Local search for the probabilistic traveling salesman problem: a proof of the incorrectness of Bertsimas' proposed 2-p-opt and 1-shift algorithms

Leonora Bianchi*, Joshua Knowles†

August 21, 2002

Abstract

The probabilistic traveling salesman problem concerns the best way to visit a set of customers located in some metric space, where each customer requires a visit only with some known probability. A solution to this problem is an a priori tour which visits all customers, and the objective is to minimize the expected length of the a priori tour over all customer subsets, assuming that customers in any given subset must be visited in the same order as they appear in the a priori tour.

This problem belongs to the class of stochastic vehicle routing problems, a class which has received increasing attention in recent years, and which is of major importance in real world applications.

Several heuristics have been proposed and tested for the probabilistic traveling salesman problem, many of which were a straightforward adaptation of heuristics for the classical traveling salesman problem. In particular, two local search algorithms (2-p-opt and 1-shift) were introduced by Bertsimas and also published in this journal. Local search algorithms repeatedly perform simple modifications of a solution (moves), accepting a move only when the cost of the move is negative. In this paper we show that the expressions for the cost evaluation of 2-p-opt and 1-shift moves, as proposed by Bertsimas, are not correct. This fact has two consequences. First, results published in several papers, which used the 2-p-opt and 1-shift local search, should be re-examined because they may be wrong. Second, it poses the question whether a new fast local search operator for the probabilistic traveling salesman problem can be derived.

Keywords: Combinatorial optimization, probabilistic traveling salesman problem, heuristics, local search

1 Statement of the problem

The probabilistic traveling salesman problem (PTSP) is a traveling salesman problem (TSP) where each customer $i = 1, 2, ..., n$ has a given probability p_i of requiring a visit. A feasible solution to the PTSP is an a priori tour τ visiting all customers. The a priori tour may be then adapted to any particular subset of customers, with the strategy of following customers in the same order as they appear in the a priori tour, just skipping those which do not require a visit. The objective is to minimize the expected length of the a

*IDSIA, Galleria 2, Strada Cantonale, CH-6928 Manno, Switzerland. E-mail address: leonora@idsia.ch

†IRIDIA, CP 194/6, Université Libre de Bruxelles, Avenue Franklin Roosevelt 50, 1050 Brussels, Belgium. E-mail address: jknowles@ulb.ac.be

priori tour over all customer subsets. In this paper we restrict our attention to the homogeneous PTSP, that is, a PTSP where each customer has the same probability p of requiring a visit, independently of the others.

Given an a priori tour $\tau = \tau(1), \tau(2), ..., \tau(n)$, its expected length (that is, the objective function), is computed in $O(n^2)$ by the following expression [5]:

$$E[L(\tau)] = \sum_{j=1}^{n} \sum_{r=1}^{n-1} p^2 (1-p)^{r-1} d(\tau(j), \tau(j+r)). \tag{1}$$

Throughout the paper we use the convention that, $\forall i \in \{1, 2, ..., n\}$ and $\forall r \in \{1, 2, ..., n-1\}$,

$$\tau(i \pm r) = \begin{cases} \tau((i \pm r) \mathrm{mod} n), & \text{if } i \pm r \neq 0 \text{ and if } i \pm r \neq n \\ \tau(n), & \text{otherwise .} \end{cases} \tag{2}$$

The meaning of equation (1) is as follows. In each term of the sum the distance between the j^{th} city $\tau(j)$ and the $(j+r)^{\text{th}}$ city $\tau(j+r)$ is weighted by the probability that the two nodes require a visit (p^2) while the $r-1$ nodes between them do not require a visit ($(1-p)^{r-1}$). In other words, each arc $(\tau(j), \tau(j+r))$ is weighted by a probability coefficient $p^2(1-p)^{r-1}$.

Using the full evaluation function (equation (1)) to evaluate the cost of moves in a local search is very expensive and so it would be desirable to have a 'delta' evaluation function (to calculate the cost difference between two neighboring tours) of a much lower complexity. To this end, Bertsimas [3] has previously proposed delta evaluation expressions for two local search moves, namely 2-p-opt and 1-shift, described below, which compute the cost evaluations of the entire neighborhood of a tour in $O(n^2)$ time. Unfortunately, derivations for these expressions were not given in [3], and it has since come to our attention that they do not correctly calculate the change in expected tour length. It is the purpose of this paper to demonstrate the error in these expressions.

The remainder of the paper is organized as follows. In the next section we describe the 2-p-opt move and give the delta evaluation expressions proposed in [3], followed by a proof of their incorrectness. Section 3 follows a similar course but with respect to the 1-shift operator. In section 4 we conclude by reflecting on the impact of our proofs. Two appendices to this paper present a full calculation of the cost of local search moves for a simple PTSP tour, when the 2-p-opt and 1-shift operators are applied, respectively. The calculations demonstrate, concretely, the different results obtained using Bertsimas' delta expressions, compared with the full evaluation. These calculations were generated by a computer program written in C which is also available at http://www.idsia.ch/~leo/.

2 The 2-p-opt

The 2-p-opt local search was introduced by Bertsimas in [1] and later published by this journal in an article by Bertsimas and Howell [3]. Given an a priori tour τ, its 2-p-opt neighborhood is the set of tours obtained by reversing a section of τ (that is, a set of consecutive nodes) and adjusting the arcs adjacent to the reversed section, as for example in Figure 1.

2.1 Expressions proposed in [3] for the cost of a 2-p-opt move

Consider, without loss of generality, a tour $\tau = 1, 2, ..., i, i+1, ..., j, ..., n$ and a tour $\tilde{\tau} = 1, 2, ..., j, j-1, ..., i, j+1, ..., n$ obtained by reversing a section $(i, i+1, ..., j)$ of τ. Let $\Delta E_{i,j}$ denote the change in the expected length

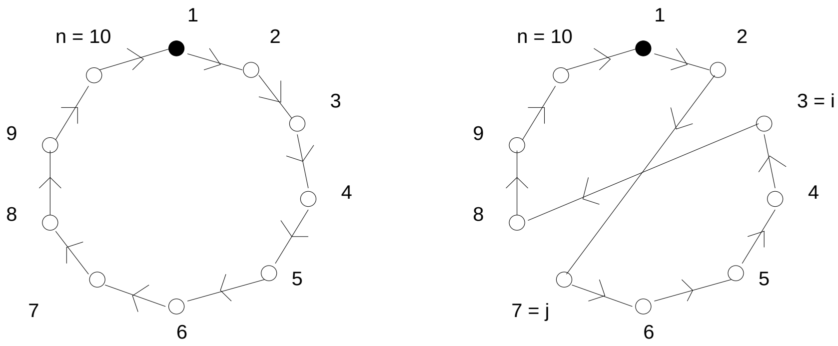

Figure 1: Tour $\tau = 1, 2, ..., i, i+1, ..., j, ..., n$ (left) and tour $\tilde{\tau} = 1, 2, ..., i-1, j, j-1, ..., i, j+1, ..., n$ (right) obtained from τ by reversing the section $(i, i+1, ..., j)$, with $n = 10$, $i = 3$, $j = 7$.

$E[L(\tilde{\tau})] - E[L(\tau)]$. In the following we reproduce verbatim the expressions given in [3] for the computation of $\Delta E_{i,j}$. Let $j = i + k$. Given the two dimensional matrices of partial results A and B:

$$A_{i,k} = \sum_{r=k}^{n-1} (1-p)^{r-1} d(i, i+r) \quad \text{and} \quad B_{i,k} = \sum_{r=k}^{n-1} (1-p)^{r-1} d(i-r, i), \quad 1 \leq k \leq n-1, \quad 1 \leq i \leq n, \tag{3}$$

$\Delta E_{i,j}$ is computed by means of the following recursive equations. Letting $q = 1 - p$, for $k = 1$,

$$\Delta E_{i,i+1} = p^3[q^{-1}A_{i,2} - (B_{i,1} - B_{i,n-1}) - (A_{i+1,1} - A_{i+1,n-1}) + q^{-1}B_{i+1,2}], \tag{4}$$

and for $k \geq 2$,

$$\begin{aligned} \Delta E_{i,j} = \Delta E_{i+1,j} \quad & +p^2[(q^{-k} - 1)A_{i,k+1} + (q^k - 1)(B_{i,1} - B_{i,n-k}) \\ & +(q^k - 1)(A_{j,1} - A_{j,n-k}) + (q^{-k} - 1)B_{j,k+1} \\ & +(q^{n-k} - 1)(A_{i,1} - A_{i,k}) + (q^{k-n} - 1)B_{i,n-k+1} \\ & +(q^{k-n} - 1)A_{j,n-k+1} + (q^{n-k} - 1)(B_{j,1} - B_{j,k})]. \end{aligned} \tag{5}$$

The 2-p-opt local search proposed proceeds in two phases. The first phase consists of computing $\Delta E_{i,i+1}$ for every value of i, and at the same time accumulating the two matrices of partial results A and B. Note that since the computation of $\Delta E_{i,i+1}$ only involves two rows of A and B, one can proceed to the next pair of nodes without recomputing each entire matrix. Each time a negative $\Delta E_{i,i+1}$ is encountered, one immediately switches the two nodes involved. The n calculations of phase one require $O(n)$ time apiece, or $O(n^2)$ time in all. At the end of this phase, an a priori tour is reached for which every $\Delta E_{i,i+1}$ is positive, and the matrices A and B are complete and correct for that tour. The second phase of the local search consists of computing $\Delta E_{i,j}$ recursively by means of equation (5). Since each $\Delta E_{i,j}$ in phase two is computed in $O(1)$ time, this phase, and thus the entire 2-p-opt checking sequence, is performed in $O(n^2)$.

Note that this procedure would be much faster than using the full evaluation function (equation (1)) for the cost of each 2-p-opt move. In fact, the full evaluation would require $O(n^2)$ time for the cost calculation of each move, and $O(n^4)$ time for the entire 2-p-opt checking sequence (since the neighborhood size of 2-p-opt is $O(n^2)$). Unfortunately, as we show in the following section, the delta evaluation expressed by equation (5) is not correct.

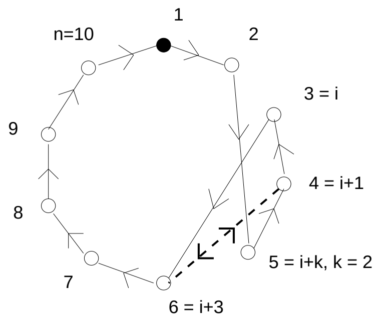

Figure 2: Tour $\tilde{\tau}$, with $\tilde{\tau}(i) = i+2$, $\tilde{\tau}(i+1) = i+1$, $\tilde{\tau}(i+2) = i$, and $\tilde{\tau}(i+3) = i+3$. The dashed line represents arcs $(i+1, i+3)$ and $(i+3, i+1)$.

2.2 Incorrectness of the proposed 2-p-opt recursive equation

Theorem 1 *Given an instance of the homogeneous PTSP, an a priori tour $\tau = (1, 2, ..., i, i+1, ..., i+k, ..., n)$, and an a priori tour $\tilde{\tau} = (1, 2, ..., i-1, i+k, i+k-1, ..., i, i+k+1, ..., n)$, then $E[L(\tilde{\tau})] - E[L(\tau)] \neq \Delta E_{i,i+k}$ for $k \geq 2$.*

The proof is by example. Let $k = 2$, then, according to (5)

$$\begin{aligned}\Delta E_{i,i+2} = \Delta E_{i+1,i+2} \quad & +p^2[(q^{-2}-1)A_{i,3} + (q^2-1)(B_{i,1} - B_{i,n-2}) \\ & +(q^2-1)(A_{i+2,1} - A_{i+2,n-2}) + (q^{-2}-1)B_{i+2,3} \\ & +(q^{n-2}-1)(A_{i,1} - A_{i,2}) + (q^{2-n}-1)B_{i,n-1} \\ & +(q^{2-n}-1)A_{i+2,n-1} + (q^{n-2}-1)(B_{i+2,1} - B_{i+2,2})].\end{aligned} \tag{6}$$

We show that in this expression there are arcs with an incorrect probability coefficient. Consider for instance arc $(i+1, i+3)$ and arc $(i+3, i+1)$. We examine the probability coefficients of these arcs, as computed by the recursive equation (6). The terms inside square brackets of equation (6) only take into account arcs joining respectively i and $i+2$ with other nodes. Therefore, the arcs $(i+1, i+3)$ and $(i+3, i+1)$ are only computed in the first term of equation (6):

$$\Delta E_{i+1,i+2} = p^3[q^{-1}A_{i+1,2} - (B_{i+1,1} - B_{i+1,n-1}) - (A_{i+2,1} - A_{i+2,n-1}) + q^{-1}B_{i+2,2}]. \tag{7}$$

In the above expression, arc $(i+1, i+3)$ is computed by the term corresponding to $r = 2$ of $A_{i+1,2} = \sum_{r=2}^{n-1}(1-p)^{r-1}d(i+1, i+1+r)$, and arc $(i+3, i+1)$ is computed by the term corresponding to $r = n-2$ of $B_{i+1,1} = \sum_{r=1}^{n-1}(1-p)^{r-1}d(i+1-r, i+1)$. From this observation and from equation (7), we see that arc $(i+1, i+3)$ is computed as follows:

$$\frac{p^3}{(1-p)}(1-p)d(i+1, i+3) = p^3 d(i+1, i+3), \tag{8}$$

and arc $(i+3, i+1)$ is computed as follows:

$$p^3(1-p)^{n-3}d(i+3, i+1). \tag{9}$$

Now we consider the objective function (1), and we write the probability coefficients that arcs $(i+1, i+3)$ and $(i+3, i+1)$ must have in $E[L(\tau)]$ and in $E[L(\tilde{\tau})]$. When the a priori tour is τ, arc $(i+1, i+3) =$

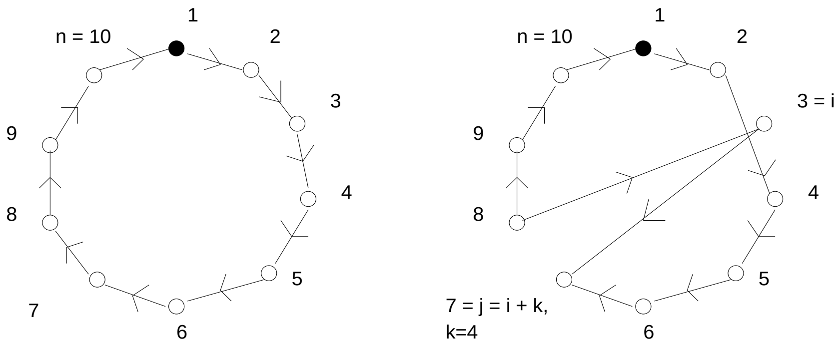

Figure 3: Tour $\tau = 1, 2, ..., i, i+1, ..., j, ..., n$ (left) and tour $\tilde{\tau} = 1, 2, ..., i-1, i+1, i+2, ..., j, i, j+1, ..., n$ (right) obtained from τ by moving node i to position j and shifting backwards one space the nodes $i+1, ..., j$, with $n = 10$, $i = 3$, $j = 7$.

$(\tau(i+1), \tau(i+3))$ and it has a probability coefficient of $p^2(1-p)$, that is, the probability that $\tau(i+1)$ and $\tau(i+3)$ require a visit while the single node $\tau(i+2)$ between them does not require a visit. The same coefficient is valid when the a priori tour is $\tilde{\tau}$, because $(i+1, i+3) = (\tilde{\tau}(i+1), \tilde{\tau}(i+3))$, and there is only one node $(\tilde{\tau}(i+2))$ between $\tilde{\tau}(i+1)$ and $\tilde{\tau}(i+3)$. Figure 2 may be of help in visualizing this. Similarly, arc $(i+3, i+1)$ has probability coefficient $p^2(1-p)^{n-3}$ both in τ and in $\tilde{\tau}$ (since $i+3 = \tilde{\tau}(i+1-(n-2))$). Therefore arc $(i+1, i+3)$ is computed as follows in $E[L(\tilde{\tau})] - E[L(\tau)]$:

$$[p^2(1-p) - p^2(1-p)]d(i+1, i+3) = 0, \qquad (10)$$

in contrast with (8), and the arc $(i+3, i+1)$ is computed as follows in $E[L(\tilde{\tau})] - E[L(\tau)]$:

$$[p^2(1-p)^{n-3} - p^2(1-p)^{n-3}]d(i+3, i+1) = 0, \qquad (11)$$

in contrast with (9). ■

Appendix A presents the step-by-step evaluation of a 2-p-opt move, first using the full evaluation of the objective function (1), and second using Bertsimas' proposed expression (5).

3 The 1-shift

The 1-shift local search was introduced by Bertsimas in [1] and later published by this journal in an article by Bertsimas and Howell [3]. Given an a priori tour τ, its 1-shift neighborhood is the set of tours obtained by moving a node i to position j of the tour, with the intervening nodes being shifted backwards one space accordingly, as for example in Figure 3.

3.1 Expressions proposed in [3] for the cost of a 1-shift move

Consider, without loss of generality, a tour $\tau = 1, 2, ..., i, i+1, ..., j, ..., n$ and a tour $\tilde{\tau} = 1, 2, ..., i-1, i+1, i+2, ..., j, i, j+1, ..., n$ obtained from τ by moving node i to position j and shifting backwards one space the nodes $i+1, ..., j$. Let $\Delta E'_{i,j}$ denote the change in the expected length $E[L(\tilde{\tau})] - E[L(\tau)]$. In the following we reproduce verbatim the expressions given in [3] for the computation of $\Delta E'_{i,j}$.

Let $j = i + k$. Note that, for $k = 1$, the tour $\tilde{\tau}$ obtained by 1-shift is the same as the one obtained by 2-p-opt. Therefore, for $k = 1$, $\Delta E'_{i,i+1} = \Delta E_{i,i+1}$ and the computation is done according to (4). For $k \geq 2$, the proposed expression is:

$$\begin{aligned}\Delta E'_{i,j} = \Delta E'_{i,j-1} \quad & +p^2[(q^{-k} - q^{-(k-1)})A_{i,k+1} + (q^k - q^{k-1})(B_{i,1} - B_{i,n-k}) \\ & +(q-1)(A_{j,1} - A_{j,n-k}) + (q^{-1} - 1)B_{j,k+1} \\ & +(q^{n-k} - q^{n-(k-1)})(A_{i,1} - A_{i,k}) + (q^{k-n} - q^{(k-1)-n})B_{i,n-k+1} \\ & +(1 - q^{-1})A_{j,n-k+1} + (1-q)(B_{j,1} - B_{j,k})],\end{aligned} \tag{12}$$

where A and B are the same matrices of partial results considered for the 2-p-opt and defined by (3). This algorithm follows the same lines as the 2-p-opt algorithm: all phase one computations, including the accumulation of matrices A and B, proceed in the same way. The second phase of the local search consists of computing $\Delta E'_{i,j}$ recursively by means of equation (12). Like 2-p-opt, since each $\Delta E'_{i,j}$ in phase two is computed in $O(1)$ time, this phase, and thus the entire 1-shift checking sequence, is performed in $O(n^2)$. Similarly to the 2-p-opt, note that this procedure would be much faster than using full evaluation function (equation (1)) for the cost of each 1-shift move. In fact, the full evaluation would require $O(n^2)$ time for the cost calculation of each move, and $O(n^4)$ time for the entire 1-shift checking sequence (since the neighborhood size of 1-shift is $O(n^2)$). Unfortunately, as we show in the following section, the delta evaluation expressed by equation (12) is not correct.

3.2 Incorrectness of the proposed 1-shift recursive expression

Theorem 2 *Given an instance of the homogeneous PTSP, an a priori tour* $\tau = (1, 2, ..., i, i+1, ..., i+k, ..., n)$*, and an a priori tour* $\tilde{\tau} = (1, 2, ..., i-1, i+1, i+2, ..., i+k, i, i+k+1, ...n)$*, then* $E[L(\tilde{\tau})] - E[L(\tau)] \neq \Delta E'_{i,i+k}$ *for* $k \geq 2$.

The proof is by example. Let $k = 2$, then, according to (12)

$$\begin{aligned}\Delta E'_{i,i+2} = \Delta E'_{i,i+1} \quad & +p^2[(q^{-2} - q^{-1})A_{i,3} + (q^2 - q)(B_{i,1} - B_{i,n-2}) \\ & +(q-1)(A_{i+2,1} - A_{i+2,n-2}) + (q^{-1} - 1)B_{i+2,3} \\ & +(q^{n-2} - q^{n-1})(A_{i,1} - A_{i,2}) + (q^{2-n} - q^{1-n})B_{i,n-1} \\ & +(1 - q^{-1})A_{i+2,n-1} + (1-q)(B_{i+2,1} - B_{i+2,2})].\end{aligned} \tag{13}$$

We show that in this expression there are arcs with an incorrect probability coefficient. Consider for instance arc $(i, i+1)$ and arc $(i+1, i)$. We examine the probability coefficients of these arcs, as computed by the recursive equation (13). The first term of (13) is $\Delta E'_{i,i+1}$ and it is equal to (4). By an inspection of the terms A and B in equation (4) it is not difficult to see that arcs $(i, i+1)$ and $(i+1, i)$ are not computed by $\Delta E'_{i,i+1}$. Arc $(i, i+1)$ is computed by the term $(q^{n-2} - q^{n-1})(A_{i,1} - A_{i,2})$ of equation (13), that is,

$$p^2((1-p)^{n-2} - (1-p)^{n-1})d(i, i+1), \tag{14}$$

and arc $(i+1, i)$ is computed by the term $(q^{2-n} - q^{1-n})B_{i,n-1}$ of equation (13), that is,

$$\frac{p^3}{(1-p)}d(i+1, i). \tag{15}$$

Now we consider the objective function (1), and we write the probability coefficients that arcs $(i, i+1)$ and $(i+1, i)$ must have in $E[L(\tau)]$ and in $E[L(\tilde{\tau})]$. When the a priori tour is τ, arc $(i, i+1) = (\tau(i), \tau(i+1))$

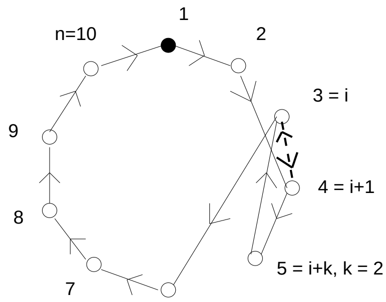

Figure 4: Tour $\tilde{\tau}$, with $\tilde{\tau}(i) = i+1$, $\tilde{\tau}(i+1) = i+2$, and $\tilde{\tau}(i+2) = i$. The dashed line represents arcs $(i, i+1)$ and $(i+1, i)$.

and it has a probability coefficient of p^2, that is, the probability that $\tau(i)$ and $\tau(i+1)$ require a visit. When the a priori tour is $\tilde{\tau}$, we have $(i, i+1) = (\tilde{\tau}(i+2), \tilde{\tau}(i)) = (\tilde{\tau}(i-(n-2)), \tilde{\tau}(i))$, which leads to the probability coefficient $p^2(1-p)^{n-3}$ (see Figure 4). The difference of the probability coefficients of $(i, i+1)$ in $\tilde{\tau}$ and τ leads to

$$p^2((1-p)^{n-3} - 1)d(i, i+1), \tag{16}$$

which is different from (14). Let us now focus on arc $(i+1, i)$. When the a priori tour is τ, arc $(i+1, i) = (\tau(i-(n-1)), \tau(i))$ and it has a probability coefficient of $p^2(1-p)^{n-2}$. When the a priori tour is $\tilde{\tau}$, we have $(i+1, i) = (\tilde{\tau}(i), \tilde{\tau}(i+2))$ (see Figure 4), and the probability coefficient is $p^2(1-p)$. The difference of the probability coefficients of $(i+1, i)$ in $\tilde{\tau}$ and τ leads to:

$$p^2(1-p-(1-p)^{n-2})d(i+1, i), \tag{17}$$

which is different from (15). ■

Appendix B presents the step-by-step evaluation of a 1-shift move, first using the full evaluation of the objective function (1), and second using Bertsimas' proposed expression (12).

4 Conclusion

We have shown that the equations proposed by Bertsimas in [3] for the evaluation of the cost of 2-p-opt and 1-shift local search moves are not correct. This fact has at least two consequences. First, results published in [2], [3] and [4], which used the 2-p-opt and 1-shift local search, should be re-examined because they may be wrong. Second, it poses the question whether an exact move cost evaluation exists, that would allow the local search to run in $O(n^2)$ time per move. Using the $O(n^2)$ full evaluation of the objective function for the cost of a move gives an $O(n^4)$ local search algorithm (since the neighborhood size of 1-shift and 2-p-opt is $O(n^2)$), so any improvement on this would be most useful.

5 Acknowledgments

The authors wish to thank Marco Dorigo and Luca Maria Gambardella for their useful comments. This research has been partially supported by the Swiss National Science Foundation project titled "On-line fleet management", grant 16R10FM, and by the "Metaheuristics Network", a Research Training Network funded by the Improving Human Potential programme of the CEC, grant HPRN-CT-1999-00106. Joshua Knowles gratefully acknowledges the support of a Marie Curie postdoctoral research fellowship of the CEC, contract number HPMF-CT-2000-00992. The information provided in this paper is the sole responsibility of the authors and does not reflect the Community's opinion. The Community is not responsible for any use that might be made of data appearing in this publication.

References

[1] D. J. Bertsimas. *Probabilistic Combinatorial Optimization Problems.* PhD thesis, MIT, Cambridge, MA, 1988.

[2] D. J. Bertsimas, P. Chervi, and M. Peterson. Computational approaches to stochastic vehicle routing problems. *Transportation Sciences*, 29(4):342–352, 1995.

[3] D. J. Bertsimas and L. Howell. Further results on the probabilistic traveling salesman problem. *European Journal of Operational Research*, 65:68–95, 1993.

[4] D. J. Bertsimas, P. Jaillet, and A. Odoni. A priori optimization. *Operations Research*, 38:1019–1033, 1990.

[5] P. Jaillet. *Probabilistic Traveling Salesman Problems.* PhD thesis, MIT, Cambridge, MA, 1985.

A Example of a 2-p-opt move evaluation

We place the nodes of the graph (customers) on the vertices of a regular hexagon in the Euclidean plane of side 1 so that all distances between pairs of nodes are from the set, $\{1, 2, \sqrt{3}\}$. We set $p = 0.5$.

Let us calculate the difference in expected length of the tours $\tau = 1, 2, 3, 4, 5, 6$ and $\tilde{\tau} = 1, 5, 4, 3, 2, 6$. We shall first do this by calculating, by equation (1), the expected length of each and subtracting one from the other. Then we shall compare this result with the calculation using Bertsimas' recursive equation (5).

The expected length of the tour $\tau = 1, 2, 3, 4, 5, 6$ is given by:

$$
\begin{aligned}
E[L(\tau)] \quad = \quad & (0.25 \times 0.5^{(1-1)} \times d(1,2) = 0.25 \times 1 = 0.25) + \\
& (0.25 \times 0.5^{(2-1)} \times d(1,3) = 0.125 \times \sqrt{3} = 0.216506) + \\
& (0.25 \times 0.5^{(3-1)} \times d(1,4) = 0.0625 \times 2 = 0.125) + \\
& (0.25 \times 0.5^{(4-1)} \times d(1,5) = 0.031250 \times \sqrt{3} = 0.054127) + \\
& (0.25 \times 0.5^{(5-1)} \times d(1,6) = 0.015625 \times 1 = 0.015625) + \\
& (0.25 \times 0.5^{(1-1)} \times d(2,3) = 0.25 \times 1 = 0.25) + \\
& (0.25 \times 0.5^{(2-1)} \times d(2,4) = 0.125 \times \sqrt{3} = 0.216506) + \\
& (0.25 \times 0.5^{(3-1)} \times d(2,5) = 0.0625 \times 2 = 0.125) + \\
& (0.25 \times 0.5^{(4-1)} \times d(2,6) = 0.031250 \times \sqrt{3} = 0.054127) + \\
& (0.25 \times 0.5^{(5-1)} \times d(2,1) = 0.015625 \times 1 = 0.015625) + \\
& (0.25 \times 0.5^{(1-1)} \times d(3,4) = 0.25 \times 1 = 0.25) + \\
& (0.25 \times 0.5^{(2-1)} \times d(3,5) = 0.125 \times \sqrt{3} = 0.216506) +
\end{aligned}
$$

$$
\begin{aligned}
& (0.25 \times 0.5^{(3-1)} \times d(3,6) = 0.0625 \times 2 = 0.125) + \\
& (0.25 \times 0.5^{(4-1)} \times d(3,1) = 0.031250 \times \sqrt{3} = 0.054127) + \\
& (0.25 \times 0.5^{(5-1)} \times d(3,2) = 0.015625 \times 1 = 0.015625) + \\
& (0.25 \times 0.5^{(1-1)} \times d(4,5) = 0.25 \times 1 = 0.25) + \\
& (0.25 \times 0.5^{(2-1)} \times d(4,6) = 0.125 \times \sqrt{3} = 0.216506) + \\
& (0.25 \times 0.5^{(3-1)} \times d(4,1) = 0.0625 \times 2 = 0.125) + \\
& (0.25 \times 0.5^{(4-1)} \times d(4,2) = 0.031250 \times \sqrt{3} = 0.054127) + \\
& (0.25 \times 0.5^{(5-1)} \times d(4,3) = 0.015625 \times 1 = 0.015625) + \\
& (0.25 \times 0.5^{(1-1)} \times d(5,6) = 0.25 \times 1 = 0.25) + \\
& (0.25 \times 0.5^{(2-1)} \times d(5,1) = 0.125 \times \sqrt{3} = 0.216506) + \\
& (0.25 \times 0.5^{(3-1)} \times d(5,2) = 0.0625 \times 2 = 0.125) + \\
& (0.25 \times 0.5^{(4-1)} \times d(5,3) = 0.031250 \times \sqrt{3} = 0.054127) + \\
& (0.25 \times 0.5^{(5-1)} \times d(5,4) = 0.015625 \times 1 = 0.015625) + \\
& (0.25 \times 0.5^{(1-1)} \times d(6,1) = 0.25 \times 1 = 0.25) + \\
& (0.25 \times 0.5^{(2-1)} \times d(6,2) = 0.125 \times \sqrt{3} = 0.216506) + \\
& (0.25 \times 0.5^{(3-1)} \times d(6,3) = 0.0625 \times 2 = 0.125) + \\
& (0.25 \times 0.5^{(4-1)} \times d(6,4) = 0.031250 \times \sqrt{3} = 0.054127) + \\
& (0.25 \times 0.5^{(5-1)} \times d(6,5) = 0.015625 \times 1 = 0.015625) + \\
= \; & 3.967548. \qquad (18)
\end{aligned}
$$

The expected length of the tour $\tilde{\tau} = 1, 5, 4, 3, 2, 6$ is given by:

$$
\begin{aligned}
E[L(\tilde{\tau})] = \; & (0.25 \times 0.5^{(1-1)} \times d(1,5) = 0.25 \times \sqrt{3} = 0.433013) + \\
& (0.25 \times 0.5^{(2-1)} \times d(1,4) = 0.125 \times 2 = 0.25) + \\
& (0.25 \times 0.5^{(3-1)} \times d(1,3) = 0.0625 \times \sqrt{3} = 0.108253) + \\
& (0.25 \times 0.5^{(4-1)} \times d(1,2) = 0.03125 \times 1 = 0.03125) + \\
& (0.25 \times 0.5^{(5-1)} \times d(1,6) = 0.015625 \times 1 = 0.015625) + \\
& (0.25 \times 0.5^{(1-1)} \times d(5,4) = 0.25 \times 1 = 0.25) + \\
& (0.25 \times 0.5^{(2-1)} \times d(5,3) = 0.125 \times \sqrt{3} = 0.216506) + \\
& (0.25 \times 0.5^{(3-1)} \times d(5,2) = 0.0625 \times 2 = 0.125) + \\
& (0.25 \times 0.5^{(4-1)} \times d(5,6) = 0.03125 \times 1 = 0.031250) + \\
& (0.25 \times 0.5^{(5-1)} \times d(5,1) = 0.015625 \times \sqrt{3} = 0.027063) + \\
& (0.25 \times 0.5^{(1-1)} \times d(4,3) = 0.25 \times 1 = 0.25) + \\
& (0.25 \times 0.5^{(2-1)} \times d(4,2) = 0.125 \times \sqrt{3} = 0.216506) + \\
& (0.25 \times 0.5^{(3-1)} \times d(4,6) = 0.0625 \times \sqrt{3} = 0.108253) + \\
& (0.25 \times 0.5^{(4-1)} \times d(4,1) = 0.03125 \times 2 = 0.0625) + \\
& (0.25 \times 0.5^{(5-1)} \times d(4,5) = 0.015625 \times 1 = 0.015625) +
\end{aligned}
$$

$$
\begin{aligned}
& (0.25 \times 0.5^{(1-1)} \times d(3,2) = 0.25 \times 1 = 0.25) + \\
& (0.25 \times 0.5^{(2-1)} \times d(3,6) = 0.125 \times 2 = 0.25) + \\
& (0.25 \times 0.5^{(3-1)} \times d(3,1) = 0.0625 \times \sqrt{3} = 0.108253) + \\
& (0.25 \times 0.5^{(4-1)} \times d(3,5) = 0.03125 \times \sqrt{3} = 0.054127) + \\
& (0.25 \times 0.5^{(5-1)} \times d(3,4) = 0.015625 \times 1 = 0.015625) + \\
& (0.25 \times 0.5^{(1-1)} \times d(2,6) = 0.25 \times \sqrt{3} = 0.433013) + \\
& (0.25 \times 0.5^{(2-1)} \times d(2,1) = 0.125 \times 1 = 0.125) + \\
& (0.25 \times 0.5^{(3-1)} \times d(2,5) = 0.0625 \times 2 = 0.125) + \\
& (0.25 \times 0.5^{(4-1)} \times d(2,4) = 0.03125 \times \sqrt{3} = 0.054127) + \\
& (0.25 \times 0.5^{(5-1)} \times d(2,3) = 0.015625 \times 1 = 0.015625) + \\
& (0.25 \times 0.5^{(1-1)} \times d(6,1) = 0.25 \times 1 = 0.25) + \\
& (0.25 \times 0.5^{(2-1)} \times d(6,5) = 0.125 \times 1 = 0.125) + \\
& (0.25 \times 0.5^{(3-1)} \times d(6,4) = 0.0625 \times \sqrt{3} = 0.108253) + \\
& (0.25 \times 0.5^{(4-1)} \times d(6,3) = 0.03125 \times 2 = 0.0625) + \\
& (0.25 \times 0.5^{(5-1)} \times d(6,2) = 0.015625 \times \sqrt{3} = 0.027063) \\
= \; & 4.144431. \qquad (19)
\end{aligned}
$$

Thus, the difference between the expected lengths of the tours is:

$$
E[L(\tilde{\tau})] - E[L(\tau)] = 3.967548 - 4.144431 = 0.176883. \qquad (20)
$$

In the following we calculate $\Delta E_{2,5}$ using Bertsimas' recursive equation (5), and we show that $\Delta E_{2,5}$ is not equal to (20). By symmetry we have that $\forall k \in \{1, ..., n-1\}, A_{i,k} = A_{j,k}, i,\ j \in \{1, ..., n\}$. Furthermore, $\forall k \in \{1, ..., n-1\}$, $i \in \{1, ..., n\}$, $B_{i,k} = A_{i,k}$. Thus, $\forall i \in \{1, ..., n\}$, we have:

$$
\begin{aligned}
A_{i,1} = B_{i,1} &= 1 \times 1 + 1/2 \times \sqrt{3} + 1/4 \times 2 + 1/8 \times \sqrt{3} + 1/16 \times 1 \\
&= 5/8 \times \sqrt{3} + 25/16 \\
&= 1/16(10 \times \sqrt{3} + 25) \\
&= 2.645031755 \qquad (21)
\end{aligned}
$$

$$
\begin{aligned}
A_{i,2} = B_{i,2} &= 1/2 \times \sqrt{3} + 1/4 \times 2 + 1/8 \times \sqrt{3} + 1/16 \times 1 \\
&= 5/8 \times \sqrt{3} + 9/16 \\
&= 1/16(10 \times \sqrt{3} + 9) \\
&= 1.645031755 \qquad (22)
\end{aligned}
$$

$$
\begin{aligned}
A_{i,3} = B_{i,3} &= 1/4 \times 2 + 1/8 \times \sqrt{3} + 1/16 \times 1 \\
&= 1/8 \times \sqrt{3} + 9/16 \\
&= 1/16(2 \times \sqrt{3} + 9) \\
&= 0.77900635 \qquad (23)
\end{aligned}
$$

$$
\begin{aligned}
A_{i,4} = B_{i,4} &= 1/8 \times \sqrt{3} + 1/16 \times 1 \\
&= 1/8 \times \sqrt{3} + 1/16 \\
&= 1/16(2 \times \sqrt{3} + 1) \\
&= 0.27900635 \qquad (24)
\end{aligned}
$$

$$
\begin{aligned}
A_{i,5} = B_{i,5} &= 1/16 \times 1 \\
&= 0.0625. \qquad (25)
\end{aligned}
$$

Using these and Bertsimas' recursive equation (5), we will now compute $\Delta E_{2,5}$ on the tour $\tau = 1, 2, 3, 4, 5, 6$. That is, the difference in expected length when we move from the tour $\tau = 1, 2, 3, 4, 5, 6$ to the tour $\tilde{\tau} = 1, 5, 4, 3, 2, 6$.

First, we have: $i = 2$, $j = 5$, $k = 3$, $p = 0.5$, and $q = 0.5$. This gives us

$$
\begin{aligned}
\Delta E_{2,5} &= \Delta E_{3,5} + (p^2 = 0.25) \times [\\
&\quad (q^{-3} - 1 = 7.0) \times (A_{2,4} = 0.279006) + \\
&\quad (q^3 - 1 = -0.875) \times ((B_{2,1} = 2.645032) - (B_{2,3} = 0.779006)) + \\
&\quad (q^3 - 1 = -0.875) \times ((A_{5,1} = 2.645032) - (A_{5,3} = 0.779006)) + \\
&\quad (q^{-3} - 1 = 7.0) \times (B_{5,4} = 0.279006) + \\
&\quad (q^3 - 1 = -0.875) \times ((A_{2,1} = 2.645032) - (A_{2,3} = 0.779006)) + \\
&\quad (q^{-3} - 1 = 7.0) \times (B_{2,4} = 0.279006) + \\
&\quad (q^{-3} - 1 = 7.0) \times (A_{5,4} = 0.279006) + \\
&\quad (q^3 - 1 = -0.875) \times ((B_{5,1} = 2.645032) - (B_{5,3} = 0.779006))] \\
&= \Delta E_{3,5} + 0.25 \times [\\
&\quad 1.953044 + \\
&\quad -1.632772 + \\
&\quad -1.632772 + \\
&\quad 1.953044 + \\
&\quad -1.632772 + \\
&\quad 1.953044 + \\
&\quad 1.953044 + \\
&\quad -1.632772] \\
&= \Delta E_{3,5} + 0.320272. \qquad (26)
\end{aligned}
$$

And we can calculate $\Delta E_{3,5}$ in a similar way. We have $i = 3$, $j = 5$, $k = 2$, $p = 0.5$, and $q = 0.5$. This gives us:

$$
\begin{aligned}
\Delta E_{3,5} &= \Delta E_{4,5} + (p^2 = 0.25) \times [\\
&\quad (q^{-2} - 1 = 3) \times (A_{3,3} = 0.779006) + \\
&\quad (q^2 - 1 = -0.75) \times ((B_{3,1} = 2.645032) - (B_{3,4} = 0.279006)) + \\
&\quad (q^2 - 1 = -0.75) \times ((A_{5,1} = 2.645032) - (A_{5,4} = 0.279006)) +
\end{aligned}
$$

$$
\begin{aligned}
& (q^{-2}-1=3)\times(B_{5,3}=0.779006)+\\
& (q^{4}-1=-0.9375)\times((A_{3,1}=2.645032)-(A_{3,2}=1.645032))+\\
& (q^{-4}-1=15)\times(B_{3,5}=0.0625)+\\
& (q^{-4}-1=15)\times(A_{5,5}=0.0625)+\\
& (q^{4}-1=-0.9375)\times((B_{5,1}=2.645032)-(B_{5,2}=1.645032))]\\
= & \Delta E_{4,5}+0.25\times[\\
& 2.337019+\\
& -1.774519+\\
& -1.774519+\\
& 2.337019+\\
& -0.9375+\\
& 0.9375+\\
& 0.9375+\\
& -0.9375]\\
= & \Delta E_{4,5}+0.28125. \qquad (27)
\end{aligned}
$$

And finally, we can calculate $\Delta E_{4,5}$ using equation (4) for switching two adjacent nodes with $i=4$, $j=5$, $k=1$, $p=0.5$, and $q=0.5$. This gives us:

$$
\begin{aligned}
\Delta E_{4,5} = & (p^3=0.125)\times[\\
& (A_{4,2}=1.645032)/(1-p=0.5)-\\
& ((B_{4,1}=2.645032)-(B_{4,5}=0.0625))-\\
& (A_{5,1}=2.645032)-(A_{5,5}=0.0625))+\\
& (B_{5,2}=1.645032)/(1-p=0.5)]\\
= & 0.125\times 1.415064\\
= & 0.176883. \qquad (28)
\end{aligned}
$$

Thus, combining (26),(27), and (28) we have:

$$
\begin{aligned}
\Delta E_{2,5} = & 0.320272+0.28125+0.176883\\
= & 0.778405. \qquad (29)
\end{aligned}
$$

Thus comparing, (29) with (20) we have demonstrated the error in Bertsimas' equation (5).

B Example of a 1-shift move evaluation

As in appendix A, we place the nodes of the graph (customers) on the vertices of a regular hexagon in the Euclidean plane of side 1 so that all distances between pairs of nodes are from the set, $\{1,2,\sqrt{3}\}$. We set $p=0.5$.

Let us calculate the difference in expected length of the tours $\tau=1,2,3,4,5,6$ and $\tilde{\tau}=1,3,4,5,2,6$. We shall first do this by calculating, by equation (1), the expected length of each and subtracting one from the other. Then we shall compare this result with the calculation using the Bertsimas' recursive equation (12). The length of the tour $\tau = 1,2,3,4,5,6$ is $E[L(\tau)] = 3.967548$, as calculated by (18) in Appendix A. The length of the tour $\tilde{\tau} = 1,3,4,5,2,6$ is given by:

$$\begin{aligned}
E[L(\tilde{\tau})] &= (0.25 \times 0.5^{(1-1)} \times d(1,3) = 0.25 \times \sqrt{3} = 0.433013) + \\
&\quad (0.25 \times 0.5^{(2-1)} \times d(1,4) = 0.125 \times 2 = 0.25) + \\
&\quad (0.25 \times 0.5^{(3-1)} \times d(1,5) = 0.0625 \times \sqrt{3} = 0.108253) + \\
&\quad (0.25 \times 0.5^{(4-1)} \times d(1,2) = 0.03125 \times 1 = 0.03125) + \\
&\quad (0.25 \times 0.5^{(5-1)} \times d(1,6) = 0.015625 \times 1 = 0.015625) + \\
&\quad (0.25 \times 0.5^{(1-1)} \times d(3,4) = 0.25 \times 1 = 0.250) + \\
&\quad (0.25 \times 0.5^{(2-1)} \times d(3,5) = 0.125 \times \sqrt{3} = 0.216506) + \\
&\quad (0.25 \times 0.5^{(3-1)} \times d(3,2) = 0.0625 \times 1 = 0.0625) + \\
&\quad (0.25 \times 0.5^{(4-1)} \times d(3,6) = 0.03125 \times 2 = 0.0625) + \\
&\quad (0.25 \times 0.5^{(5-1)} \times d(3,1) = 0.015625 \times \sqrt{3} = 0.027063) + \\
&\quad (0.25 \times 0.5^{(1-1)} \times d(4,5) = 0.25 \times 1 = 0.25) + \\
&\quad (0.25 \times 0.5^{(2-1)} \times d(4,2) = 0.125 \times \sqrt{3} = 0.216506) + \\
&\quad (0.25 \times 0.5^{(3-1)} \times d(4,6) = 0.0625 \times \sqrt{3} = 0.108253) + \\
&\quad (0.25 \times 0.5^{(4-1)} \times d(4,1) = 0.03125 \times 2 = 0.0625) + \\
&\quad (0.25 \times 0.5^{(5-1)} \times d(4,3) = 0.015625 \times 1 = 0.015625) + \\
&\quad (0.25 \times 0.5^{(1-1)} \times d(5,2) = 0.25 \times 2 = 0.5) + \\
&\quad (0.25 \times 0.5^{(2-1)} \times d(5,6) = 0.125 \times 1 = 0.125) + \\
&\quad (0.25 \times 0.5^{(3-1)} \times d(5,1) = 0.0625 \times \sqrt{3} = 0.108253) + \\
&\quad (0.25 \times 0.5^{(4-1)} \times d(5,3) = 0.03125 \times \sqrt{3} = 0.054127) + \\
&\quad (0.25 \times 0.5^{(5-1)} \times d(5,4) = 0.015625 \times 1 = 0.015625) + \\
&\quad (0.25 \times 0.5^{(1-1)} \times d(2,6) = 0.25 \times \sqrt{3} = 0.433013) + \\
&\quad (0.25 \times 0.5^{(2-1)} \times d(2,1) = 0.125 \times 1 = 0.125) + \\
&\quad (0.25 \times 0.5^{(3-1)} \times d(2,3) = 0.0625 \times 1 = 0.0625) + \\
&\quad (0.25 \times 0.5^{(4-1)} \times d(2,4) = 0.03125 \times \sqrt{3} = 0.054127) + \\
&\quad (0.25 \times 0.5^{(5-1)} \times d(2,5) = 0.015625 \times 2 = 0.03125) + \\
&\quad (0.25 \times 0.5^{(1-1)} \times d(6,1) = 0.25 \times 1 = 0.25) + \\
&\quad (0.25 \times 0.5^{(2-1)} \times d(6,3) = 0.125 \times 2 = 0.25) + \\
&\quad (0.25 \times 0.5^{(3-1)} \times d(6,4) = 0.0625 \times \sqrt{3} = 0.108253) + \\
&\quad (0.25 \times 0.5^{(4-1)} \times d(6,5) = 0.03125 \times 1 = 0.03125) + \\
&\quad (0.25 \times 0.5^{(5-1)} \times d(6,2) = 0.015625 \times \sqrt{3} = 0.027063) + \\
&= 4.285056. \qquad (30)
\end{aligned}$$

Thus, the difference between the expected lengths of the tours is:

$$E[L(\tilde{\tau})] - E[L(\tau)] = 4.285056 - 3.967548 = 0.317508. \qquad (31)$$

In the following we calculate $\Delta E'_{2,5}$ using Bertsimas' recursive equation (12). Using equations (22)...(25) for the values of A and B and Bertsimas' recursive equation (12), we will now compute $\Delta E'_{2,5}$ on the tour $\tau = 1,2,3,4,5,6$. That is, the difference in expected length when we move from the tour $\tau = 1,2,3,4,5,6$ to the tour $\tilde{\tau} = 1,3,4,5,2,6$.

First, we have: $i = 2$, $j = 5$, $k = 3$, $p = 0.5$, and $q = 0.5$. This gives us

$$\begin{aligned}
\Delta E'_{2,5} &= \Delta E'_{2,4} + (p^2 = 0.25) \times [\\
&\quad (q^{-3} - q^{-2} = 4) \times (A_{2,4} = 0.279006) +\\
&\quad (q^3 - q^2 = -0.125) \times ((B_{2,1} = 2.645032) - (B_{2,3} = 0.779006)) +\\
&\quad (q - 1 = -0.5) \times ((A_{5,1} = 2.645032) - (A_{5,3} = 0.779006)) +\\
&\quad (q^{-1} - 1 = 1) \times (B_{5,4} = 0.279006) +\\
&\quad (q^3 - q^4 = 0.0625) \times ((A_{2,1} = 2.645032) - (A_{2,3} = 0.779006)) +\\
&\quad (q^{-3} - q^{-4} = -8) \times (B_{2,4} = 0.279006) +\\
&\quad (1 - q^{-1} = -1) \times (A_{5,4} = 0.279006) +\\
&\quad (1 - q = 0.5) \times ((B_{5,1} = 2.645032) - (B_{5,3} = 0.779006))]\\
&= \Delta E'_{2,4} + 0.25 \times [\\
&\quad 1.116025 +\\
&\quad -0.233253 +\\
&\quad -0.933013 +\\
&\quad 0.279006 +\\
&\quad 0.116627 +\\
&\quad -2.232051 +\\
&\quad -0.279006 +\\
&\quad 0.933013]\\
&= \Delta E'_{2,4} - 0.308163. \qquad (32)
\end{aligned}$$

And we can calculate $\Delta E'_{2,4}$ in a similar way. We have $i = 2$, $j = 4$, $k = 2$, $p = 0.5$, and $q = 0.5$. This gives us:

$$\begin{aligned}
\Delta E'_{2,4} &= \Delta E'_{2,3} + (p^2 = 0.25) \times [\\
&\quad (q^{-2} - q^{-1} = 2) \times (A_{2,3} = 0.779006) +\\
&\quad (q^2 - q = -0.25) \times ((B_{2,1} = 2.645032) - (B_{2,4} = 0.279006)) +\\
&\quad (q - 1 = -0.5) \times ((A_{4,1} = 2.645032) - (A_{4,4} = 0.279006)) +\\
&\quad (q^{-1} - 1 = 1) \times (B_{4,3} = 0.779006) +\\
&\quad (q^4 - q^5 = 0.03125) \times ((A_{2,1} = 2.645032) - (A_{2,2} = 1.645032)) +\\
&\quad (q^{-4} - q^{-5} = -16) \times (B_{2,5} = 0.0625) +\\
&\quad (1 - q^{-1} = -1) \times (A_{4,5} = 0.0625) +\\
&\quad (1 - q = 0.5) \times ((B_{4,1} = 2.645032) - (B_{4,2} = 1.645032))]\\
&= \Delta E'_{2,3} + 0.25 \times [\\
&\quad 1.558013 +
\end{aligned}$$

$$
\begin{aligned}
& -0.591506 + \\
& -1.183013 + \\
& 0.779006 + \\
& 0.03125 + \\
& -1 + \\
& -0.0625 + \\
& 0.5] \\
= \quad & \Delta E'_{2,3} + 0.007812. \qquad (33)
\end{aligned}
$$

And finally, we can calculate $\Delta E'_{2,3}$ using equation (4) for switching two adjacent nodes with $i = 2$, $j = 3$, $k = 1$, $p = 0.5$, and $q = 0.5$. This gives us:

$$
\begin{aligned}
\Delta E'_{2,3} \quad = \quad & (p^3 = 0.125) \times [\\
& (A_{2,2} = 1.645032)/(1 - p = 0.5) - \\
& ((B_{2,1} = 2.645032) - (B_{2,5} = 0.0625)) - \\
& (A_{3,1} = 2.645032) - (A_{3,5} = 0.0625)) + \\
& (B_{3,2} = 1.645032)/(1 - p = 0.5)] \\
= \quad & 0.125 \times 1.415064 \\
= \quad & 0.176883. \qquad (34)
\end{aligned}
$$

Thus, combining (32), (33), and (34) we have:

$$
\begin{aligned}
\Delta E'_{2,5} \quad = \quad & -0.308163 + 0.007812 + 0.176883 \\
= \quad & -0.123468. \qquad (35)
\end{aligned}
$$

Thus comparing, (35) with (31) we have demonstrated the error in Bertsimas' equation (12).