A Discrete On-Line Monotone Estimation Algorithm

Katerina P. Papadaki[1,2]
Warren B. Powell[1]

[1]Department of Operations Research and Financial Engineering,
Princeton University, Princeton, NJ 08544

[2]Department of Operational Research
London School of Economics & Political Science
Houghton Street, London WC2A 2AE

Working Paper LSEOR 03.73 ISBN: 07530 1695 8

First published in Great Britain in 2003
by the Department of Operational Research
London School of Economics and Political Science

Typeset, printed and bound by:
The London School of Economics and Political Science
Houghton Street
London WC2A 2AE

Working Paper No: LSEOR 03.73 ISBN No: 07530 1695 8

A Discrete On-Line Monotone Estimation Algorithm

Katerina P. Papadaki
Warren B. Powell

Department of Operations Research and Financial Engineering,
Princeton University, Princeton, NJ 08544

August 14, 2003

Abstract

In the paper Papadaki & Powell (2002) we introduced an adaptive dynamic programming algorithm to estimate the monotone value functions for the problem of batch service of homogeneous customers at a service station. The algorithm uses an updating scheme that takes advantage of the monotone structure of the function by imposing a monotonicity-preserving step. In this paper we introduce an algorithm (DOME) that uses this monotonicity-preserving step to approximate discrete monotone functions. Our algorithm requires sampling a discrete function and using Monte Carlo estimates to update the function. It is a known result that sampling a discrete function on each point of its domain infinitely often converges to the correct function as long as standard requirements on the stepsize are maintained. Imposing a monotonicity-preserving step raises anew the question of convergence. We prove convergence of such an algorithm.

Estimating discrete monotone functions arises in the study of batch service queues. When modeling these problems as discrete dynamic programs it can be shown that the value functions are discrete monotone functions. Estimation of the value functions is the key to estimating the optimal decision policies. Backward dynamic programming techniques suffer from the curse of dimensionality, so there has been considerable attention given recently to forward adaptive dynamic programming methods (see, for example, Bertsekas & Tsitsiklis (1996)) which use Monte Carlo methods to estimate the value of a function at each state. In our paper Papadaki & Powell (2002) we introduced an adaptive dynamic programming algorithm to estimate the monotone value functions for the problem of batch service of homogeneous customers at a service station. The algorithm is based on a discrete representation of the value function and uses an updating scheme that takes advantage of the monotone structure of the value function by imposing a monotonicity-preserving step. In this scalar setting, we showed experimentally that when we maintain the monotonicity of the function after each Monte Carlo update, the overall performance of the adaptive dynamic programming algorithm improves dramatically. At 500 iterations the monotone-preserving algorithm is 3% from optimal, whereas the classical algorithm is 30% from optimal. In this paper we introduce and prove convergence of an algorithm (DOME) that uses this monotonicity-preserving step to approximate discrete monotone functions when information about the function arrives in an on-line fashion.

DOME deals with the problem of estimating the means of a finite number of random variables $\{W_i\}$, $i \in \{1, ..., S\}$, whose means $\{\mu_i\}$, $i \in \{1, ..., S\}$, follow a monotone structure and thus constitute our monotone function f: $f(i) = \mu_i$, $f(i) > f(j)$ for $i > j$. For the rest of the paper we refer to the points in the discrete domain $\{1, ..., S\}$ as cells. The DOME algorithm consists of two steps. In the first step a single cell i is picked with a positive probability and a realization of the corresponding random variable W_i is observed. The value of the estimate at cell i is updated with the new observation using an averaging technique. However, the monotonicity property might be violated between the recently updated estimate of a cell and the estimates of the neighboring cells. In the second step a projection method is used to recover monotonicity of the estimates by projecting the value of the newly updated cell onto the value of violated neighboring cells. If we were estimating general functions using solely the first step, the problem would separate in $|S|$ subproblems, each one converging according to the law of large numbers under the right averaging technique. In that case, the algorithm converges to the desired function. This raises the question of convergence when the monotonicity-preserving step is imposed. In this paper we prove almost sure convergence of DOME under certain assumptions (see section 1).

Other work on estimating monotone functions has been done in the context of isotonic regression. Consider an off-line version of the problem described above, where we are given at once all sample realizations of the random variables whose means we are estimating. Barlow & Brunk

(1972) provide a simple example of the isotonic regression problem:

$$min_{\hat{\mu}_1,...,\hat{\mu}_n} \sum_{k=1}^{n} (\bar{x}_i - \hat{\mu}_i)^2$$
$$subject\ \ to\ \ \hat{\mu}_1 \leq ... \leq \hat{\mu}_n$$

where $\hat{\mu}_i$'s are our estimates and $\bar{x}_i$'s are the means of the sample realizations of the values we are trying to estimate: $\{\mu_i\}$, $i \in \{1, ..., n\}$. Barlow, Bartholomew, Bremner & Brunk (1972) give an extensive description of isotonic regression and related algorithms. They describe an algorithm called the 'Pool Adjacent Violators Algorithm' that was first proposed by Ayer, Brunk, Ewing, Reid & Silverman (1955). Other algorithms have been used such as the Minimum Lower Sets algorithm given in Brunk, Ewing & Utz (1957) and in Brunk (1955). Wright (1978) proposed a method for estimating strictly increasing regression functions. Later, Mammen (1991) proposed a two step algorithm for estimating smooth regression functions where the first step is a smoothing step performed by a kernel estimator and the second step is an isotonisation step performed by the pool adjacent violator algorithm. The isotonisation step is a projection of the estimate onto the monotone functions. The second step of the DOME algorithm performs a similar projection step to ensure monotonicity of the estimates at each iteration.

The DOME algorithm falls in the class of stochastic approximation algorithms. The pioneers of these methods are Robbins & Monro (1951), who introduce an algorithm to find the solution $x = \theta$ of the equation $M(x) = \alpha$, where $M(x)$ is the expected value at level x of the response of a certain experiment. The experimenter does not know the value of $M(x)$ but is allowed to perform the experiment at different levels of x and observe sample realizations of the response of the experiment. They prove convergence in mean square of their algorithm under certain assumptions. In the proof of convergence of the DOME algorithm we encounter similar issues as the ones in the paper by Robbins and Monro. The difference between the typical stochastic approximation algorithm and our algorithm is that DOME constructs sequences that converge to the function values whereas the stochastic approximation algorithms search for a value in the domain of the function (for example a root or the point where the function attains its maximum). However, the first step of the DOME algorithm, of updating the function value estimates using new observations, is very similar to the updating scheme of the Robbins Monro process.

Wolfowitz (1952) generalizes the results by Robbins and Monro, and Kiefer & Wolfowitz (1952) solve a similar problem in the case when $M(x)$ has a maximum at $x = \theta$. Blum (1954*a*) proves a stronger form of convergence, namely convergence in probability, of a similar result under weaker conditions than those imposed by Wolfowitz, Kiefer and Wolfowitz. Gladyshev (1965) uses martingale convergence theorems to prove almost sure convergence of stochastic approximation methods. Also, Blum (1954*b*) extends the problem to finding the maximum of a multidimensional regression function and provides conditions for which the multidimensional stochastic approximation schemes

converge almost surely. These results and other early research were unified and generalized by Dvoretzky (1956). Reviews of stochastic approximation methods during the 1950's and 1960's can be found in Wasan (1969).

Mukerjee (1981) uses isotonic regression for the stochastic approximation problem of estimating the roots of a non-decreasing regression function. Dupac (1987) extends Mukerjee's method by using quasi-isotonic regression. Later, Hanson & Mukerjee (1990) alter the Robbins Monro process of estimating a root θ of a regression function, by separating the estimator of θ at each iteration from the design/control setting of the process at the next step. They use isotonic regression to define a new estimator of θ at each iteration and they combine this isotonic estimator with the Robbins Monro process.

In the second step of the DOME algorithm the current function estimate is projected onto the feasible region, which in this case is the set of all monotone functions defined on the domain $\{1, ..., S\}$. The constraint to preserve monotonicity of the estimates at each iteration classifies the problem as a constrained optimization problem, where the objective is to minimize the expected error between the estimates of the function and the random variables $\{W_i\}$.

Ermoliev (1988) describes methods for the solution of stochastic constrained optimization problems, called stochastic quasigradient methods. Stochastic quasigradient methods generalize stochastic approximation methods for unconstrained optimization of the expectation of random functions, to problems involving general constraints and nondifferentiable functions. Ermoliev (1968) studies the stochastic analog of deterministic projection methods in the context of constrained stochastic optimization and proves the convergence. The projection method, which is a method to minimize convex functions, was proposed and studied by Ermoliev & Nekrylova (1967), Ermoliev & Shor (1968), Ermoliev (1969). This method parallels the projection step of the DOME algorithm.

Furthermore, Andradottir (1995) provides methods for solving discrete stochastic optimization problems where the objective function cannot be evaluated exactly but has to be estimated or measured. In the proposed method, at each iteration the estimates of neighboring cells are compared with the current cell and the one with the best observed function value becomes the current cell in the next iteration.

The contribution of this paper is the proof of convergence of an on-line algorithm for estimating discrete monotone functions that preserves monotonicity of the estimates at each iteration.

We start the paper with the definitions in section 1. In section 2 we provide the formal statement of the algorithm. Section 3 provides the proof of convergence starting with some known results and including a sketch of the proof. We conclude the paper in section 4.

1 Definitions

In this section we define the parameters, variables and stochastic sequences needed to describe the problem and the algorithm. A precise description of the algorithm is found in section 2.

Let $\{1, ..., S\}$ be the discrete finite domain of the deterministic strictly increasing function f, where S is a constant positive integer. Then f is defined as follows: $f(i) = \mu_i \in \mathcal{R}$, where $\{\mu_1, ..., \mu_S\}$ are the function values that we want to estimate. We receive information about these values in the form of sample realizations of random variables whose means are in the set $\{\mu_1, ..., \mu_S\}$. This information arrives in an on-line fashion as we are estimating the μ's. Let $\{W_1, ..., W_S\}$ be random variables of unknown distribution that take values in $\mathcal{R}$ and whose respective means are $\{\mu_1, ..., \mu_S\}$. At each iteration one of the indices $\{1, ..., S\}$ gets picked at random, index i, and then the random variable corresponding to that index, W_i, is sampled and the observation is denoted by w_i. We let S^k be the random variable that takes values in the index set $\{1, ..., S\}$ and determines which one of the random variables $\{W_1, ..., W_S\}$ is going to be sampled at iteration k. For notational purposes we assume that all random variables $\{W_1, ..., W_S\}$ are sampled at each iteration but S^k determines which one we observe. Thus we define the multidimensional stochastic process $\{(W_1^k, ..., W_S^k)\}_{k\geq 1}$ whose realization at iteration k is $(w_1^k, ..., w_S^k) \in \mathcal{R}^S$. We also define the stochastic process $\{S^k\}_{k\geq 1}$ whose realization at iteration k is $s^k \in \{1, ..., S\}$. Now we define our outcome space Ω to be the set of all $\omega \in \Omega$ such that:

$$\omega = ((s^1, w_1^1, ..., w_S^1), (s^2, w_1^2, ..., w_S^2), ...),$$

and so we have:

$$\begin{aligned}
&(W_1^k, ..., W_S^k) : \Omega \to \mathcal{R}^S \\
&(W_1^k, ..., W_S^k)(\omega) = (w_1^k, ..., w_S^k) \\
&S^k : \Omega \to \{1, ..., S\} \\
&S^k(\omega) = s^k
\end{aligned}$$

Thus the outcome space is completely defined by these two processes. For ease of notation, we denote the probabilities associated with the random variables $\{S^k\}$ as follows: $p_i^k = P(S^k = i)$. We make the following assumptions on these processes:

Assumptions

1. The random variables $\{(W_1^k, ..., W_S^k)\}_{k\geq 1}$ are independent, identically distributed of unknown distribution.

2. For all $i \in \{1, ..., S\}$, $k \geq 1$, we have $E[W_i^k] = \mu_i$, $\|\mu_i\| < \infty$, $Var[W_i^k] = \sigma_i^2 < \infty$ where all means and variances are unknown and the means satisfy:

$$\mu_1 < \mu_2 < ... < \mu_S \tag{1}$$

3. The processes $\{(W_1^k, ..., W_S^k)\}_{k\geq 1}$ and $\{S^k\}_{k\geq 1}$ are independent.
4. For all $i \in \{1, ..., S\}$ and $k \geq 1$, there exists an $\epsilon > 0$ such that $P(S^k = i) = p_i^k > \epsilon$.

We can now define a standard probability space $(\Omega, \mathcal{F}, \mathcal{P})$, where $\mathcal{F}$ is the σ-algebra defined over Ω, and $\mathcal{P}$ is a probability measure defined over $\mathcal{F}$. $\mathcal{F}$ is the set of all possible events, however we go further and define $\mathcal{F}_k$ to be the set of events up to iteration k. $\mathcal{F}_k$ contains all the events that are completely determined by the following observations:

$$((s^1, w_1^1, ..., w_S^1), ..., (s^k, w_1^k, ..., w_S^k)),$$

Since we have $\mathcal{F}_k \subset \mathcal{F}_{k+1}$, the process $\{\mathcal{F}_k\}_{k\geq 1}$ is a filtration.

In order to understand the above definitions we briefly describe their motivation. The problem is of estimating a discrete monotone function defined on the set of discrete points $\{1, ..., S\}$, which we call cells. At iteration k we receive information on the value of the function at one discrete point in its domain: cell s^k. The cell is picked according to the stochastic process $\{S^k\}_{k\geq 1}$. The actual value of the function at cell s^k is μ_{s^k}. The information we receive, however, is a sample realization of a random variable that has mean μ_{s^k}. The random variable sampled at iteration k is $W_{s^k}^k$. Thus each cell i has a stochastic process $\{W_i^k\}_{k\geq 1}$ of i.i.d. random variables associated with it.

We proceed to define some other stochastic processes that depend on the two defined above. Let $\{w^k\}_{k\geq 1}$ be a stochastic process defined as follows:

$$w^k(\omega) = w_{s^k}^k$$

and also define $\{w_i^k\}_{k\geq 1}$ as follows:

$$w_i^k(\omega) = w_i^k$$

$\{w^k\}_{k\geq 1}$ is the process that contains at each iteration the observed value of the cell selected at that iteration. The process $\{w_i^k\}_{k\geq 1}$ contains the observed values of cell i at each iteration.

To proceed we introduce our stepsize sequence $\{\alpha^k\}_{k\geq 1}$ which has the following properties:

$$\alpha^k \rightarrow 0 \tag{2}$$

$$\sum_{k\geq 1} \alpha^k = \infty \tag{3}$$

$$\sum_{k\geq 1} (\alpha^k)^2 < \infty \tag{4}$$

For each $i \in \{1, ..., S\}$ and $n \in \{i, ..., S\}$, we are going to define two processes, which we refer to as

the y^k and $x^{n,k}$ processes. We define the stochastic process $\{y_i^k\}_{k\geq 0}$ recursively, for all $i \in \{1,...,S\}$:

$$\begin{aligned} y_i^{k+1} &= \alpha^{k+1} w_i^{k+1} + (1-\alpha^{k+1}) y_i^k \\ y_i^0 &= w^0 \end{aligned} \tag{5}$$

where the initial value w^0 is a deterministic constant. If the DOME algorithm consisted only of the first updating step, and did not require monotonicity of its estimates, then the estimates of the DOME algorithm would be the same as the y^k processes. As we see later (theorem 2) the y_i^k process converges to μ_i. This is a well known convergence result that we use in order to prove convergence of DOME.

We define the $x^{n,k}$ process, for $i \in \{1,...,n\}$ and $n \in \{i,...,S\}$ as follows:

$$x_i^{n,k+1} = \begin{cases} \alpha^{k+1} w^{k+1} + (1-\alpha^{k+1}) x_i^{n,k} & \text{if } S^{k+1} = i \\ x_i^{n,k} \wedge \left\{ \alpha^{k+1} w^{k+1} + (1-\alpha^{k+1}) x_j^{n,k} \right\} & \text{if } j = S^{k+1},\ j \leq n, j > i \\ x_i^{n,k} \vee \left\{ \alpha^{k+1} w^{k+1} + (1-\alpha^{k+1}) x_j^{n,k} \right\} & \text{if } j = S^{k+1},\ j \leq n, j < i \\ x_i^{n,k} & \text{if } j = S^{k+1},\ j > n \end{cases} \tag{6}$$

where $x_i^{n,0} = w^0$ and for $j, i \in \{1,...,n\}$. First, note that the processes $\{x_i^{n,k}\}$, $i = 1,...,n$ are only updating cells $1,...,n$ while maintaining monotonicity of the estimates of these cells. Further, they ignore iterations where S^k is greater than n, by keeping all estimates the same at such iterations. In this paper, we refer to the problem of providing convergent monotone estimates of the first n cells as the n-cell problem, and we show that these processes solve the n-cell problem. From the definition of the $\{x^{n,k}\}$ processes we can see that they follow the two-step updating scheme of the DOME algorithm (as we described it this far) for cells $i,...,n$. Thus the estimates at each iteration are monotone.

Let us investigate in more detail how these processes get updated. The four cases described above are explained as follows: (1) If at iteration $k+1$ the cell i is chosen then we update that cell as usual (as in the case of the y sequence). (2) If cell j is chosen that is within the range $\{1,...,n\}$ and j is greater than i then we update the value at cell i by taking the minimum of the value at i at the previous iteration and the currently iteration value at j. This ensures that if the new value at j was too small and has violated monotonicity, we correct by this by lowering the value at i. (3) This is similar to case two only in the other direction. (4) In the case that j is outside of the set $\{1,...,n\}$ we do not update any cells. This last case is characteristic of the x^n sequences which update only values at cells 1 through n.

For convenience of notation we define the following random variables:

$$U_{ji}^{n,k+1}(\omega) = \begin{cases} 1 & \text{if } S^{k+1}(\omega) = j, \text{ where } i,j \in \{1,...,n\} \text{ AND} \\ & \text{either } j > i,\ x_i^{n,k} > \alpha^{k+1}w^{k+1} + (1-\alpha^{k+1})x_j^{n,k} \\ & \text{or } j < i,\ x_i^{n,k} < \alpha^{k+1}w^{k+1} + (1-\alpha^{k+1})x_j^{n,k} \\ 0 & \text{otherwise} \end{cases} . \tag{7}$$

$U_{ji}^{n,k}$ is the indicator for the event that cell j is sampled and updated using the sequence $x^{n,k}$ at iteration k, and the new value at cell j violates monotonicity of the estimates of cell i in the n-cell problem.

To summarize, at each iteration we receive a sample realization of the value of the function f at the chosen cell. The processes $\{y_i^k\}_{k\geq 0}$ for $i \in \{1,...,S\}$ just average the new sample realization with the old estimate of the function value at that cell. Since the stepsize sequence α^k satisfies properties (2), (3) and (4) we have from a well known result that the sequence $\{y_i^k\}_{k\geq 0}$ converges to μ_i (see theorem 1). When $\alpha^k = 1/k$, then we have the same convergence result from the strong law of large numbers. These processes treat each cell independently ignoring any violations of monotonicity this new update might have on the estimates of neighboring cells. They all converge to the function values but at each iteration there is no guarantee of monotonicity, since each cell is treated independently.

Our aim is to produce an estimate of the entire monotone function f and thus we want to preserve monotonicity at each iteration. That is why we introduce the processes $\{x_i^{n,k}\}_{k\geq 0}$, for $i \in \{1,...,S\}$ and $n \in \{i,...,S\}$, which are not independent of each other and monitor the monotone structure of the estimates of the first n cells. Thus after updating the chosen cell i, they project the value of that cell to neighboring cells whose last iteration estimates violate the monotone structure of the cells $\{1,...,n\}$. These processes with $n = S$ are used by our algorithm to estimate the function values while preserving monotonicity at each iteration. We later prove that these processes, for $n = S$, converge to the values of the function f.

2 Algorithm for estimating monotone functions

This section describes the algorithm for estimating the discrete monotone function f, maintaining monotone structure of the function estimate at each iteration. The formal statement of the algorithm is as follows:

Discrete On-line Monotone Estimation (DOME) algorithm

Step 1 Set $x_i^{S,0} = w^0$, for all $i \in \{1,...,S\}$. Pick an $\omega \in \Omega$ and set $k = 0$.

Step 2 We pick cell $s^{k+1} = S^{k+1}(\omega)$. For each $i \in \{1,...,S\}$, update the processes $\{x_i^{S,k}\}_{k\geq 0}$

using:

$$x_i^{S,k+1} = \begin{cases} \alpha^{k+1} w^{k+1} + (1-\alpha^{k+1}) x_i^{S,k} & \text{if } S^{k+1} = i \\ x_i^{S,k} \wedge \left\{ \alpha^{k+1} w^{k+1} + (1-\alpha^{k+1}) x_j^{S,k} \right\} & \text{if } j = S^{k+1},\ j \leq S, j > i \\ x_i^{S,k} \vee \left\{ \alpha^{k+1} w^{k+1} + (1-\alpha^{k+1}) x_j^{S,k} \right\} & \text{if } j = S^{k+1},\ j \leq S, j < i \\ x_i^{S,k} & \text{if } j = S^{k+1},\ j > S \end{cases}$$

Step 3 If $k >$ maximum number of iterations, then go to step 4. Otherwise set $k := k+1$ and go to step 2.

Step 4 The monotone estimates of the function f values are: $\{x_1^{S,k}, ..., x_S^{S,k}\}$.

We assume that w^0 is a finite deterministic constant. Also, note that the sequence used in the algorithm is equivalent to the sequence $\{x_i^{n,k}\}_{k \geq 0}$ defined in the previous section with $n = S$.

3 Convergence of the algorithm

This section describes the proof of convergence of the DOME algorithm. In section 3.1 we present some technical results that are needed for the proof of convergence. One of our technical results is a variation of a well known convergence result, where we use stochastic stepsize sequences. In section 3.2 we provide a sketch of the proof of convergence and in section 3.3 we provide the proof.

3.1 Some technical results

In this section we state a well known convergence result, which is a generalization of the law of large numbers (theorem 1), and we prove a variation of it. Theorem 1 establishes the convergence of a sequence of random variables that is created by averaging another sequence of i.i.d. random variables. Their averaging is done according to a stepsize sequence $\{\alpha^k\}$ that satisfies properties (2), (3) and (4). Our variation of the result comes from averaging a sequence of i.i.d. random variables, where the averaging is done using a stochastic stepsize sequence $\{\beta_k\}_{k \geq 1}$ which satisfies certain properties. This result is essential to the upcoming proof of the algorithm. We prove this result using the supermartingale convergence theorem.

We state the well known convergence result (see Bertsekas & Tsitsiklis (1996), example 4.3, p.143):

Theorem 1 *Let $\{\alpha_k\}_{k \geq 1}$ be a sequence such that $\alpha^k < 1$ for all $k \geq 1$ and:*

$$\sum_{k=1}^{\infty} \alpha^k \quad = \quad \infty$$

$$\sum_{k=1}^{\infty} [\alpha^k]^2 \quad < \quad \infty$$

Also, let $\{\hat{z}^k\}_{k\geq 0}$ *be i.i.d. random variables with mean* $\|\mu\| < \infty$ *and variance* $\sigma^2 < \infty$*. Then the sequence* $\{z^k\}_{k\geq 0}$ *defined by the recursion:*

$$z^{k+1} = \alpha^{k+1}\hat{z}^{k+1} + (1-\alpha^{k+1})z^k$$

with any deterministic initial value z^0*, converges to* μ *almost surely.*

Before we proceed to the next theorem, we state a technical lemma from Bertsekas and Tsitsiklis ((Bertsekas & Tsitsiklis 1996) p.116).

Lemma 3.1 *Suppose that* e_t *and* δ_t *are nonnegative sequences,* $\delta_t \leq 1$ *and* c *is a positive constant such that*

$$e_{t+1} \leq (1-\delta_t)e_t + c\delta_t^2$$

for all $t = 0, 1, ...,$ *and*

$$\delta_t \to 0, \quad \textstyle\sum_{t=0}^{\infty} \delta_t = \infty.$$

Then $e_t \to 0$*.*

Using the above lemma we can introduce and prove the following theorem:

Theorem 2 *Let* $\{\alpha_k\}_{k\geq 1}$ *be a sequence such that* $\alpha^k < 1$ *for all* $k \geq 1$ *and:*

$$\sum_{k=1}^{\infty} \alpha^k \quad = \quad \infty$$
$$\sum_{k=1}^{\infty} [\alpha^k]^2 \quad < \quad \infty$$

For each $i \in \{1, ..., S\}$*, we define* $\{\beta_i^k\}_{k\geq 1}$ *to be a stochastic sequence defined by:*

$$\beta_i^k(\omega) \quad = \quad \begin{cases} \alpha^k & \text{if } S^k(\omega) = i \\ 0 & \text{otherwise} \end{cases}$$

where $\{S^k\}_{k\geq 0}$ *is defined in section 1. Also, for each* $i \in \{1,...,S\}$*, let* $\{\hat{z}_i^k\}_{k\geq 0}$ *be i.i.d. random variables with mean* μ_i*,* $\|\mu_i\| < \infty$*, and variance* $\sigma_i^2 < \infty$*. Assume that the sequences* $\{S^k\}_{k\geq 0}$*,* $\{\hat{z}_i^k\}_{k\geq 0}$ *are independent for all* $i \in \{1,...,S\}$*. Then the sequences* $\{z_i^k\}_{k\geq 0}$ *defined by the recursion:*

$$z_i^{k+1} = \beta_i^{k+1}\hat{z}_i^{k+1} + (1-\beta_i^{k+1})z_i^k$$

with any finite deterministic initial value z_i^0*, converge to* μ_i *almost surely, for all* $i \in \{1,...,S\}$

Proof: We define Z_i^k for all $k \geq 0$ and all $i \in \{1,...,S\}$ as follows:

$$Z_i^k \equiv (z_i^k - \mu_i)^2$$

Then we have:

$$\begin{aligned}
E[Z_i^{k+1}|\mathcal{F}_k] &= E[(z_i^{k+1}-\mu_i)^2|\mathcal{F}_k] \\
&= E[(\beta_i^{k+1}\hat{z}_i^{k+1} + (1-\beta_i^{k+1})z_i^k - \mu_i)^2|\mathcal{F}_k] \\
&= E[[\beta_i^{k+1}(\hat{z}_i^{k+1}-\mu_i) + (1-\beta_i^{k+1})(z_i^k-\mu_i)]^2|\mathcal{F}_k] \\
&= E[(\beta_i^{k+1})^2(\hat{z}_i^{k+1}-\mu_i)^2|\mathcal{F}_k] + (z_i^k-\mu_i)^2 E[(1-\beta_i^{k+1})^2|\mathcal{F}_k] \\
&\quad +2(z_i^k-\mu_i)E[\beta_i^{k+1}(1-\beta_i^{k+1})(\hat{z}_i^{k+1}-\mu_i)|\mathcal{F}_k]
\end{aligned}$$

Since β_i^{k+1} is completely determined by S^{k+1}, it is independent of $\hat{z}_i^{k+1}$ and also of $\mathcal{F}_k$. Thus we can rewrite the above expression:

$$\begin{aligned}
E[Z_i^{k+1}|\mathcal{F}_k] &= E[(\beta_i^{k+1})^2]E[(\hat{z}_i^{k+1}-\mu_i)^2|\mathcal{F}_k] + (z_i^k-\mu)^2 E[(1-\beta_i^{k+1})^2]] + \\
&\quad 2(z_i^k-\mu_i)E[\beta_i^{k+1}(1-\beta_i^{k+1})]E[(\hat{z}_i^{k+1}-\mu_i)|\mathcal{F}_k] \qquad (8)
\end{aligned}$$

By definition, $p_i^k = P(S^k = i)$ and so we have:

$$\begin{aligned}
&E[\hat{z}_i^{k+1}-\mu_i|\mathcal{F}_k] = E[\hat{z}_i^{k+1}-\mu_i] = E[\hat{z}_i^{k+1}] - \mu_i = 0 \\
&E[(\hat{z}_i^{k+1}-\mu_i)^2|\mathcal{F}_k] = E[(\hat{z}_i^{k+1}-\mu_i)^2] = \sigma_i^2 \\
&E[(\beta_i^{k+1})^2] = (\alpha^{k+1})^2 p_i^{k+1} + 0(1-p_i^{k+1}) = (\alpha^{k+1})^2 p_i^{k+1} \\
&E[(1-\beta_i^{k+1})^2] = (1-\alpha_i^k)^2 p_i^{k+1} + (1-p_i^{k+1})
\end{aligned}$$

Hence we can write (8) as:

$$\begin{aligned}
E[Z_i^{k+1}|\mathcal{F}_k] &= (\alpha^{k+1})^2 p_i^{k+1}\sigma_i^2 + \left[(1-\alpha^{k+1})^2 p_i^{k+1} + (1-p_i^{k+1})\right] Z_i^k \qquad (9) \\
&\leq (\alpha^{k+1})^2\sigma_i^2 + Z_i^k \qquad (10)
\end{aligned}$$

Let $M_i^k \equiv Z_i^k + \sigma_i^2 \sum_{j=k+1}^{\infty} (\alpha^j)^2$, then we have:

$$\begin{aligned} E[M_i^{k+1}|\mathcal{F}_k] &= E[Z_i^{k+1}|\mathcal{F}_k] + \sigma_i^2 \sum_{j=k+2}^{\infty} (\alpha^j)^2 \\ &\leq (\alpha^{k+1})^2\sigma_i^2 + Z_i^k + \sigma_i^2 \sum_{j=k+2}^{\infty} (\alpha^j)^2 \\ &= M_i^k \end{aligned} \tag{11}$$

Since M_i^k is $\mathcal{F}_k$ measurable and it satisfies (11) it only remains to show that $E[|M_i^k|] < \infty$ in order to prove that $\{M_i^k\}$ is a supermartingale. From equation (9) we have:

$$\begin{aligned} E[|Z_i^{k+1}|] &= E[Z_i^{k+1}] = E[E[Z_i^{k+1}|\mathcal{F}_k]] \\ &= E\left[(\alpha^{k+1})^2 p_i^{k+1}\sigma_i^2 + \left((1-\alpha^{k+1})^2 p_i^{k+1} + (1-p_i^{k+1})\right) Z_i^k\right] \\ &= (\alpha^{k+1})^2 p_i^{k+1}\sigma_i^2 + \left((1-\alpha^{k+1})^2 p_i^{k+1} + (1-p_i^{k+1})\right) E[Z_i^k] \end{aligned}$$

Thus if $E[Z_i^k] < \infty$ then $E[Z_i^{k+1}] < \infty$ for all $k \geq 0$, since $\sigma_i^2 < \infty$. We also have:

$$E[Z_i^0] = E[(z_i^0 - \mu_i)^2] = (z_i^0 - \mu_i)^2 < \infty$$

since $\mu_i < \infty$ and the initial value z_i^0 is finite. Therefore, we can conclude that $E|Z_i^k| < \infty$ for all $k \geq 0$. By assumption we have that $\sum_{j=k+1}^{\infty} (\alpha^j)^2 < \infty$. So by the following inequality:

$$E|M_i^k| \leq E|Z_i^k| + \sigma_i^2 \sum_{j=k+1}^{\infty} (\alpha^j)^2 < \infty$$

we have $E[|M_i^k|] < \infty$ and thus $\{M_i^k\}$ is a supermartingale. Also note that $\{M_i^k\}$ is a nonnegative supermartingale. By the supermartingale convergence theorem we have that $\{M_i^k\}$ converges almost surely to a nonnegative random variable M_i^*. From the definition of M_i^k and from the fact that $\sum_{j=k+1}^{\infty} (\alpha^j)^2 \to 0$ when $k \to \infty$ we conclude that $\{Z_i^k\}$ converges to $Z_i^* \equiv M_i^*$ almost surely.

Now, we show that $Z_i^* = 0$. If we let $m_i^k = E[Z_i^k]$ and take expectations of (9) we get:

$$m_i^{k+1} = (\alpha^{k+1})^2 p_i^{k+1}\sigma_i^2 + \left((1-\alpha^{k+1})^2 p_i^{k+1} + (1-p_i^{k+1})\right) m_i^k \tag{12}$$

$$= (\alpha^{k+1})^2 p_i^{k+1}\sigma_i^2 + \left(1 - p_i^{k+1}\alpha^{k+1}(2-\alpha^{k+1})\right) m_i^k \tag{13}$$

$$\leq (\alpha^{k+1})^2 p_i^{k+1}\sigma_i^2 + \left(1 - p_i^{k+1}\alpha^{k+1}\right) m_i^k \tag{14}$$

$$\leq \epsilon(\alpha^{k+1})^2(\sigma_i^2/\epsilon) + \left(1 - \epsilon\alpha^{k+1}\right) m_i^k \tag{15}$$

Inequality (15) comes from assumption 4: $\epsilon < p_i^k < 1$ for all i and k. With the above inequalities we satisfy all the conditions of lemma 3.1. Thus, we have $m_i^k \to m_i^* = 0$ as $k \to \infty$. By Fatou's

lemma we have the inequality:

$$E[Z_i^*] = E[\lim_{k\to\infty} Z_i^k] \leq \lim_{k\to\infty} E[Z_i^k] = \lim_{k\to\infty} m_i^k = 0$$

Since Z_i^* is a nonnegative random variable with $E[Z_i^*] \leq 0$, then we must have $Z_i^* = 0$. Thus, $\{Z_i^k\}_{k\geq 0}$ converges to zero almost surely. So, $\{z_i^k\}_{k\geq 0}$ converges to μ_i almost surely for all $i \in \{1, ..., S\}$. *Q.E.D.*

3.2 Sketch of the convergence proof

In section 3.3 we provide the proof of convergence of the DOME algorithm. The sketch of the proof is described below:

Sketch of the proof:

Step 1 We establish the following inequalities amongst the sequences $\{x_i^{n,k}\}_{k\geq 0}$, $\{x_i^{n+1,k}\}_{k\geq 0}$ and $\{y_i^k\}_{k\geq 0}$ (see equations (6) and (5)) in proposition 1:

$$\begin{aligned} x_i^{n+1,k} &\leq x_i^{n,k} \quad for \;\; i < n+1 \\ x_{n+1}^{n+1,k} &\geq y_{n+1}^k \end{aligned}$$

for all $n \geq 1$, all $k \geq 0$.

Step 2 In proposition 2 we prove that if $\sum_{k=0}^{\infty} U_{i,j}^{n,k} < \infty$ almost surely for $i, j \in \{1, ..., n\}$ and for all $n \geq 1$, then the sequence $\{x_i^{n,k}\}_{k\geq 0}$ converges to μ_i almost surely for all $i \in \{1, ..., n\}$ and for all $n \geq 1$. To prove this result we use theorem 2 which is a variation of a well known Robbins-Monro type of convergence result (which we state in theorem 1).

Step 3 We use the inequalities of step 1 to prove that $\sum_{k=0}^{\infty} U_{i,j}^{n,k} < \infty$ almost surely for $i, j \in \{1, ..., n\}$ and for all $n \geq 1$. This means that updates to neighboring cells occur a finite number of times. We prove this by induction on n, the number of cells: (a) In lemma 3.2 we prove the result for $n = 2$. (b) In proposition 3 we use lemma 3.2 as the initial step of the induction. We make the induction hypothesis that $\sum_{k=0}^{\infty} U_{i,j}^{n,k} < \infty$ almost surely for $i, j \in \{1, ..., n\}$ and using lemmas 3.3, 3.4 and 3.5 we prove that $\sum_{k=0}^{\infty} U_{i,j}^{n+1,k} < \infty$ almost surely for $i, j \in \{1, ..., n+1\}$.

Step 4 By propositions 1, 2 and 3 we get the result that the sequences $\{x_i^{n,k}\}_{k\geq 0}$ converge to μ_i almost surely for all $i \in \{1, ..., n\}$, and $n \in \{1, ..., S\}$. From step 2 of the DOME algorithm we can observe that the algorithm reaches its estimates by following the sequences $\{x_i^{S,k}\}_{k\geq 0}$. And thus we establish convergence of the algorithm.

3.3 Our proof of convergence

We start by proving some inequalities between the sequences that we described in section 1.

Proposition 1 *For any $n \geq 1$ and all $k \geq 0$ we have:*

$$x_i^{n+1,k} \leq x_i^{n,k} \quad for \;\; i < n+1 \tag{16}$$
$$x_{n+1}^{n+1,k} \geq y_{n+1}^k \tag{17}$$

almost surely.

We first consider the simple case of proposition 1 where the random variable S^k takes values in the domain $\{1, 2\}$, that is when $n = 1$. Then we proceed to prove these inequalities for the general case where S^k takes values in $\{1, ..., n\}$.

Proof of case $n = 1$:
The inequalities (16) and (17) become:

$$x_1^{2,k} \leq x_1^{1,k} \tag{18}$$
$$x_2^{2,k} \geq y_2^k \tag{19}$$

for all $k \geq 0$. Note that in equation (18) $x_1^{1,k}$ is equal to y_1^k, since for $n = 1$ the sequences $\{x_i^{n,k}\}$ become the same as the sequences $\{y_i^k\}$. This is because in the one cell case there are no updates from neighboring cells. Thus (18) and (19) can be written as:

$$x_1^{2,k} \leq y_1^k \tag{20}$$
$$x_2^{2,k} \geq y_2^k \tag{21}$$

The above inequalities are for the 2-cell problem and we prove each of them using induction on k. We start with (20):

Initially we have $x_1^{2,0} = y_1^0 = x^0$ which means that (20) is satisfied for $k = 0$. Now we assume that (20) holds for k iterations and we prove that it holds for $k + 1$ iterations. At iteration $k + 1$, if cell 1 is picked at iteration $k + 1$ then by the induction hypothesis we get:

$$(1 - \alpha^{k+1})x_1^{2,k} \leq (1 - \alpha^{k+1})y_1^k.$$

Adding $\alpha^{k+1}w_1^{k+1}$ to both sides:

$$\alpha^{k+1}w_1^{k+1} + (1 - \alpha^{k+1})x_1^{2,k} \leq \alpha^{k+1}w_1^{k+1} + (1 - \alpha^{k+1})y_1^k$$

gives us:

$$x_1^{2,k+1} \leq y_1^{k+1}.$$

Now, suppose that cell 2 is picked at iteration $k+1$; then we have $y_1^{k+1} = y_1^k$ and the following two cases:
Case (a): $x_2^{2,k+1} > x_1^{2,k}$, which means that cell 1 is not updated, and by the induction hypothesis we have:

$$x_1^{2,k+1} = x_1^{2,k} \leq y_1^k = y_1^{k+1}.$$

Case (b): $x_2^{2,k+1} \leq x_1^{2,k}$, which means that cell 1 is updated to $x_2^{2,k+1}$, and we have:

$$x_1^{2,k+1} = x_2^{2,k+1} \leq x_1^{2,k} \leq y_1^k = y_1^{k+1}$$

Thus, (20) holds for $k+1$ iterations, which means that it holds for all k.

The proof of (21) closely mirrors that of (20) and thus we omit it. *Q.E.D.*

Now we prove proposition 1 for the general case where S^k takes values in the domain $\{1, ..., n\}$.

Proof of case $n \geq 2$:

We first prove (16) by induction on k. Initially we have $x_i^{n,0} = x_i^{n+1,0}$, which proves (16) for $k = 0$. Now, we assume (16) holds for k iterations and we prove it holds for $k+1$.

At iteration $k+1$, if cell i is picked then both estimates $x_i^{n,k}$ and $x_i^{n+1,k}$ are updated by the same value w_i^{k+1} and thus the inequality (16) still holds.

Now, suppose that cell $n+1$ is picked at iteration $k+1$. Then we have $x_i^{n,k+1} = x_i^{n,k}$ and the following two cases:
Case (a): $x_i^{n+1,k} < x_{n+1}^{n+1,k+1}$, which means that cell i is not updated. Thus, by the induction hypothesis:

$$x_i^{n+1,k+1} = x_i^{n+1,k} \leq x_i^{n,k} = x_i^{n,k+1}$$

Case (b): $x_i^{n+1,k} \geq x_{n+1}^{n+1,k+1}$, which means that cell i is updated. So, we have:

$$x_i^{n,k+1} = x_i^{n,k} \geq x_i^{n+1,k} \geq x_{n+1}^{n+1,k+1} = x_i^{n+1,k+1}$$

So, if cell $n+1$ was picked at iteration $k+1$ the inequality (16) holds for $k+1$.

Finally, suppose that cell j was picked at iteration $k+1$, where $j < i$ and $j \neq n+1$. Then we have the following four cases:

Case (a): $x_j^{n,k+1} < x_i^{n,k}$ and $x_j^{n+1,k+1} < x_i^{n+1,k}$, which means that both estimates $x_i^{n,k}$ and $x_i^{n+1,k}$ are not updated. Since the estimates remain the same at iteration $k+1$ the inequality (16) is satisfied for $k+1$.

Case (b): $x_j^{n,k+1} \geq x_i^{n,k}$ and $x_j^{n+1,k+1} \geq x_i^{n+1,k}$, which means that both estimates $x_i^{n,k}$ and $x_i^{n+1,k}$ are updated to $x_j^{n,k+1}$ and $x_j^{n+1,k+1}$ respectively. By the induction hypothesis we have:

$$x_j^{n,k} \geq x_j^{n+1,k}$$

and since the estimates $x_j^{n,k}$, $x_j^{n+1,k}$ are both updated by the same value we have at iteration $k+1$:

$$x_j^{n,k+1} \geq x_j^{n+1,k+1}$$

and thus:

$$x_i^{n+1,k+1} = x_j^{n+1,k+1} \leq x_j^{n,k+1} = x_i^{n,k}$$

Case (c): $x_j^{n,k+1} \geq x_i^{n,k}$ and $x_j^{n+1,k+1} < x_i^{n+1,k}$, which means that $x_i^{n,k}$ is updated to $x_j^{n,k+1}$ and $x_i^{n+1,k}$ is not updated. So we have:

$$x_i^{n,k+1} = x_j^{n,k+1} \geq x_i^{n,k} \geq x_i^{n+1,k} = x_i^{n+1,k+1}$$

Case (d): $x_j^{n,k+1} < x_i^{n,k}$ and $x_j^{n+1,k+1} \geq x_i^{n+1,k}$, which means that $x_i^{n,k}$ is not updated and $x_i^{n+1,k}$ is updated to $x_j^{n+1,k+1}$. So we have:

$$x_i^{n,k+1} = x_i^{n,k} > x_j^{n,k+1} \geq x_j^{n+1,k+1} = x_i^{n+1,k+1}$$

Thus (16) holds for $k+1$ iterations in the case that cell j is picked, where $j < i$.

The argument for the case that cell j is picked and $i < j < n+1$ is exactly symmetric to the case $j < i$ that we proved above. Therefore, we proved (16) by induction on k for all $n \geq 2$.

The proof of (17) closely mirrors that of (21) and thus we omit it. *Q.E.D.*

Proposition 2 *If* $\sum_{k=0}^{\infty} U_{i,j}^{n,k} < \infty$ *almost surely for all* $i, j \in \{1, 2, ..., n\}$ *and all* $n \leq S$*, then the sequence of estimates* $\{x_i^{n,k}\}_{k \geq 0}$ *converges to* μ_i *for all* $i \in \{1, 2, ..., n\}$ *and all* $1 \leq n \leq S$.

Proof: Let $i \in \{1, 2, ..., n\}$, then for almost every ω there exists $N_i^n(\omega) < \infty$ such that for $k \geq N_i^n$ no more updates occur in cell i from neighboring cells (in the n cell problem). Thus we have for $k \geq N_i^n$:

$$x_i^{n,k+1} = \begin{cases} \alpha^{k+1} w_i^{k+1} + (1 - \alpha^{k+1}) x_i^{n,k} & \text{if } S^{k+1} = i \\ x_i^{n,k} & \text{otherwise} \end{cases}$$

In the iterations greater than N_i^n that cell i is not chosen, cell i gets updated using a stepsize zero. So, we can replace the sequence of stepsizes $\{\alpha^k\}$ by the stochastic stepsize sequence $\{\beta_i^k\}$ which at iteration k is completely determined by the random variable S^k:

$$\beta_i^k(\omega) \quad = \quad \begin{cases} \alpha^k & \text{if } S^k = i \\ 0 & \text{otherwise} \end{cases}$$

Then we can write x_i^{k+1} explicitly:

$$x_i^{k+1} = \beta_i^{k+1} w_i^{k+1} + (1 - \beta_i^{k+1}) x_i^k \tag{22}$$

for all $k > N_i^n$. Now we use theorem 2 to prove convergence of the sequence $\{x_i^{n,k}\}_{k \geq N_i}$. By the definition of the stepsize sequences $\{\beta_i^k\}$, the properties of the stepsize sequence $\{\alpha^k\}$ (see equations (2), (3), (4)), equation (22) and assumption 3 of independence of $\{(W_1^k, ..., W_S^k)\}_{k \geq 1}$ and $\{S^k\}_{k \geq 1}$, all conditions of theorem 2 are satisfied. Thus $\{x_i^{n,k}\}_{k \geq N_i^n}$ converges to μ_i almost surely for all $i \in \{1, ..., n\}$. And thus we have the desired result. *Q.E.D.*

The next step is to prove that the events described by the indicator variables $\{U_{ij}^{n,k}\}_{k \geq 0}$ occur finitely often: $\sum_{k=0}^{\infty} U_{ij}^{n,k} < \infty$ almost surely. This means that the updates from neighboring cells occur finitely many times. We prove this by induction on the number of cells. We start with the case of $n = 2$ as the initial induction step in lemma 3.2, and then we proceed to the general case in proposition 3.

Lemma 3.2 $\sum_{k=0}^{\infty} U_{ij}^{n,k} < \infty$ *almost surely, for* $i, j = 1, 2$ *and* $i \neq j$.

Proof:
We prove the case where $i = 1, j = 2$ and the other case is symmetric.

$$\begin{aligned} \{U_{12}^{2,k+1} = 1\} &= \{\omega : x_1^{2,k+1}(\omega) \geq x_2^{2,k}(\omega)\} && (23) \\ &= \{\omega : \alpha^{k+1} w_1^{k+1}(\omega) + (1 - \alpha^{k+1}) x_1^{2,k}(\omega) \geq x_2^{2,k}(\omega)\} \\ &= \{\omega : \alpha^{k+1} w_1^{k+1}(\omega) \geq x_2^{2,k}(\omega) - (1 - \alpha^{k+1}) x_1^{2,k}(\omega)\} \\ &\subset \{\omega : \alpha^{k+1} w_1^{k+1}(\omega) \geq y_2^k(\omega) - (1 - \alpha^{k+1}) y_1^k(\omega)\} && (24) \\ &= \{\omega : \alpha^{k+1} w_1^{k+1}(\omega) + (1 - \alpha^{k+1}) y_1^k(\omega) \geq y_2^k(\omega)\} \\ &= \{\omega : y_1^{k+1}(\omega) \geq y_2^k(\omega)\} \end{aligned}$$

The equality (23) comes directly from the definition of the indicator variables $U_{ij}^{n,k}$ (equation (7)), and (24) follows from proposition 1. The above set of calculations imply:

$$\{U_{12}^{2,k+1} = 1\} \subset \{\omega : y_1^{k+1}(\omega) \geq y_2^k(\omega)\} \tag{25}$$

By theorem 1, we have that $\{y_i^k\}_{k\geq 0}$ converges to μ_i almost surely, for $i = 1, 2$. Thus,

$$y_2^k - y_1^{k+1} \to \mu_2 - \mu_1 \quad a.s.$$

From the definition of almost sure convergence and $\mu_2 - \mu_1 > 0$, we have:

$$\sum_{k=0}^{\infty} 1_{(\mu_2-\mu_1,\infty)} |(y_2^k - y_1^{k+1}) - (\mu_2 - \mu_1)| < \infty \quad a.s. \tag{26}$$

We still maintain boundedness even if we remove the absolute values from (26):

$$\sum_{k=0}^{\infty} 1_{(\mu_2-\mu_1,\infty)} \left\{(\mu_2 - \mu_1) - (y_2^k - y_1^{k+1})\right\} < \infty \quad a.s.,$$

which is equivalent to:

$$\sum_{k=0}^{\infty} 1_{\{\omega: y_1^{k+1}(\omega) \geq y_2^k(\omega)\}} < \infty \quad a.s..$$

Using (25), the above inequality becomes:

$$\sum_{k=0}^{\infty} 1_{\{U_{12}^{2,k+1}=1\}} < \infty \quad a.s.$$

which is equivalent to:

$$\sum_{k=0}^{\infty} U_{12}^{2,k+1} < \infty \quad a.s.$$

Q.E.D.

The following three technical lemmas are important to prove the nth case of lemma 3.2.

Lemma 3.3 *Suppose the following:*

$$\sum_{k=0}^{\infty} U_{ij}^{n,k} < \infty \quad \textit{almost surely for all } i, j \in \{1, ..., n\}. \tag{27}$$

Then $\sum_{k=0}^{\infty} U_{i,n+1}^{n+1,k} < \infty$ *almost surely for all* $i < n + 1$.

Proof:
We start with the following set relationships:

$$\begin{aligned}
\{U_{i,n+1}^{n+1,k+1}=1\} &= \{\omega : x_i^{n+1,k+1}(\omega) \geq x_{n+1}^{n+1,k}(\omega)\} && (28)\\
&= \{\omega : \alpha^{k+1} w_i^{k+1}(\omega) \geq x_{n+1}^{n+1,k}(\omega) - (1-\alpha^{k+1}) x_i^{n+1,k}(\omega)\} \\
&\subset \{\omega : \alpha^{k+1} w_i^{k+1}(\omega) \geq y_{n+1}^{k}(\omega) - (1-\alpha^{k+1}) x_i^{n,k}(\omega)\} && (29)\\
&= \{\omega : x_i^{n,k+1}(\omega) \geq y_{n+1}^{k}(\omega)\}
\end{aligned}$$

The equality (28) comes directly from the definition of the indicator variables $U_{ij}^{n,k}$, and (29) follows from proposition 1. The above set of calculations imply:

$$\{U_{i,n+1}^{n+1,k+1}=1\} \subset \{\omega : x_i^{n,k+1}(\omega) \geq y_{n+1}^{k}(\omega)\} \tag{30}$$

From assumption (27) and proposition 2 we have that $\{x_i^{n,k}\}_{k\geq 0}$ converges to μ_i almost surely. Also, by theorem 1, we have that $\{y_{n+1}^{k}\}_{k\geq 0}$ converges to μ_{n+1} almost surely. Thus,

$$y_{n+1}^{k} - x_i^{n,k+1} \rightarrow \mu_{n+1} - \mu_i \quad a.s.$$

From the definition of almost sure convergence, the fact that $\mu_{n+1} - \mu_i > 0$ and using a similar argument as in the proof of lemma 3.2 we get:

$$\sum_{k=0}^{\infty} 1_{\{\omega : x_i^{n,k+1}(\omega) \geq y_{n+1}^{k}(\omega)\}} < \infty \quad a.s.$$

Using (30), the above inequality becomes:

$$\sum_{k=0}^{\infty} U_{i,n+1}^{n+1,k+1} < \infty \quad a.s.$$

Q.E.D.

Lemma 3.4 *Suppose the following:*

$$\sum_{k=0}^{\infty} U_{ij}^{n,k} < \infty \quad \textit{almost surely for all } i,j \in \{1,...,n\}. \tag{31}$$

Then $\sum_{k=0}^{\infty} U_{n+1,j}^{n+1,k} < \infty$ *almost surely for all* $j < n+1$.

Proof:
The line of proof is very similar to the proof of lemma 3.3, so we provide a brief version of the proof:

$$\begin{aligned}
\{U_{n+1,j}^{n+1,k+1}=1\} &= \{\omega : x_{n+1}^{n+1,k+1}(\omega) \leq x_j^{n+1,k}(\omega)\} \\
&= \{\omega : \alpha^{k+1} w_{n+1}^{k+1}(\omega) \leq x_j^{n+1,k}(\omega) - (1-\alpha^{k+1}) x_{n+1}^{n+1,k}(\omega)\} \\
&\subset \{\omega : \alpha^{k+1} w_{n+1}^{k+1}(\omega) \leq x_j^{n,k}(\omega) - (1-\alpha^{k+1}) y_{n+1}^{k}(\omega)\} \qquad (32) \\
&= \{\omega : y_{n+1}^{k+1}(\omega) \leq x_j^{n,k}(\omega)\}
\end{aligned}$$

where (32) follows from proposition 1. Thus, we have:

$$\{U_{n+1,j}^{n+1,k+1}=1\} \subset \{\omega : y_{n+1}^{k+1}(\omega) \leq x_j^{n,k}(\omega)\} \qquad (33)$$

From assumption (31) and theorem 1 we have that:

$$y_{n+1}^{k+1} - x_j^{n,k} \to \mu_{n+1} - \mu_i \;\; as \;\; k \to \infty \quad a.s.$$

which implies

$$\sum_{k=0}^{\infty} 1_{\{\omega : y_{n+1}^{k+1}(\omega) \leq x_j^{n,k}(\omega)\}} < \infty \quad a.s.$$

From (33) we have:

$$\sum_{k=0}^{\infty} U_{n+1,j}^{n+1,k} < \infty \quad a.s. \qquad (34)$$

Q.E.D.

Lemma 3.5 *Suppose the following:*

$$\sum_{k=0}^{\infty} U_{ij}^{n,k} < \infty \;\; almost \;\; surely \;\; for \;\; all \;\; i,j \in \{1,...,n\}. \qquad (35)$$

Then $\sum_{k=0}^{\infty} U_{ij}^{n+1,k} < \infty$ *almost surely for all* $i,j \in \{1,...,n\}$.

Proof:
In lemma 3.4, specifically equation (34), we proved that, given (35), then, for almost every ω, out of infinitely many times that cell $n+1$ is sampled, only finitely many times result in updates of

cell j when $j < n+1$. So, for almost every ω there exists an $N_j^{n+1} < \infty$ such that for $k > N_j^{n+1}$ cell j, in the $n+1$ cell problem, receives no more updates from cell $n+1$.

Then, for almost every ω there exists $N^{n+1} = \max_{j=1,...,n} N_j^{n+1}$ such that for $k \geq N^{n+1}(\omega)$ no more updates occur in cells $1, ..., n$ from cell $n+1$.

If we start the algorithm for the $n+1$ problem at iteration N^{n+1}, then the updating scheme of the sequences $\{x_i^{n+1,k}\}$ for cells $i \in \{1, .., n\}$ is the exactly the same as the updating scheme of the sequences $\{x_i^{n,k}\}$ for cells $i \in \{1, .., n\}$, since no more updates occur on cells $\{1, .., n\}$ from cell $n+1$. The only difference between the $\{x_i^{n+1,k}\}_{k>N^{n+1}}$ sequence and the $\{x_i^{n,k}\}_{k>N^{n+1}}$ sequence is that they have different initial values. Thus we conclude that the random variables $\{U_{i,j}^{n,k}\}_{k>N^{n+1}}$ have the same distribution as the variables $\{U_{i,j}^{n+1,k}\}_{k>N^{n+1}}$ for $i, j \in \{1, ..., n\}$ since their distribution depends only on the random variables $\{S^k\}$ and the updating scheme of the sequences $\{x^{n,k}\}$, $\{x^{n+1,k}\}$.

Since assumption (35) gives us that $\sum_{k=N^{n+1}}^{\infty} U_{i,j}^{n,k} < \infty$ almost surely for all $i, j < n+1$, we also have that $\sum_{k=N^{n+1}}^{\infty} U_{i,j}^{n+1,k} < \infty$ almost surely for all $i, j < n+1$. *Q.E.D.*

In the following proposition we use lemma 3.2 as an initial step to prove inductively that there are finitely many updates to neighboring cells when updating the estimates according to the process $x^{n,k}$, for all $n \in \{1, ..., S\}$. In order to complete the proof we use lemmas 3.3, 3.4 and 3.5.

Proposition 3 $\sum_{k=0}^{\infty} U_{ij}^{n,k} < \infty$ *almost surely for all* $n > 1$, *and* $1 \leq i, j \leq n$, $i \neq j$.

Proof:
We prove this by induction on n. From lemma 3.2 we have the case for $n = 2$. We assume that the proposition holds for the problem with n cells and then we prove that it holds for $n+1$ cells. Now have three cases: 1) $j = n+1$, 2) $i = n+1$, and 3) $i, j < n+1$. Cases 1,2 and 3 follow by the induction hypothesis and lemmas 3.3, 3.4, 3.5 respectively.

Q.E.D.

We use the result of proposition 3 along with the result of proposition 2 to prove the convergence of the sequences $\{x_i^{n,k}\}_{k\geq 0}$ in the theorem that follows.

Theorem 3 *The sequences* $\{x_i^{S,k}\}_{k\geq 0}$ *converge to* μ_i *almost surely for all* $i \in \{1, ..., S\}$.

Proof: By proposition 3 and proposition 2 we have that for all $n \in \{1, ..., S\}$ the sequences $x_i^{n,k}$ converge almost surely to the respective mean values, μ_i. Thus, by letting $n = S$ we get that:

$$x_i^{S,k} \to \mu_i, \quad as \quad k \to \infty \quad a.s.$$

Q.E.D.

The above theorem proves the almost sure convergence of the DOME algorithm since the algorithm estimates the mean values of the discrete monotone function f using the convergent sequences $x^{S,k}$.

4 Conclusions

In this paper we have presented an algorithm for the on-line estimation of monotone functions. The algorithm falls in the class of stochastic approximation algorithms, which were pioneered by Robbins and Monro DOME has the useful property that the estimates of the function are monotone at each iteration. The monotonicity of the estimates is preserved by a projection step at each iteration.

The DOME algorithm is convergent but can also be used as a method for approximating monotone functions at a small number of iterations. Relaxing some of our assumptions, such as having a strictly positive probability of visiting each state, could also result in an approximation algorithm for estimating monotone functions.

References

Andradottir, S. (1995), 'A method for discrete stochastic optimization', *Management Science* **41**(12), 1946–1961.

Ayer, M., Brunk, H. D., Ewing, G. M., Reid, W. T. & Silverman, E. (1955), 'An empirical distribution function for sampling with incomplete information', *Annals of Mathematical Statistics* **26**, 641–647.

Barlow, R. E. & Brunk, H. D. (1972), 'The isotonic regression and its dual', *Journal of the American Statistical Association* **67**(337), 140–147.

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. & Brunk, H. D. (1972), *Stastistical Inference under Order Restrictions; The Theory and Application of Isotonic Regression*, John Wiley & Sons Ltd., chapter Isotonic Regression, pp. 1–55.

Bertsekas, D. & Tsitsiklis, J. (1996), *Neuro-Dynamic Programming*, Athena Scientific, Belmont, MA.

Blum, J. (1954*a*), 'Approximation methods which converge with probability one', *Annals of Mathematical Statistics* **25**, 382–386.

Blum, J. (1954*b*), 'Multidimensional stochastic approximation methods', *Annals of Mathematical Statistics* **25**, 737–744.

Brunk, H. (1955), 'Maximum likelihood estimates of monotone parameters', *Annals of Mathematical Statistics* **26**, 607–616.

Brunk, H., Ewing, G. & Utz, W. (1957), 'Minimizing integrals in certain classes of monotone functions', *Pacific journal of mathematics* **7**, 833–847.

Dupac, V. (1987), 'Quasi-isotonic regression and stochastic approximation', *Metrika* **34**, 117–123.

Dvoretzky, A. (1956), On stochastic approximation, *in* J. Neyman, ed., 'Proc. 3^{rd} Berkeley Sym. on Math. Stat. and Prob.', Berkeley: University of California Press, pp. 39–55.

Ermoliev, Y. (1968), 'On the stochastic quasi-gradient methods and stochastic quasi-Feyer sequence', *Kibernetika*.

Ermoliev, Y. (1969), 'On the stochastic quasigradient method and stochastic quasi-feyer sequences', *Kibernetika*.

Ermoliev, Y. (1988), Stochastic quasigradient methods, *in* Y. Ermoliev & R. Wets, eds, '*Numerical Techniques for Stochastic Optimization*', Springer-Verlag, Berlin.

Ermoliev, Y. M. & Nekrylova, Z. V. (1967), The method of stochastic subgradients and its applications, *in* 'Notes, Seminar on the Theory of Optimal Solution. Academy of Sciences of the U.S.S.R.', Kiev.

Ermoliev, Y. M. & Shor, N. Z. (1968), 'Method of random walk for two-stage problem and its generation', *Kibernetika*.

Gladyshev, E. G. (1965), 'On stochastic approximation', *Theory of Prob. and its Appl.* **10**, 275–278.

Hanson, D. & Mukerjee, H. (1990), 'Isotonic estimation in stochastic approximation', *Statistics and Probability Letters* **9**, 279–287.

Kiefer, J. & Wolfowitz, J. (1952), 'Stochastic estimation of the maximum of a regression function', *Annals Math. Stat.* **23**, 462–466.

Mammen, E. (1991), 'Estimating a smooth monotone regression function', *The Annals of Statistics* **19**(2), 724–740.

Mukerjee, H. (1981), 'A stochastic approximation by observations on a discrete lattice using isotonic regression', *The Annals of Statistics* **9**(5), 1020–1025.

Papadaki, K. & Powell, W. B. (2002), 'A monotone adaptive dynamic programming algorithm for a stochastic batch service problem', *European Journal of Operational Research* **142**(1), 108–127.

Robbins, H. & Monro, S. (1951), 'A stochastic approximation method', *Annals of Math. Stat.* **22**, 400–407.

Wasan, M. (1969), Stochastic approximations, *in* 'Cambridge Transactions in Math. and Math. Phys. 58', Cambridge University Press, Cambridge.

Wolfowitz, J. (1952), 'On the stochastic approximation method of Robbins and Monro', *Annals Math. Stat.* **23**, 457–461.

Wright, F. (1978), 'Estimating strictly increasing regression functions', *Journal of the American Statistical Association* **73**(363), 636–639.