} \\ P &= \{(V_1,V_2,...,V_m) \in \mathcal{R}^m : G \in \mathcal{G}\} \end{aligned}$$

Using these definitions we can rewrite the equilibrium defined in (7) in terms of (11) as follows:

$$\min_{\mathbf{S}\in Q}\max_{\mathbf{V}\in P} \mathcal{M}(\mathbf{S},\mathbf{V}) = \max_{\mathbf{V}\in P}\min_{\mathbf{S}\in Q} \mathcal{M}(\mathbf{S},\mathbf{V}). \qquad (15)$$

Let us denote the pair achieving the equilibrium of (15) by $(\mathbf{S}^*,\mathbf{V}^*)$. However, this restatement of (7) does not reduce the computational complexity of the optimization. Hence we need a simpler characterization of the sets Q and P which is provided by the following theorem. The next theorem is a straightforward extension of the results derived in [13], to multi-dimensional Euclidean spaces.

Theorem 3. *If $M(\mathbf{x},\mathbf{z})$ is of the form (10), then we can claim:*

(i) There is at least one pair of optimal cumulative distribution functions (F^*,G^*) achieving the equilibrium defined by (7) such that $F^*$and G^* are step functions of at most n and m steps, respectively.

(ii) $Q = D$ and $P = H$; i.e., every cdf $F \in \mathcal{F}$ (alternatively $G \in \mathcal{G}$) corresponds to a point in the convex set D (alternatively H) and vice versa. Moreover, each point in the set D (alternatively H) can be represented by a distribution function of at most n (alternatively m) steps.

Proof: Using Kakutani's fixed point theorem ([16]) it is known that there exist at least one pair of cumulative distribution functions $(F^*,G^*) \in \mathcal{F}\times\mathcal{G}$ which achieves the equality in (7). Hence claim (ii) implies (i). We thus start by proving claim (ii) of the theorem.

Assume that there exists $\mathbf{S}^0 := [S_1^0, S_2^0, ..., S_n^0]^T \in Q$ such that $\mathbf{S}^0 \notin D$. Let $F_0(\mathbf{x})$ be the distribution function which produces $\mathbf{S}_0$, i.e. $\begin{bmatrix} S_1^0 \\ S_2^0 \\ . \\ . \\ S_n^0 \end{bmatrix} = \begin{bmatrix} \int_{R_x} s_1(\mathbf{x}) dF_0(\mathbf{x}) \\ \int_{R_x} s_2(\mathbf{x}) dF_0(\mathbf{x}) \\ . \\ . \\ \int_{R_x} s_n(\mathbf{x}) dF_0(\mathbf{x}) \end{bmatrix}$. Since D is convex, there exists a hyperplane separating $\mathbf{S_0}$ and D, i.e. for some $\delta > 0$, there exists a constant vector $\mathbf{w} \in \mathcal{R}^n$ such that

$$\mathbf{w}^T\mathbf{S}_0 - \mathbf{w}^T \begin{bmatrix} s_1(\mathbf{x}) \\ s_2(\mathbf{x}) \\ . \\ . \\ s_n(\mathbf{x}) \end{bmatrix} > \delta \text{ for all } \mathbf{x} \in R_x$$

Integrating both sides with respect to $dF_0(\mathbf{x})$, we get

$$\begin{aligned} &\mathbf{w}^T\mathbf{S}_0 - \mathbf{w}^T \begin{bmatrix} \int_{R_x} s_1(\mathbf{x}) dF_0(\mathbf{x}) \\ \int_{R_x} s_2(\mathbf{x}) dF_0(\mathbf{x}) \\ . \\ . \\ \int_{R_x} s_n(\mathbf{x}) dF_0(\mathbf{x}) \end{bmatrix} > \delta \\ \implies &\mathbf{w}^T\mathbf{S}_0 - \mathbf{w}^T\mathbf{S}_0 > \delta \end{aligned}$$

which implies a contradiction since $\delta > 0$.

Conversely, let $\mathbf{d}_0 \in D$. Then there exists a vector $\boldsymbol{\gamma} = [\gamma_1....\gamma_n] \in \mathcal{R}^n$ with $\gamma_i \geq 0$ and $\sum_i \gamma_i = 1$, and n vectors $\mathbf{x}^1,...,\mathbf{x}^n \in R_x$ with corresponding $(s_1(\mathbf{x}^i), s_2(\mathbf{x}^i), ..., s_n(\mathbf{x}^i)) \in C$ for all $i = 1,...,n$, such that $\mathbf{d}_0 = \gamma_1 \begin{bmatrix} s_1(\mathbf{x}^1) \\ s_2(\mathbf{x}^1) \\ . \\ . \\ s_n(\mathbf{x}^1) \end{bmatrix} + \gamma_2 \begin{bmatrix} s_1(\mathbf{x}^2) \\ s_2(\mathbf{x}^2) \\ . \\ . \\ s_n(\mathbf{x}^2) \end{bmatrix} + ... + \gamma_n \begin{bmatrix} s_1(\mathbf{x}^n) \\ s_2(\mathbf{x}^n) \\ . \\ . \\ s_n(\mathbf{x}^n) \end{bmatrix}$. It is easy to see that if we choose $F_0(\mathbf{x}) = \gamma_1 I_{\mathbf{x}^1}(\mathbf{x}) + \gamma_2 I_{\mathbf{x}^2}(\mathbf{x}) + ... + \gamma_n I_{\mathbf{x}^n}(\mathbf{x})$ then we recover $\mathbf{d}_0 = \int_{R_x} \begin{bmatrix} s_1(\mathbf{x}) \\ s_2(\mathbf{x}) \\ . \\ . \\ s_n(\mathbf{x}) \end{bmatrix} dF_0(\mathbf{x})$. Hence $\mathbf{d}_0 \in Q$.

This shows that every cdf corresponds to a point in the convex set D and vice versa. Moreover, by noting the form

of $F_0(\mathbf{x})$ above, it can be concluded that each point in the set D can be represented by a distribution function of at most n steps. Since the existence of the optimal distribution function is guaranteed by Kakutani's result, it follows that it is equivalent to a step function of at most n steps. An identical argument holds for proving that $G^*(\mathbf{z})$ is a step function of at most m steps. ■

The above result reduces the problem of computation of the optimal distribution functions to a finite dimensional optimization over convex sets. This can be seen by using claim (2) of theorem 3 in (15). Since $Q = D$ and $P = H$, (15) reduces to

$$\min_{\mathbf{S}\in D}\max_{\mathbf{V}\in H}\mathcal{M}(\mathbf{S},\mathbf{V}) = \max_{\mathbf{V}\in H}\min_{\mathbf{S}\in D}\mathcal{M}(\mathbf{S},\mathbf{V}). \qquad (16)$$

Here the sets D and H are convex subsets of n and m dimensional Euclidean spaces, respectively. Moreover, the sets D and H are easy to compute using (13) and (14). Once we have calculated D and H, we can compute the values $(\mathbf{S}^*, \mathbf{V}^*) \in D \times H$ which achieves the equilibrium in (16), using, e.g., the IER algorithm described in Section IV-A. But claim (2) of theorem 3 also asserts that any point in D corresponds to a distribution function of at most n steps. Hence we can recover the equivalent n step cdf $F^* \in \mathcal{F}$ corresponding to $\mathbf{S}^*$ and the m step cdf $G^* \in \mathcal{G}$ corresponding to $\mathbf{V}^*$.

IV. Computation of the Optimal CDF

A. The IER algorithm

Various researchers have studied efficient numerical algorithms for solving min-max problems similar to (16). (e.g. see [17] and the references therein). In this article we propose to use a method of outer approximation first proposed in [17] and modified using an entropic regularization in [15]. We follow the authors of the latter paper in calling this algorithm as Iterative Entropic Regularization (or IER). This method has been shown to have guaranteed convergence properties to the min-max solution. However, in using this method one needs to solve a sequence of maximization and/or minimization problems with linear cost function but arbitrary convex constraint sets. First we briefly describe the IER algorithm to make this presentation self-contained.

Iterative Entropic Regularization:: Let us rewrite Problem 1 in terms of (16) simply as: Find $(\mathbf{S}^*, \mathbf{V}^*) \in D \times H$ such that $\mathcal{M}(\mathbf{S}^*, \mathbf{V}^*) = \min_{\mathbf{S}\in D}\max_{\mathbf{V}\in H}\sum_{j=1}^{m}\sum_{i=1}^{n} a_{ij}S_iV_j$ where $\mathcal{M}(\mathbf{S},\mathbf{V})$ is defined as in (11). Further define that $\mathcal{M}^*(\mathbf{S}) = \max_{\mathbf{V}\in H}\mathcal{M}(\mathbf{S},\mathbf{V})$.

Next approximate H by a finite subset $H^m := \{\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_m\}$ of m points. Correspondingly $\mathcal{M}^*(\mathbf{S})$ is approximated by $\mathcal{M}^*_m(\mathbf{S}) = \max_{\mathbf{V}\in H^m}\mathcal{M}(\mathbf{S},\mathbf{V})$. However this approximate function is still not differentiable and hence we use the following smoothed version using entropic regularization ([17] and [15]):

$$\mathcal{M}^*_{m,p}(\mathbf{S}) = (1/p)\log\left\{\sum_{\mathbf{V}\in H^m}\exp\{p\mathcal{M}(\mathbf{S},\mathbf{V})\}\right\}$$

where $p > 0$. Then the algorithm can be written as follows:

1) Select $\mathbf{V}_1 \in H$ and let $H^1 := \{\mathbf{h}_1\}$;$m = k = 1$. Choose $\delta \in (0,1)$, and $p > 0$.
2) Find $\mathbf{S}_{m,p} \in D$ satisfying $\mathcal{M}^*_{m,p}(\mathbf{S}_{m,p}) \leq \min_{\mathbf{S}\in D}\mathcal{M}^*_{m,p}(\mathbf{S}) + \delta^k$. Increase the iteration count k by one.
3) If (i) $\mathcal{M}^*(\mathbf{S}_{m,p}) \leq \mathcal{M}^*_{m,p}(\mathbf{S}_{m,p})$ and (ii)$\delta^k + \log(m)/p$ is below a desired tolerance, then stop. If (i) is violated then choose any $\mathbf{V}_{m+1} = arg\max_{\mathbf{V}\in H}\mathcal{M}(\mathbf{S}_{m,p},\mathbf{V})$, set $H^{m+1} = H^m \cup \{\mathbf{V}_{m+1}\}$, increase m by one, select $p \geq (\log(m))^2$ and go to step 2. If (ii) is violated increase p by a constant factor and go to step 2.

This algorithm is guaranteed to converge to the global solution at step 3 with at most a $(\delta^k + \log(m)/p)$ error; or it produces an infinite sequence of solutions $\mathbf{S}_m \in D(m \to \infty)$, any cluster point of which is a global solution.

B. Numerical Example

In this example, we study the effects of a single control variable (x) and a single noise variable (z) on the response y for repeated runs. A 2^2 factorial design with four added center runs is used. The response values for this example are simulated using the R (Version 2.7.1) software. The interest is in determining the value of x which minimizes the response for all settings of z. The regions R_x and R_z are considered to be the closed intervals $[-1, 1]$. The fitted model is $\hat{y}(x,z) = 4 + x - x^2 - 0.05z + 3xz$. For simplicity and ease of computation, we assume that the target is to keep the response $\hat{y}$ as small as possible. Clearly, this scenario is a subclass of the theory developed in this article. For example, one may choose T to be a extremely low value to fit this problem into Problem 1. Equivalently, the error criterion of (2) and (10) can be chosen simply as, $M(x,z) = \hat{y}(x,z) = 4 + x - x^2 - 0.05z + 3xz$. We would like to apply the theory developed in Section III for repeated process runs in this example. Clearly $M(x,z)$ is a separable function and hence it meets the requirements for applying theorem 3. Using the notation of (11), let $s_1(x) = x$, $s_2(x) = x^2$ and $v_1(z) = z$. The expected value of $M(x,z)$ is,

$$\begin{aligned}\mathcal{M}(F(\mathbf{x}), G(\mathbf{z})) &= 4 + S_1 - S_2 - 0.05V_1 + 3S_1V_1,\\ &(=: \mathcal{M}(S_1, S_2, V_1))\end{aligned}$$

where

$$S_1 = \int_{-1}^{1} x\, dF(x), \quad S_2 = \int_{-1}^{1} x^2\, dF(x), .$$

$$\text{and } V_1 = \int_{-1}^{1} z\, dG(z)$$

Here

$$Q = \{(S_1, S_2) \in \mathcal{R}^2 : F \in \mathcal{F}\} \text{ and } P = \{V_1 \in \mathcal{R} : G \in \mathcal{G}\}$$

We aim to find the equilibrium solution $(S_1^*, S_2^*) \in Q$ and $V_1^* \in P$ such that

$$\begin{aligned}\mathcal{M}(S_1^*, S_2^*, V_1^*) &= \min_{(S_1,S_2)\in Q}\max_{V_1\in P}\mathcal{M}(S_1, S_2, V_1)\\ &= \max_{V_1\in P}\min_{(S_1,S_2)\in Q}\mathcal{M}(S_1, S_2, V_1)\end{aligned}$$

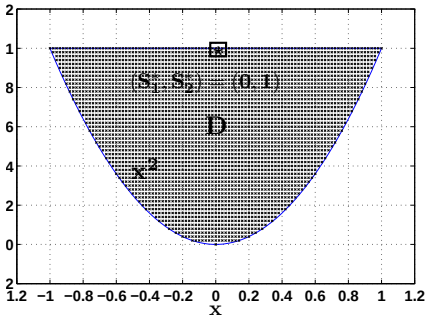

Figure 1: The set D

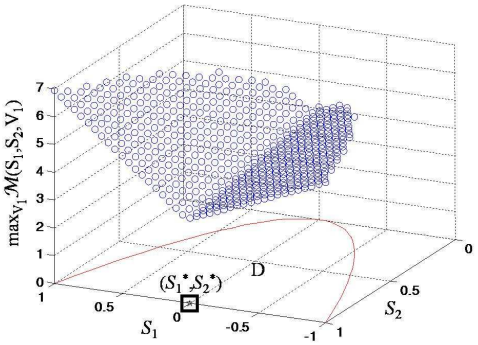

Figure 2: $\max_{V_1} \mathcal{M}(S_1, S_2, V_1)$ vs $S_1$and S_2 and the optimal choice $(S_1^*, S_2^*) = (0, 1)$

Even for this simple example, it is seen that these sets Q and P are difficult to compute. However, using Theorem 3 we know that the search of the equilibrium solution over the entire sets of distribution can be reduced to a search over the following simpler sets:

$$\begin{aligned} D &= \text{conv } \{(x, x^2) : -1 \le x \le 1\} \text{ and} \\ H &= \{z : -1 \le z \le 1\} \end{aligned}$$

Hence,

$$\begin{aligned} \mathcal{M}(S_1^*, S_2^*, V_1^*) &= \min_{(S_1,S_2)\in D} \max_{V_1 \in H} \mathcal{M}(S_1, S_2, V_1) \\ &= \max_{V_1 \in H} \min_{(S_1,S_2)\in D} \mathcal{M}(S_1, S_2, V_1) \end{aligned}$$

The set D is shown as the shaded region in figure 1. Set D is the convex closure of C. Using the IER algorithm, the optimal choice in terms of $(S_1, S_2) \in D$ which solves:

$$\min_{(S_1,S_2)\in D} \max_{V_1 \in H} [4 + S_1 - S_2 - 0.05V_1 + 3S_1V_1].$$

is $(S_1^* = 0, S_2^* = 1)$. The optimal choice $(S_1^*, S_2^*) = (0, 1)$ is shown in figure 1. Figure 2 plots the values of $\max_{V_1} \mathcal{M}(S_1, S_2, V_1)$ against $S_1$and S_2. Note from figure 2 that the minimum value of $\max_{V_1} \mathcal{M}(S_1, S_2, V_1)$ with respect to $S_1$and S_2 corresponds to $(S_1^*, S_2^*) = (0, 1) \in D$. We know from theorem 3 each point in D can be recovered from a cumulative distribution function of at most 2 steps. In this case it is easy to see that, (S_1^*, S_2^*) can be expressed as $\begin{bmatrix} S_1^* \\ S_2^* \end{bmatrix} = 0.5 \begin{bmatrix} s_1(x^a) \\ s_2(x^a) \end{bmatrix} + 0.5 \begin{bmatrix} s_1(x^b) \\ s_2(x^b) \end{bmatrix}$ for $x^a = 1$ and $x^b = -1$.

This implies that the optimal mixed strategy F^* is a step function of two steps, $F^*(x) = 0.5\, I_1(x) + 0.5\, I_{-1}(x)$, where

$$I_1(x) = \begin{cases} 0 & \text{if } x < 1 \\ 1 & \text{if } x \ge 1 \end{cases} \quad \text{and} \quad I_{-1}(x) = \begin{cases} 0 & \text{if } x < -1 \\ 1 & \text{if } x \ge -1 \end{cases}$$

V. Conclusion

A general methodology for computing the best possible sequence of control settings for an infinitely repeated production process is introduced. It is shown, using an idea from game theory, that the control parameters of such a process should be selected according to an optimal probability distribution. The polynomial structure of the fitted model is then exploited to present an efficient algorithm to compute the corresponding optimal cumulative distribution function. The proposed method is then demonstrated on a simulated example process.

References

[1] G. Taguchi, *Introduction to Quality Engineering*. White Plains, NY: UNIPUB/Kraus International, 1986.

[2] G. G. Vining and R. H. Myers, "Combining Taguchi and response surface philosophies: A dual response approach," *Journal of Quality Technology*, vol. 22, pp. 38–45, 1990.

[3] R. H. Myers, A. I. Khuri, and G. G. Vining, "Response surface alternatives to the Taguchi robust parameter design approach," *The American Statistician*, vol. 46, pp. 131–139, 1992.

[4] G. Taguchi and Y. Wu, *Introduction to Off-Line Quality Control*. Nagoya, Japan: Central Japan Quality Control Association, 1985.

[5] G. Taguchi, *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Cost*. White Plains, NY: UNIPUB/Kraus International, 1987.

[6] M. S. Phadke, *Quality Engineering Using Robust Design*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[7] C. F. J. Wu and M. Hamada, *Experiments: planning, analysis, and parameter design optimization*. New York: Wiley-Interscience, 2000.

[8] R. H. Myers and D. C. Montgomery, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, 2nd ed. New York: Wiley, 2002.

[9] W. J. Welch, T. K. Yu, S. M. Kang, and J. Sacks, "Computer experiments for quality control by parameter design," *Journal of Quality Technology*, vol. 22, pp. 15–22, 1990.

[10] K. A. F. Copeland and P. R. Nelson, "Dual response optimization via direct function minimization," *Journal of Quality Technology*, vol. 28, pp. 331–336, 1996.

[11] S. K. S. Fan, "A generalized global optimization algorithm for dual response systems," *Journal of Quality Technology*, vol. 32, pp. 444–456, 2000.

[12] J. von Neumann and O. Morgenstern, *Theory of games and economic behavior*. Princeton University Press, 1947.

[13] M. Dresher, *Games of strategy : theory and applications*. Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1961.

[14] P. A. Parrilo, "Polynomial games and sum of squares optimization," *Proceedings of 45th IEEE Conference on Decision and Control*, Dec 13-15 2006, san Diego, USA.

[15] R. L. Sheu and J. Y. Lin, "Solving continuous min-max problems by an iterative entropic regularization method," *Journal of Optimization Theory and Applications*, vol. 121, pp. 597–612, 2004.

[16] S. Kakutani, "A generalization of brouwer's fixed point theorem," *Duke Mathematical Journal*, vol. 8, pp. 457–459, 1941.

[17] E. Polak, *Optimization: Algorithms and Consistent Approximations*, 2nd ed. New York: Springer Verlag, 1994.