NTNU
Norwegian University of
Science and Technology

Sequential value information for Markov random field

Tommy Sneltvedt

Master of Science in Physics and Mathematics
Submission date: June 2011
Supervisor: Jo Eidsvik, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Contents

1 Preface

This thesis was written as a part of my master degree at NTNU (Norwegian University of Science and technology). I would like to thank my supervisor, Jo Eidsvik, for his insightful comments and helpful advices. Another thanks goes to Gabriele Martinelli and Haakon Michael Austad who kindly allowed me to use part of his code in my thesis.

2 Abstract

This papers deals with exploration of binary Markov random fields. Several different strategies have been tried out, among them the myopic and exact approach. The latter one computes all paths, which becomes quite computer intensive for large grids. Due to this a blockwise procedure has been tried out, conditioning future steps on outcomes in a few adjacent former blocks. The blockwise partitioning of the field have been tested out on different cases. While it often performs quite good, the limited memory of the method can be a source of problems.

3 Introduction

This paper is concerned with exploration of a Markov Random fields by means of different methods, sequential and non sequential. This notion of spatial correlation between different objects (i.e. oil, gas et cetera) is of huge importance in many industries [6, 7] The goal is to perform as accurate predictions as possible. Due to the innate complexity in many such spatially correlated fields, finding the optimal strategy is often impossible. To deal with this issue many different approaches have been developed. In the case of Markov Random Field even relatively small grids demands a lot of computing to perfom optimal sequential routines. To deal with this issue a blockwise partitioning of the field will be performed. We will start on small grids and later use them as building blocks for larger grids and make use of modified sequential strategies to deal with these new and larger grids.

4 Markov Random Field (MRF)

This section presents the Markov Random Field model and two methods for sampling realizations from such a field. The first method is outlined in Section 4.2 and was developed by Reeves and Pettitt [8]. It applies a recursive forward and backward routine to obtain realizations and compute the marginals respectively. The second method (Section 4.3) gives an approximation to binary Markov random fields by means of treating a set of interaction parameters defined later on. This method is used to draw realizations from large grids (which cannot be done with the other routine due to the large computing time involved in exact computations).

4.1 Model formulation

The topic of this text is the Markov Random Field model. Such models can be viewed as an extension of the simpler Markov chain model. This is a model for a stochastic process where a future state only depends on the present state. This memoryless feature is known as a Markov property. For a Markov sequence of random variables $\{X_1, X_2, X_3, \dots\}$ this implies that:

$$p(x_n|x_1, x_2, \dots, x_{n-1}) = p(x_n|x_{n-1}) \tag{1}$$

Usually the sequence of random variables depends on a temporal or spatial property. The Markov random field model can be viewed as an extension of this model. Instead of having a one dimensional line of progress, it is extended to a network of interconnected nodes, each one representing a stochastic variable.

For the purpose of this thesis only a two dimensional rectangular grid will be considered. The indexing of cells in such grids will be similar to the one depicted in Figure 1. The number of rows and columns will br n_1 and

1	4	7
2	5	8
3	6	9

Figure 1: A 3×3 grid. The indexing above will be used to denote cells.

n_2 respectively. The total number of cells is defined as $n = n_1 n_2$. The set of all indices of a grid is denoted $\mathcal{I}$. The outcome of the grid at point i is denoted x_i. The whole grid can be described as $\mathbf{x} = (x_1, \dots, x_n)$. The full

conditional distributions (also known as local characteristics) are:

$$p(x_i|x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = p(x_i|x_j; j \in \partial_i) \tag{2}$$

For all $i \in \{1, 2, \ldots, n\}$. ∂_i is some sort of neighborhood structure. There are many possible choices of neighborhood structures, but commonly they consists of the four, eight or twelve closest neighbors. A proper neighborhood system $\{\partial_i; i \in \mathcal{I}\}$ must fulfill two requirements:

$$i \notin \partial_i \tag{3}$$

$$j \in \partial_i \Leftrightarrow i \in \partial_j \tag{4}$$

These two condition implies that a cell can never be a neighbor to itself and that there is a mutual neighboring relationship between points in a neigborhood. For this paper the neighborhood will be chosen such that the neighborhood for any site is the adjacent cells. If such a neighborhood is chosen the neighborhood will take nine different forms depending on the boundaries. This is shown in Figure 2. The grey colored cells represents a neighborhood with respect to the adjacent black cells. It is tedious to show

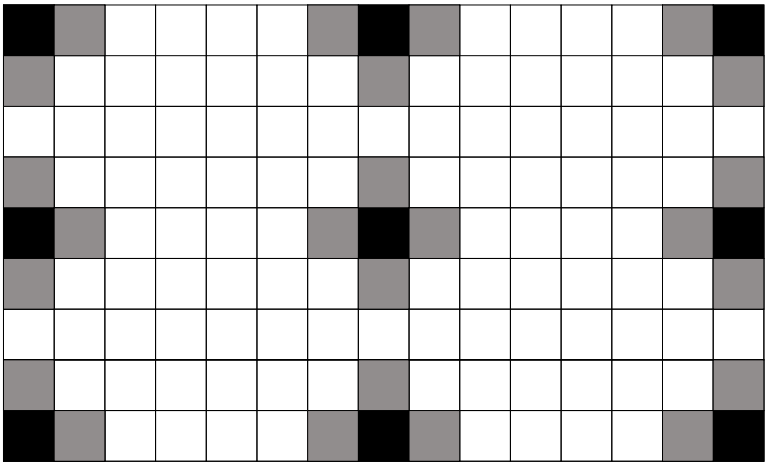


Figure 2: Depicts nine different neighborhoods (grey) with respect to nine sites (black). These nine spatial configurations are the only ones used for a rectangular grid.

that the resulting set of conditional distributions defines a legitimate joint probability distribution. The consistency requirements can be checked in the theory, but it is easier to use the Hammersley-Clifford theorem to arrive at a proper Markov Random Field. Before introducing this theorem the notion of cliques must be explained. A clique is a subset c of the grid $\mathcal{I}$ such that

every pair of sites are neighbors, or in a more mathematical form:

$$\left.\begin{matrix} u \in \partial_v \\ v \in \partial_i \end{matrix}\right\} \quad \text{for all } u, v \in c \text{ and } u \neq v$$

The Hammersley-Clifford theorem is defined as:

Hammersley-Clifford 1. *A joint probability distribution $p(\mathbf{x})$, such that $p(\mathbf{x}) > 0$ for all values of $x \in \chi$, is a Markov random field if, and only if, $p(\mathbf{x})$ can be written on the following form:*

$$p(\mathbf{x}) = \frac{1}{Z} exp(-\sum_{c \in C} \Phi_c(x_c)) \tag{5}$$

and the normalising constant is equal to:

$$Z = \sum_{x \in \chi} exp(-\sum_{c \in C} \Phi_c(x_c)) < \infty \tag{6}$$

C is the set of all cliques and the functions Φ_c is known as potential functions and can be chosen arbitarily. χ is the of the set of all possible configurations of x.

This satisifies the Markov property discussed earlier. This theorem can be used two ways: (a) defining new Markov fields through appropriate choices of potential functions or (b) to validate that a certain set of conditional distributions defines a valid joint probability distribution. For a more comprehensive account on this theorem and related topics see [1].

While many highly complex models can be defined by some set of potential functions, the one to be studied closer in this text is the Ising model. This is obtained by defining the potential functions as follows:

$$\Phi_c(\mathbf{x}) = \begin{cases} -\beta & \text{If cells in click } c \text{ has the same outcome} \\ 0 & \text{else} \end{cases} \tag{7}$$

This leads to the following joint probability distribution

$$p(\mathbf{x}) = \frac{1}{Z} exp(\beta \sum_{i \sim j} I(x_i = x_j)) \tag{8}$$

The normalising constant, Z, is equal to:

$$Z = \sum_{\mathbf{x} \in \chi} exp(\beta \sum_{i \sim j} I(x_i = x_j)) \tag{9}$$

$I(*)$ is an indicator function defined as follows:

$$I(S) = \begin{cases} 1 & \text{if } S \text{ is true} \\ 0 & \text{else} \end{cases} \tag{10}$$

The first expression in the exponent, $\sum_{i\sim j}$ indicates that every pair of x_i and x_j standing adjacent to each other must be summed with respect to the indicator function. In short this implies that neighbors having the same outcome will increase the value of the summation. Using the notation from the grid above this gives $n_1(n_2-1)+n_2(n_1-1)$ pairs of cells (or cliques) being checked.

Notice that information can be encoded into the field through a point-wise prior function $\alpha_i(x_i)$. This is straightforward from Equation 8 since the exponent can be rewritten as: $\sum_{c\in C}\Phi_c(x_c)=\sum_{c\in C}(\phi_c^1(x_c)+\phi_c^2(x_c))$. The two functions can be chosen arbitarily. Since there are more cliques than cells for an ordinary rectangular grid (with minimum two rows and columns), one unique clique (c) can be picked for every cell i (assuming that $i\in c$) and the corresponding $\phi_c^2(x_c)$Âă can be defined to only depend on the value of i $\phi_c^2(x_c)=\phi_c^2(x_i)$. Call this set of cliques for G. Define $\phi_c^2(x_c)=0$ $\forall c\notin G$. If $\phi_c^1(x_c)$ is defined in similar manner as in Equation 7 the new joint distribution becomes:

$$p(\mathbf{x})=\frac{1}{Z}exp(\beta\sum_{i\sim j}I(x_i=x_j)+\sum_{i=1}^{n}\alpha_i(x_i)) \tag{11}$$

where $\phi_c^2(x_i)$ has been renamed $\alpha_i(x_i)$. This is the general form that will be used in this thesis.

Assuming that $\beta>0$ it follows that an increase in the β-value will increase the likelihood of outcomes where neighboring cells are the same. This known as the spatial dependency paramter. The second sum in Equation 11 is the sum of the point-wise prior function ($\alpha_i(x_i)$). In general it differs from each cell i and depends only on the outcome of that cell, x_i. This function can be used to add prior information about the field into the model. Say some information is available from a group of experts, a series of measurements or some other kind of prior knowledge, this can encoded into the model by defining appropriate values of $\alpha_i(x_i)$. It should be noted that if $\alpha_i(x_i)$ is zero for all cells independent of the outcome, then the marginal probability of each state is equally likely for all cells. This is also known as the uninformative case.

It should be stressed that the MRF-model is not limited to rectangualar grids nor the particular neighborhood relation chosen her. For more on the model see [5, 4].

4.2 Exact computation

The method to be applied in this section [8] is useful for small grids, but becomes unfeasible for larger grids due to rapid increase in computational time. As mentioned the method was developed by Reeves and Pettitt (2004) and is valid for any factorisable model. A general factorisable model is one

where the discrete probabaility distribution ($p(\mathbf{x}) = (x_1, x_2, \ldots, x_n)$) can be written as a product of terms taking the following form: $q_i(y_i, y_{i+1}, \ldots, y_{i+j})$. This is the case for any discrete Markov random field. The Reeves/Pettitt model has earlier been applied to the Ising model by by Bhattacharjya et al [2].

The probability function of the Ising model (Equation 11) can be rewritten in the following form:

$$p(\mathbf{x}) = \frac{q_1(x_1|x_2, \ldots, x_{1+n_1})q_2(x_2|x_3, \ldots, x_{2+n_1}) \ldots q_{n-1}(x_{n-1}|x_n)q_n(x_N)}{z} \tag{12}$$

These n functions of q_i above take four different forms depending on position in the grid (must be adjusted with respect to the boundary conditions. In short this means that cells not confined to either the easternmost colum nor the bottom row will take the following form: $q_i(x_i|x_{i+1}, \ldots, x_{i+n_1}) = exp(\beta[I(x_i = x_{i+1}) + I(x_1 = x_{1+n_1})] + a_i(x_i))$. If it belong to the last row but not the last column it takes a simplified form: $q_i(x_i|x_i, \ldots, x_{i+n_1}) = exp(\beta I(x_i = x_{i+n_1}) + a_i(x_i))$. Easternmost column bot not on bottom row: $q_i(x_i|x_i, x_{i+1}) = exp(\beta I(x_i = x_{i+1}) + a_i(x_i))$. And then the last case: Bottom row and last column: $q_i(x_n) = a_i(x_n))$. According to the method outlined by Reeves and Pettitt for general factorisable models the normalizing constant can be archieved by applying a recursive formula. The first value computed is defined as:

$$z_1(x_2, \ldots, x_{1+n_1}) = \sum_{x_1=0}^{d-1} q_i(x_1|x_2, \ldots, x_{1+n_1}) \tag{13}$$

It continues by computing:

$$z_i(x_2, \ldots, x_{1+n_1}) = \sum_{x_1=0}^{d-1} q_i(x_i|x_{i+1}, \ldots, x_{i+n_1}) \tag{14}$$

These forumlas are adjusted when reaching the final column since the buffer configuration isn't defined outside x_n. The final value, z_n, is equal to the normalizing constant, i.e.:

$$z = z_n = \sum_{x_n=1}^{d} q(x_n)z_{n-1}(x_n) \tag{15}$$

The process above is just a clever way of summing all possible outcomes of the nominator of the fraction in 12. The computational efficiency of computing the values of $z_i \; \forall i \in \{1, \ldots, n\}$ depends on the row length, number of categorical variables, d and the size of the grid. The process outlined above

[0]For each $z_i(\cdot)$ a number of possible buffer configurations are available depending on the value of d, n and n_1. In the case where $i \leq n - n_1$ there are d^{n_1} buffer configurations

uses a forward tecnique (Starting in the uppermost left corner, x_1 continuing to cell x_2, x_3 et cetera), contrary to the next tecnique working backwards. It utilizes the z_i values obtained earlier. Realizations from the grid is obtained by sampling from the first marginal distribution which is obtained as follows:

$$p_n(x_n) = \frac{1}{z}\sum_{x_1=1}^{d}\cdots\sum_{x_{n-1}=1}^{d}\{\prod_{i=1}^{n}q_i(\mathbf{x})\} = \frac{1}{z}q_n(x_n)z_{n-1}(x_n) \tag{16}$$

This process is continued and another value is obtinaed from the distribution of x_{n-1} conditioned on x_n, i.e.

$$\begin{aligned} p(x_{n-1}|x_n) &= \frac{1}{z}\frac{\sum_{x1=1}^{d}\cdots\sum_{x_{n-2}=1}^{d}\{\prod_{i=1}^{n}q_i(\mathbf{x})\}}{p(x_n)} \\ &= \frac{q_2(x_{n-1}|x_n)z_{n-2}(x_n,x_{n-1})}{z_{n-1}(x_n)}s \end{aligned}$$

And it continues in the fashion until a whole realization is obtained. In short this means solving:

$$p(x_1,x_2,\ldots,x_n) = p(x_1|x_2,\ldots,x_n)\ldots p(x_{n-1}|x_n)p(x_n) \tag{17}$$

Note how efficiently this is done by simply storing the values of $z_i(x_{i+1},\ldots,x_{i+n_1-1}x_{i+n_1})$ for all buffer configurations. The marginals are obtained straightforwardly by noticing that:

$$\begin{aligned} p_i(x_i) &= \frac{1}{z}\sum_{x_1=1}^{d}\cdots\sum_{x_{i-1}=1}^{d}\sum_{x_{i+1}=1}^{d}\cdots\sum_{x_{x_n}=1}^{d}\{\prod_{i=1}^{n}q_i(\mathbf{x})\} \\ &= \frac{1}{z}\sum_{x_n=1}^{d}\cdots\sum_{x_{n-i}=1}^{d}q_n(x_n)\cdots q_i(x_i|x_{i+1},\ldots x_{i+n_1})z_{i-1}(\mathbf{x}) \end{aligned}$$

This procedure allows us to obtain the conditional probabilities by encoding information about the cells into the α-function. For example, assume that cell j is known to have outcome r, theb this information can be encoded into the distribution by setting the $\alpha_j(r)$ to a high value. Computing the marginal probability for any grid point i will now for all practical application be equal to $p(x_i|x_j = r)$. This follows since it is extremely unlikely that

- $(x_{i+1},\ldots,x_{i+n_1})$ - to be considered for each i. When $i > n - n_1$ the number of buffer configurations equals: $d^(n-i)$. From this consideration it is obvious that some constraint must be set on the size of n_1, else the procedure becomes to computionally costly. The paper [ref] advices using a value of n_1 less than 20 if $d = 2$. Note that the effect of increasing the number of categorical variables have a monementus effect on the running time

$x_j \neq r$. A simple proof follows for general d. Say that $p(x_i|x_j = r)$ is of interest where $r \in \{0, \ldots, d-1\}$ and $i, j \in \mathcal{I}$. Then $\alpha_j(x_i)$ is set equal to:

$$a_j(x_j) = \begin{cases} \infty & \text{if } x_j = r \\ 0 & \text{else} \end{cases} \tag{18}$$

This means that $p(x_j = r) \to 1$ and $p(x_j = h) \to 0 \ \forall r \in \{0, \ldots, d-1\} \setminus r$.

$$\begin{aligned} p(x_i) &= \sum_{x_j=1}^{d} p(x_i|x_j)p(x_j) = \sum_{x_j \in \{0,\ldots,d-1\}\setminus r} p(x_i|x_j)p(x_j) + p(x_i|x_j = r)p(x_j = r) \\ &\to 0 + p(x_i|x_j = r) = p(x_i|x_j = r) \end{aligned}$$

These conditional probabilities are needed to compute sequential paths later on.

4.3 Approximate computation for large grids

This method is needed in order to sample from large grids. This process becomes computationally inefficient using the exact algorithm above, hence this approximate method is utilized in solving this problem. The method applied was developed by Tjelmland and Austad [9]. For a probability distribution that can be written on the following form:

$$p(x) = c \exp(-U(x)) \tag{19}$$

where c is a normalizing constant, can be rewritten in form of interaction parameters:

$$U(x) = \sum_{\Lambda \in \mathcal{P}(S)} \beta_U(\Lambda) \prod_{k \in \Lambda} x_k \tag{20}$$

Here $\mathcal{P}(S)$ is the powerset, that is $\mathcal{P}(S) = \{\Lambda | \Lambda \subseteq S\}$. Many of these interaction parameters $\beta = \{\beta(\Lambda), \Lambda \subset S\}$. Using this form of representation in the forward-backward algorithm allows can simplify the computational effort by removing interaction parameters close to zero. This can be used to simulate realization from large grids. The details of this is quite complicated and not relevant apart from the sampling business.

5 Strategies for small grids

This section deals with strategies for small grids as opposed to large grids which is the topic of the next Section. To make the principles more tangible the model is assumed to be a statistical description of an oil field with two possible outcome ($d = 2$) for each cell. The cells are either dry or with oil, the goal is to develope methods for picking cells containing oil. The following values are assumed known:

- The cost of failure, C_i. This value can vary across the grid. Obviously some areas might be more expensive to drill due to differences in depth, geology, logistics etc.
- The profit if oil is found, R_i. This number also includes costs related to drilling, so potentially the number can be negative.
- The spatial dependency parameter β. This parameter defines the spatial correlation in the field. If set to zero this would imply that there is no correlation between grid points.
- $\alpha_i(x_i)$ contains prior knowledge about each cell point/region.

The two possible outcomes are $x_i = 0$ (dry) and $x_i = 1$ (oil).

For a drilling scenario several strategies are possible. One strategy is to pick all the sites to be drilled simultaneously (joint strategy). Compared to sequential strategies (i.e. cells are being drilled in a consecutive manner, either one at a time or in groups) this method is inferior. The only exception is when there is no spatial relation between points, i.e. that the spatial dependency parameter is equal to zero. In some cases performing a non-sequential strategy might be the only option, for example if some time constrain is imposed.

The squential procedures differs from the joint strategy due to their temporal aspect. For each drilling performed more information is available, information that can be utilized to improve on the selection of the next drilling site. This allows for multiple strategies ranging from an exact optimal path procedure (see Section 5.3) to more approximate approaches. One such heuristic strategy is the naive strategy [6] which ignores the spatial relationship between points. Another approximate sequential algorithm is the myopic strategy which selects the best region of drilling (highest expected income) dependent on earlier outcomes. In other words for every single step two possibilities remains, oil or no oil, which affects the next step. In that way a tree of possible paths opens up. The same goes for the optimal procedure which similarly offers instructions on which cell to pick conditioned on earlier outcomes.

It should be stressed that while the methods outlined in this Section (as well as the next) deals with only two outcomes, they are easily extendable to more general cases.

5.1 Joint strategy

The following formulas are a slightly modified form of the formulas presented in [2]. If all the drilling procedures are performed simultaneously, the expected value of a drilled cell is:

$$V_i = \max\{R_i p_i(1) - C_i p_i(0), 0\} \tag{21}$$

$p_i(x_i)$ is the marginal probability of cell i. The expected income from all cells, which will be called joint value (JV) from this point on, is hence:

$$JV = \sum_{i=1}^{n} \max\{R_i p_i(1) - C_i p_i(0), 0\} \tag{22}$$

If the decision maker could get information concerning the outcome at each site (for example by a clairvoyant person), this would radically change the situation. If this information can be obtained for free the expected income is known as the value of free clairvoyance or simply VFC. Since we know beforehand which cells are with oil and which are without the value of free clairvoyance becomes:

$$VFC = \sum_{i=1}^{n} p_i(1) \max\{R_i, 0\} \tag{23}$$

In the this latter case perfect information is available. To establish how much this information is worth (i.e. how much one is willing to pay for it) one has to substract the value of free clairvoyance from the joint value (in the case of a joint strategy). value. This difference is known as the value of perfect information (VOPI) and takes the following mathematical form: which takes the following form: $VOPI = VFC - JV$. This is the upper limit to the price one is willing to pay for perfect information if the alternative approach is a joint strategy. This choice of $VOPI$ depends on the utility function being a straight-line or exponential.

It is never possible to find a strategy that on average beats the VFC-value. This is of course because no strategy deals with absolute certainty, and the penalty for not knowing is that mistakes are done (or likely to be done). No mistakes are possible when all the outcomes are known. Hence this value can be used as an upper limit to any strategy, sequential or non-sequential. Later on in Section 5.5 another boundary is introduced. This boundary represents the maximum value of any sequential procedure.

5.2 Triplets or duplets

Another strategy is drilling triplets or duplets in order to maximize income. For a triplet this implies finding three sites such that the information that becomes available (together with the outcomes of the cells)

maximizes the expected income. Obviously the joint value for the remaining cells change after the drilling has performed changes due to more information being available. Let the triplets of cells drilled be named evidence, or e. Evidence is defined as $e = \{x_i, x_j, x_k\}$ for some i,j and k such that $i, j, k \in \{1, 2, \ldots, n\}$, The joint value of the remaining cells are hence $JV = \sum_{u \in \mathcal{I} \setminus \{i,j,k\}} \max(R_u p_u(1|e) + C_u p_u(0|e), 0)$. However the evidence obtained after drilling will also affect the expected income of the strategy (selection of triplets). This is accounted for by computing the expected value of the triplet itself. The expected value for some choice of evidence hence become:

$$GS_{i,j,k} = \sum_{e \in \{0,1\}^3} \sum_{u \in \mathcal{I} \setminus \{i,j,k\}} \max(R_u p_u(1|e) + C_u p_u(0|e), 0) p(e) + \sum_{e \in \{0,1\}^3} \Phi(e) p(e)$$

The last function $\Phi(e)$ simply computes the profit or loss when drilling the three cells. This is $\Phi(e) = \sum_{r \in \{i,j,k\}} R_r I(x_r = 1) + C_r I(x_r = 0)$. This a strategy for finding the three best cells offering the best prospects for income. The best group strategy will be the one(s) satisfying the following condition:

$$\{i, j, k\} \arg\max_{\{i,j,k\}} GS_{i,j,k} \tag{24}$$

The same procedure as above is applied to duplets, with appropriate modifications. It should be noted that the number of possible configurations of cells for a p-tuplet collection is equal to $\binom{n}{p} = \frac{n!}{p!(n-p)!}$.

5.3 Exact sequential procedure

The most comprehensive apprach is computing the optimal path. This is done following the dynamic programming to compute all possible paths. The description here follows the paper of Bickel et al [3].

A state vector $\omega = (\omega_1, \omega_2, \cdots, \omega_n)$ is defined where n is the number of possible drilling sites and w_i is the present state at grid i. If oil has been found at grid i the value of w_i is 1, while it is equal to 0 if has been drilled and found dry and denoted '*' if no drilling have yet occured. For example a state vector $\omega = (*, *, *, 1, 0, *)$ implies that only cell four and five have been drilled, furthermore only grid four resulted in oil being found. The expected value of the future cash flow with respect to present state is denoted $v(\omega)$. If every well has been subject to drilling (i.e. ω consists only of zeros and ones) then the expected income must be equal to zero. At state ω the expected future income, granted the next drilling site is at grid i becomes:

$$v_i(\omega) = \frac{\mu(\omega_i^1)}{\mu(\omega)}(R_i + \delta v(\omega_{\mathbf{i}}^{\mathbf{1}}))) + \frac{\mu(\omega_i^0)}{\mu(\omega)}(C_i + \delta v(\omega_i^0)) \tag{25}$$

δ is a discount factor chosen in the range $[0, 1]$. So if two paths are judged equally good with $\delta = 1$, a lower δ will make the algorithm pick the path

that is expected to offer the income earlier in the stepwise procedure than the competing paths. $\mu(\omega)$ is the probability associated with ω, meaning that the non-drilled sites are summed out of the probability distribution. If $\omega = (0, 1, *, *, 1, *)$ then $\mu(\omega) = \sum_{w_3=0}^{1} \sum_{w_4=0}^{1} \sum_{w_6}^{1} p(0, 1, w_3, w_4, 1, w_6)$. In the equation above ω_i^1 means that ω_i^1 is similar to ω apart from site i being $w_i = 1$. In other words the fraction $\frac{\mu(\omega_i^j)}{\mu(\omega)}$ is a conditional probability distribution. Using $\omega = (0, 1, *, *, 1, *)$ means $\frac{\mu(\omega_3^1)}{\mu(\omega)} = p(x_3 = 1 | x_1 = 0, x_2 = 1, x_5 = 1)$.

To obtain the conditional distributions for each cell, one can, as the author, use the recursive computations in Section 4.2 to obtain the appropriate conditional probabilities by changing the α-values. The process follows the philosophy of Equation: 19.

The optimal continuation value for state ω then becomes $\{ v_i(\omega), 0\}$. The trick is to compute all the values of v_i with respect to all states, then compare these to obtain the optimal path. This leads to a tree structure parting into two different brances for every cell (depending on the outcome of that cell, oil or no oil). Hence the outcome at a cell will affect the future cells to be selected, illustrating how the information gained from previous outcomes affect the way new drilling sites are picked.

Note that for possible outcomes for a cell, each selection will lead to d different branches. From this the maximum number of possible sequences equals: $\sum_{i=1}^{n} d^{i-1}$. But to get this tree all possible combinations must be computed before arriving at a final desicion tree that shows all the prefered paths during the sequential run.

The algorithm below can be used to compute all such paths. The input is a matrix called hashTable. Every row contains $n + 2$ cells. The first n cells tells the situation for the present state, i.e. each cell represents a cell following the indexing from Figure 1 - If one of the cells have the value $d+1$ this implies that no drilling have yet occured. The $(n + 1)$'th cell in a row represents the value v for this particular state (i.e. the expected income conditioned on starting in the state specified in the same row). The $(n + 2)$'th value points in the optimal cell to pick next from the present state. 0 means that no more cells should be drilled. For a 2×2 grid with $d = 2$ a row $(3, 1, 3, 3, 2.3, 1)$ means that the present state is $\left(\begin{smallmatrix} * & * \\ 1 & * \end{smallmatrix}\right)$ where '*' marks an undrilled site. Furthermore the expected income from this state with exact approach is 2.3. The next cell to be drilled is 1. From this list it is obvious how to compute the next cell to be drilled (if any). The list is computed dynamically in the algorithm below. The starting point should be $(3, 3, 3, 3)$ as one is interested in the complete tree. The table input is a vector containing all the conditional probabilities obtained using the method from 4.2.

$\omega = (3, 3, 3, 3)$

[hashTable,v] = computeSequenceExact(ω, table, revenue, cost, δ, hashTable)
if ω has no more undrilled cells **then**
$v = 0$
$site = 0$

Note that $v = 0$ means that there is no further path to take, at least that is economically feasible.
if ω is not stored in hashTable **then**
Store ω, v and $site$ in hashTable.
end if
else
if ω is already stored in hashTable **then**
Get v and $site$ that corresponds to ω in the hashTable.
end if
else
$length \leftarrow$ number of undrilled cells
for $i = 1$ to length **do**
Compute $v_i(\omega)$ for the i'th undrilled cell. This is done by calling the function dynamically. For the i'th undrilled cell compute:

[hashTable,oil]=computeSequence(ω_i^1,table,revenue,cost,δ,hashTable)
[hashTable,dry]=computeSequence(ω_i^0,table,revenue,cost,δ,hashTable)

Get the probability for oil, probOil and probability for no oil, probNoOil from table.

From this compute vNew(i)=probOil*(rev+δ*oil)+probNoOil*(cost+δ*dry)

See equation 25 for reference.
end for
Find the largest vNew(i) computed in the for loop. The value of ω, v and i is stored in the hashTable. If all values of vNew is negative store in hashTable with v set to zero.
end if

At this point hashTable contains all $v(\omega)$ $\forall\omega$. Thus one can just traverse the hashTable to obtain the best possible path.

5.4 Myopic sequential procedure

The myopic search computes the best path by stepwise procedure where the best continuation is chosen at each step conditioned on the previous outcomes. This means computing a tree with $\sum_{i=1}^{n} d^{i-1}$ brances. The first

drilling site is found simply by picking the cell offering the largest expected income. This site is:

$$x_{(1)} = \arg\max_i \{p_i(1)R_i - p_i(0)C_i, 0\} \tag{26}$$

The outcome at $x_{(1)}$ will determine the future path. The second step must be conditioned on both outcomes (oil or dry). Hence two paths must be computed (more generally d paths): $x_{(2)}|x_{x_{(1)}} = 0$ and $x_{(2)}|x_{(1)} = 1$.

$$x_{(2)}|x_{x_{(1)}} = 0 = \arg\max_{i \setminus x_{x_{(1)}}} \{p(x_i = 1|x_{x_{(1)}} = 0)R_i - p(x_i = 0|x_{x_{(1)}} = 0)C_i, 0\} \tag{27}$$

Likewise:

$$x_{(2)}|x_{x_{(1)}} = 1 = \arg\max_{i \setminus x_{x_{(1)}}} \{p(x_i = 1|x_{x_{(1)}} = 1)R_i - p(x_i = 0|x_{x_{(1)}} = 1)C_i, 0\} \tag{28}$$

And the third step is conditioned on the outcome of the three other steps (4 configurations if $d = 2$). This gives the following formulas for expected income v_i at the i'th step:

$$v_1 = \max\{p_i(1)R_i - p_i(0)C_i, 0\} \tag{29}$$

Since the next step is computed conditioned on the outcome of $x_{(1)}$ this becomes:

$$v_2 = \sum_{j=0}^{1} [\max\{p(x_{x_{(2)}} = 0|x_{x_{(1)}} = j)R_i + p(x_{x_{(2)}} = 1|x_{x_{(1)}} = j)C_i\})p(x_{x_{(1)}} = j)] \tag{30}$$

v_i for $i > 2$ is computed in an analogious manner. δ is the discount factor. The expected income is hence:

$$v = \sum_{i=1}^{n} \delta^{i-1} v_i \tag{31}$$

5.5 Upper limit for sequential approaches

It is is possible to find an upper limit for any drilling scheme by noticing that the conditional probability of any cell only depends on the outcome of the neighboring cells (can either be 4,3 or 2 neighboring cells depending on the position of the cell in question). This is an essential feature of the Markov Random field model. There are all nine spatially distinct neighborhood structures. These were depicted earlier in Figure **??**.

This suggests a natural estimate for the upper bound of sequential algorithms, in short the more neighbors we know the better descision can be

taken concerning the outcome of a cell. It is the spatial relationship between points that makes the sequential procedures reasonable as each new cell adds important information about the field. Without any spatial correlation the joint strategy would produce the same results as any of the two sequential algorithms defined so far. But as the spatial dependency increases it allows the sequential algorithms to make more informed choices. Deciding to drill conditioned on the outcome of all neighboring cell is the upper limit for any sequential algorithm. This limit can be obtained through Monte Carlo simulation. If M realizations are obtained by sampling and x^m denotes realization m, then the approximation of upper boundary is obtained from Monte Carlo by computing the following expression:

$$\text{UP} = \frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{n}\max\{R_i p(x_i = 1|x_{-i}^m) - C_i p(x_i = 0|x_{-i}^m), 0\} \qquad (32)$$

Here n is the number of cells in the grid and x_{-i}^m means the outcome of the neighboring cells to i (known!). The value of $p(x_i = 1|x_{-i}^m)$ is:

$$\begin{aligned} p(x_i = 1|x_{-i}^m) &= p(x_i = 1|x_j^m; j \in \partial_i) \\ &= \frac{exp(\beta \sum_{j\in\partial_i} I(1 = x_j^m) + \alpha_i(1))}{exp(\beta \sum_{j\in\partial_i} I(1 = x_j^m) + \alpha_i(1)) + exp(\beta \sum_{j\in\partial_i} I(0 = x_j^m) + \alpha_i(0))} \end{aligned}$$

5.6 Cases

5.7 asf

In this Section four cases will be studied closer, the first two cases are defined in Table 1. These are both 3×3 grids offering different revenues at each cell, but with the same cost for drilling at a dry site.

Case:	$\alpha_i(0)$	$\alpha_i(1)$	Revenue	Cost
C1	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1.1 & 2 & 0.9 \\ 0.2 & 1.4 & 1.9 \\ 0.7 & 0.1 & 0.4 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$
C2	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1.1 & 0.6 & 0.1 \\ 0.4 & 0.3 & 1.5 \\ 0.2 & 0.2 & 1.7 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

Table 1: Case C1 and C2

The first case to be treated is C1. It is uninformative in the sense that the α-value is zero for any cell and outcome. This leads to the marginal

probabilities having values $p_i(0) = p_i(1) = 1, 5$ for all cells in $\mathcal{I}$. This makes the joint value quite easy to compute. It follows from Equation 22 that a cell i is drilled conditioned on $R_i > C_i$. From this one can conclude that cells numbered 1, 4, 5 and 8 will be subjected to drilling. Since the marginal probability for each cell i is $p_i(0) = p_i(1) = 0.5$, both the joint value (JV) and value of free clairvoayance (VFC) is independent of the β-value. This property follows directly from Equation 22 and 23, none of them depend on the spatial dependency parameter. With any choice of β the expected value of the joint strategy is $\frac{1}{2}\sum_{i=1}^{n} \max(R_i - 1, 0) = 1.2$.

The expected income would be $\frac{1}{2}\sum_{i=1}^{n} \max(R_i, 0) = 4.35$ (VFC). The value of this information is hence: VOPI=VFC-JV=3.15. This is confirmed by performing Monte Carlo sampling. The average income earned at each cell for $\beta = 0.0$ and $\beta = 0.5$ and $\beta = 1.5$ is depicted in table 2. The number of realizations used is 50000 in all three cases.

$\beta = 0.0$

$$\begin{pmatrix} 0.0540 & 0.5148 & 0 \\ 0 & 0.1944 & 0.4507 \\ 0 & 0 & 0 \end{pmatrix}$$

$\beta = 0.5$

$$\begin{pmatrix} 0.0466 & 0.4949 & 0 \\ 0 & 0.2046 & 0.4450 \\ 0 & 0 & 0 \end{pmatrix}$$

$\beta = 1.5$

$$\begin{pmatrix} 0.0516 & 0.5068 & 0 \\ 0 & 0.2072 & 0.4550 \\ 0 & 0 & 0 \end{pmatrix}$$

Table 2: Expected income from each cell for a joint strategy. The obtained values are Monte Carlo estimates obtained from running 50000 samples. Case run: C1.

The joint value (JV), the value of free clairvoyance (VFC) and value of information (VOI) are all given in table 3. These results are coherent with

	$\beta = 0.0$	$\beta = 0.5$	$\beta = 1.5$
JV	1.2139	1.1910	1.2205
VFC	4.3559	4.3447	4.3676
VOI	3.1420	3.1537	3.1471

Table 3: Joint value, value of free clairvoyance and value of information for three different values of β (0.0, 0.5 and 1.5). The obtained values are Monte Carlo estimates obtained from running 50000 samples. Case run: C1.

the results predicted from Equation 22 and 23. Though the value of β has no direct effect on the expected value of the joint strategy, it does affect the

variance of the strategy. This is clearly seen by comparing the income of all realizations following the joint strategy as shown in histograms in figure 5a ($\beta = 0$), 5b ($\beta = 0.5$) and 5c ($\beta = 1.5$). The variance seem to increase as the value of β becomes larger. This property is caused by the spatial correlation forcing neighboring cells to have the same values, in short making it easier to get more extreme results.
When the value β becomes large, only two outcomes are likely to happen, i.e. either all points dry or all points with oil. Then the limit approached is:

$$\sigma_{JV}^2 \to \frac{1}{2}((\sum_{i \in I} R_i - E[JV])^2 + (\sum_{i \in I} C_i - E[JV])^2) \tag{33}$$

This means that the σ_{JV}^2 will approach $\frac{1}{2}((6.4-1.2)^2+(-4-1.2)^2) = 27.0400$. The variance of the data is obtained for different values of β using Monte Carlo:

$$\hat{\sigma_{JV}} = \frac{1}{B-1} \sum_{i=1}^{B} (JV_i - \overline{JV})^2 \tag{34}$$

Where JV_i is the value of the joint strategy for sample i, $\overline{JV} = \frac{1}{B} \sum_{i=1}^{B} JV_i$ and B is equal to the number of samples. The result on running this for 10000 samples over 31 different values of β evenly spaced between 0 and 3 is depicted in Figure 3. This show that the the standard deviation of the joint strategy converge towards the value in equation 33 when β values are increased.

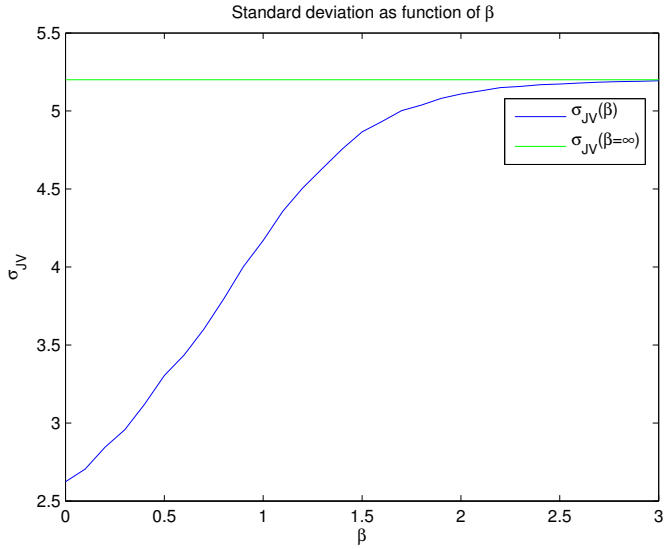


Figure 3: Shows how the standard deviation of the joint value increases in case C1 as a function of β. The values are obtained using Monte Carlo estimation using 10000 samples for each value of β. The upper line represents the value of 33.

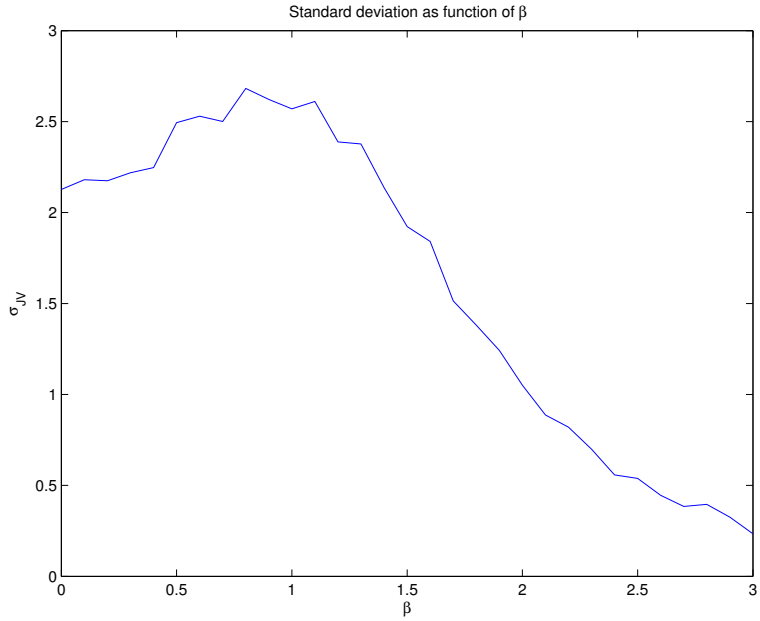
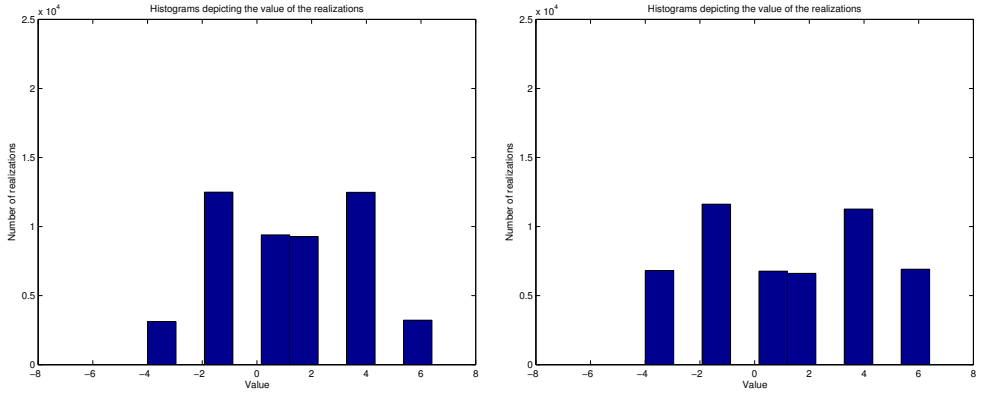
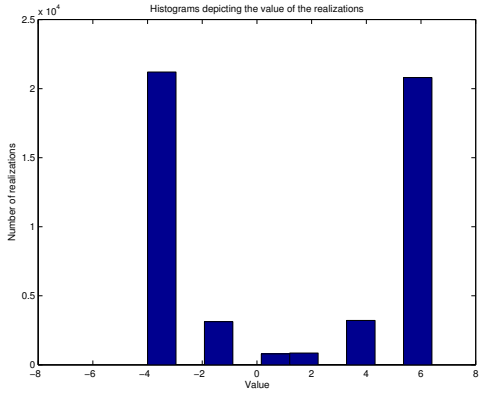


Figure 4: Shows how the standard deviation of the joint value decreases in case C2 as a function of β. The values are obtained using Monte Carlo estimation using 10000 samples for each value of β.

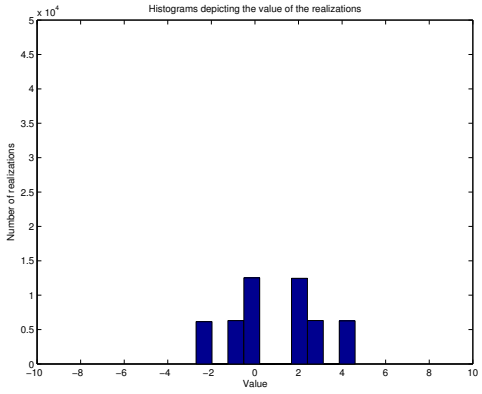


(a) C1, $\beta = 0$

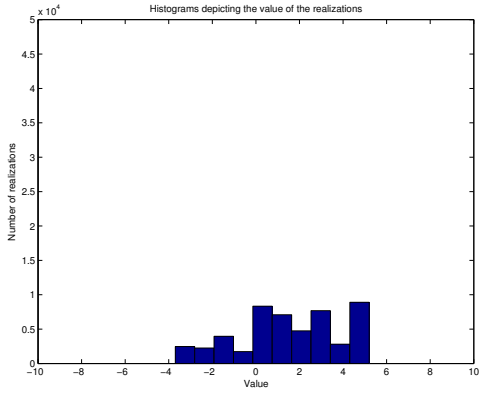
(b) C1, $\beta = 0.5$



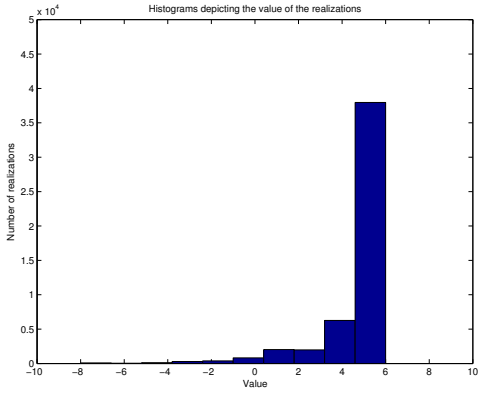
(c) C1, $\beta = 1.5$



(d) C2, $\beta = 0$



(e) C2, $\beta = 0.5$



(f) C2, $\beta = 1.5$

Figure 5: Depicts histograms for the joint strategy value running 50000 realizations for cases C1 and C2. Figure 5a, 5b, 5c is case C1 with $\beta = 0$, $\beta = 0.5$ and $\beta = 1.5$ respectively. Figure 5d, 5e, 5f similarily case C2 for $\beta = 0$,untitled $\beta = 0.5$ and $\beta = 1.5$.

Another case to be studied is C2. A noticable property of this field is the high value of $\alpha_5(1)$, which indicates that that this particular spot is higly likely to be an oil reservior. No other prior information is given through the $\alpha_i(x_i)$ parameter. For very low β-values one would expect cell 1, 5, 8 and 9 to be the only cells drilled in a joint strategy: cell 5 due to the α-parameter, the other ones because $R_i > C_i$ (for small values of β the value of $p(x_i)$ for all cells but the middle one will be fairly close to $\frac{1}{2}$ as this is the limiting case for $\beta = 0$). This tendency is confirmed by Table 4 which depicts the average income earned for different cells as β increases (0,0.5,1.5). It shows that cell four is included in the drilling strategy when β reaches 0.5. As β reaches 1.5 all cells but the one in the upper right corner is drilled. Knowing with almost certainty that the middle cell indeed is an oil site, increases the likelihood of the neighboring cells cointaining oil too.

$\beta = 0.0$	$\beta = 0.5$
$\begin{pmatrix} 0.0512 & 0 & 0 \\ 0 & 0.2999 & 0.2465 \\ 0 & 0 & 0.3474 \end{pmatrix}$	$\begin{pmatrix} 0.1854 & 0.0218 & 0 \\ 0 & 0.2999 & 0.5959 \\ 0 & 0 & 0.5187 \end{pmatrix}$

$\beta = 1.5$
$\begin{pmatrix} 0.8935 & 0.5087 & 0 \\ 0.3213 & 0.3000 & 1.3553 \\ 0.0831 & 0.1339 & 1.4352 \end{pmatrix}$

Table 4: Expected income from each cell for a joint strategy. The obtained values are Monte Carlo estimates obtained from running 50000 samples. Case run: C2.

The joint value (JV), the value of free clairvoyance (VFC) and value of information (VOI) are all given in Table 5. For $\beta = 0$ all cells, apart from the middle one has a equal probability of being eiter dry or with oil. The middle cell is however known to be with oil, hence the expected outcome for $\beta = 0$ is $JV = \frac{1}{2}((R_1 + C_1) + (R_8 + C_8) + (R_9 + C_9) + R_5) = 0.95$ This is in corresponcende with both Table 4 and 5. The joint value increases as a function of β. As the spatial correlation increases the oil field in the middle becomes paramount, in short it is very unlikely that neighboring cells are dry.

Interestingly the value with free clairvoyance is close to zero. The value of JV is showed as a function of β in Table **??**, likewise with VFC. For high values of β they approach each other. In other words the value of information becomes zero. It can phrased as saying that the economic difference between knowing something for sure and knowing something with almost certainty is about the same. It could be compared to playing an inverse lottery where

	$\beta = 0.0$	$\beta = 0.5$	$\beta = 1.5$
JV	0.9450	1.6217	5.0293
VFC	3.1974	3.7685	5.6416
VOI	2.2524	2.1468	0.6123

Table 5: Joint value, value of free clairvoyance and value of information for three different values of β (0.0, 0.5 and 1.5. The obtained values are Monte Carlo estimates obtained from running 50000 samples. Case run: C2.

you have to pick a number that's not picked in the lottery. The likelihood for doing so is staggering. The value of knowing the exact number picked beforehand compared to not knowing is insignificant. The mathematical reason for this is evident from equation 22 and 23 since $\max\{0, R_i p_i(1)\} \rightarrow p_i(1)\max\{0, R_i\}$ when $p_i(1) \rightarrow 1$.

From C1 and C2 a few interesting observations should be noted:

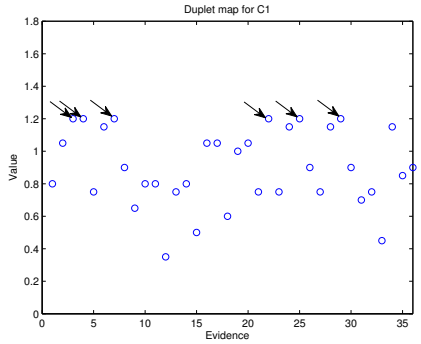
- If the prior information in zero, i.e. $\alpha_i(x_i) = 0 \ \forall i \in I, x_i \in \{0, 1\}$, changes in β have no effect on the expected outcome following a joint strategy. However it affects the variability of the outcomes, in turn increasing the risk of huge reward or huge loss. If the drilling company is risk averse, this is certainly a factor that needs to be taken into account. For large values of β the method can be likened to throwing a coin, either all cells are with oil or without oil. This view offers a clear insight into why the sequential method is preferable. In that case, picking one cell and investigating the outcome would be enough to make a descision to continue drilling or not.

- If prior information is available, as in the case of C2, this information will affect the outcome as the spatial dependency between points grow. In C2 one point is known to contain oil (or known with great certainty). The spatial dependency forces the adjacent points to be similar in outcome, in turn reducing the variability of the joint strategy. This is shown in figure 3 and 4. This is simply because the cells were forced toward one particular outcome, while in C1 the outcomes for all cells fluctuated between two extremes (offering either max income or max loss).

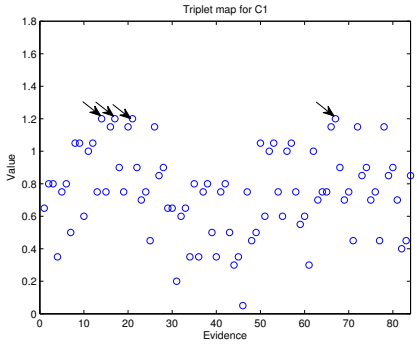
- It can be shown by performing a mirror image of C2 (that is middle cell is known to be dry) that the variance is reduced, but also that JV and VFC decreases as a function of β, eventually becoming close to zero.

Another interesting approach is computing duplets and triplets of cells in accordance with Equation 5.2. These are depicted graphically below for the

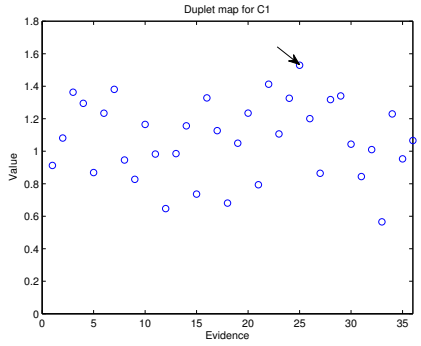
case C1 (Figure: 6) and C2 (7). The value of β is chosen equal to 0 and 0.5 in both configurations.



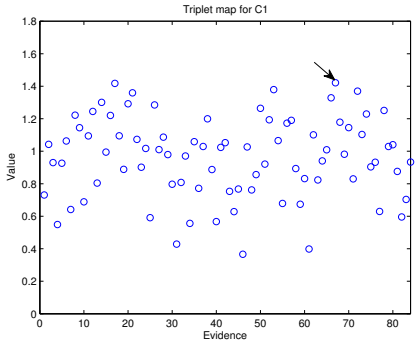
(a) $\beta = 0$, Duplet



(b) $\beta = 0$, Triplet

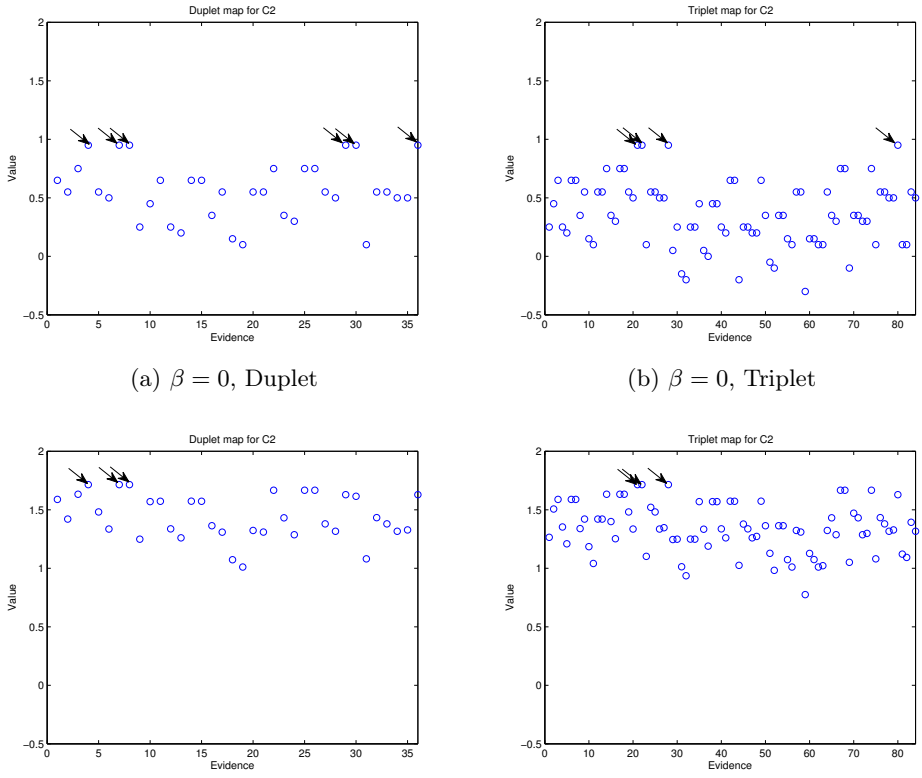


(c) $\beta = 0.5$, Duplet



(d) $\beta = 0.5$, Triplet

Figure 6: Triplets and duplets for C1 for β equal to 0 and 0.5.



(a) $\beta = 0$, Duplet

(b) $\beta = 0$, Triplet

(c) $\beta = 0.5$, Duplet

(d) $\beta = 0.5$, Triplet

Figure 7: Triplets and duplets for C2 for β equal to 0 and 0.5.

In the case of C1 with $\beta = 0$ all the cells that offers the maximum prospect of income are combinations of the cells indexed 1,4,5 and 8. It doesn't matter which three are chosen, this peculiarity is caused by β being equal to zero, in other words there are no spatial relationships between points. The probability of outcome at one point won't affect another point, in other words the chances for earning future money on the remaining cells is unaffected by the two (duplet) or three (triplet) first choices, i.e. no information is gained nor lost with respects to the unknown cells. This limits the number of optimal choices to be a subset of all points i such that $(R_i p_i(1) - C_i p_i(0)) > 0$. In C1 four different cells have this property. If five points were chosen (quintuplet!) this would have a negative effect on the expected value as opposed to only picking four (this is because a fifth value would have to be picked on a point with negative expected value). In practice, this means performing a joint strategy (no spatial information is obtained with respect to non-selected points), but including one point such that $R_i - C_i < 0$. The highest expected value for both a duplet and a triplet is 1.2 as is the case for a joint strategy

on similar input.

For C2 a similar logic can be applied for $\beta = 0$. Apart from this case the middle cell is know to contain oil. If it is picked immediately as part of the duplet/triplet, or not, doesn't matter. The number of points that gives the best expected income is hence any combination of cells such that $R_i - C_i > 0$ and cells known to contain oil. In case Let N denote the number of cells such that $R_i p_i(1) - C_i(1)p_i(0) > 0$. Assume that the number of cells picked (n_c) is less than N, then the number of optimal evidence groups equal $\binom{N}{n_c}$. This is valid when $\beta = 0$. N is 4 for both C1 and C2. This means that $\binom{4}{2} = 6$ and $\binom{4}{3} = 4$ for duplets and triplets respectively for $\beta = 0$. This is confirmed in figure 6a, 6b, 7a and 7b. The particular groups of evidence that maximizes the outcome is given in Table 7. Note that Table 7 confirms that the groups of cells offering the highest expected income when β equals zero indeed is combinations of cells picked using the joint approach.

For increasing values of β both cases get a higher expected income, see Table 6. For the uninformative case this is due to stronger correlation between points such that duplets and triplets reveal more information about the field. For C2 there is an additional positive effect since the middle cell is with oil, forcing this property over on the other cells as well.

	Duplet			Triplet		
	$\beta = 0.0$	$\beta = 0.5$	$\beta = 1.5$	$\beta = 0.0$	$\beta = 0.5$	$\beta = 1.5$
C1	1.2000	1.5293	2.8856	1.2000	1.4213	2.5433
C2	0.9500	1.7153	5.1066	0.9500	1.7153	5.1284

Table 6: Shows the optimal value for duplets and triplets for case C1 and C2 with varying β (0, 0.5 and 1.5).

It is also possible to run a case similar to C2, but with middle cell being dry. In such a case an opposite effect will be observed. Triplets will then be worse than duplets for high β-values, since it in most cases means picking three dry cell instead of two.

The best outcome for different values of triplet and duplets are depicted in Figure 8. Not that in case of C running triplets and duplets give about the same results. Note how both value converges for high values of β. For C1 there are two alternating outcomes that dominates; all cells are either with oil or dry. Hence the duplets are preferred to triplets. I.e. if three cells are picked showing oil, all the remaining cells will be picked afterwards. Likewise is the case for duplets. However, if the cells picked are dry, there is no continuation in the drilling. In other words triplet forces the user to drill thrice instead of twice. The best strategy would of course be to drill only once, and from this single result decide to drill all remaining points or not.

	Duplet			Triplet		
C1	$\beta = 0.0$	$\beta = 0.5$	$\beta = 1.5$	$\beta = 0.0$	$\beta = 0.5$	$\beta = 1.5$
	$(1, 4)$	$(4, 8)$	$(1, 5)$	$(1, 4, 5)$	$(4, 5, 8)$	$(3, 4, 8)$
	$(1, 4)$			$(1, 4, 8)$		
	$(1, 8)$			$(1, 5, 8)$		
	$(4, 5)$			$(4, 5, 8)$		
	$(4, 8)$					
	$(5, 8)$					
C2		Duplet			Triplet	
	$\beta = 0.0$	$\beta = 0.5$	$\beta = 1.5$	$\beta = 0.0$	$\beta = 0.5$	$\beta = 1.5$
	$(1, 5)$	$(1, 5)$	$(4, 9)$	$(1, 5, 8)$	$(1, 5, 8)$	$(2, 4, 9)$
	$(1, 8)$	$(1, 8)$		$(1, 5, 9)$	$(1, 5, 9)$	
	$(1, 9)$	$(1, 9)$		$(1, 8, 9)$	$(1, 8, 9)$	
	$(5, 8)$			$(5, 8, 9)$		
	$(5, 9)$					
	$(8, 9)$					

Table 7: Shows the best pick of cells for triplets and duplets for three different values of β (0, 0.5 and 1.5).

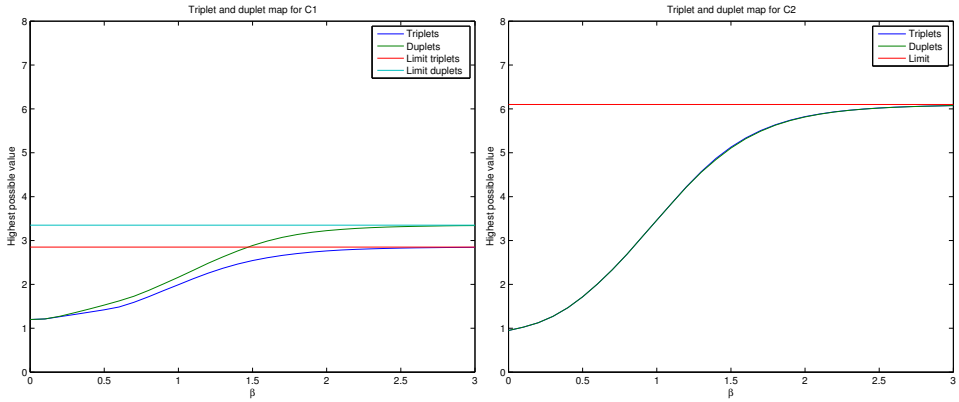
From this the two limiting cases for C1 (for high β's) is:

$$GS(p, \beta = \infty) = \frac{1}{2}(\sum_{i=1}^{n} R_i - \sum_{i=1}^{p} C_i^*) \tag{35}$$

Where C_i^* denotes the cost for the i'th cell smallest cost. p is the number of cells picked in group. This is valid as long as β is large and $\alpha_i(x_i) = 0$. The limiting case for duplets are hence: 3,375 for duplets, 2.875 for triplets. In the special case of C2 there will only be one limiting outcome when $\beta \to \infty$, i.e. all cells are with oil. In this respect the outcome for both the duplet and triplet strategy is the same: pick points (with oil) and pick the rest afterwards. In other words the limiting case is simply: $\sum_{i=1}^{n} R_i$=6.1.

The upper limit, complete search algorithm and myopic algorithm is applied to the two cases above for different values of β. The results are given in the Table 8 (These are Monte Carlo estimates obtained with 25000 realizations). The upper limit used is the one defined by equation 32. No sequential procedure utilizing the spatial information available can give an average expected value larger than this. points alone is enough to determine the outcome of a cell with extremely good probability.

For β equal to zero both method reduces to a joint strategy. Each point is picked independent of every other points picked earlier, in other words the sequential procedure cannot add more information about the remaining



(a) Duplet (b) Triplet

Figure 8: Value of duplets and triplets for C1 (left) and C2 (right) for different values of beta. The value given is the expected value for the best duplet og triplet of cells.

C1	Myopic (MC)	Myopic (Th)	Exact (MC)	Exact (Th)	Upper Limit (UL)
$\beta = 0.0$	1.1792	1,2	1.1792	1.2	1.1792
$\beta = 0.5$	1.5113	1.5303	1.5236	1.5425	1.8072
$\beta = 1.5$	3.3128	3.2928	3.3424	3.3270	3.8993

Table 8: Shows values the myopic, exact and upper limit for C1 when $\beta = 0$, $\beta = 0.5$ and $\beta = 1.5$. The Monte Carlo estimates are obtained for all three cases. The theoretical values for 31 and 25 are also given (denoted by th). δ is set equal to 1.0.

undrilled cells. Hence the result is the same of both myopic and exact. The upper limit condition also becomes simplified by only (in the univariate case) picking points such that $R_i > C_i$. This also becomes similar to a joint procedure. All methods discussed so far will hence converge toward the same sum, this being a limiting case for all procedures duscussed so far. As expected all the three approaches show increasing values as the β-value increase. This effect will happen for all uninformative cases. Table 9 for C2 (Monte Carlo, 25000 realizations) gives estimates for the myopic, exact and upper limit. Comparing the expected values for different β-values shows a slight difference in expected values for myopic and exact approach for both C1 and C2. But compared to the joint strategy there is considerable difference. While the expected value remains 1.2 (C1) for all values of β, the myopic and exact approach shows considerable improvement compared to this. For $\beta = 1.5$ the expected values are almost thrice as good. Clearly this

kind of strategy is far superior to the joint strategy.

C2	Myopic (MC)	Exact (MC)	Upper Limit (MC)
$\beta = 0.0$	0.9107	0.9107	0.9107
$\beta = 0.5$	1.7127	1.7127	1.8642
$\beta = 1.5$	5.1510	5.1766	5.2260

Table 9: Shows values the myopic, exact and upper limit for C2 when $\beta = 0$, $\beta = 0.5$ and $\beta = 1.5$. The Monte Carlo estimates are obtained for all three cases. δ is set equal to 1.0.

To illustrate the sequential effect of optimal path algorithm, it is more convenenient to use a smaller grid, not because of the computational efficiency (A 3×3 grid can be solved for optimal solution within a reasonable timeframe), but it becomes very hard to depict graphically by a tree structure. Hence in the interest of clarity two smaller grids are used to depict the effect of the exact tree method. These grids belong to Table 10.

Case:	$\alpha_i(0)$	$\alpha_i(1)$	Revenue	Cost
C3	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 \\ 10 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0.4 & 0.9 & 0.7 \\ 0.5 & 0.6 & 0.8 \end{pmatrix}$	$\begin{pmatrix} 1.2 & 1.1 & 1 \\ 0.9 & 1 & 1.1 \end{pmatrix}$
C4	$\begin{pmatrix} 0 & 0 & 0 \\ 10 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 7 & 5 & 2 \\ 4 & 6 & 7 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

Table 10: Case C3 and C4

C3 offers small rewards if oil is found, but on the other hand it is almost certain to be oil at cell 2. By computing an exact procedure it is possible to say something about which path is preferable, that is, which path utilizes the information obtained at each step in the most profitable manner. In case C4 all cells offers great economic rewards in the case of oil. The case is a mirrorimage of C4: One cell is highly unlikely to contain oil according to the prior information available (the α-function). The second method for computing sequential paths is the myopic approach. The results for different beta values are computed (25000 samples) results are given in Table 11.

For $\beta = 0$ and $\beta = 0.2$ too little spatial information is available for there to be feasible to perform on the other points. This obvious for $\beta = 0$ as there are no spatial correlation between points, but even for β-values as high as 0.2, no strategy will be performed apart from picking the south-eastern cell known to contain oil (with high probability) from prior information.

C3	Exact (MC)	Exact (Th)	Myopic (MC)	Myopic (Th)
$\beta = 0.0$	0.4999	0.5000	0.4999	0.5000
$\beta = 0.2$	0.4999	0.5000	0.4999	0.5000
$\beta = 0.3$	0.5445	0.5468	0.5000	0.5000
$\beta = 0.5$	0.8075	0.8154	0.8029	0.8096
$\beta = 1.5$	2.8998	2.8987	2.8946	2.8913
	Joint(MC)	Joint (Th)	Upper limit	
$\beta = 0.0$	0.4999	0.5000	0.4999	
$\beta = 0.2$	0.4999	0.5000	0.5825	
$\beta = 0.3$	0.5000	0.5000	0.7091	
$\beta = 0.5$	0.5740	0.5742	1.0017	
$\beta = 1.5$	2.7106	2.7003	3.0751	

Table 11: This tables depicts theoretical values for the exact, myopic and joint approach. Similar values are also obtained through Monte Carlo estimation including the upper value limit. Number of realizations used is 25000. δ is equal to 0.99.

When β reaches 0.3 the exact procedure starts to pick cells outside cell 2. This is however not done for the myopic approach, only cell 2 is picked. For the exact approach the nighbor adjacent to cell 2 is picked as a second step. Conditioned on the outcome of this cell, more cell might be visited. For $\beta = 0,5$ and higher values of β the myopic approach starts to perform similar runs through the grid (defined as.

The average number of visits to each cell is given below for the two sequential approaches, the cell number being the ratio of visits, i.e. $\frac{\text{Number of times drilling at cell i}}{\text{Total number realizations}}$. Table 12 shows how many times different cells are visited with respect to the exact and myopic strategy. Note that that the myopic and exact approach is superior to the joint strategy when β is equal to 0.5. For very low and very high values of β the methods approach each other in value. It should be noted that for very high values of the spatial dependency parameter the value of the joint strategy will converge towards $\sum_{i=1}^{6} R_i = 3.9$ which actually is better than the sequential algorithms due to the δ parameter.

Similar values for C4 is given in Table 13.

For $\beta = 0$ the joint method is preferred. This is obvious because no spatial correlation exists and that a penalty paramter $\delta = 0.99$ is enforced on the sequential procedures. In other words the two sequential procedures are reduced to a joint strategy with the added disadvantage of being 'punished' for every cell being picked. For all other values of β apart from $\beta = 1.5$ the exact approach is better than the myopic. Interestingly at $\beta = 0.5$ the myopic method is incapable of finding the best path as opposed to the other

$\beta \leq 0.2$

$$\begin{pmatrix} 0 & 0 & 0 \\ 1.0000 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 & 0 \\ 1.0000 & 0 & 0 \end{pmatrix}$$

$\beta = 0.3$

$$\begin{pmatrix} 0 & 1.0000 & 0.1956 \\ 1.0000 & 0.5224 & 0.3384 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 & 0 \\ 1.0000 & 0 & 0 \end{pmatrix}$$

$\beta = 0.5$

$$\begin{pmatrix} 0.4127 & 0.6284 & 0.2641 \\ 1.0000 & 1.0000 & 0.6284 \end{pmatrix} \qquad \begin{pmatrix} 0.5624 & 1.0000 & 0.4138 \\ 1.0000 & 0.5624 & 0.4127 \end{pmatrix}$$

$\beta = 1.5$

$$\begin{pmatrix} 0.8477 & 0.8960 & 0.7733 \\ 1.0000 & 1.0000 & 0.8960 \end{pmatrix} \qquad \begin{pmatrix} 0.8730 & 1.0000 & 0.7823 \\ 1.0000 & 0.8730 & 0.8730 \end{pmatrix}$$

Table 12: Shows the number of times cells are visited with respect to myopic (left) and exact (right) algorithm for varying β-values in the case of C3.

C4	Exact (MC)	Exact (Th)	Myopic (MC)	Myopic (Th)
$\beta = 0.0$	0.2084	0.1985	0.2084	0.1985
$\beta = 0.2$	0.2007	0.1929	0.1924	0.1827
$\beta = 0.3$	0.1986	0.2101	0.1644	0.1838
$\beta = 0.5$	0.2429	0.2378	0.0000	0.0000
$\beta = 1.5$	0.0000	0.0000	0.0000	0.0000
	Joint(MC)	Joint (Th)	Upper limit	
$\beta = 0.0$	0.2101	0.2000	0.2101	
$\beta = 0.2$	0.0714	0.0683	0.3144	
$\beta = 0.3$	0.0158	0.0282	0.3591	
$\beta = 0.5$	0.0000	0.0000	0.4937	
$\beta = 1.5$	0.0000	0.0000	0.3233	

Table 13: This tables depicts theoretical values for the exact, myopic and joint approach. Similar values are also obtained through Monte Carlo estimation including the upper value limit. Number of realizations used is 25000. δ is equal to 0.99.

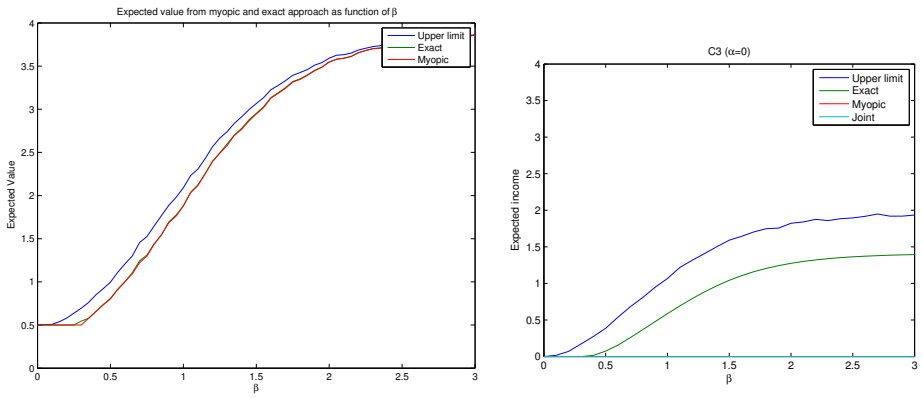
method. For $\beta = 1.5$ the effect of the lower right cell is too strong, forcing neighbors to be negative with very high probability. At this level neither the myopic nor exact approach work.

The expected value from for the myopic, exact and the joint approach is depicted in Figure 10 and **??** for C3 and C4 respectively.. The upper limit

$\beta = 0$	
$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
$\beta = 0.2$	
$\begin{pmatrix} 0.4886 & 1.0000 & 0.7375 \\ 0 & 0.3692 & 0.8890 \end{pmatrix}$	$\begin{pmatrix} 0.3899 & 0.7755 & 0.4978 \\ 0 & 0.1502 & 1.0000 \end{pmatrix}$
$\beta = 0.3$	
$\begin{pmatrix} 0.2740 & 0.4936 & 1.0000 \\ 0 & 0.1542 & 0.4936 \end{pmatrix}$	$\begin{pmatrix} 0.1542 & 0.2729 & 0.4823 \\ 0 & 0.1542 & 1.0000 \end{pmatrix}$
$\beta = 0.5$	
$\begin{pmatrix} 0.2750 & 0.4773 & 1.0000 \\ 0 & 0.1714 & 0.4773 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$
$\beta = 1.5$	
$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

Table 14: Shows the number of times cells are visited with respect to myopic (left) and exact (right) algorithm for varying βvalues in the case of C4.

is obtained from Monte Carlo estimation using 10000 samples.



(a) Case: C3

(b) Case: C3 ($\alpha_i(x_i) = 0$)

Figure 9: Two plots showing the expected value of myopic, exact and joint strategy for different values of β ranging from zero to three. The upper limit (UL) is obtained through Monte Carlo using 10000 samples for each value of β. The figure on the left is running case C3, right is case C4.

Judging from the graph, case C3 have the greatest divergence in expected income for small values of β. This is caused by the fact that the complete search starts processing other cells earlier than the myopic algorithm, in turn offering better rewards. But as β increases these values converge. Then it converges with the other two methods. This being forced by the fact that every cell is forced to be equal to cell i as the spatial dependence between points increase. Both methods (and the upper limit) have the same value for zero spatial correlation.

Figure 11 and 11b shows the effect of running with running on a field with prioer information encoded into it through the $\alpha_i(x_i)$ parameter.
The variability of the data can be deduced from the histograms of outcomes for the realizations. These are depicted for both C3 and C4 in Figure 12 and 13. Note how the realization gets less and less varied as β increases. This is of course what one should expect when all cells are of one type.

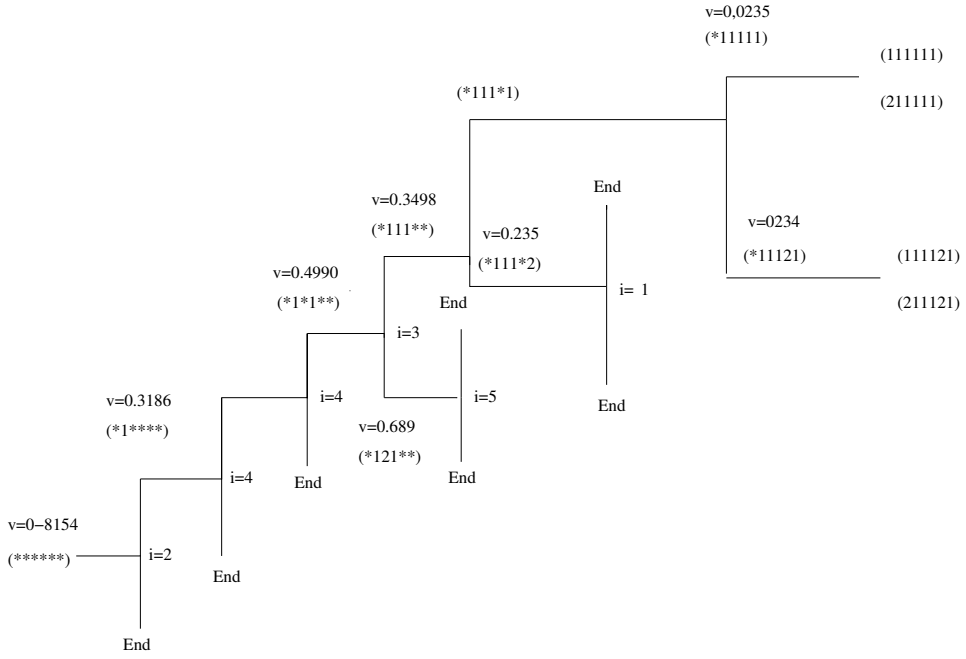
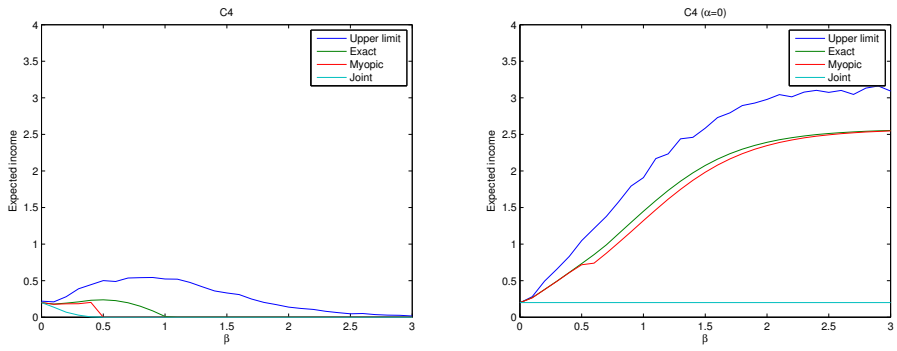


Figure 10: Shows the exact sequence tree for C3, β set to 0.5. This is the optimal path choosing cells.'*' denotes undrilled cells. i denotes the index to be drilled. 1 represents oil and 2 represents lack of oil.



(a) Case: C4

(b) Case: C4 ($\alpha_i(x_i) = 0$)

Figure 11: The expected value of myopic and exact path are given for the C4 as a function of beta in both graps, though the right plot is modified by setting $\alpha_i(x_i) = 0 \; \forall i, x_i$.

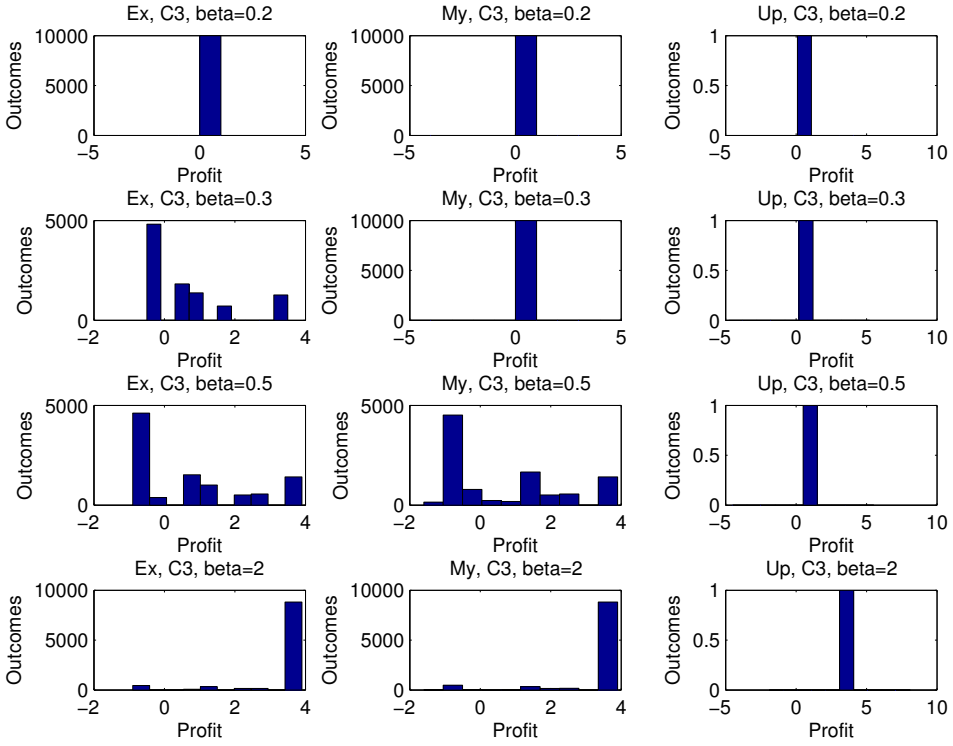


Figure 12: The plot above shows the outcome of running Monte Carlo simulations. Each row represents case C3 for some specified β (0.2, 0.3, 0.5 and 2) for exact algorithm (left), myopic (middle) and upper limit etsimate (right)

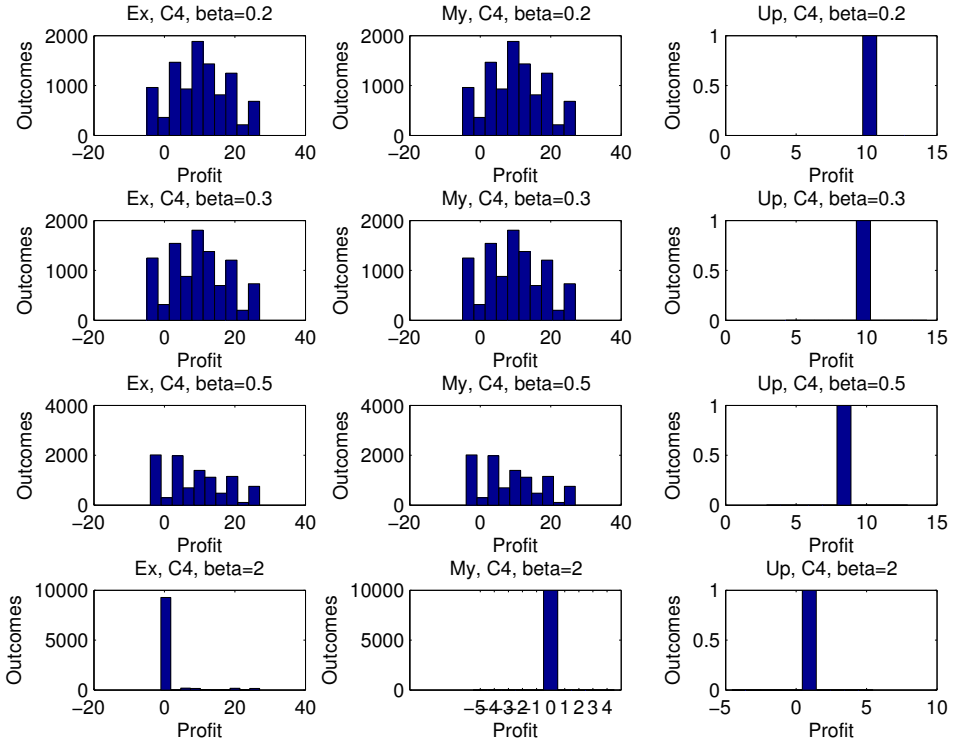


Figure 13: The plot above shows the outcome of running Monte Carlo simulations. Each row represents case C3 for some specified β (0.2, 0.3, 0.5 and 2) for exact algorithm (left), myopic (middle) and upper limit etsimate (right)

6 Strategies for large grids

This Section introduces different methods for dealing with the problems related to dealing with large grids. For smaller grids it is possible to use use the sequential and joint approaches, but for larger grids heuristic approaches are needed.

6.1 Horizontal partitioning of blocks

So far methods used have been limited to rather small cases, running a 2×2 and 3×3 grid. Running on even larger grids becomes increasingly difficult. This is particularly true for the case of the exact approach. Computing every possible path becomes a hugely limiting factor. Due to this three heuristic approaches to this problem will be considered, all are concerned with partitioning the grid into blocks. The two first cases to be intoduced will be used in the special case where the grid have many columns but few rows (or opposite!). Such a grid can be partioned into smaller blocks along the horizontal direction. This can be illustrated by a 2×6 grid: $\left(\begin{smallmatrix}1&3&5&7&9&11\\2&4&6&8&10&12\end{smallmatrix}\right)$. This can be partitioned into three (or two) smaller blocks of equal length; $\left(\begin{smallmatrix}1&3\\2&4\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}5&7\\6&8\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}9&11\\10&12\end{smallmatrix}\right)$. Within reasonably small blocks both the exact and myopic sequential approach can be computed, likewise with the joint strategy. There are different ways in which these blocks can be solved. The two methods used here are:

Independent blockwise approach This approach starts at lefternmost block. It runs the exact, myopic and joint strategy in this block alone. After completion it continues in the same manner of the adjacent block to the right. In this way the spatial relationship between blocks are ignored. For the grid above this means running the sequential and joint strategies on $\left(\begin{smallmatrix}1&3\\2&4\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}5&7\\6&8\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}9&11\\10&12\end{smallmatrix}\right)$ in turn as if they were independent fields. It should be duly noted that the expected outcome of any strategy is not the sum of the expected outcome of each block independently. This is obvious as the original field itself has no blockwise boundaries in the spatial correlation.

Dependent blockwise approach The first block to be solved is the lefternmost. This is solved independently of all the other blocks. When the next block is computed all probabilities are conditioned on the outcomes of the former block. For example the probability of $p(x_5)$ in the independent procedure would be analogious to $p(x_5|x_1, x_4)$ if say cell one and cell four was picked in the former block. In this way former information is propegated forward. It is possible to condition on more blocks than the last one, but only one block have been used in this thesis.

Both metods are shown graphically in figure 14 (independent) and 15 (dependent).

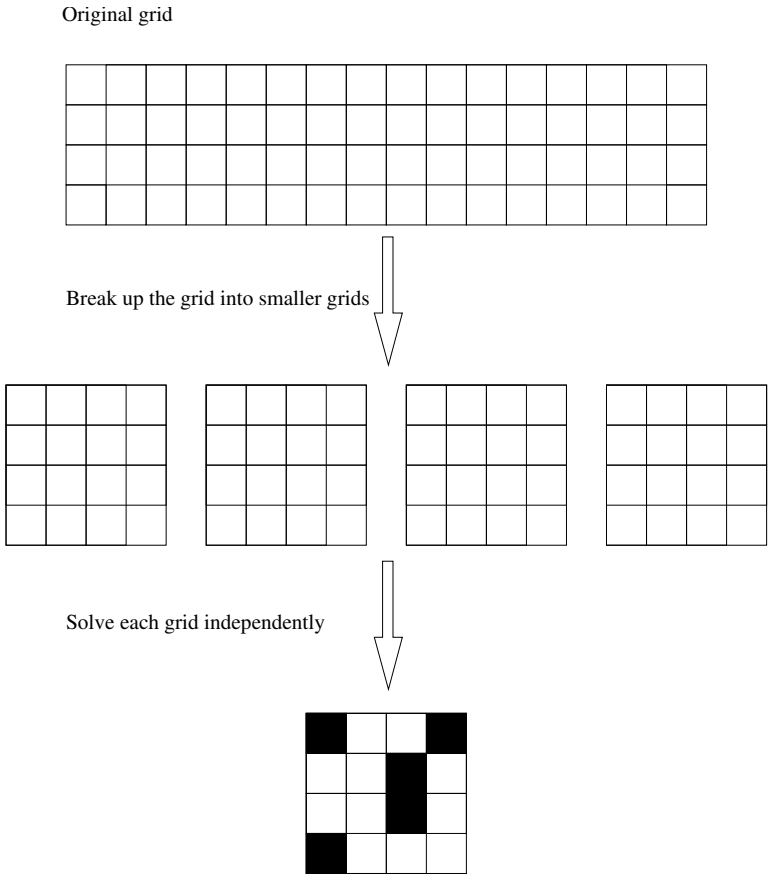


Figure 14: The graph above shows how the grid is partioned horizontally into smaller grids, each being solved independently of each other. Different drilling strategies can then be applied to every such block independently.

In short smaller blocks are used to solve a larger grid. The case above was for a matrix consisting of two rows and six columns, but it can also be applied to cases where there are many rows but only two or three columns. This is done by simple rotating the grid 90 degrees or using the same approach starting at the bottom block and working upwards.

Both methods have advantages and disadvantages. The independent strategy has the advantage of being quite fast compared to the dependent strategy. This is obvious since only one descision tree (sequential) or group of cells (joint) is used for every block. However the dependent approch is more complex as the descisions taken in a future block depends on the outcomes of the former block. Let b denote the number of grid points in each block. In short this means that there are $(d+1)^b$ possible sequential paths that

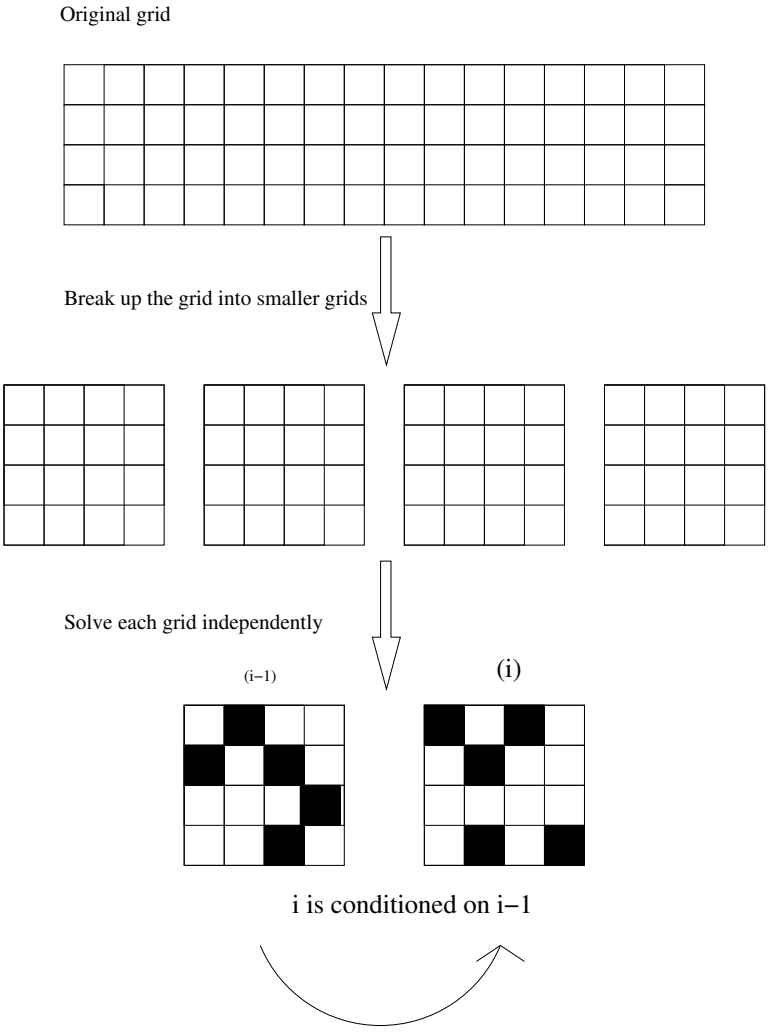


Figure 15: The graph above shows how the grid is partioned horizontally into smaller grids, each being solved dependent of each other. Different drilling strategies are then applied to every grid conditioned on the outcome of the adjacent block to the left, can also include more blocks in this conditioning

needs to be computed. Anoter option is of course not to store the decision paths beforehand, but then one would need to compute each path anew for every realization, quickly becoming expensive as the number of realization increase. Another disadvantage with the dependent approach is that the probabilities needs be computed in a grid that is $2b$ cells larger than the case for the independent approach. The advantage of the dependent approach is that information from earlier blocks are carried over to new blocks, allowing to make more informed choices. In short the dependent approach is likely

(not always) to give better estimates but at a higher computational cost.

In the future exact, myopic and joint strategies will be referred to as conditional exact, conditional myopic and conditional joint when used in the context of solving partitions of the grid. This should be stressed since it is fully possible to use these strategies for larger grids too, if only in theory. Whether conditional exact/myopic/joint strategy refers to an independent or dependent blockwise approach should be clear from the context it is used in.

6.2 Large grid, many rows and columns

The limitation of the method presented in the preceding Subsection was the need for having either few rows or columns in the grid to be studied. The soulution to this problem is to partition not only horizontally but also vertically, creating a two dimensional grid of blocks. While both a dependent and independent approach can be studied in this case too, the focus here will be on the dependent version. Two questions need answering: (a) in which order will the blocks be solved and (b) how do the blocks depend on eachother. The order of the blocks chosen here is from left to right. Starting at the upper row at the lefternmost corner and working all the way to the end of the row. After finishing the first row the second row below is chosen, continuing all the way to the last column. And it continues in this fashion until all rows have been traversed. The ending point is then the bottom right corner. As for block dependency, the following rules have been chosen:

- First block (upper left) will be solved independently of other blocks.
- All blocks on the upper row (apart from the one in upper left corner) will depend on the adjacent block to the left. This is in essence doing the same as in the dependent blockwise approach.
- All blocks on the first column (apart from the one in the upper left corner) will depende on the adjacent block above. Rotating the two blocks 90 degrees and they can be solved using the dependent blockwise approach.
- All blocks in neither the first row or first column will depend on the adjacent block to the left and above.

The whole procedure is shown graphically in figure 16.

The most computer intensive part of the computations is computing blocks that are neither belonging to the first row nor column. While the dependent blockwise approach demanded $(d+1)^b$ decision paths for the conditional exact strategy, the blocks depending on two other blocks demands a total of $(d+1)^2 b$, not to mention that the (conditional) probabilities needed for computation of sequential paths must be computed on a larger grid with

at least $3b$ cells (Can choose between a square grid - upper leftern corner unknown - or a 'shoe' formed grid, only including the block of interest and the two adjacent blocks that's been solved in advance). The code used in this thesis is square formed.

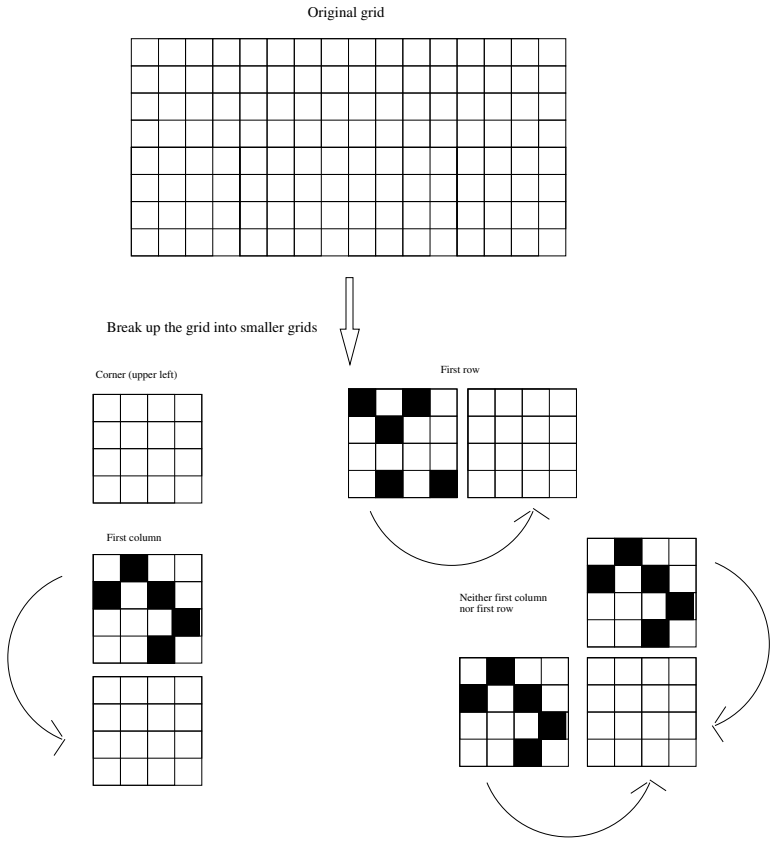


Figure 16: The graph above shows how the grid is partioned horizontally and vertically into smaller grids, each being solved dependent of each other. Different drilling strategies are then applied to every grid conditioned on the outcome of the adjacent block to the left and above

6.3 Cases

The two cases to be studied closer in this section are the ones defined below, C5 and C6. Both cases have the same revenues and costs associated with each cell. The difference is that C5 is uninformative while C6 is known to be dry at the bottom left corner. As β increases in value the cells in C6 are also expected to be dry, offering an interesting situation for both the dependent and independent block approach as the one of the method only retains a memory from the last block while the other is completely memoryless with respect to earlier outcomes in other blocks.

Case:	C5
$\alpha_i(j)$, $j \in \{0,1\}$	$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$
Revenue	$\begin{pmatrix} 0.5 & 0.2 & 0.3 & 0.2 & 1.2 & 0.7 & 1.1 & 0.3 & 1.5 & 3 \\ 0.9 & 1.1 & 0.9 & 0.8 & 0.6 & 0.5 & 0.6 & 0.9 & 0.1 & 0.5 \end{pmatrix}$
Cost	$\begin{pmatrix} 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 \\ 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 \end{pmatrix}$

Case:	C6
$\alpha_i(j)$, $j \in \{0,1\}$	$\alpha_i(0) = 0 \ \forall \ i \in \{1,3,4,\ldots,20\}$, $\alpha_2(0) = 10$, $\alpha_i(1) = 0 \ \forall \ i \in \{1,2,\ldots,20\}$
Revenue	$\begin{pmatrix} 0.5 & 0.2 & 0.3 & 0.2 & 1.2 & 0.7 & 1.1 & 0.3 & 1.5 & 3 \\ 0.9 & 1.1 & 0.9 & 0.8 & 0.6 & 0.5 & 0.6 & 0.9 & 0.1 & 0.5 \end{pmatrix}$
Cost	$\begin{pmatrix} 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 \\ 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 \end{pmatrix}$

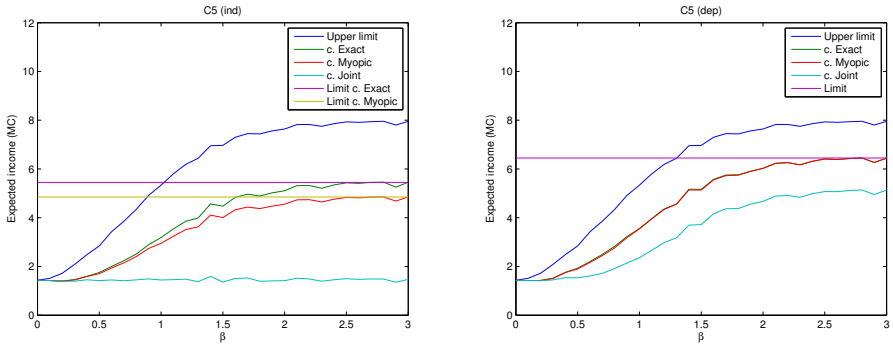
C5 and C6 gives values for the blockwise independent and dependent procedure for for case C5 and C6 respectively. The values are Monte Carlo estimates obtained running 25000 samples. C5 shows that the exact approach is better than myopic, but both estimates are very good compared to the joint strategy. While the joint values remains about constant for all values of β. This is exactly what one would expect for an uninformative case. This is because all the univariate marginal probabilites are equally likely to contain oil as not containing oil. Since equation 22 is independent of β the expected outcome also have to be independent of the increase in β. Note that the values of the joint strategy varies more and more as β increases, this is caused only extreme results being available, increasing the variance of JV as discussed in Section 5.7. This argument is valid for all the estimates.

Comparing the sequential and joint estimates from the dependent with the independent approach shows that the dependent version is the preferable method, and the difference becomes larger as β increases. This is expected as stronger spatial correlation means that the outcomes in the former block adds increasingly important information to the present block. As β becomes larger only two outcomes are likely; all cells oil or all cells dry. Thus when $\beta \to \infty$ the value is expected value of the exact approach for the independent method is $\frac{1}{2}(\sum_{i=1}^{n} R_i - n_b) = 5.4500$. Here n_b denotes the number of blocks.

Ind, C5	Exact	Myopic	Joint	Upper Limit
$\beta = 0$	1.4405	1.4405	1.4405	1.4405
$\beta = 0.2$	1.4508	1.4508	1.4327	1.7847
$\beta = 0.5$	1.7710	1.7614	1.4524	2.8894
$\beta = 1.0$	3.1264	2.8942	1.3875	5.3096
$\beta = 1.5$	4.5291	4.0560	1.4253	7.0123
$\beta = 2.0$	5.0527	4.5018	1.3780	7.5970
Dep, C5	Exact	Myopic	Joint	Upper Limit
$\beta = 0$	1.4897	1.4897	1.4897	1.4897
$\beta = 0.2$	1.4749	1.4749	1.4569	1.7529
$\beta = 0.5$	1.9562	1.9194	1.5630	2.8877
$\beta = 1.0$	3.5341	3.5164	2.3474	5.3148
$\beta = 1.5$	5.3080	5.2649	3.8652	7.0784
$\beta = 2.0$	6.0809	6.0738	4.7215	7.6905

Table 15: Estimates for the expected income for three different methods and the upper limit using Monte Carlo simulation with 25000 samples. The case being run is C5 both for the dependent and independent approach.

For myopic this becomes $\frac{1}{2}(\sum_{i=4}^{n} R_i + \sum_{i=9}^{n} R_i - (n_b - 1)) = 4.8500$ because points is never picked in the second block ($R_i - C_i < 0$ for all i in the second block). For the dependent case information is carried from one block to the other. If all cells are contains oil this information is propegated from the first cell to the last. Hence offering $\sum_{i=1}^{n} R_i$ in income for both sequential methods. If the first cell drilled is dry no more cells will be drilled in the first and second block. However when the third block is reached no information is available in the former block since no drilling have occured there. Hence it will start to drill once more. Finding this cell to be dry to it will stop drilling. No drilling will occur in the fourth block since it is known that a cell in the third block is dry. Fifth block will br drilled once only to drill a dry cell. This leads to the following formula: $\frac{1}{2}(\sum_{i=1}^{n} R_i - \lceil n_b/2 \rceil) = 6.4500$. These limits when $\beta \to \infty$ are depicted in figure 17 together with estimates for conditional exact, myopic and joint strategy and upper limit. Table 16 shows quite clearly that the dependent approach is superior to the dependent one. The bad (even negative values) for high β is because the likelihood of the whole field to be dry increases. The memoryless property of the independent approach removes important information. If a cell in the earlier block is dry there is good reason to believe that cells in the present block is dry too. This kind of information is snapped up by the dependent procedure. The combination of a high value of β and (a) dry cell(s) in block i will often lead to a strategy in block $i + 1$ where no cells are drilled. However when block $i+2$ is drilled block $i+1$ offers no particular information about the field and



(a) Independent

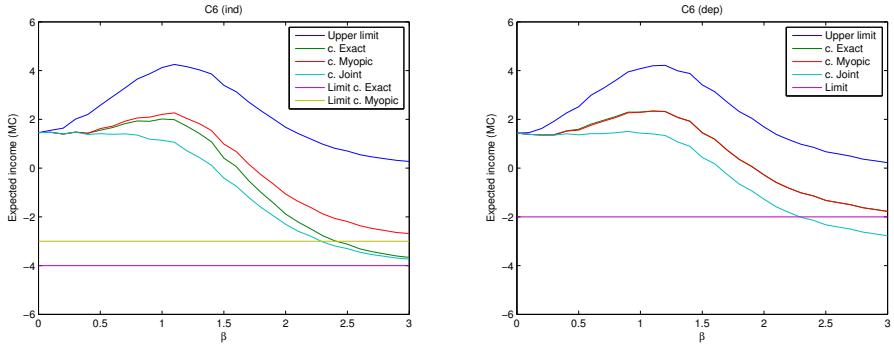
(b) Dependent

Figure 17: Shows the estimates for the upper limit, conditional exact, myopic and joint for varying β-values in the range 0 to 3 for the case of C5. These values are obtained from Monte Carlo estimation running 100000 samples for each value of β. The limiting value of conditional exact and conditional myopic is also shown. The figure on the left represents the independent procedure, the one on the right the dependent one. In the case of dependent blocks the conditional exact and myopic have almost the same values. Plots obtained running 10000 realization for each value of β.

Ind, C6	Exact	Myopic	Joint	Upper Limit
$\beta = 0$	1.4395	1.4395	1.4395	1.4395
$\beta = 0.2$	1.3877	1.3877	1.3877	1.6837
$\beta = 0.5$	1.4963	1.5443	1.3532	2.5215
$\beta = 1.0$	1.9972	2.1843	1.1112	4.1203
$\beta = 1.5$	0.5402	1.0850	0.3487	3.4772
$\beta = 2.0$	-1.7961	-0.9879	-2.2541	1.7458
Dep, C6	Exact	Myopic	Joint	Upper Limit
$\beta = 0$	1.4480	1.4480	1.4480	1.4480
$\beta = 0.2$	1.3927	1.3927	1.3927	1.6606
$\beta = 0.5$	1.6075	1.5832	1.3727	2.5148
$\beta = 1.0$	2.3049	2.3013	1.4323	4.0725
$\beta = 1.5$	1.5200	1.5125	0.5098	3.4545
$\beta = 2.0$	-0.27586	-0.27789	-1.2799	1.7045

Table 16: Estimates for the expected income for three different methods and the upper limit using Monte Carlo simulation with 25000 samples. The case being run is C6 both for the dependent and independent approach.

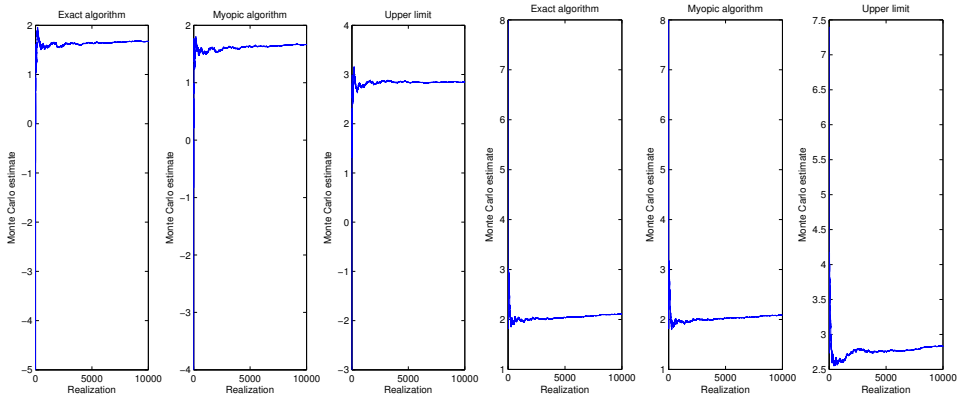
the knowledge of dry cells in i is lost. Hence the dependent strategy is likely to drill cells in $i+2$ even though the dry cells in i suggests dry cells also in block $i+2$. This is quite contrary to the effect of having oil in the first block for field with strong spatial correlation. In that case oil on one block suggest drilling in the next block, the result of which might suggest further drilling in the next block et cetere. The limit for the conmditional exact and myopic approach in the case of C5 is $\lceil -n_b/2 \rceil = -2$. That is every odd numbered block will be drilled (apart from the first block) resulting in dry cell being drilled. The conditional myopic is even worse for an independent procedure with $n_b - 2 = -3$ (the middle block will never be drilled according to a myopic strategy, nor the first). The conditional exact procedure will have an expected value equal to $n_b - 1 = -4$ (only the first block will not be drilled). In other words an interesting situation appears for the limiting case. This tendency for the conditional exact to be worse than the myopic strategy for high values of β when running an independent approach is confirmed by table 16. Likewise it is worth noticing that the value expected income of the sequential methods are quite close to eachother when the spatial correlation is strong. This seems reasonable considering that the limiting case has the same values. Figure 18 gives a graphical display of all three methods for varying β in the range [0 3].



(a) Independent

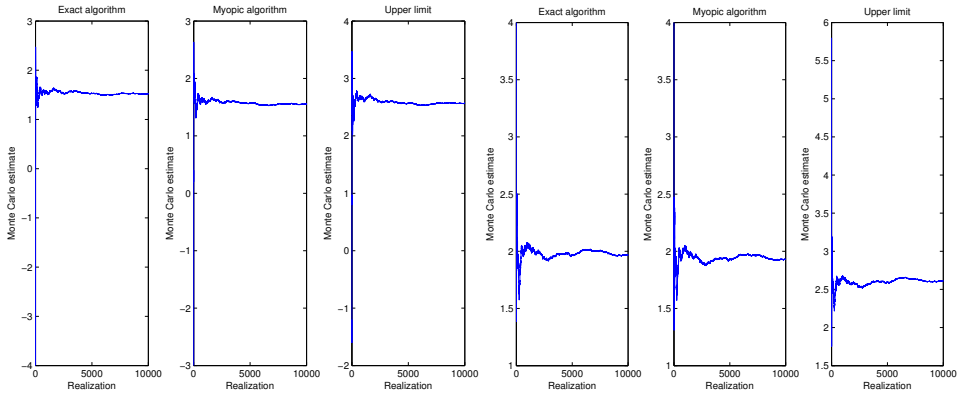
(b) Dependent

Figure 18: Shows the estimates for the upper limit, conditional exact, myopic and joint for varying β-values in the range 0 to 3 for the case of C6. These values are obtained from Monte Carlo estimation running 100000 samples for each value of β. The limiting value of conditional exact and conditional myopic is also shown. The figure on the left represents the independent procedure, the one on the right the dependent one. For the dependent approach both the conditional exact and myopic have almost the same values. Plots obtained running 10000 realization for each value of β.



(a) Independent (b) Dependent

Figure 19: Depicts the convergence running case C5 with $\beta = 0.5$.



(a) Independent (b) Dependent

Figure 20: Depicts the convergence running case C6 with $\beta = 0.5$.

The next thing to be studied more is the use of conditional myopic, exact and joint to a grid consisting of many rows and columns. The method to be used is the one outlined in Section 6.2. The field to be study closer is a univariate 10×10 case. The cost for failure is set to minus -1 for every cell. The potential revenue matrix consists of the revenue in case C5 copied five times in downward direction. Let the cells in the revenue matrix for C5 be indexed by i, j where i denotes row ($\{1, 2\}$) and j denotes column ($\{1, 2, \ldots, 10\}$). Then the revenue matrix, R^* of this new larger field

$$R^*_{i,j} = R_{i-2\lfloor \frac{i-1}{2} \rfloor, j} \tag{36}$$

This method was applied to for a 10×10 matrix The are promising due

	Large Matrix, dependent			
β	Exact	Myopic	Joint	Upper Limit
$\beta = 0.0$	7.2776	7.2776	7.2776	7.2776
$\beta = 0.5$	10.4500	10.0070	8.0581	16.2985
$\beta = 0.6$	12.1812	11.6461	8.7465	18.9362
$\beta = 0.8$	17.6801	17.3020	13.6841	25.8553

Table 17: Shows the value of Exact, myopic, joint and upper limit

to the spatial freedom of the problem, i.e. that there are many unknown adjacent blocks, still the outcome seem to outperform the joint strategy. And the difference between the upper limit and the sequential strategies is relatively small. Running on larger grids (30×30) with smaller samples seem to suggest the sequential strategies to have a value about half that of the upper limit.

7 Concluding remarks

The different methods applied so far shows both weaknesses and strengths of the blockwise approach. While the blockwise approach do manage to propegate information about oil rather successfully, one of the main weaknesses is the memoryless property of the blocks, in the sense that important information is lost. This was particularly clear in the case of C6 where all three methods showed negative values. It is not obvious how to deal with this issue. It is clear that a block approach would suffer from the same problem as C6 in cases of large β-values.

Both the sequential strategies behaved good on univariate cases. It is also interesting to note that the sequential method improved relative to the joint strategy on the larger grid. As β increased the tendency seemed to favour the sequential approach more than the joint strategy. That even the conditional myopic approach could outperform the conditional blockwise only (C6, independent) confirms that every method needs to be used with care.

One possibility for future work on this subject might be to adopat a more flexible blockwise routine, in the sense that new information in future blocks also can be added to investigate former blocks. It would also be interesting to run the blockwise method on a non-univariate field.

References

[1] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the royal statistical society. Series B (methodological)*, 36:192–236, 1974.

[2] D. Bhattacharjya, J. Eidsvik, and T. Mukerji. The value of information in spatial decision making. *Math geisciences*, 42:141–163, 2010.

[3] J. E. Bickel and J. E. Smith. Optimal sequential exploration: A binary learning model. *Decision Analysis*, 3:16–32, 2006.

[4] Merrilee A. Hurn, Oddvar K. Husby, and HÃěvard Rue. A tutorial on image analysis.

[5] Ross Kindermann and J Laurie Snell. *Turbulence*. Amer Mathematical Society, 1980.

[6] G. Martinelli, J. Eidsvik, and James E. Smith.

[7] Gabriele Martinelli, Jo Eidsvik, Ragnar Hauge, and Maren Drange FÃÿrland.

[8] R. Reeves and A. N. Pettitt. Efficient recursions for general factorisable models. *Biometrika*, 91:751–757, 2004.

[9] H. Tjelmland and H. M. Austad. Exact and approximate recursive calculations for binary markov random fields defined on graphs. 2011.