Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 12, No. 2, pp. 399–414, April 2016

ADAPTED CROSS ENTROPY METHOD TO INVESTIGATE COSTLY PRICE-CHANGES IN PRICING AND PRODUCTION PLANNING

Elham Mardaneh and Louis Caccetta

Abstract: In this paper, we propose an adaptation of the cross entropy (CE) method, called ACE, to solve an integrated production planning and pricing problem in which price change is not costless. More specifically, we consider a firm with a common production capacity shared amongst multiple products. We show that by using a chance constrained approach we can convert the case of uncertain price dependent demand to a model similar to the deterministic case. Both systems of fixed and variable price change costs are studied. This problem arises in the management of manufacturing systems where it is necessary to find a policy that is both economical and operational from the production perspective. The above problem is mathematically formulated as a mixed integer nonlinear program. Solving such problems is algorithmically very challenging. In fact, commercial codes fail to solve or even find a feasible solution to realistic size problems. The challenge originates from the fact that an optimum should be found despite the difficulty of finding even a feasible solution. The ACE method shows promise in solving optimisation problems regardless of continuity or other assumptions. In our approach, we sample the integer variables using the CE mechanism, and solve simplified nonlinear programming problems (NLP) to obtain corresponding continuous variables. Numerical results, on a range of test problems with sufficient complexity to reflect the difficulty of practical size problems demonstrate the effectiveness of our methodology.

Key words: *pricing, price-change cost, production planning, uncertain demand, cross entropy*

Mathematics Subject Classification: *90B30, 90C11, 90C90*

---

1 Background and Introduction

In recent years, companies benefit from the development of new technologies for innovative pricing strategies to improve their operations and optimise their outputs. Various tools have been developed and utilised for this purpose including: dynamic pricing over time; pricing to understand customer demand and target pricing for different classes of customers. The advantages of these tools include a potential increase in market share or profit and a decrease in demand or production variability. A recent review of the approaches to modelling and solving different variants of the joint pricing and inventory/production planning problem is provided by [5, 8, 20].

At present, a common assumption made in the joint pricing and inventory / production planning literature is that price changes are costless. However, different studies clearly show that price change costs are not negligible [1, 2, 11, 21]. Zbaracki et al. [23] study the pricing practices of a one billion dollar industrial firm and observe that the price change costs comprise 1.22% of the companys revenue and 20.03% of the companys net margin. In

In the economics literature, two kinds of price change costs are introduced, managerial costs and physical costs. The managerial costs are defined as the time and attention required of managers to gather the relevant information and to make and implement decisions. Physical costs result from changing thousands of shelf prices or producing, printing, and distributing price books and catalogues.

Despite the emphasis in the economics literature on the inclusion of price change costs in the pricing strategies, very few integrated pricing and production / inventory systems have explicitly incorporated such costs. There are only four papers [1,4,5,7] that are closely related to our work in which cost of price change is considered in joint pricing and production / inventory planning.

Motivated by the phenomenon in practice of large periods without nominal price changes and short periods with low prices, Aguirregabiria [1] proposes an inventory and pricing model in retailing firms that incorporates both fixed ordering costs and fixed price change costs but focuses more on empirical studies. The author considers a model where a monopolistically competitive retailer decides price and inventories, and assumes lump-sum costs when placing orders or changing nominal prices.

Celik et al. [4] focus on a continuous time revenue management problem with costly price changes. They characterize the optimal pricing policies for the case of ample inventory and develop several heuristics based on a corresponding fluid model. However, their model does not take into account the inventory replenishment decision.

Chen et al. [5] consider a single product periodic review inventory model with both fixed and variable price change costs. Demand was stochastic and price dependent. The problem is to coordinate the pricing and inventory replenishment decisions in each period to maximize the total discounted profit over a finite planning horizon. They develop a general model and characterize the optimal policies for two special scenarios, a model with backorders and no fixed price change costs and a model with fixed price change costs and no backorders. For the general problem with fixed plus linear variable price adjustment costs, they develop an intuitive heuristic policy to control inventory and set selling prices. Compared with the optimal policy, which is difficult to compute, their heuristic policy is amenable to implementation, as they showed for examples up to 12 periods in the planning horizon.

Chen and Hu [7] assume that if price in a period is different from the previous period, then a cost of changing the price tags, independent of the magnitude of the price change, is incurred. Considering both fixed ordering cost and the price change cost makes their model more complicated. However, under the assumption of no capacity constraint, they develop exact algorithms to find the optimal order quantities and selling prices for the single product. Their idea is to partition the planning horizon into some consecutive fixed priced periods. The total profit is then appropriately allocated to each member of the partition. For each member they solve a joint inventory and fixed pricing problem to find a single constant price maximizing its allocated profit. Based on this result, authors develop an acyclic network so that solving their inventory and pricing model with price change cost is equivalent to finding a longest path in the acyclic network. However, all of the above named papers consider a single product setting.

To our knowledge, there is no contribution in the literature dealing with the joint pricing and production planning of multiple products with costly price change. We contribute towards addressing this gap in the literature by incorporating price change costs into integrated production and pricing models for a multiple product setting. More specifically, we consider a firm with a common production capacity shared amongst multiple products. The system under study has a multi period planning horizon with both deterministic and uncertain price dependent demand. The objective of the firm is to coordinate the pricing and production planning decisions in each period to maximize its expected total profit over a finite horizon. First, the case of deterministic demand is studied through developing mathematical programming models for only variable or fixed and variable price change cost components. The fixed cost component is independent of the price change magnitude and the variable cost component relies on the amount of adjustment. We next bring into the models the uncertainty of the price dependent demand.

It turns out that our problem is computationally difficult, because both continuous and integer variables appear in a nonlinear objective function and some nonlinear constraints. The problem falls into the category of mixed integer nonlinear programming (MINLP) Problems. There are two general strategies for dealing with MINLPs: (1) relaxation of integer variables (by considering them as continuous or by fixing them) and (2) separation of the nonlinearity (to obtain easier sub problems) which leads to upper and lower bounds of the optimal objective value. Constructing linear or convex underestimations of overstated functions at certain points results in mixed integer linear or convex problems. Applying these strategies leads to algorithms such as Branch and Bound, Generalised Bender Decomposition, Outer Approximation and so forth (see [13, 14]). The shortcoming of these mentioned algorithms is that they, by their nature, do not tackle well optimisation problems that have non convex objective functions. Unfortunately, some of the MINLPs encountered in the joint pricing and production planning sector suffer from non convexity which makes it harder to find even a suboptimal solution. In addition to the above mentioned difficulties, the concept of large scale realistic size problems for MINLP is different from usual continuous programming problems. To overcome these complexities, our strategy is a novel use and adaption of the CE method.

The CE method was proposed by Rubinstein [16] for solving rare event simulation problems and extended to solving combinatorial problems. It has been successfully applied in several engineering problem [9, 17] and in this paper an extension of this approach for solving a MINLP is proposed and applied to solve the problem of joint pricing and production planning with cost of price adjustment.

The CE approach is a variant of Monte Carlo methods, which use the importance sampling (IS). Two of the main reasons why we use Monte Carlo methods are because of their anti aliasing properties and their ability to approximate quickly an answer that would otherwise be very time consuming. This last point refers to the fact that Monte Carlo methods are used to simulate problems that are too difficult and time consuming when using other methods. An example is the use of Monte Carlo techniques in integrating very complex multidimensional integrals. This is a task that other techniques cannot handle well, but which Monte Carlo can.

In our approach, we develop an effective computational methodology by using CE innovatively so as to eliminate either the nonlinearity or the integer restriction whilst maintaining feasibility. Our developed algorithm has inherited the easy programmable characteristic of CE with far few parameters compared to other random search methods. We incorporate the uncertainty of demand into the problem and develop a chance constrained model. Interestingly, by using the chance constrained programming approach the uncertain demand is treated without the creation of too much complexity into the model.

The rest of this paper is organized as follows. In Section 2, we describe the general problem and introduce the notation and terminology used in our models. The problem of joint pricing and production planning with deterministic demand is discussed first as two possible cases, without or with fixed price-change cost. Uncertain demand is next incorporated into the problem by developing a chance constrained model. Section 3 presents an overview of the CE method for optimisation. In Section 4 we bring the detail of the adapted CE method for the problem of joint pricing and production planning with costly price adjustments. Section 5 implements a numerical example for a range of algorithm parameters.

2 Models

This section formulates mathematically the problem of coordinated pricing and production planning of multiple products with costly price changes. Some important features of the problem, which we model, are:

- The finite planning horizon consists of T discrete periods.
- The firm produces n non-perishable differentiated products.
- Demand of each product is dependent on its price.
- The production capacity is limited and shared among different products.
- Products use the same amount of capacity; here each product uses one unit of capacity.

  Remark. This feature is not limiting as the case of different capacity use can only change the coefficient of the related function and is still solvable with the proposed algorithm.
- The production setup cost is negligible.

  Remark. Note that this feature of the problem can be extended to the case in which the setup cost is considerable. Our solution methodology can also cater for this category of problems.

In this paper, settings with an initial product price are investigated, where a specific price of each item either has been inherited from a previous planning horizon or obtained by communication with the marketing decision makers. In this kind of settings, the firm has the option of keeping the current price at the beginning of the new selling period, or changing it to some other pricing level.

Because of the complexity of the problem we first present the deterministic demand / price function. Then, randomness is added to the induced demand of each item in each period.

2.1 Parameters and Decision Variables

We consider situations in which different products are presented to different market sectors. Hence there is no interaction between the price of one product and the demand of one another product or in other words, the cross price elasticity among various products is zero. The firm controls the price and production amount of each item at all time periods. We assume that backorders are not allowed.

We make use of the following parameters and decision variables in our models.

Parameters:
c_{jt} = the production cost of one unit of item j in period t.
h_{jt} = the holding cost of one unit of item j in inventory in period t.
a_{jt} = the variable cost of price change of item j in period t.
a_j = the fixed cost of price change of item j.

L_j = the minimum acceptable price change of item j.
U_j = the maximum acceptable price change of item j.
K_t = the total amount of shared production capacity in period t.
I_j^0 = the initial inventory of item j.
p_j^0 = the initial price of item j.

Variables:
p_{jt} = the price of item j in period t.
x_{jt} = the amount of item j produced in period t.

Functions:
$D_{jt}(p)$= correlation between demand and price of item j in period t.
I_{jt}= the inventory level of item j at the beginning of period t.
Note that, here only the relationship between demand and price is known, but demand is the result of price decision.

2.2 Variable Price-Change Cost

In the economics literature, several forms of variable cost have been used, including piecewise linear functions [21, 22] and quadratic functions [15]. In this paper, we assume a piecewise linear function for the convex variable cost of price change. Defining

$$y_{jt} = \begin{cases} 1, & \text{if } |p_{jt} - p_{j,t-1}| \neq 0 \\ 0, & \text{if } |p_{jt} - p_{j,t-1}| = 0 \end{cases}$$

we express the problem of jointly determining the price and production plan for the case of variable price change cost. We introduce X=(x_{jt}), P=(p_{jt}) and Y=(y_{jt}) as the $n \times T$ production amount, price and price change indicator matrix, respectively.

Model 1:

$$max_{X,P,Y\geq 0} f(X,P,Y) = \sum_{t=1}^{T}\sum_{j=1}^{n} p_{jt}D_{jt}(p) - c_{jt}x_{jt} - h_{jt}I_{jt} - a_{jt}|p_{jt} - p_{j,t-1}| \tag{2.1}$$

subject to

$$\sum_{j=1}^{n} x_{jt} \leq K_t; \quad t \in \{1,2,...,T\}, \tag{2.2}$$

$$I_{jt} = I_j^0 + \sum_{s=1}^{t-1}(x_{js} - D_{js}(p)); \quad t \in \{1,2,...,T\} \; and \; j \in \{1,...,n\}, \tag{2.3}$$

$$I_j^0 + \sum_{s=1}^{T}(x_{js} - D_{js}(p)) = 0; \quad j \in \{1,...,n\}, \tag{2.4}$$

$$L_j y_{jt} \leq |p_{jt} - p_{j,t-1}| \; \leq U_j y_{jt}; \quad t \in \{1,2,...,T\} \; and \; j \in \{1,...,n\}, \tag{2.5}$$

$$x_{jt}, p_{jt}, I_{jt} \; \geq 0; \quad y_{jt} \in \{0,1\}; \quad t \in \{1,...,T\} \; and \; j \in \{1,...,n\}. \tag{2.6}$$

In (2.1) the objective is to maximise the profit consisting of sales revenue, production, inventory holding and price change costs respectively. Constraint (2.2) ensures that capacity in period t is sufficient to allow all of the production that is planned for the period for all n items. Constraint (2.3) is a set of flow balance equations that ensure that all of the induced demand is satisfied. We impose the zero inventories at the end of planning horizon by Constraint (2.4). In Constraint (2.5), the bounds of price change of each item at all time periods is incorporated into the model. The requirement in (2.6) that inventory be nonnegative assures that demand is satisfied in all periods $t = 1, 2, ..., T$ with no backorders. We formulate the upcoming models with fixed price change component based on Model 1.

2.3 Fixed and Variable Price-Change Cost

Now, we bring into account the fixed price change cost which is independent of the price change magnitude. In this case the convexity of the cost is not valid anymore. The cost of price change of item j with magnitude $|p_{jt} - p_{j,t-1}|$in period t is denoted by A_{jt}, which is given as:

$$A_{jt} = a_j y_{jt} + a_{jt}|p_{jt} - p_{j,t-1}|.$$

Where the fixed cost a_j represents the menu cost or physical cost associated with a price change, and the variable cost $a_{jt}|p_{jt} - p_{j,t-1}|$ represents the managerial or customer cost depending on the magnitude of the price change.
At the beginning of the planning horizon, given the problem parameters, initial inventory level and price suggestion, the firm's aim is to maximise the total profit. The mathematical programming problem can be expressed as follows:

Model 2:

$$max_{X,P,Y\geq 0} f(X,P,Y) = \sum_{t=1}^{T}\sum_{j=1}^{n} p_{jt}D_{jt}(p) - c_{jt}x_{jt} - h_{jt}I_{jt} - A_{jt} \tag{2.7}$$

subject to (2.2) – (2.6).

The objective function is again maximising the sales revenue minus production, inventory holding and price change costs respectively. As can be seen both models formulate their corresponding problem as a mixed integer nonlinear programming problem with the nonlinearity in both objective function and absolute value terms in constraints.

Interestingly, the introduction of fixed price change cost does not change the size of problem in Model 1 compared to Model 2 (both constraints and variables). The reason is that the strict non zero lower bound on the price change (L_j) requires introduction of the binary variable y_{jt} for the constraints (2.5). A special case of Model 1 is with zero lower bound ($L_j = 0, \forall j$) which leads to elimination of binary variables and the problem formulations of Model 1 diminishes to the category of nonlinear programming problems instead of mixed integer nonlinear programing problems. However, a zero lower bound does not seem appropriate in practice because it may result in tiny insignificant price changes.

2.4 Uncertainty Consideration

We consider again a firm selling n differentiated products. The time horizon is finite and time is viewed discretely. The firm controls the price p_{jt} and production amount x_{jt} of each product at all times with fixed and variable cost of price change. We assume that no backorder is allowed and the uncertain demand model is additive. Given the uncertain as well as the dynamic nature of the problem, we consider an open-loop approach to the optimisation formulation. The model incorporates uncertainty not only in the objective function but also in the constraints. In the stochastic framework, we deal with constraint uncertainty by using the modelling tool of chance constrained.

We denote $\widetilde{D}_{jt}(p) = D_{jt}(p) + \varepsilon_{jt}$ as the realized demand, involving realized noise ε_{jt} which follows normal distribution $(0,\sigma_{jt}^2)$. Normally distributed noises are frequently used in the literature [3,12,18]. In a setting with uncertainty, notice that the inventory constraint is on the realized inventory level $\tilde{I}_{jt} = I_j^0 + \sum_{s=1}^{t-1}(x_{js} - \widetilde{D}_{js}(p))$.

Similarly, realized holding costs involve the realized inventory levels, and realized profits involve realized demand. Assuming that the probability distributions of the random variables ε_{jt} and the failure margins α_{jt} are known, the following chance constrained model arises:

Model 3:

$$max_{X,P,Y\geq 0} E[f(X,P,Y)] = E[\sum_{t=1}^{T}\sum_{j=1}^{n} p_{jt}\widetilde{D}_{jt}(p) - c_{jt}x_{jt} - h_{jt}\tilde{I}_{jt} - A_{jt}] \tag{2.8}$$

subject to (2.2), (2.5) and

$$\tilde{I}_{jt} = I_j^0 + \sum_{s=1}^{t-1}(x_{js} - \widetilde{D}_{js}(p)); \quad t \in \{1,2,...,T\} \textit{ and } j \in \{1,...,n\}, \tag{2.9}$$

$$I_j^0 + \sum_{s=1}^{T}(x_{js} - \widetilde{D}_{js}(p)) = 0; \quad j \in \{1,...,n\}, \tag{2.10}$$

$$Pr(\tilde{I}_{jt} \geq 0) \geq 1 - \alpha_{jt}; \quad t\{1,2,...,T\} \textit{ and } j \in \{1,...,n\}, \tag{2.11}$$

$$x_{jt}, p_{jt} \geq 0, y_{jt} \in \{0,1\}; \quad t \in \{1,...,T\} \textit{ and } j \in \{1,...,n\}. \tag{2.12}$$

Because of the linearity of the uncertain data relating to the inventory levels and profit, the expected profit (2.8) equals the nominal profit (the mean value which is the same as deterministic case). Constraint (2.9) is a set of realised flow balance equations ensuring that all of the induced demand is satisfied. We impose the zero realised inventories at the end of planning horizon by Constraint (2.10). The requirement that the realised inventory be nonnegative assures that the induced demand is satisfied in all periods with no backorders. The complexity of solving this problem depends essentially on the structure of the chance constraints (2.11). In particular, for this problem, these constraints involve the convolution of distributions of uncertain parameters. Therefore, the problem becomes more complex if $\varepsilon_{jt}s$ follow different distributions or if the probability distribution is not stable (i.e. a linear combination does not necessarily follow the same type of distribution).Here, we investigate the chance constrained:

$$Pr(\tilde{I}_{jt} \geq 0) = Pr(I_j^0 + \sum_{s=1}^{t-1}(x_{js} - \widetilde{D}_{js}(p)) \geq 0) = Pr(\sum_{s=1}^{t-1}\widetilde{D}_{js}(p) \leq I_j^0 + \sum_{s=1}^{t-1} x_{js}).$$

With the assumption that the random parts of uncertain demands of different products in any period are mutually independent and identically distributed (i.i.d), by using $\widetilde{D}_{jt} \sim$

$N(D_{jt}(p), \sigma_{jt}^2)$ and the convolution of the normal distributions we have:

$$\sum_{s=1}^{t-1} \widetilde{D}_{js} \sim N(\sum_{s=1}^{t-1} D_{js}(p), \sum_{s=1}^{t-1} \sigma_{js}^2).$$

Let $\varphi(z)$ be the cumulative distribution function of the standard normal random variable. As a result we rewrite the chance constrained as

$$Pr(\tilde{I}_{jt} \geq 0) = \varphi\left(\frac{I_j^0 + \sum_{s=1}^{t-1}(x_{js} - D_{js}(p))}{\sqrt{\sum_{s=1}^{t-1} \sigma_{js}^2}}\right) \geq 1 - \alpha_{jt}.$$

Or equivalently:

$$\sum_{s=1}^{t-1}(x_{js} - D_{js}(p)) \geq \varphi^{-1}(1 - \alpha_{jt})\sqrt{\sum_{s=1}^{t-1} \sigma_{js}^2} - I_j^0. \tag{2.13}$$

Then a corresponding formulation of Model 3 is:
Model 4:

$$max_{X,P,Y \geq 0} f(X, P, Y) = \sum_{t=1}^{T} \sum_{j=1}^{n} p_{jt} D_{jt}(p) - c_{jt} x_{jt} - h_{jt} I_{jt} - A_{jt} \tag{2.14}$$

subject to $(2.2) - (2.5), (2.12)$ *and* (2.13).

As can be seen, by the above discussion we change the chance constrained (2.11) to the linear constraint (2.13) by using the cumulative distribution function of the normal distribution.
In the next Section we will present the main CE method for optimisation and its adaption to handle our mentioned MINLP in Model 2 and Model 4.

3 The Cross Entropy Method for Optimisation

The cross entropy method was introduced by Rubinstein [16] as an adaptive algorithm based on variance minimization for estimating the probabilities of rare events in complex stochastic networks. It was soon realised that by implementing simple modifications, CE could be used not only for estimating the probabilities of rare events, but also for solving difficult combinatorial optimisation problems. This is done by transforming the deterministic optimisation problem into a related stochastic optimisation problem and then using the rare event simulation techniques proposed by [16]. Several recent applications demonstrate the power of the CE method as a generic and practical tool for solving NP-hard problems [10,19].

The CE method is an iterative method involving the following two phases:

- Generation of a sample of random data (trajectories, vectors, etc.) according to a specified random mechanism.
- Updating the parameters of the random mechanism, on the basis of the data collected, to produce a better sample in the next iteration.

The main idea of CE in the optimisation setting is to start with an initial probability distribution function over a feasible region (typically, a uniform distribution), and then update it adaptively based on a random sample collected from the feasible region. This revision process should converge (ideally) to a discrete uniform distribution that assigns equal positive probabilities to global optima and 0 elsewhere. Although the CE method is still evolving, its performance has been promising in a variety of applications [10,19]. Useful materials related to CE can be found in [17].

Now, let us consider a general optimisation where $X \subset \mathrm{R}^n$ and $S(x)$ is a real value objective function. The basic idea of CE for this problem is to use a stochastic approach based on importance sampling in order to get a good evaluation of $\hat{x}$

$$\hat{x} = argmax_{x \in X}(S(x)) \tag{3.1}$$

We start from a priori probability distribution function (PDF) to draw samples and iteratively compute series of PDF that increase the probability to draw samples near the optimal solutions. With respect to the main CE method the event that is used to iteratively compute the PDF is not given by the problem but has to be chosen.
The PDF updating step translates into the following: Given a random sample of size M, the parameters are updated based on the $M^{elite} = \varrho M$ best performing samples.These are called the elite samples. The updated parameters are found to be the maximal likelihood estimates (MLEs) of the elite samples [17].

In particular, for the normal distribution, the parameter updating is especially simple. In the next section we will adapt the idea of using CE method in optimisation to solve the MINLPs specified by Model 2 or Model 4 where $X \subset \mathrm{R}^{(3T-1)n} \times \{0,1\}^{nT}$.

4 Adapted Cross Entropy method for Pricing and Production Planning

As we have seen earlier the problem of joint pricing and production planning of multiple products becomes a mixed integer nonlinear programming problem, if we consider a non-zero lower bound on the amount of price change. Our intention to simplify the above mentioned problem with the nonlinearity in both objective function and some constraints is to use CE method on a partial set of variables (price change indicators). By this adaption, for a specific vector of price change indicator (Y) the optimisation problem becomes a simplified nonlinear programming problem instead of a large mixed integer nonlinear programming problem, in which all constraints are linear.

To explain our approach, note that any of the objective functions (2.1), (2.7) *and* (2.14) can be written as:

$$max_{X,P,Y \geq 0} f(X,P,Y) = max_{Y \geq 0} max_{X,P \geq 0} f(X,P,Y) \tag{3.2}$$

where maximizations are taken with respect to the related constraints to each problem setting. For any fixed and feasible vector Y, the inner maximization on the right-hand side of (3.2) can be performed using any nonlinear programming algorithm and results in the accompanying X and P variables. The outer maximization is performed by the proposed adapted CE method. Our approach is to develop an algorithm that involves a sequence of generating random vectors of y_{jt} variables ($j \in \{1,...,n\}$ *and* $t \in \{1,...,T\}$) and solving the corresponding nonlinear programming problem. Through CE, we construct a sequence of transition probability matrices whose entries will, ultimately, be concentrated on the optimal pricing strategies. It is likely to encounter infeasibility in nonlinear programming problems for some Y candidates, which we will explain later how to overcome this difficulty.

We assume that the k^{th} binary variable, the k^{th} entry of Y, is a discrete random variable with a PDF defined by $Pr\{Y_k = 0\} = Pr\{Y_k = 1\} = 0.5$ for $k = 1, ..., nT$. These PDF's define a $nT \times 2$ initial distribution matrix P^0 as

$$P^0 = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \\ \vdots & \vdots \\ 0.5 & 0.5 \end{bmatrix}.$$

At the end of each iteration of the algorithm, the maximum of objective function value at the points (Y_i, X_i, P_i) becomes a lower bound for the optimal value of the problem. At each point (Y_i, X_i, P_i), Y_i is a sampled binary variable vector, X_i and P_i are accompanying production and price vectors found by solving the corresponding nonlinear programming problem. The stop criterion for the algorithm is to find a probability distribution matrix very close to a certainty matrix (this defines a very high probability to one possible value of 0 *or* 1). This is specified by the Euclidian norm of the distribution matrix. In each iteration k, we check if $|\|C\|_2 - \|P^K\|_2| \leq \varepsilon$, where C represents the certainty distribution matrix and ε is a given tolerance.

In matrix C each row has two components, one of them one and the other one zero, because it tells by certainty each y_{jt} is zero or one, for instance:

$$C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ \vdots & \vdots \\ 0 & 1 \end{bmatrix},$$

$$\|C\|_2 = \sqrt{0^2 + 1^2 + 1^2 + 0^2 + ... + 0^2 + 1^2} = \sqrt{nT}$$

It can be easily demonstrated that among those matrices we consider for the probability distribution of y_{jt}s, matrix P^0 has the minimum Euclidian norm whilst matrix C has the maximum Euclidian norm. As mentioned earlier, one possible difficulty that arises is when the current Y vector is infeasible with respect to the nonlinear programming problem's constraints. If this occurs, we search neighbourhood of Y to find a feasible binary vector. As our problem consists of only binary variables for the integer part, we limit our neighbourhood search to those points which bask in $\{-1, 0, 1\}$ apart from Y. The resulting vector should remain binary, so we keep track of original problem's feasibility by projection of newly generated vector to the nearest binary values before going any further. The Adapted Cross Entropy algorithm is detailed in Table 1.

Step 1 As suggested in [17], we have used the discrete uniform distribution for the initial probability matrix.

Step 2 When the y_{jt} vector is fixed, integer restrictions of the problem disappear. The degree of nonlinearity in the objective function and constraints reduces. As a result the main mixed integer nonlinear programming problem is simplified to a nonlinear programming problem with x_{jt} and p_{jt} variables and mostly linear constraints. If the resulting nonlinear programming problem is infeasible for a Y_i candidate, we conduct a local search to find a feasible neighbour. Assume that Y_0 is an infeasible candidate, we generate a random vector of size nT, called U_0 with entries in $-1, 0, 1$. A new Y vector is calculated as $Y_1 = U_0 + Y_0$, if Y_1 is not feasible again, another vector U_1

Table 1: Adapted Cross Entropy (ACE) for Costly Price-Change Pricing and Production Planning.

Step 0	Initialise by giving sample size and elite sample size, M and M^{elite} respectively; obviously, $0 < M^{elite} \leq M$. Set $z = 0$ *and* $L_0 = 0$ (initial lower bound). Initialise α, a parameter from the interval $[0, 1]$. Initialise r, an integer parameter (e.g. $r = 5$). Initialise ε , a tolerance parameter (e.g. 10^{-6}).
Step 1	Generate initial sample of binary vectors $\{Y_1, ..., Y_M\}$ using P^0 as the initial distribution matrix. Note that each $Y_i; i \in \{1, ..., M\}$ is a vector of y_{jt}s; $j \in \{1, ..., n\}$ *and* $t \in \{1, ..., T\}$.
Step 2	For each sample vector $Y_i; i \in \{1, ..., M\}$, calculate the related continuous part X_i and P_i by solving the resulting nonlinear programming problem. If the point is infeasible, do a local search among neighbours of Y_i until a feasible tuple (Y'_i, X'_i, P'_i) is obtained.
Step 3	Evaluate the performance of each tuple (Y_i, X_i, P_i), called (Y_i) by checking the objective value. Then sort Y_is in descending order with respect to values of (X_i, P_i, Y_i). This creates a sequence of statistics $(Y_{(1)}), f(Y_{(2)}), ..., f(Y_{(M)})$ with the property: $Pr(f(Y_{(1)}) \geq f(Y_{(2)}) \geq ... \geq f(Y_{(M)}))) = 1.$ Now, consider the best M^{elite} generated solutions $Y_{(1)}, Y_{(2)}, ...Y_{(M^{elite})}$, called the elite sample.
Step 4	If $f(Y_{(1)}) < L_z$ then $L_{z+1} = f(Y_{(1)})$ and $Y^* = Y_{(1)}$, where Y^* represents the current best performing set of y_{jt} variables; otherwise $L_{(z+1)} = L_z$.
Step 5	If $z \geq r$ and $L_{z+1} = L_z = ... = L_{z-r}$, then STOP and claim that Y^* is an optimal set of y_{jt} variables. Otherwise, go to step 6.
Step 6	If $\|\|C\|_2 - \|P^z\|_2\| \leq \varepsilon$ then STOP and claim that Y^* is an optimal set of y_{jt} variables. Otherwise, go to step 7.
Step 7	Use the elite sample to update the probability distribution matrix by applying the following equations: $\tilde{P}^{z+1}_{y_{jt}}(x) = \frac{Ny_{jt}(x)}{M^{elite}}$, $P^{z+1}_{y_{jt}}(x) = (1-\alpha)P^z_{y_{jt}}(x) + \alpha\tilde{P}^{z+1}_{y_{jt}}(x)$ *for* $x = 0, 1,$ where, $Ny_{jt}(x)$ is the number of times among elite sample such that discrete random variable y_{jt} takes the value of x.
Step 8	Generate a new sample using the updated probability matrix P^{z+1}, set $z = z + 1$ and go to step 2.

is generated with entries in $-1, 0, 1$ and then the vector $Y_2 = U_1 + Y_1$ is checked for feasibility. This process continues until a feasible vector of binary variables is found. As the U vectors are generated randomly, we don't get trapped in a sick loop. Also, the feasible region for binary variables is finite which assures getting the feasible vector in finite number of steps.

Step3 As the objective function (profit) is maximized, we sort the performance of each sample by the value found for the nonlinear programming problem's objective function. Now in this step, the elite sample can be derived for the current iteration, which consists of the best performing samples.

Step4-6 In these steps the stop criterion is checked. If the value of the best performing objective function is better than the current lower bound, lower bound will be updated, otherwise it remains the same. Then, lower bound is compared to some previous iteration's lower bounds, if any improvement has been made, more investigation is needed, so the algorithm proceeds to step 6. If lower bound has not changed for the last r iterations, stop.

Step7 Based on elite sample, the probability matrix is updated. For each y_{jt} variable, a secondary PDF, $\tilde{P} = Pr\{y_{jt} = (x)\}$ is calculated by evaluating the fraction that the variable took the x value ($x = 0$ *or* 1). Then, the main probability matrix is obtained by using a smoothing parameter to incorporate the historical matrices into play. One desirable consequence of using smoothing parameter is to reduce the probability that some components of the probability matrix will be zero or one in the early iterations.

Step8 The next iteration of the algorithm starts here. A new sample is generated based on the updated matrix which is closer to the optimal solution. To obtain a new binary vector at the iteration z , as a single member of the new sample, we use the matrix P^z in the following manner. First, the cumulative probability matrix, $C^z_{nT\times 2}$, is calculated using the probability matrix P^z as follows:
$c^z_{i0} = P^z_i(0); \quad \forall\, i \in \{y_{jt}; j = 1, ..., n \textit{ and } t = 1, ..., T\}$,
$c^z_{i1} = P^z_i(0) + P^z_i(1) = 1; \quad \forall\, i \in \{y_{jt}; j = 1, ..., n \textit{ and } t = 1, ..., T\}$.
In other words, $C^z_{(nT\times 2)}$ is a matrix in the following form:

$$C = \begin{bmatrix} P^z_{y_{11}}(0) & 1 \\ P^z_{y_{12}}(0) & 1 \\ \vdots & \vdots \\ P^z_{y_{nT}}(0) & 1 \end{bmatrix} = \begin{bmatrix} Pr\{y_{11} = 0\} & 1 \\ Pr\{y_{12} = 0\} & 1 \\ \vdots & \vdots \\ Pr\{y_{nT} = 0\} & 1 \end{bmatrix}$$

Next, a random vector $(u_1, ..., u_{nT})$, where each u_i is sampled from $U[0,1]$ is generated. A new binary vector $\tilde{Y} = (\tilde{y}_{11}, \tilde{y}_{12}, ..., \tilde{y}_{nT})$ is constructed by applying the rule: For each $i \in \{y_{jt}; j = 1, ..., n \textit{ and } t = 1, ..., T\}$ if $u_i \leq c^z_{i0}$, then $[\tilde{Y}]_i = 0$; otherwise $[\tilde{Y}]_i = 1$. This process is repeated M times to produce the next sample, which is likely to be better than previous samples, as it has more knowledge about high values of the objective function that was inherited from the previous elite samples.

Remark. A similar approach that fixes the continuous part (price variable) and finds the integer part of each feasible solution is also investigated. A big difference between the two approaches is that in the latter one, a linear programming problem needs to be solved at each sample point rather than a non-linear program. In fact, the linear sub-problem is a simple production planning problem. Moreover, more options are available to choose as the probability distributions such as the normal, beta, or some others. However, because of the infinite set of selections for the price variables, the joint pricing and production planning problem is not successful in converging to an optimal solution.

5 Numerical Results

To investigate the performance of the proposed method, we have chosen a series of problems with $n = 5$ products and $T = 12$ periods. As we consider situations in which different class of products are presented to different market sectors, selection of $n = 5$ products over a horizon of $T = 12$ periods (seasonal: 3 years, monthly: 1 year, weekly: 3 months) can cover a good range of realistic size problems.

We developed a set of test problems using uniformly distributed random parameters. The problems vary in complexity based on the chosen parameters, as the feasible area may become very tight. The solving time of generated examples is calculated by using a standard personal computer with an Intel 2 Core' CPU and 2.00 GB of RAM. So, even a small improvement in the computation time can be significant for tight problems which take a long time to be solved. We used MAPLE 14 software to code the algorithm and implement its NLPSolve procedure to solve nonlinear sub problems.

In order to compare the efficiency of the ACE method with the commercial MINLP solvers, we used the commercial software AIMMS in which, the algorithm used in handling the MINLP is Outer Approximation (see [14]). For Nonlinear sub-problems, the solver CONOPT and for the master Mixed-Integer Linear problems the solver CPLEX is called. By comparing the AIMMS results and our developed algorithm in MAPLE, we found out that there is at least 11% improvement in the local optimal objective value if we choose the ACE algorithm.

To report the numerical potential of the ACE method in solving Costly Price Adjustments in Joint Pricing and Production Planning, we run the algorithm with different settings to solve a randomly generated problem. Here, the algorithm is tested with different values for α, the parameter that determines how much information we extract from the history, different values for the elite sample size, M^{elite} and different value for the sample size M. Figure 1 presents the performance of the ACE with different values of α. The parameter plays a critical role in the performance of the algorithm by influencing the probability matrix used at each iteration:

$$P^{z+1}_{y_{jt}}(x) = (1-\alpha)P^{z}_{y_{jt}}(x) + \alpha \tilde{P}^{z+1}_{y_{jt}}(x)$$

In this equation, the quantity $\alpha \in [0, 1]$ is the coefficient of the most recent probability matrix, and $(1 - \alpha)$ is for the previous probability matrix which captures, to some extent, the history of searches among feasible points. Therefore, by assigning a smaller α, the method concentrates more on the old matrix, and hence behaving as a global search. In contrast, by assigning higher values for α, the algorithm uses less of the history and, thereby plays more of a role of a local search.

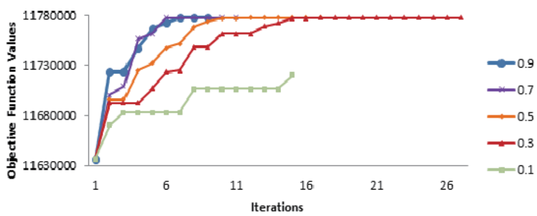

Figure 1: ACE performance for different α values

We run the ACE method for values of α from $\{0.1, 0.3, 0.5, 0.7, 0.9\}$ while the sample size was fixed at $M = 100$ and the elite sample size was fixed at $M^{elite} = 10$. The main observation to be drawn from Figure 1 is that the algorithm rapidly increases the objective functions value. Another important observation is that as the value of α decreases, the algorithm needs to evaluate more points with more iteration and hence takes longer to converge.

Another important parameter which has considerable impact on ACE performance is the size of the elite sample. Using elite sample, the algorithm decides how much information it should use from the best solutions found so far. We run the ACE method for different values of $M^{elite} = \{10, 20, 50\}$while we have already chosen the best value for $\alpha = 0.9$ and fixed the sample size at $M = 100$. In Figure 2, the algorithm's performance is demonstrated with different choices of elite sample size.

As can be observed, for a larger elite sample, the number of iterations needed to converge to the optimal solution is more, the performance gets worse with larger values of elite sample

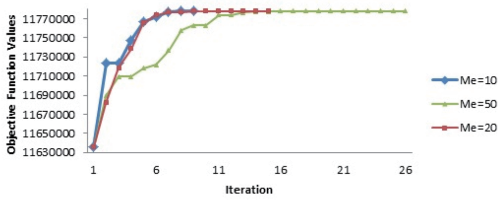

Figure 2: ACE performance for different elite sample size, M^{elite}

size. This should be the case, as intuitively is clear that by having more contribution from the not great solutions we actually postpone reaching to the optimal solution.
We also investigated the impact of the sample size on the performance of the ACE method. We test the algorithm for different sample sizes $M = \{100, 200, 500, 700, 1000\}$ while $\alpha = 0.9$ and fixed the elite sample size at $M^{elite} = 10$. Figure 3 illustrates the performance of the algorithm.

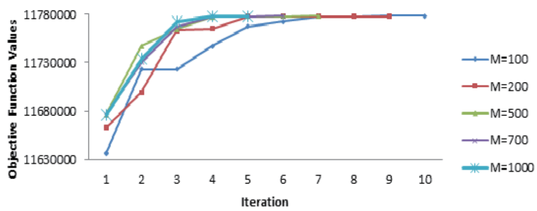

Figure 3: ACE performance for different sample size, M

As expected, the greater the sample size, the more of the feasible region is explored at each iteration. It is important to note that we may have to compromise between the sample size and the running time, generally larger samples need more of the CPU time. But, in our example, the greater the sample size, the less iteration needed to converge to the solution. Interestingly in our problem, each iteration is not highly affected by the sample size, so it's recommended to choose a bigger sample size to benefit from the less number of iterations.

The numerical results for 5 products, 12 periods problem may suggest an acceptable tuning for the ACE method. Our computational experiments suggest that the best choice of parameters for the problem under study are $\alpha = 0.9$, $M^{elite} = 10$ *and* $M = 1000$.

6 Conclusion

In this paper we incorporated price-change costs into integrated production and pricing models for a multiple-product setting. More specifically, we considered a firm with a common production capacity shared amongst multiple products. We showed that by using a chance constrained approach we can convert the computationally difficult case of uncertain price-dependent demand to a model similar to the deterministic case. Both systems of fixed and variable price-change costs were studied, and we focussed our effort on the adaption of the CE method to the more difficult model. Our developed Adapted Cross Entropy (ACE) algorithm tackles the formulated Mixed Integer Nonlinear Program in an iterative manner until a good optimal solution is found. The ACE method has a strong search property, and in a small number of iterations approached the optimal solution much better than the existing commercial packages. As numerical results showed, there is promise in the CE method to solve such large optimisation problems.

References

[1] V. Aguirregabiria, The dynamics of mark-ups and inventories in retailing firms, *Rev. Econ. Stud.* 66 (2001) 275–308.

[2] M. Bergen, M. Ritson, S. Dutta, D. Levy and M. Zbaracki, Shattering the myth of costless price change, *Euro. Management J.* 21 (2003) 63–669.

[3] Ben-Daya and M. Raouf, Inventory models involving lead time as a decision variable, *J. Oper. Res. Society* 45 (1994) 579–582.

[4] S. Celik, A. Muharremoglu and S. Savin, Revenue management with costly price adjustments, *Oper. Res.* 57 (2009) 1206–1219.

[5] X. Chen and D. Simchi-Levi, Pricing and inventory management, in *Handbook of Pricing Management*, Ozer Philips (edi.), Oxford University Press, 2011.

[6] X. Chen, S. Zhou and F. Chen, Integration of inventory and pricing decisions with costly price-changes, *Oper. Res.* 59 (2011) 1144–1158.

[7] X. Chen and P. Hu, Joint pricing and inventory management with deterministic demand and costly price-change, *Oper. Res. Lett.* 40 (2012) 385–389.

[8] W. Elmaghraby and P. Keskinocak, Dynamic pricing in the presence of inventory considerations: research overview, current practices, and future directions, *Management Sci.* 49 (2003) 1287–1309.

[9] D. Ernst, M. Glavic, G.B. Stan, S. Mannor and L. Wehekel, The cross-entropy method for power system combinatorial optimisation problems in *Proc. 7th IEEE Power Eng. Society*, Lausanne, Switzerland, 2007.

[10] A. Eshragh, J. Filar and A. Nazari, A projection-adapted cross entropy (PACE) method for transmission network planning, *Energy Sys.* 2 (2011) 189–208.

[11] D. Levy, M. Bergen, S. Dutta and R. Venable, The magnitude of menu costs: direct evidence from large U.S. supermarket chains, *Quarterly Journal of Economics* 113 (1997) 791–825.

[12] C.J. Liao, C.H. Shyu, An analytical determination of lead time with normal demand, *Int. J.Oper. Prod. Management* 11 (1991) 72–78.

[13] D. Li, X.L. Sun and K. McKinnon, An exact solution method for reliability optimisation in complex systems, *Ann. Oper. Res.* 133 (2005) 129–148.

[14] D. Li and X. Sun, *Nonlinear Integer Programming*, Springer Science and Business Media LLC, Springer, 2006.

[15] J. Roberts, Evidence on price-change costs in U.S. manufacturing industry. *Econ. Inquiry* 30 (1992) 399–417.

[16] R.Y. Rubinstein, Optimisation of computer simulation models with rare events. *Eur. J. Oper. Res.* 99 (1997) 89–112.

[17] R.Y. Rubinstein and D.P. Kroese, The cross-entropy method, A unified approach to combinatorial optimisation, Monte-Carlo simulation, and machine learning, *Inf. Sci. Statist.*, 2004.

[18] N. Rudi, S. Kapur and D.F. Pyke, A two-location inventory model with transhipment and local decision making, *Management Sci.* 47 (2001) 1668–1680.

[19] K. Sebaa, H. Gueguen and M. Boudour, Mixed integer non-linear programming via the cross-entropy approach for power system stabilisers location and tuning. *IET Generation, Trans. Dist.* 4) (2010) 928–939.

[20] D. Simchi-Levi, S.D. Wu and Z.J. Shen, *Handbook of Qualitative Supply Chain Analysis: Modelling in the E-business Era*, Kluwer Academic Publishers, Boston, Dordrecht, London, 2004, pp. 335–392.

[21] M.E. Slade, Optimal pricing with costly adjustment: Evidence from retail-grocery prices, *Rev. Econ. Stud.* 65 (1998) 87–107.

[22] D. Tsiddon, On the stubbornness of sticky prices, *Int. Econ. Rev.* 32 (1991) 69–75.

[23] M.J. Zbaracki, M. Ritson, D. Levy, S. Dutta and M. Bergen, Managerial and customer costs of price-change: direct evidence from industrial markets. *Rev. Econ. Statist.* 86 (2004) 512–533.

Manuscript received 19 October 2014
revised 19 August 2015
accepted for publication 2 October 2015

Elham Mardaneh
Department of Mathematics and Statistics
Curtin University, Australia
E-mail address: elham.mardaneh@curtin.edu.au

Louis Caccetta
Department of Mathematics and Statistics
Curtin University, Australia
E-mail address: l.caccetta@curtin.edu.au