The probabilistic vehicle routing problem
Dimitris Bertsimas
Sloan W.P. No. 2067-88 September 1988


#### Abstract

The probabilistic vehicle routing problem (PVRP) is a natural probabilistic variation of the classical vehicle routing problem (VRP), in which demands are probabilistic. The goal is to determine an a priori route of minimal expected total length, which corresponds to the expected total length of the route plus the expected value of the extra distance that might be required because demand on the route may occasionally exceed the capacity of the vehicle and force it to go back to the depot before continuing on its route. In this paper we analyze the PVRP using a variety of theoretical approaches. We find closedform expressions and algorithms to compute the expected length of an a priori route under various probabilistic assumptions. Based on these expressions we find upper and lower bounds for the PVRP and the VRP re-optimization strategy, in which we find the optimal route at every instance. We propose heuristics and analyze their worst-case performance. Moreover, we perform probabilistic analysis for the case that customer locations are random in the unit square and succeed in proving some sharp asymptotic theorems for the PVRP and the VRP re-optimization strategy, in which we find the optimal route at every instance. We further propose some asymptotically optimal algorithms. It is quite surprising to find that the PVRP and the strategy of re-optimization are asymptotically equivalent in terms of performance. Our results suggest that the PVRP is a strong and useful alternative to the strategy of re-optimization in capacitated routing problems.


Key words:Probabilistic vehicle routing problem, re-optimization strategy, probabilistic analysis, worst-case analysis of heuristics.

## Introduction

We consider a standard vehicle routing problem, i.e. a routing problem with a vehicle of capacity $Q$, a single depot and an objective function which minimizes total distance traveled, but with demands which are probabilistic in nature rather than deterministic. The problem to determine a fixed route of minimal expected total length, which corresponds to the expected total length of the fixed set of routes plus the expected value of the extra distance that might be required by a particular realization of the random variables, is called the probabilistic vehicle routing problem (PVRP). The extra distances will be due to the fact that demand on the route may occasionally exceed the capacity of the vehicle and force it to go back to the depot before continuing on its route. This class of problems differs from the stochastic vehicle routing problems described in [13] in the sense that here we are concerned only with routing costs without the introduction of additional parameters.

Depending on the time information becomes available, Jaillet and Odoni [8] define several alternative versions of the PVRP. Two examples are shown in Figure 1.

Under strategy $a$ the vehicle visits all the points in the same fixed order as under the a priori route, but serves only customers requiring service that day. The total expected distance traveled corresponds to the fixed length of the a priori route plus the expected value of the additional distance that must be covered whenever the demand on the route exceeds vehicle capacity. Strategy $b$ is defined similarly to $a$ with the sole difference that


4 (i) An apriori route through 6 customers (each with a demand of zero or one unit) by a vehicle of capocity 2

Class A.


4(ii) The two strategies when only the second, third and fifth customers have a non-zero demand.

Figure 1: The PVRP methodology
customers with no demand on a particular instance of the vehicle tour are simply skipped.

The difference between the two strategies (strategy $a$ and $b$ ) is essentially due to the time the information on demand becomes available. Strategy $a$ models situations in which there is no a priori information on customers' demand. The demand (if any) of any particular customer becomes known only when the customer is visited. The vehicle is then forced to return to the depot when its capacity is reached. Under strategy $b$, however, the actual demand is known before the tour starts, so that savings can occur by skipping customer locations with zero demand.

The reason for not re-optimizing the route on every problem instance could be that the system's operator does not have the resources for doing so; or, it may be decided that such redesign of tours is not sufficiently important to justify the required effort and cost. Even more importantly, the operator may have other priorities, such as regularity and personalization of service by having the same vehicle and driver visit a particular customer every day. These priorities can best be attained by following a probabilistic vehicle routing strategy.

The PVRP has important applications in the fields of logistics and of goods distribution. In general, PVRPs arise in practice whenever a company, on any given day, is faced with the problem of collections (deliveries) from (to) a random subset of its (known) global set of customers in an area, does not wish to or, simply, cannot redesign the routes every day and the capacity of the vehicles used is a major constraint. Consider for example
a problem in which a central bank has to collect money on a daily basis from several but not all of its branches. The capacity $Q$ of the vehicle used may not correspond to any physical constraint but to an upper bound on the amount of money that a vehicle might carry because of safety reasons. There is a certain probability that a certain branch requires a visit depending on the amount of money it handles. In the same way there is a similar problem, when the bank wishes to deliver money to the automatic teller machines that are located in several locations in each area. In both cases the problem of designing the routes can very well be modeled as a PVRP.

Similarly, the distribution of packages from a post office can be modeled as a PVRP, where the probability that a certain building requires a visit is given and the capacity $Q$ corresponds to the physical constraint that a mailman can carry only a fixed weight. Other areas of application include transportation and strategic planning.

The scientific literature concerning the VRP has been expanding at a very rapid pace, see for example the three excellent review volumes on the traveling salesman problem [11], on routing and scheduling [3] and on vehicle routing [5], each of which offers several hundreds of references. Except for an isolated result in the 1970's ([14]), VRPs with stochastic elements in their definitions have received attention only recently. Stewart and Golden [13], Dror and Trudeau [4], Laporte and Louveau [9] and Laporte et al. [10] use techniques from stochastic programming to solve optimally small problems and find bounds for the problems, the definitions of which are different from the ones we are considering in this paper.

The idea of using an a priori tour for the solution of traveling salesman problems when instances are modified probabilistically was first introduced in the Ph.D thesis of Jaillet [7]. This idea was generalized to other combinatorial optimization problems in the Ph.D thesis of the author [2], in which the probabilistic minimum spanning tree, the probabilistic traveling salesman problem, the probabilistic vehicle routing problem and facility location problems were analyzed.

The paper is organized as follows. In section 1 we address the question of finding closed-form expressions and algorithms to compute the expected length of an a priori route under various probabilistic assumptions. In section 2 we examine some combinatorial properties of the problem by proving bounds for the PVRP and the VRP re-optimization strategy, in which we find the optimal route at every instance. In section 3 we exploit the bounds derived in section 2 and propose some heuristics with provable worst-case performance. In section 4 we perform a probabilistic analysis for the case that customer locations are random in the unit square, prove some sharp asymptotic theorems for the PVRP and the VRP re-optimization strategy and also propose some asymptotically optimal algorithms. In the final section we summarize the contributions of the paper.

## 1 The Expected Length of an a Priori Route

The PVRP defines an efficient strategy for updating vehicle routes when problem instances are modified probabilistically in response to customers
not having any demand. Given an a priori route $\tau$ we define $L_{\tau}^{i}(S) \triangleq$ the length of the route which will result under strategy $i=a, b$ if only customers in the set $S$ have a unit demand. For example in Figure $1, S=\{2,3,5\}$ and $L_{\tau}^{a}(S), L_{\tau}^{b}(S)$ are the lengths of the routes shown in Figure 1(i), 1(ii) respectively.

Then if $p(S)$ is the probability that only customers in the set $S$ have a unit demand, the problem can be defined formally as follows:

## Problem definition:

Given a complete graph $G=(V, E)$, a cost function $d: E \rightarrow R$, an integer capacity $Q$ and a probability function $p: 2^{V} \rightarrow[0,1]$, we want to find a route $\tau_{i}$ that minimizes the expected length $E_{i}\left[L_{\tau}\right]$ :

$$
E_{i}\left[L_{\tau}\right]=\sum_{S \subseteq V} p(S) L_{\tau}^{i}(S)
$$

where the summation is taken over all subsets of $V$, the instances of the problem. Note that at this level of generality we can model dependencies among the probabilities of zero demand of sets of customers.

An alternative strategy would be the re-optimization strategy $\Sigma_{V R}$, in which we find the minimum length route of the set of customers with nonzero (unit) demand in every instance. Let $R(S)$ be the length of the optimal route when only customers in the set $S$ have a unit demand. We then define the expectation of this re-optimization strategy as follows:

$$
E\left[\Sigma_{V R}\right] \triangleq \sum_{S \subseteq V} p(S) R(S)
$$

The probabilistic traveling salesman problem (PTSP) defined in [7] and further explored in [2] provides a related strategy for the problem. In the
context of the PVRP, the PTSP can be viewed as a special case of the PVRP under strategy $b$, for which the capacity $Q$ is equal to $n$, i.e. the capacity of the vehicle is not a binding constraint. Similarly, the usual TSP can be viewed as a special case of the PVRP under strategy $a$, if the capacity $Q=$ $n$. Related to the vehicle routing re-optimization strategy is the traveling salesman re-optimization strategy, in which we find the optimum traveling salesman tour at every instance. We denote with $E\left[\Sigma_{T S P}\right]$ the expectation of the TSP re-optimization strategy, defined completely analogously with the vehicle routing re-optimization strategy.

Our initial goal then is to compute $E_{i}\left[L_{\tau}\right], i=a, b$ efficiently for a given a priori route $\tau$.

Let $p_{i}$ be the probability that customer $i$ has demand of one unit and $1-p_{i}$ of not having any demand independently of any other customer. Then we can compute the expected length of an a priori route as follows:

## Theorem 1

If the a priori route is $\tau=(0,1, \ldots, n, n+1 \triangleq 0)$ then

$$
\begin{gather*}
E_{a}\left[L_{r}\right]=\sum_{i=0}^{n} d(i, i+1)+\sum_{i=1}^{n} \gamma_{i} s(i, i+1)  \tag{1}\\
E_{b}\left[L_{\tau}\right]=\sum_{i=1}^{n} d(0, i) p_{i} \prod_{r=1}^{i-1}\left(1-p_{r}\right)+\sum_{i=1}^{n} d(i, 0) p_{i} \prod_{r=i+1}^{n}\left(1-p_{r}\right)+ \\
\sum_{i=1}^{n} \sum_{j=i+1}^{n} d(i, j) p_{i} p_{j} \prod_{r=i+1}^{j-1}\left(1-p_{r}\right)+\sum_{i=1}^{n} \sum_{j=i+1}^{n} s(i, j) \gamma_{i} p_{j} \prod_{r=i+1}^{j-1}\left(1-p_{r}\right) \tag{2}
\end{gather*}
$$

where

$$
s(i, j) \triangleq d(i, 0)+d(0, j)-d(i, j)
$$

$$
\begin{gather*}
\gamma_{i}=0, \quad i=0, \ldots Q-1 \\
\gamma_{r Q+i}=p_{r Q+i} \sum_{k=1}^{r} f(r Q+i-1, k Q-1), \quad 0 \leq i<Q \tag{3}
\end{gather*}
$$

and $f(m, r) \triangleq \operatorname{Pr}$ \{exactly $r$ customers among the customers $1, \ldots, m$ have non-zero demand $\}$ are computed from the recursion

$$
\begin{equation*}
f(m, r)=p_{m} f(m-1, r-1)+\left(1-p_{m}\right) f(m-1, r), \tag{4}
\end{equation*}
$$

with the initial conditions

$$
f(m, m)=\prod_{i=1}^{m} p_{i}, \quad f(m, 0)=\prod_{i=1}^{m}\left(1-p_{i}\right)
$$

Proof:
Consider first strategy $a$. The expected length of the route is a summation of the length of the a priori route plus the expected value of the extra distance when the vehicle reaches its capacity. To evaluate this second term, let $i$ be a node on the route where the vehicle reaches its capacity. The vehicle will then go to the depot before going back to the following node in the route, which is $i+1$ under strategy $a$, even if node $i+1$ has no demand. The extra distance traveled is then $s(i, i+1)=d(i, 0)+d(0, i+1)-d(i, i+1)$.

In (1) $\gamma_{i}$ is the probability that the vehicle reaches its capacity $Q$ at node i. Clearly $\gamma_{i}=0$ for $i=0, \ldots, Q-1$. Consider now node $r Q+i$. Then in order for the vehicle to reach its capacity at node $r Q+i(0 \leq i<Q)$, node $r Q+i$ must have a unit demand and from the previous $r Q+i-1$ nodes exactly $k Q-1$ must be present for some $k=1, \ldots r$, so that with the addition of node $r Q+i$ the capacity is reached. From this observation
(3) follows. The probabilities $f(m, r)$ are computed recursively from (4) by conditioning on the event that node $m$ has a demand.

Under strategy $b$, the first three terms in (2) are simply the expected length of the tour $\tau$ in the PTSP sense. The fourth term is identical with strategy $a$, except that when the vehicle reaches its capacity at node $i$, it goes back, after a visit to the depot, to the first node $j$ with a non-zero demand, skipping nodes $i+1, i+2, \ldots, j-1$ with no demand. $\bullet$

As an application of expressions (1) and (2) we find the closed-form expressions derived in [8] for the case in which all points have the same probability $p$ of requiring a visit. Then expressions (4) imply that $f(m, r)=$ $\binom{m}{r} p^{r}(1-p)^{m-r}$, and thus

$$
\begin{gather*}
E_{a}\left[L_{\tau}\right]=L_{\tau}+\sum_{i=1}^{n} s(i, i+1) \sum_{k=1}^{\left\lfloor\frac{i}{\square}\right\rfloor}\binom{i-1}{k Q-1} p^{k Q}(1-p)^{i-k Q}  \tag{5}\\
E_{b}\left[L_{\tau}\right]=E\left[L_{\tau}\right]+\sum_{i=1}^{n} \sum_{k=1}^{\left\lfloor\frac{i}{6}\right\rfloor}\binom{i-1}{k Q-1} p^{k Q}(1-p)^{i-k Q} \sum_{j=i+1}^{n} s(i, j) p(1-p)^{j-i-1} \tag{6}
\end{gather*}
$$

An important consequence of expressions (1), (2) is that they provide an algorithm of $\mathrm{O}\left(n^{2}\right)$ to compute $E_{a}\left[L_{\tau}\right], E_{b}\left[L_{\tau}\right]$ for the general case of unequal probabilities, because the computation of the probabilities $f(m, r)$ can be done recursively from (4) in $\mathrm{O}\left(n^{2}\right)$, and there are $n-Q$ non-zero probabilities $\gamma_{i}$. The computation of each one of these probabilities from (3) requires the evaluation of a sum of at most $\left\lceil\frac{n}{Q}\right\rceil$ terms. Thus we can compute all the $\gamma_{i}$ in $\mathrm{O}\left(\frac{(n-Q) n}{Q}+n^{2}\right)=O\left(n^{2}\right)$. Finally, the expectation of the length of the route, given that we have already computed the probabilities
$\gamma_{i}$ is done in $\mathrm{O}(n)$ for strategy $a$ and $\mathrm{O}\left(n^{2}\right)$ for strategy $b$, which means that we can compute the expected length of an a priori route in $\mathrm{O}\left(n^{2}\right)$ for both strategies.

In the next section we exploit the expressions found in this section for $E_{a}\left[L_{\tau}\right]$ and $E_{b}\left[L_{\tau}\right]$ to prove some interesting combinatorial properties of the PVRP.

## 2 Some Combinatorial Properties of the PVRP

Let $\tau_{a}, \tau_{b}$ be the optimal routes for strategies $a, b$ respectively of the PVRP and let also $\tau_{p}, \tau_{T S P}$ be the optimal tours for the PTSP and the TSP respectively. For $Q=n$, clearly $\tau_{a}=\tau_{T S P}, \tau_{b}=\tau_{p}$.

In this section we concentrate on understanding the relation among the expected lengths of the optimal solutions for the PVRP under strategies $a$, $b\left(E_{a}\left[L_{\tau_{a}}\right], E_{b}\left[L_{\tau_{b}}\right]\right)$, the expected length of the optimal tour for the PTSP ( $E\left[L_{\tau_{p}}\right]$ ), the length of the optimal deterministic tour $\left(L_{T S P}\right)$ and the expectation of the re-optimization strategies $E\left[\Sigma_{V R}\right], E\left[\Sigma_{T S P}\right]$.

### 2.1 Relation among the Different Strategies

The probabilistic strategies are useful in a routing context, mainly in the case where the triangle inequality holds. In this case we can find the following relation among the strategies.

Proposition 2

Under the triangle inequality

$$
\begin{equation*}
E\left[\Sigma_{V R}\right] \leq E_{b}\left[L_{\tau_{b}}\right] \leq E_{a}\left[L_{\tau_{a}}\right] \tag{7}
\end{equation*}
$$

Proof:
Consider an a priori route $\tau$. Then

$$
L_{\tau}^{b}(S) \leq L_{\tau}^{a}(S)
$$

because under strategy $b$ we skip customers with zero demand and thus because of the triangle inequality the length of the resulting route is smaller. Note that the breakpoints in the routes occur at the same customers under both strategies. As a result,

$$
E_{b}\left[L_{\tau}\right] \leq E_{a}\left[L_{\tau}\right]
$$

Consider now the tour $\tau_{a}$. The above inequality gives

$$
E_{b}\left[L_{\tau_{a}}\right] \leq E_{a}\left[L_{\tau_{a}}\right]
$$

But, because of the optimality of the route $\tau_{b}$ for strategy $b$,

$$
E_{b}\left[L_{\tau_{b}}\right] \leq E_{b}\left[L_{\tau_{a}}\right]
$$

from which the right inequality of (7) follows. Also, since $R(S) \leq L_{\tau_{b}}^{b}(S)$ in every instance the left inequality follows. $\bullet$

### 2.2 Lower Bounds

In this subsection we derive some lower bounds for the different strategies. For convenience we assume that the distance matrix is symmetric.


Figure 2: The optimal route at instance $S$

## Proposition 3

If the probability that customer $i$ has a unit demand is $p_{i}$, then under the triangle inequality

$$
\begin{equation*}
E\left[\Sigma_{V R}\right] \geq \max \left(\frac{2}{Q} \sum_{i=1}^{n} d(0, i) p_{i}, E\left[\Sigma_{T S P}\right]\right) \tag{8}
\end{equation*}
$$

Proof:
Consider an instance $S$ of the problem. Under the re-optimization strategy, a vehicle starts from the depot, visits a subset $X_{j} \subseteq S$ of customers $\left(\left|X_{j}\right| \leq\right.$ $Q)$, returns to the depot and then continues to the next subset $X_{j+1}$. See also Figure 2. Then, if $L_{j}$ is the length of the route for visiting the subset $X_{j}$ of customers in the optimal solution at instance $S$, we have

$$
L_{j} \geq 2 \max _{i \in X_{j}} d(0, i) \geq 2 \frac{\sum_{i \in X_{j}} d(0, i)}{\left|X_{j}\right|} \geq 2 \frac{\sum_{i \in X_{j}} d(0, i)}{Q}
$$

As a result,

$$
R(S)=\sum_{j} L_{j} \geq 2 \frac{\sum_{i \in S} d(0, i)}{Q}
$$

Consequently,

$$
E\left[\Sigma_{V R}\right] \geq \frac{2}{Q} \sum_{S \subseteq V} p(S) \sum_{i \in S} d(0, i)=\frac{2}{Q} \sum_{i=1}^{n} d(0, i) \sum_{S, i \in S} p(S)=\frac{2}{Q} \sum_{i=1}^{n} d(0, i) p_{i}
$$

since $\sum_{S, i \in S} p(S)=p_{i}$. Also, because of triangle inequality $R(S) \geq L_{T S P}(S)$ and thus $E\left[\Sigma_{V R}\right] \geq E\left[\Sigma_{T S P}\right]$. •

We then use (7) and (8) to find some lower bounds on $E_{a}\left[L_{\tau_{a}}\right]$, and $E_{b}\left[L_{\tau_{b}}\right]$.

Proposition 4
If the probability that customer $i$ has a unit demand is $p_{i}$, then under the triangle inequality

$$
\begin{align*}
& E_{a}\left[L_{\tau_{a}}\right] \geq \max \left(\frac{2}{Q} \sum_{i=1}^{n} d(0, i) p_{i}, L_{T S P}\right)  \tag{9}\\
& E_{b}\left[L_{\tau_{b}}\right] \geq \max \left(\frac{2}{Q} \sum_{i=1}^{n} d(0, i) p_{i}, E\left[L_{\tau_{p}}\right]\right) \tag{10}
\end{align*}
$$

Proof:
From the triangle inequality $s(i, i+1) \geq 0$. Therefore, from (1) $E_{a}\left[L_{\tau_{a}}\right] \geq$ $L_{\tau_{a}} \geq L_{T S P}$. From (7) and (8), (9) follows.
Similarly, $E_{b}\left[L_{\tau_{b}}\right] \geq E\left[L_{\tau_{b}}\right] \geq E\left[L_{\tau_{p}}\right]$, and hence from (7) and (8), (10) follows.

In propositions 3 and 4, if the distance matrix is asymmetric, then we should replace the term $\frac{2}{Q} \sum_{i=1}^{n} d(0, i) p_{i}$ in (8), (9) and (10) with $\frac{1}{Q} \sum_{i=1}^{n}[d(0, i)+$ $d(i, 0)] p_{i}$.

If we do not assume the triangle inequality, we can still find lower bounds on the $E_{a}\left[L_{\tau_{a}}\right], E_{b}\left[L_{\tau_{b}}\right]$.

## Proposition 5

If $\tau_{a}, \tau_{b}$ are the optimal routes for strategies $a, b$ respectively, then

$$
\begin{equation*}
E_{a}\left[L_{\tau_{a}}\right] \geq z_{a}^{*} \tag{11}
\end{equation*}
$$

where $z_{a}^{*}$ is the optimal solution to the transportation problem:

$$
\begin{gather*}
z_{a}^{*}=\min \sum_{i, j} x_{i, j} d(i, j) \\
\sum_{i=0}^{n} x_{i, 0}=1+\sum_{i=1}^{n} \gamma_{i}, \quad \sum_{i=0}^{n} x_{i, j}=1, \quad j \geq 1 \\
\sum_{j=0}^{n} x_{0, j}=1+\sum_{i=1}^{n} \gamma_{i}, \quad \sum_{j=0}^{n} x_{i, j}=1, \quad i \geq 1  \tag{12}\\
x_{i, j} \geq 0 .
\end{gather*}
$$

Similarly,

$$
\begin{equation*}
E_{b}\left[L_{r_{b}}\right] \geq z_{b}^{*} \tag{13}
\end{equation*}
$$

where $z_{b}^{*}$ is the optimal solution to the transportation problem:

$$
\begin{gather*}
z_{b}^{*}=\min \sum_{i, j} x_{i, j} d(i, j) \\
\sum_{i=0}^{n} x_{i, 0}=1+\sum_{i=1}^{n} \gamma_{i}\left(1-\prod_{r=i+1}^{n}\left(1-p_{r}\right)\right), \quad \sum_{i=0}^{n} x_{i, j}=p_{j}, \quad j \geq 1 \\
\sum_{j=0}^{n} x_{0, j}=1+\sum_{i=1}^{n} \gamma_{i}\left(1-\prod_{r=i+1}^{n}\left(1-p_{r}\right)\right), \quad \sum_{j=0}^{n} x_{i, j}=p_{i}, \quad i \geq 1  \tag{14}\\
x_{i, j} \geq 0 .
\end{gather*}
$$

The numbers $\gamma_{i}$ appearing in (12) and (14) are computed from (3).
Proof:

Let $d_{1}(i, j) \triangleq d(i, j)-u_{i}-v_{j}$. By renumbering the nodes let $\tau_{a}=(0,1,2, \ldots, n, 0)$ be the optimum route under strategy $a$ with respect to the original distances $d(i, j)$. Let $E_{a}\left[L_{\tau_{a}}\right], E_{a}\left[L_{\tau_{a}}^{1}\right]$ be the expected lengths of the route $\tau_{a}$ with respect to distances $d(i, j), d_{1}(i, j)$ respectively. Then

$$
E_{a}\left[L_{\tau_{a}}^{1}\right]=E_{a}\left[L_{\tau_{a}}\right]-\sum_{i=1}^{n}\left(u_{i}+v_{i}\right)-\left(u_{0}+v_{0}\right)\left(1+\sum_{i=1}^{n} \gamma_{i}\right)
$$

If we demand that $d_{1}(i, j)=d(i, j)-u_{i}-v_{j} \geq 0$ we have the trivial bound $E_{a}\left[L_{\tau_{a}}^{1}\right] \geq 0$. As a result, we obtain

$$
E_{a}\left[L_{\tau_{a}}\right] \geq \sum_{i=1}^{n}\left(u_{i}+v_{i}\right)+\left(u_{0}+v_{0}\right)\left(1+\sum_{i=1}^{n} \gamma_{i}\right)
$$

Since we want the best possible bound, we wish to choose $u_{i}, v_{i}$ satisfying:

$$
\begin{gather*}
\max \sum_{i=1}^{n}\left(u_{i}+v_{i}\right)+\left(u_{0}+v_{0}\right)\left(1+\sum_{i=1}^{n} \gamma_{i}\right), \\
d(i, j)-u_{i}-v_{j} \geq 0 \tag{15}
\end{gather*}
$$

The dual of the linear program (15) is (12) and hence (11) follows from strong duality. Using the same ideas the bound (13) follows. $\bullet$

### 2.3 Upper Bounds

For the upper bounds we concentrate on the case $p_{i}=p$ and for convenience we assume that the distance matrix is symmetric. Consider the following heuristic.

## Cyclic Heuristic

1. Given an initial route $\tau \triangleq \tau_{1}=(0,1,2, \ldots, n, 0)$, consider the routes $\tau_{i}=(0, i, \ldots, n, 1, \ldots, i-1,0), i=2, \ldots n$.
2. Compute $E_{a}\left[L_{\tau_{i}}\right]$ for all $i=1, \ldots, n$,
3. The route with the minimum expectation is the proposed solution $\tau_{H}$ to the PVRP under strategy $a$.

## Proposition 6

Let the probability that customer $i$ has a unit demand be $p$. If the initial route to the cyclic heuristic is the optimal deterministic tour and $\tau_{H}$ is the tour proposed by the heuristic, then under the triangle inequality

$$
\begin{equation*}
E_{a}\left[L_{\tau_{a}}\right] \leq E_{a}\left[L_{\tau_{H}}\right] \leq L_{T S P}\left(1-\frac{2}{n}-\frac{p}{Q}\right)+2\left(2+\frac{n p}{Q}\right) \frac{\sum_{i=1}^{n} d(0, i)}{n} \tag{16}
\end{equation*}
$$

Proof:
If the initial route to the cyclic heuristic is the optimal deterministic tour, then let
$L \triangleq \sum_{i=1}^{n-1} d(i, i+1)+d(n, 1)$. With this definition the lengths of the routes $\tau_{i}$ become:

$$
\begin{gathered}
L_{\tau_{1}}=L+d(0,1)+d(0, n)-d(1, n) \\
L_{\tau_{i}}=L+d(0, i)+d(0, i-1)-d(i, i-1), \quad i=2, \ldots n
\end{gathered}
$$

As a result, $\sum_{i=1}^{n} L_{\tau_{i}}=2 \sum_{i=1}^{n} d(0, i)+(n-1) L$. Clearly,

$$
\begin{aligned}
& E_{a}\left[L_{\tau_{a}}\right] \leq E_{a}\left[L_{\tau_{H}}\right] \leq \frac{1}{n} \sum_{i=1}^{n} E_{a}\left[L_{\tau_{i}}\right]= \\
& \frac{1}{n}\left[\sum_{i=1}^{n} L_{\tau_{i}}+\left(\sum_{i=1}^{n} \gamma_{i}\right)\left(\sum_{i=1}^{n} s(i, i+1)\right)\right]
\end{aligned}
$$

since the probability $\gamma_{i}$ multiplies every term $s(i, i+1)$ in $\sum_{i=1}^{n} E_{a}\left[L_{\tau_{i}}\right]$. Therefore,

$$
E_{a}\left[L_{\tau_{a}}\right] \leq \frac{1}{n}\left[2 \sum_{i=1}^{n} d(0, i)+(n-1) L+\left(\sum_{i=1}^{n} \gamma_{i}\right)\left(\sum_{i=1}^{n} d(0, i)+d(0, i+1)-d(i, i+1)\right)\right]=
$$

$$
\begin{equation*}
\frac{1}{n}\left[2 \sum_{i=1}^{n} d(0, i)+(n-1) L+\left(\sum_{i=1}^{n} \gamma_{i}\right)\left(2 \sum_{i=1}^{n} d(0, i)-L\right)\right] \tag{17}
\end{equation*}
$$

Our goal is then to find an upper bound for the term $\sum_{i=1}^{n} \gamma_{i}$. We define the random variable
$N \triangleq$ the number of breakpoints in the route, where the capacity is reached.

## Let also

$X_{i} \triangleq$ the indicator random variable taking the value 1 if a breakpoint occurs at customer $i$ and 0 otherwise. With these definitions, $N=\sum_{i=1}^{n} X_{i}$. But since $\operatorname{Pr}\left\{X_{i}=1\right\}=\gamma_{i}$ then

$$
E[N]=\sum_{i=1}^{n} E\left[X_{i}\right]=\sum_{i=1}^{n} \gamma_{i}
$$

If $W \triangleq$ the number of customers with non-zero demand then

$$
E[N]=E\left[\left\lceil\frac{W}{Q}\right\rceil\right] \leq E\left[\frac{W}{Q}+1\right] \leq \frac{n p}{Q}+1
$$

and hence,

$$
\begin{equation*}
\sum_{i=1}^{n} \gamma_{i} \leq \frac{n p}{Q}+1 \tag{18}
\end{equation*}
$$

Using (18) in (17) and since $L \leq L_{T S P}$, (16) follows
We now prove an upper bound for strategy $b$.

## Proposition 7

If the probability that customer $i$ has a unit demand is $p_{i}$, then under the triangle inequality

$$
\begin{equation*}
E_{b}\left[L_{\tau_{b}}\right] \leq E_{b}\left[L_{\tau_{p}}\right] \leq E\left[L_{\tau_{p}}\right]+2 \sum_{i=1}^{n} d(0, i) p_{i} \tag{19}
\end{equation*}
$$

Proof:
Consider the optimal tour for the PTSP $\tau_{p}=(0,1, \ldots, n, 0)$ as a solution
to the PVRP under strategy $b$. Then,

$$
E_{b}\left[L_{\tau_{b}}\right] \leq E_{b}\left[L_{\tau_{p}}\right]=E\left[L_{\tau_{p}}\right]+\sum_{i=1}^{n} \sum_{j=i+1}^{n} \gamma_{i} p_{j} \prod_{r=i+1}^{j-1}\left(1-p_{r}\right) s(i, j)
$$

Since $s(i, j)=d(i, 0)+d(0, j)-d(i, j) \leq 2 d(0, i)$ from the triangle inequality and $\sum_{j=i+1}^{n} p_{j} \prod_{r=i+1}^{j-1}\left(1-p_{r}\right)=\left(1-\prod_{r=i+1}^{n}\left(1-p_{r}\right)\right)$,

$$
E_{b}\left[L_{\tau_{b}}\right] \leq E\left[L_{\tau_{p}}\right]+2 \sum_{i=1}^{n} \gamma_{i} d(0, i) \leq E\left[L_{\tau_{p}}\right]+2 \sum_{i=1}^{n} p_{i} d(0, i)
$$

because $\gamma_{i}=p_{i} \operatorname{Pr}\left\{W_{i-1}\right.$ is a multiple of $\left.Q\right\} \leq p_{i}$, where $W_{i-1}$ is the number of non-zero demand customers among the customers $1, \ldots, i-1 . \bullet$
We finally find an upper bound on the re-optimization strategy $E\left[\Sigma_{V R}\right]$.

## Proposition 8

If the probability that customer $i$ has a unit demand is $p$, then under the triangle inequality

$$
\begin{equation*}
E\left[\Sigma_{V R}\right] \leq E\left[\Sigma_{T S P}\right]\left(1-\frac{1}{Q}\right)+2\left(\frac{1}{n}+\frac{p}{Q}\right) \sum_{i=1}^{n} d(0, i) \tag{20}
\end{equation*}
$$

Proof:
We use a result of [6] that

$$
R(S) \leq 2\left\lceil\frac{|S|}{Q}\right\rceil \frac{\sum_{i \in S} d(0, i)}{|S|}+L_{T S P}(S)\left(1-\frac{1}{Q}\right)
$$

As a result,

$$
\begin{aligned}
E\left[\Sigma_{V R}\right]= & \sum_{S} p(S) R(S) \leq E\left[\Sigma_{T S P}\right]\left(1-\frac{1}{Q}\right)+2 \sum_{S} p(S)\left\lceil\frac{|S|}{Q}\right\rceil \frac{\sum_{i \in S} d(0, i)}{|S|} \leq \\
& \leq E\left[\Sigma_{T S P}\right]\left(1-\frac{1}{Q}\right)+2 \sum_{S}\left(\frac{1}{Q}+\frac{1}{|S|}\right) \sum_{i \in S} d(0, i) p(S)
\end{aligned}
$$

But since $\sum_{S} \sum_{i \in S} d(0, i) p(S)=p \sum_{i=1}^{n} d(0, i)$ and

$$
\begin{gathered}
\sum_{S} \frac{1}{|S|} \sum_{i \in S} d(0, i) p(S)=\sum_{k=1}^{n} \frac{1}{k} \sum_{S,|S|=k} \sum_{i \in S} d(0, i) p(S)= \\
=\sum_{k=1}^{n} \frac{1}{k} p^{k}(1-p)^{n-k} \sum_{i=1}^{n} d(0, i)\binom{n-1}{k-1}=\frac{1}{n}\left(1-(1-p)^{n}\right) \sum_{i=1}^{n} d(0, i),
\end{gathered}
$$

(20) follows. •

From (8) and (20) we can establish a relation between the re-optimization strategies $E\left[\Sigma_{V R}\right]$ and $E\left[\Sigma_{T S P}\right]$.

## Theorem 9

If the probability that customer $i$ has a unit demand is $p$, then under the triangle inequality

$$
\begin{equation*}
1 \leq \frac{E\left[\Sigma_{V R}\right]}{E\left[\Sigma_{T S P}\right]} \leq 2-\frac{1}{Q}+\frac{Q}{n p}=2+O\left(\frac{1}{n}\right) \tag{21}
\end{equation*}
$$

## 3 Heuristics for the PVRP

In this section we exploit the bounds derived in the previous section to propose some heuristics with good worst-case performance. In section 2.3 we introduced the cyclic heuristic. In the following theorem we prove that the heuristic is within a constant factor from the optimal route under strategy $a$.

## Theorem 10

Let the probability that customer $i$ has a unit demand be $p$. Then if the initial route given to the cyclic heuristic is $\tau_{T S P}$ and the tour found by the
cyclic heuristic is $\tau_{H}$, then under the triangle inequality

$$
\begin{equation*}
\frac{E_{a}\left[L_{\tau_{H}}\right]}{E_{a}\left[L_{\tau_{a}}\right]} \leq 2-\frac{p}{Q}+\frac{1}{n}\left(\frac{2 Q}{p}-1\right)=2-\frac{p}{Q}+\mathrm{O}\left(\frac{1}{n}\right) . \tag{22}
\end{equation*}
$$

Proof:
From (16) and (9) we obtain

$$
\begin{aligned}
& \frac{E_{a}\left[L_{\tau_{H}}\right]}{E_{a}\left[L_{\tau_{a}}\right]} \leq \frac{L_{T S P}\left(1-\frac{2}{n}-\frac{p}{Q}\right)+\frac{2}{n}\left(2+\frac{n p}{Q}\right) \sum_{i=1}^{n} d(0, i)}{\max \left(\frac{2 p}{Q} \sum_{i=1}^{n} d(0, i), L_{T S P}\right)} \leq \\
& \leq 1-\frac{2}{n}-\frac{p}{Q}+\frac{Q}{n p}\left(2+\frac{n p}{Q}\right) \leq 2-\frac{p}{Q}+\frac{1}{n}\left(\frac{2 Q}{p}-2\right) .
\end{aligned}
$$

Theorem 10 says that the cyclic heuristic produces a solution to the PVRP under strategy $a$, which, for large enough $n$, is within a factor of $2-\frac{p}{Q}$ of the optimal route. If instead of the optimal deterministic tour we give to the cyclic heuristic the Christofides tour then the guarantee will be $\frac{5}{2}-\frac{p}{Q}+\frac{1}{n}\left(\frac{2 Q}{p}-2\right)$, and the running time of the combined heuristic (Christofides heuristic and then the cyclic heuristic) is $\mathrm{O}\left(n^{3}\right)$, since the Christofides heuristic takes $O\left(n^{3}\right)$ and the cyclic heuristic needs the evaluation of the expected length of $n$ routes, each of which takes $O\left(n^{2}\right)$. Therefore, this combined heuristic runs in polynomial time and produces solutions which are within a constant factor of the optimal route under strategy $a$.

The natural question is then to investigate if a constant-guarantee heuristic exists for the PVRP under strategy $b$.

Theorem 11
If the probability that customer $i$ has a unit demand is $p_{i}$, then under the
triangle inequality

$$
\begin{equation*}
\frac{E_{b}\left[L_{\tau_{p}}\right]}{E_{b}\left[L_{\tau_{b}}\right]} \leq Q+1 \tag{23}
\end{equation*}
$$

where $\tau_{p}$ is the optimal tour for the PTSP. Proof:
From (19) and (10) we have that

$$
\frac{E_{b}\left[L_{\tau_{p}}\right]}{E_{b}\left[L_{\tau_{b}}\right]} \leq \frac{E\left[L_{\tau_{p}}\right]+2 \sum_{i=1}^{n} d(0, i) p_{i}}{\max \left(\frac{2}{Q} \sum_{i=1}^{n} d(0, i) p_{i}, E\left[L_{\tau_{p}}\right]\right)} \leq Q+1 . \bullet
$$

Thus, for constant capacity $Q$ and even in the case of unequal probabilities $p_{i}$, the optimal tour for the PTSP (PTST) is within a constant factor of the optimal route for the PVRP under strategy $b$. For Euclidean problems, if instead of the optimal PTST, which we do not know how to compute in polynomial time, we use the spacefilling curve heuristic $\tau_{S F}$ in [12], which is found in $O(n \log n)$ time, then it was found in [2] (equation 3.30) that $E\left[L_{\tau_{S F}}\right] / E\left[L_{\tau_{p}}\right]=\mathrm{O}(\log n)$. Thus, the following guarantee follows immediately from (19):

$$
\begin{equation*}
\frac{E_{b}\left[L_{\tau_{S F}}\right]}{E_{b}\left[L_{\tau_{b}}\right]} \leq \frac{E\left[L_{\tau_{S F}}\right]+2 \sum_{i=1}^{n} d(0, i) p_{i}}{\max \left(\frac{2}{Q} \sum_{i=1}^{n} d(0, i) p_{i}, E\left[L_{\tau_{p}}\right]\right)} \leq Q+\mathrm{O}(\log n) \tag{24}
\end{equation*}
$$

Platzman and Bartholdi [12] conjecture that the spacefilling curve heuristic produces a tour $\tau_{S F}$, which is within a constant factor of the optimal TST. Bertsimas [2] proves that this conjecture implies that the expected length of $\tau_{S F}$ is within a constant factor of the optimal PTST and thus,

$$
\frac{E_{b}\left[L_{\tau_{S F}}\right]}{E_{b}\left[L_{\tau_{b}}\right]} \leq Q+\mathrm{O}(1)
$$

i.e. there exists a constant-guarantee heuristic for the PVRP under strategy $b$.

## 4 Asymptotic Theorems for the PVRP in the Random Euclidean Model

In this section we investigate under the random Euclidean model the asymptotic behavior of the PVRP and of the re-optimization strategy.

Let $X_{1}, X_{2}, \ldots$ be a sequence of independent, identically distributed random points in the unit square. Let $E[r]$ be the expected distance from the origin and let $X^{(n)}$ denote the first $n$ points of the sequence. We denote with

$$
\begin{gathered}
E\left[\Sigma_{V R}^{n}\right] \triangleq E\left[\Sigma_{V R}\left(X^{(n)}\right)\right] \\
E_{a}\left[L_{\tau_{a}}^{n}\right] \triangleq E_{a}\left[L_{\tau_{a}}\left(X^{(n)}\right)\right], \quad E_{b}\left[L_{\tau_{b}}^{n}\right] \triangleq E_{b}\left[L_{\tau_{b}}\left(X^{(n)}\right)\right]
\end{gathered}
$$

Observe that the quantities we just defined are random variables since the locations of the customers are random.

For the PTSP the following results are known. Fix the probability $p$ of unit demand. Let $E\left[\Sigma_{T S P}^{n}\right] \triangleq E\left[\Sigma_{T S P}\left(X^{(n)}\right)\right]$, and $E\left[L_{\tau_{p}}^{n}\right] \triangleq E\left[L_{\tau_{p}}\left(X^{(n)}\right)\right]$. Then with probability 1

$$
\begin{align*}
& \lim _{n \rightarrow \infty} \frac{E\left[\Sigma_{T S P}^{n}\right]}{\sqrt{n}}=\beta_{T S P} \sqrt{p}, \quad \text { Jaillet }  \tag{25}\\
& \left.\lim _{n \rightarrow \infty} \frac{E[7]}{\sqrt{n}}=L_{T P}^{n}\right]  \tag{26}\\
& \sqrt{n}
\end{align*}
$$

where $\beta_{\text {TSP }}$ is the constant appearing in the celebrated Beardwood et. al. [1] theorem.

We first investigate the asymptotic behavior of $E\left[\Sigma_{V R}^{n}\right]$ and then establish the asymptotic behavior of $E_{a}\left[L_{\tau_{a}}^{n}\right], E_{b}\left[L_{\tau_{b}}^{n}\right]$.

### 4.1 The Asymptotic Behavior of the Re-optimization Strategy

The asymptotic behavior of $E\left[\Sigma_{V R}^{n}\right]$ depends critically on the dependence of the capacity $Q$ on the number of customers $n$. This dependence is also critical for the asymptotic behavior of the VRP examined in [6]. Let $Q_{n}$ denote the capacity of the vehicle to indicate its dependence on $n$. We prove the following theorem:

## Theorem 12

Fix the probability $p$ of non-zero demand.

1. If $Q$ is a constant, then almost surely

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{E\left[\Sigma_{V R}^{n}\right]}{n}=\frac{2 E[r] p}{Q} \tag{27}
\end{equation*}
$$

2. If $\lim _{n \rightarrow \infty} \frac{Q_{n}}{\sqrt{n}}=0$, then almost surely

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{Q_{n} E\left[\Sigma_{V R}^{n}\right]}{n}=2 E[r] p \tag{28}
\end{equation*}
$$

3. If $\lim _{n \rightarrow \infty} \frac{Q_{n}}{\sqrt{n}}=\infty$, then almost surely

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{E\left[\Sigma_{V R}^{n}\right]}{\sqrt{n}}=\beta_{T S P} \sqrt{p} \tag{29}
\end{equation*}
$$

Proof:
We define $\bar{r} \triangleq \frac{\sum_{i=1}^{n} d(0, i)}{n}$.

1. Assume that $Q$ is constant. From (8) and (20)

$$
\frac{2 p n}{Q} \bar{r} \leq E\left[\Sigma_{V R}^{n}\right] \leq E\left[\Sigma_{T S P}^{n}\right]\left(1-\frac{1}{Q}\right)+\frac{2 p n \bar{r}}{Q}\left(1+\frac{Q}{n p}\right)
$$

But $E[r]<\infty$ implies that $\bar{r} \rightarrow E[r]$ almost surely by the strong law of large numbers. Furthermore, from (25) $\lim _{n \rightarrow \infty} \frac{E\left[\sum_{\left.n_{S P P}^{n}\right]}^{\sqrt{n}}\right.}{}=\beta_{T S P} \sqrt{p}$, i.e. $\lim _{n \rightarrow \infty} \frac{E\left[\sum_{S T S}^{n}\right]}{n}=0$ almost surely. As a result,

$$
\frac{2 p \bar{r}}{Q} \leq \frac{E\left[\Sigma_{V R}^{n}\right]}{n} \leq \frac{E\left[\Sigma_{T S P}^{n}\right]}{n}\left(1-\frac{1}{Q}\right)+\frac{2 p \bar{r}}{Q}\left(1+\frac{Q}{n p}\right)
$$

from which, by taking limits, we obtain (27).
2. Similarly,

$$
\begin{equation*}
2 p \bar{r} \leq \frac{Q_{n}}{n} E\left[\Sigma_{V R}^{n}\right] \leq \frac{Q_{n} E\left[\Sigma_{T S P}^{n}\right]}{n}+2 p \bar{r}\left(1+\frac{Q_{n}}{n p}\right) \tag{30}
\end{equation*}
$$

Since with probability $1, \lim _{n \rightarrow \infty} \frac{Q_{n} E\left[\Sigma_{n S P}^{n}\right]}{n}=\lim _{n \rightarrow \infty} \frac{Q_{n}}{\sqrt{n}} \frac{E\left[\Sigma_{n}^{n}\right.}{\sqrt{n}}=0$ and $\lim _{n \rightarrow \infty} \frac{Q_{n}}{n p}=0$, the result follows by taking limits in (30).
3. From (8)

$$
\frac{E\left[\Sigma_{T S P}^{n}\right]}{\sqrt{n}} \leq \frac{E\left[\Sigma_{V R}^{n}\right]}{\sqrt{n}} \leq \frac{2 p \bar{r}}{Q_{n} / \sqrt{n}}\left(1+\frac{Q_{n}}{n p}\right)+\frac{E\left[\Sigma_{T S P}^{n}\right]}{\sqrt{n}} .
$$

From (25) $\lim _{n \rightarrow \infty} \frac{E\left[\sum_{n}^{n} S_{p}\right]}{\sqrt{n}}=\beta_{T S P} \sqrt{p}$ almost surely and also

$$
\lim _{n \rightarrow \infty} \frac{2 p \bar{r}}{Q_{n} / \sqrt{n}}\left(1+\frac{Q_{n}}{n p}\right)=\lim _{n \rightarrow \infty} 2 p E[r]\left(\frac{1}{Q_{n} / \sqrt{n}}+\frac{1}{p \sqrt{n}}\right)=0
$$

from which (29) follows.e
The case $Q_{n}=\Theta(\sqrt{n})$ is not covered in theorem 12. The reason is that in this case neither of the two terms; the radial collection term, $\frac{2 p n \bar{r}}{Q_{n}}(1+$ $\frac{Q_{n}}{n p}$, and the local collection term, $E\left[\Sigma_{T S P}^{n}\right]$, dominates as was the case in theorem 12. In cases 1 and 2 the radial collection term dominated and in case 3 the local collection term dominated.

### 4.2 Asymptotic Theorems for the PVRP under Strat-

 egy $a$As in the previous subsection the dependence of the capacity $Q$ on the number of customers $n$ is critical. We assume again that each customer has the same probability $p$ of a unit demand. In the following theorem we examine the asymptotic behavior of the strategy $a$ and propose asymptotically optimal heuristics.

Theorem 13
Fix the probability $p$ of non-zero demand.

1. If $Q$ is a constant, then almost surely

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{E_{a}\left[L_{r_{a}}^{n}\right]}{n}=\frac{2 E[r] p}{Q} \tag{31}
\end{equation*}
$$

2. If $\lim _{n \rightarrow \infty} \frac{Q_{n}}{\sqrt{n}}=0$, then almost surely

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{Q_{n} E_{a}\left[L_{\tau_{a}}^{n}\right]}{n}=2 E[r] p \tag{32}
\end{equation*}
$$

In both cases 1, 2 any tour produced by the cyclic heuristic with initial tour any tour of length $\mathrm{O}(\sqrt{n})$ is asymptotically optimal.
3. If $\lim _{n \rightarrow \infty} \frac{Q_{n}}{\sqrt{n}}=\infty$, then almost surely

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{E_{a}\left[L_{\tau_{a}}^{n}\right]}{\sqrt{n}}=\beta_{T S P} \tag{33}
\end{equation*}
$$

The route produced by the cyclic heuristic with initial tour the optimal TST is asymptotically optimal.

Proof:

1. Assume that $Q$ is constant. From (9) and (16)

$$
\frac{2 p}{Q} \bar{r} \leq \frac{E_{a}\left[L_{r_{a}}^{n}\right]}{n} \leq \frac{L_{T S P}^{n}}{n}+\frac{2 p \bar{r}}{Q}+\frac{4 \bar{r}}{n}
$$

Since $\bar{r} \rightarrow E[r]$ almost surely as $n \rightarrow \infty$ and $\lim _{n \rightarrow \infty} \frac{L_{n S P}^{n}}{n}=0$, (31) follows.
2. Similarly,

$$
2 p \bar{r} \leq \frac{Q_{n}}{n} E_{a}\left[L_{\tau_{a}}^{n}\right] \leq \frac{Q_{n}}{\sqrt{n}} \frac{L_{T S P}^{n}}{\sqrt{n}}+2 p \bar{r}+\frac{4 \bar{r} Q_{n}}{n}
$$

Since with probability $1, \lim _{n \rightarrow \infty} \frac{Q_{n}}{\sqrt{n}} \frac{L_{n s p}^{n}}{\sqrt{n}}=0$, (32) follows easily.
Since in both cases 1,2 the radial collection term dominates asymptotically, if instead of the TST we used in the cyclic heuristic any tour $\tau_{0}$ with length $\mathrm{O}(\sqrt{n})$ (for example a tour produced by the strip method), then the tour produced by the cyclic heuristic would be asymptotically optimal.
3. From (9)

$$
\frac{L_{T S P}^{n}}{\sqrt{n}} \leq \frac{E_{a}\left[L_{\tau_{a}}^{n}\right]}{\sqrt{n}} \leq \frac{2 p \bar{r}}{Q_{n} / \sqrt{n}}+\frac{4 \bar{r}}{\sqrt{n}}+\frac{L_{T S P}^{n}}{\sqrt{n}}
$$

Since $\lim _{n \rightarrow \infty} \frac{L_{V S P}^{n}}{\sqrt{n}}=\beta_{T S P}$ almost surely and also $\lim _{n \rightarrow \infty} \frac{Q_{n}}{\sqrt{n}}=\infty$ (33) follows.

In this case if the initial tour to the cyclic heuristic is the TST, then the tour produced by the heuristic is asymptotically optimal.

In theorem 13 we characterized very sharply the asymptotic behavior of the PVRP under strategy $a$ and furthermore we proposed asymptotically optimal heuristics. We next consider the PVRP under strategy b.

### 4.3 Asymptotic Theorems for the PVRP under Strategy $b$

We assume again that each customer has the same probability $p$ of a unit demand. We then characterize the asymptotic behavior of strategy $b$.

## Theorem 14

Fix the probability $p$ of non-zero demand.

1. If $Q$ is a constant, then almost surely

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{E_{b}\left[L_{\tau_{b}}^{n}\right]}{n}=\frac{2 E[r] p}{Q} \tag{34}
\end{equation*}
$$

2. If $\lim _{n \rightarrow \infty} \frac{Q_{n}}{\sqrt{n}}=0$, then almost surely

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{Q_{n} E_{b}\left[L_{\tau_{b}}^{n}\right]}{n}=2 E[r] p \tag{35}
\end{equation*}
$$

3. If $\lim _{n \rightarrow \infty} \frac{Q_{n}}{\sqrt{n}}=\infty$, then almost surely

$$
\begin{equation*}
\lim _{n \rightarrow \infty} \frac{E_{b}\left[L_{\tau_{b}}^{n}\right]}{\sqrt{n}}=\beta_{T S P} \sqrt{p} \tag{36}
\end{equation*}
$$

Proof:
Cases 1 and 2 follow from (7) since

$$
\frac{E\left[\Sigma_{V R}^{n}\right]}{n} \leq \frac{E_{b}\left[L_{\tau_{b}}^{n}\right]}{n} \leq \frac{E_{a}\left[L_{\tau_{a}}^{n}\right]}{n}
$$

and hence (34) and (35) follow, because in theorems 12 and 13 both the lower and the upper bounds converge to the same limit almost surely.

For the third case, from (10), (19) we find

$$
\begin{gathered}
E\left[L_{\tau_{p}}^{n}\right] \leq E_{b}\left[L_{\tau_{b}}^{n}\right] \leq E\left[L_{\tau_{p}}^{n}\right]+2 \sum_{i=1}^{n} \gamma_{i} d(0, i) \leq E\left[L_{\tau_{p}}^{n}\right]+2 d\left(0, i_{\max }\right) \sum_{i=1}^{n} \gamma_{i} \leq \\
E\left[L_{\tau_{p}}^{n}\right]+2 \sqrt{2}\left(\frac{n p}{Q_{n}}+1\right) .
\end{gathered}
$$

Dividing with $\sqrt{n}$, taking the limit as $n \rightarrow \infty$ and using (26) that $\lim _{n \rightarrow \infty} \frac{E\left[L_{r_{p}}^{n}\right]}{\sqrt{n}}=$ $\beta_{T S P} \sqrt{p}$ (34) follows.

Note that in case 3 the PTST solves the PVRP optimally. An even more surprising consequence of theorems 11,12 and 13 is that in the cases $Q$ being a constant and $\lim _{n \rightarrow \infty} \frac{Q_{n}}{\sqrt{n}}=0$, then both the probabilistic strategies $a$ and $b$ are asymptotically equivalent to the re-optimization strategy. Furthermore, in these cases any tour produced by the cyclic heuristic with initial tour any tour of length $O(\sqrt{n})$ is asymptotically equivalent to the re-optimization strategy. In the case $\lim _{n \rightarrow \infty} \frac{Q_{n}}{\sqrt{n}}=\infty$, strategy $b$ is asymptotically the same as the strategy of re-optimization and even further the PTST is asymptotically optimal.

Finally, we only considered the case where customer locations are uniformly distributed in the unit square. Similar asymptotic theorems can be proved in the $d$-dimensional Euclidean space and furthermore, for the case that the distribution of customer locations has continuous part with density $f$. We chose the Euclidean plane in the exposition because the geometry is clearer and the uniform distribution since it is more intuitive.

## 5 An Overview of the Paper

In this paper we showed that the PVRP is a natural extension of the deterministic VRP and has a large area of potential applications. We proposed an $\mathrm{O}\left(n^{2}\right)$ method to calculate the expected length of an a priori route under various probabilistic assumptions. Based on these expressions we found upper and lower bounds for the expected length of the optimal routes under strategies $a$ and $b$, as well as for the strategy of re-optimization. Based on these bounds we proposed the cyclic heuristic for the PVRP under strategty $a$, which is within a constant factor $\left(\frac{5}{2}-\frac{p}{Q}\right)$ of the optimal solution, if the starting tour is Christofides tour. For the PVRP under strategy $b$ we found that the spacefilling curve heuristic provides good solutions (within a logarithmic factor) of the optimal solution. Moreover, we conjectured that the spacefilling curve heuristic is within a constant factor of the optimal solution.

More importantly, we proved that if customer locations are random in the unit square, strategy $b$ performs asymptotically the same with the strategy of re-optimization in all cases, and strategy $a$ performs the same with the strategy of re-optimization if the capacity is "small", which is a strong indication of the usefulness of these strategies. Furthermore, our analysis revealed some asymptotically optimal heuristics for both strategies $a, b$. As a result of our analysis, the two strategies we considered provide a strong alternative to the strategy of re-optimization, and therefore they can be both useful in practice.

As a final conclusion, we believe that the paper demonstrated that in the context of capacitated vehicle routing problems a priori strategies (PVRP) are a serious and practical alternative to re-optimization strategies. Our investigation in [2] reached the same conclusion for other combinatorial optimization problems.

## Acknowledgments

I would like to thank my thesis advisor Professor Amedeo Odoni for several useful comments and interesting discussions. This work was partially supported from the National Science Foundation under grant ECS-8717970.

## References

[1] Beardwood J., Halton J. and Hammersley J., (1959), "The Shortest Path Through Many Points", Proc. Camb. Phil. Soc., 55, 299-327.
[2] Bertsimas D., (1988), "Probabilistic Combinatorial Optimization Problems", Ph.D thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.
[3] Bodin L., Golden B., Assad A. and Ball M., (1983), "Routing and Scheduling of Vehicles and Crews, the State of the Art", Comp. Oper. Res., 10, 69-211.
[4] Dror M. and Trudeau P., (1986), "Stochastic Vehicle Routing with Modified Savings Algorithm", Eur. Jour. Oper. Res., 23, 228-235.
[5] Golden G. and Assad A., (1988), Vehicle Routing; Methods and Studies, North Holland, Amsterdam.
[6] Haimovitch M. and A. Rinnooy Kan, (1985), "Bounds and Heuristics for Capacitated Routing Problems", Math. Oper. Res., 10, 527-542.
[7] Jaillet P., (1985), "Probabilistic Traveling Salesman Problems", (Ph.D. Thesis) Technical Report No. 185, Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts.
[8] Jaillet P. and Odoni A., (1988), The Probabilistic Vehicle Routing Problem, in Vehicle Routing; Methods and Studies, edited by B.L. Golden and A.A. Assad, North Holland, Amsterdam.
[9] Laporte G. and Louveau F., (1987), "Formulations and Bounds for the Stochastic Capacitated Vehicle Routing Problem with Uncertain Supplies", G-87-23, Ecole des Hautes Etudes Commerciale, University of Montreal, Montreal, Canada.
[10] Laporte G., Louveau F. and Mercure H., (1987), "Models and Exact Solutions for a Class of Stochastic Location-Routing Problems", G-87-14, Ecole des Hautes Etudes Commerciale, University of Montreal, Montreal, Canada.
[11] Lawler E.L., Lenstra J.K., A. Rinnooy Kan and Shmoys D.B. (eds.), (1985), The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization, Wiley, Chichester.
[12] Platzman L.K. and Bartholdi J.J., (1983), "Spacefilling Curves and the Planar Traveling Salesman Problem", Report no. PDRC 83-02, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.
[13] Stewart W. and Golden B., (1983), "Stochastic Vehicle Routing: A Comprehensive Approach", Eur. Jour. Oper. Res., 14, 371-385.
[14] Tillman F., (1969), "The Multiple Terminal Delivery Problem with Probabilistic Demands", Transp. Science, 3, 192-204.

