Convergence Results for Vehicle Routing Problems with Stochastic Demands

Dennis Weyland   Roberto Montemanni   Luca Maria Gambardella

Technical Report No. IDSIA-06-11
November 2011

IDSIA / USI-SUPSI
Dalle Molle Institute for Artificial Intelligence
Galleria 2, 6928 Manno, Switzerland

Convergence Results for Vehicle Routing Problems with Stochastic Demands

Dennis Weyland Roberto Montemanni Luca Maria Gambardella

November 2011

Abstract

In this work we investigate two variants of the Stochastic Vehicle Routing Problem: The Vehicle Routing Problem with Stochastic Demands and the Vehicle Routing Problem with Stochastic Demands and Customers. We show that under some moderate conditions there is an asymptotic equivalence between the Vehicle Routing Problem with Stochastic Demands and the Traveling Salesman Problem, as well as between the Vehicle Routing Problem with Stochastic Demands and Customers and the Probabilistic Traveling Salesman Problem. Based on our results we give explanations for different observations in literature and we provide ideas for the development of new approximation algorithms and heuristics for these problems.

1 Introduction

Within the last years the field of Combinatorial Optimization under Uncertainty has received increasing attention. Using uncertain and dynamic information in the problem formulation, more realistic models of real world problems can be generated. A common way for the representation of the uncertainty is to use probability distributions as input parameters instead of single values. Combinatorial Optimization Problems that use this kind of representation for the uncertainty are called Stochastic Combinatorial Optimization Problems. A comprehensive overview about recent developments in the field of Stochastic Combinatorial Optimization is given in Bianchi et al. [5].

Since most of the Stochastic Combinatorial Optimization Problems are generalizations of (non stochastic) Combinatorial Optimization Problems they are in practice usually harder to solve than their (non stochastic) counterparts. Even the evaluation of a solution can be computational expensive for some of the problems and the design of good heuristics is usually a great challenge.

Among the Stochastic Combinatorial Optimization Problems there is the class of Stochastic Vehicle Routing Problems. The problems that belong to that class are widely investigated and have many real world applications. There is a huge variety of Stochastic Vehicle Routing Problems. Typical variants are the Probabilistic Traveling Salesman Problem (Jaillet [17], Campbell [9], Branke and Guntsch [8], Bianchi et al. [7, 6]), the Vehicle Routing Problem with Stochastic Demands (Laporte et al. [18], Novoa et al. [20], Chepuri et al. [10], Christiansen and Lysgaard [11]) and the Vehicle Routing Problem with Stochastic Demands and Customers (Gendreau et al. [15, 14, 13]). There are also a lot of variants using multiple vehicles, multiple goods and time constraints. In this work we focus on the Vehicle Routing Problem with Stochastic Demands and the Vehicle Routing Problem with Stochastic Demands and Customers, which are defined formally in the next section. In section 3 we introduce a Markov chain model for the distribution of goods in the vehicle, which is used extensively within the theoretical analyses in the sections 4 and 5.

We discuss the impacts of our results in section 6 and finish the paper with conclusions and an outlook for further research in section 7.

2 Stochastic Vehicle Routing

In this section we give a formal definition of the Vehicle Routing Problem with Stochastic Demands (VRPSD) and the Vehicle Routing Problem with Stochastic Demands and Customers (VRPSDC), both using the so called basic restocking strategy.

For those problems the task is to deliver integral amounts of identical goods from a depot to customers using one vehicle with a limited capacity. The amount of goods that the customers require depends on given probability distributions, which are independent from each other. Usually the demands are bounded from above by the vehicle capacity, but our results can easily be extended to the more general case of unlimited demands. In the case of the VRPSD the demand can also be 0, but the exact amount is not revealed until the vehicle arrives at the customer. For the VRPSDC each customer only requires goods with a certain probability, but whether a customer requires a visit or not is only revealed directly before this customer is processed. The goal for both problems is to find a so called a priori tour, starting at the depot, visiting all customers and ending at the depot, which minimizes the expected cost (or travel distance) for the a posteriori tours over all possible realizations of the random variables. A good overview about the concept of a priori optimization is given in Bertsimas et al. [3]. This concept is mainly used when problems occur on a regular basis and the computational resources are limited in a way, such that it is impossible to optimize from scratch each time. Another advantage of the a priori approach is, that customers are visited roughly at the same time in each new realization of the problem. Furthermore the order in which customers are processed does not change and this makes it easier for the drivers.

The a posteriori tour for a specific realization of the random variables can be derived from the a priori tour in the following way. Note that for the VRPSDC and a specific realization of the random variables, customers that do not require a visit are just deleted from the a priori tour, before the a posteriori tour is derived. The vehicle starts fully loaded at the depot. Then it proceeds to the first customer that requires a visit. If the demand cannot be served, the vehicle unloads all goods, returns to the depot, performs a restocking action and returns to that customer. After a customer has been fully served, a possible restocking action is performed due to the used restocking strategy and the vehicle proceeds to the next customer. This procedure is repeated until the vehicle has served all customers and finished the tour at the depot.

The cost for an a posteriori tour is the sum of the travel costs, including those of the restocking actions. So the expected cost depends on the a priori tour itself and on the restocking strategy that is used. In this work we use the so called basic restocking strategy. The vehicle only performs a restocking action, if a customer cannot be fully served or if the vehicle is empty after serving a customer. We call this strategy the basic restocking strategy, because a minimum requirement to any reasonable restocking strategy should be to restock, if the vehicle is empty after serving a customer. This restocking strategy is very common and widely used in literature (Gendreau et al. [13, 14], Laporte et al. [18], Dror et al. [12], Bertsimas [2], Chepuri et al. [10], Mak and Guo [19], Novoa et al. [20], Christiansen and Lysgaard [11]).

Given a number n of customers, a cost matrix $C = (c_{ij})_{0 \leq i,j \leq n}$, probability distributions for the demand of each customer (and in the case of the VRPSDC also for each customer the probability that a visit is required), a limit Q for the vehicle capacity, and a restocking strategy, the goal of the VRPSD and the VRPSDC is to find a permutation $\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ which represents the a priori tour, such that the expected cost of the a posteriori tour $E(cost(\pi))$ is minimized. The a priori tour starts at the depot in location 0, visits the customers in the order $\pi(1), \pi(2), \ldots, \pi(n)$ and returns to the depot at location 0. The travel costs are represented by C, which is usually symmetric and obeys the triangle inequality. For the VRPSD $E(cost(\pi))$ contains the travel costs and the expected costs for the restocking actions. Since the travel costs for the VRPSDC depend on the probabilities that the customers require a visit, $E(cost(\pi))$ contains the expected travel costs and the expected costs for the restocking actions in that case.

3 Markov Chain Model

In this section we introduce the Markov chain model for the amount of goods in the vehicle, which is fundamental for the remaining part of this paper, and in particular for the proofs in the next two sections. After that we prove some basic properties for this model using the basic restocking strategy.

Given the amount of goods in the vehicle before processing a certain customer, the amount of goods after the customer has been processed depends only on the demand distribution of this customer (and in the case of the VRPSDC additionally on the probability that this customer requires a visit) and is independent of the demands of the other customers. Thus it is clear that we can model the amount of goods in the vehicle for a given route with a time-discrete Markov chain, starting with a distribution representing a full vehicle, and using transition matrices that are based on the demand distributions of the customers and on the restocking strategy.

The different states represent the different amounts of goods in the vehicle. If the amount of goods is bounded from above by the vehicle capacity Q, and if we observe the amount of goods after a possible restocking action has been performed, it is sufficient to have different states for the amount levels $1, 2, \ldots, Q$, because a minimum requirement for a restocking strategy should be to perform a restocking action, if the vehicle is empty after processing a customer. For being able to use modular arithmetic in a straightforward way, we use in the following $1, 2, \ldots, Q-1$ for amount levels of $1, 2, \ldots, Q-1$ and 0 for an amount of Q.

That means the distribution of the amount of goods can be represented by a column vector of size Q, whose elements are all non-negative and sum up to 1. Such a vector is called stochastic. Furthermore the transition matrices are of size $Q \times Q$ and depend only on the demand distributions of the customers (and in the case of the VRPSDC additionally on the probability that the customer requires a visit), and on the restocking strategy that is used. The entry in row i and column j represents the probability that we reach state i from state j in one step. In other words, this is the probability that we have an amount of i goods after processing the customer and performing a possible restocking action, if the amount has been j goods before processing the customer. By definition the transition matrices are stochastic in its columns, i.e. that the elements in each column sum up to 1. Note that the transition matrices for the VRPSDC are convex combinations of the corresponding transition matrices for the VRPSD and the identity matrix, since with a certain probability a customer is skipped and the amount of goods does not change, and otherwise the customer is served exactly as in the VRPSD.

In this work we focus on the basic restocking strategy. This means that we perform a restocking action if we run out of goods while processing a customer, or if the vehicle contains 0 goods after processing a customer. The probability to reach state i from state j is the same as the probability to reach state $(i+k) \bmod Q$ from state $(j+k) \bmod Q$, $\forall k \in \{1, 2, \ldots, Q-1\}$. These observations lead to two additional properties for the transition matrices. Two successive rows only differ in a cyclic shift of one position, in particular for each row the following row is shifted cyclic one position to the right. The other property, which follows directly from this one for square matrices, is that the transition matrices are also stochastic in their rows, i.e. that the elements in each row sum up to 1.

Before we finish this section, we give a formal overview about the properties mentioned above, using the following definitions. For the vectors and matrices we use indices starting at 0 to allow the use of modular arithmetic in a straightforward way.

Definition 1 (Column stochastic matrix)**.** *A $m \times n$ matrix $A = (a_{ij})$ is called stochastic in its columns, if the following two properties hold:*

1. $\forall i \in \{0, 1, \ldots, m-1\}\ \forall j \in \{0, 1, \ldots, n-1\}: \ a_{ij} \geq 0$
2. $\forall j \in \{0, 1, \ldots, n-1\}: \ \sum_{i=0}^{m-1} a_{ij} = 1.$

Definition 2 (Row stochastic matrix)**.** *A $m \times n$ matrix $A = (a_{ij})$ is called stochastic in its rows, if the transposed matrix A^T is stochastic in its columns.*

Definition 3 (Doubly stochastic matrix)**.** *A matrix A is called doubly stochastic, if it is stochastic in its columns and stochastic in its rows.*

Definition 4 (Cyclic matrix)**.** *A $m \times m$ matrix $A = (a_{ij})$ is called cyclic, if the following holds:*

$$\forall i \in \{0, 1, \ldots, m-1\}\ \forall j \in \{0, 1, \ldots, m-1\}: \ a_{ij} = a_{(i+1) \bmod m, (j+1) \bmod m}.$$

Now we are able to give a formal summary of the properties of the transition matrices in the following lemma.

Lemma 1. *Using the Markov chain model introduced at the beginning of this section with the basic restocking strategy, the transition matrices are square matrices, doubly stochastic and cyclic.* □

Another important fact is that column stochastic matrices, row stochastic matrices, doubly stochastic matrices and cyclic matrices are all closed under multiplication. A proof for the last statement is given in the appendix, the other statements are well known results and easy to verify.

4 Convergence Results

We show in this section that there is an asymptotic equivalence between the Vehicle Routing Problem with Stochastic Demands (VRPSD) and the Traveling Salesman Problem (TSP), as well as between the Vehicle Routing Problem with Stochastic Demands and Customers (VRPSDC) and the Probabilistic Traveling Salesman Problem (PTSP), if the basic restocking strategy is used.

One might think that it is important for the VRPSD and the VRPSDC to optimize the (expected) length of the route as well as to optimize the route in a way, such that those customers which cause a restocking action with high probability, are located close to the depot. We show that the distribution of the amount of goods in the vehicle converges to the uniform distribution under certain conditions. In that case the probability that restocking is needed at a certain customer only depends on the distribution of its demand and not on the particular tour. In other words, it is in some cases sufficient to optimize the (expected) length of the tour.

Let us assume that the vehicle is loaded in a probabilistic way, such that the amount of goods at the beginning is uniformly distributed. Then the VRPSD and the TSP, and the VRPSDC and the PTSP are equivalent, if the basic restocking strategy is used. These equivalences can be verified easily using the Markov chain model introduced in the previous section. Together with the convergence of the amount of goods in the vehicle towards the uniform distribution, this observation is the basic idea behind our analyses.

We start with a more general mathematical result and after that we show under which conditions this result can be used to derive concrete results for the VRPSD and for the VRPSDC.

Theorem 1. *Let $(A_n)_{n\in\mathbb{N}}$ be a series of doubly stochastic $m \times m$ matrices, whose entries are all bounded from below by a constant $L > 0$. Then $\lim_{n\to\infty} \prod_{i=1}^{n} A_i = \frac{1}{m}J$, where J is a matrix of ones.*

Proof. This theorem is trivially true for $m = 1$, so we can assume $m > 1$ for the following proof. Since the product of doubly stochastic matrices is a doubly stochastic matrix and for any doubly stochastic matrix $A = (a_{ij})$ we have $\sum_{j=0}^{m-1} a_{ij} = 1 \ \forall i \in \{0, 1, \ldots, m-1\}$ and $a_{ij} \geq 0 \ \forall i, j \in \{0, 1, \ldots, m-1\}$, it is sufficient to show that the entries in each row converge to the same value. We prove this by showing that the difference between the smallest and the largest entry in each row converges to 0.

We show by induction that the difference between the smallest and the largest entry in each row of $\prod_{i=1}^{n} A_i$ is bounded from above by $(1-2L)^n$. It is clear that this expression is always non-negative, because L is trivially bounded from above by 1/2.

For $n = 1$ this is clearly the case. The smallest entry is bounded from below by L and since we have doubly stochastic matrices, the largest entry is bounded from above by $1 - L$. The difference of the the largest entry and the smallest entry can be at most $(1 - L) - L = 1 - 2L = (1 - 2L)^1$.

Now suppose the statement holds for $n \in \mathbb{N}$. Here we use $A^{(n)} = \prod_{i=1}^{n} A_i$ as an abbreviation for the product of the first n matrices. Let $i \in \{0, 1, \ldots, m-1\}$ be a fixed row index and let $l_i(n)$ and $u_i(n)$ be the smallest, respectively the largest entry in the i-th row of $A^{(n)}$. The entries in the i-th row of $A^{(n+1)}$ are convex combinations of the entries in the i-th row of $A^{(n)}$, where the corresponding coefficients of each of these convex combinations is a column of A_{n+1}. Since the entries of A_{n+1} are bounded from below by L we have

$$u_i(n+1) \leq (1-L)u_i(n) + Ll_i(n) = u_i(n) - L(u_i(n) - l_i(n)) \text{ and}$$

$$l_i(n+1) \geq (1-L)l_i(n) + Lu_i(n) = l_i(n) + L(u_i(n) - l_i(n)).$$

The difference of these values is

$$\begin{aligned} u_i(n+1) - l_i(n+1) &\leq u_i(n) - L(u_i(n) - l_i(n)) - l_i(n) - L(u_i(n) - l_i(n)) \\ &= u_i(n) - l_i(n) - 2L(u_i(n) - l_i(n)) \\ &= (1-2L)(u_i(n) - l_i(n)) \\ &\leq (1-2L)^{n+1}. \end{aligned}$$

In the last step we used the induction hypothesis and since $\lim_{n\to\infty}(1-2L)^n = 0$ we finished the proof. □

In the previous section we have seen that we can model the distribution of the amount of goods in the vehicle with a Markov chain. It is possible to identify the matrices A_i of theorem 1 with transition matrices of that Markov chain. Using the basic restocking strategy these transition matrices are doubly stochastic, but unfortunately it is a very strong assumption that their entries are all bounded from below by a constant $L > 0$. In that case each customer has a strictly positive probability for each possible demand and this assumption is obviously too strong for practical applications. In the remaining part of this section we show that we can use theorem 1 also under more moderate conditions.

The main idea here is to partition the matrices into blocks of consecutive matrices, such that the product of all matrices within a block fulfills the conditions of theorem 1. In this way we can prove convergence for the series of these products. We will see that this result can then be extended to the original series, because the convergence behavior is not affected when some additional matrices, which do not form a complete block, are multiplied at the end.

Before we proceed, we need the following definition.

Definition 5 (greatest common divisor property, gcd property)**.** *A cyclic $m\times m$ matrix $A = (a_{ij})$ with only non-negative entries and with strictly positive entries $a_{i_1 0}, a_{i_2 0}, \ldots, a_{i_l 0}$, $i_1 < i_2 < \ldots < i_l$ in the first column fulfills the greatest common divisor property (short: gcd property), if* $\gcd(i_l - i_1, i_l - i_2, \ldots, i_l - i_{l-1}, m) = 1$.

Instead of using the differences relative to i_l we could have also used differences relative to any other of the elements a_i, $i \in \{1, 2, \ldots, l\}$. The definition is also invariant with respect to the chosen column. We prove both invariances in the appendix.

At first glance this definition looks quite technical, but in section 6 we show that the gcd property is usually fulfilled by a lot of the transition matrices for practical instances of the VRPSD and the VRPSDC. In particular the greatest common divisor property is fulfilled for a transition matrix, if the corresponding customer has two demands of amount k and $k+1$ with strictly positive probabilities. For the VRPSDC we have always a strictly positive entry at the first position in the first column and the condition is even slightly weaker in that case.

We are now ready to prove the following important lemma using the definition of the gcd property and some modular arithmetic.

Lemma 2. *Let A be a $m\times m$ matrix with only non-negative entries and with k_i strictly positive entries in row i, $i \in \{0, 1, \ldots, m-1\}$, and let B be a cyclic $m \times m$ matrix with only non-negative entries fulfilling the greatest common divisor property. Then the product $C = AB$ contains only non-negative entries and has at least* $\min\{k_i + 1, m\}$ *strictly positive entries in row i, for each $i \in \{0, 1, \ldots, m-1\}$.*

Proof. The elements of the set $I = \{i_1, i_2, \ldots, i_n\} \subset \mathbb{Z}/m\mathbb{Z}$, $n \in \mathbb{N}$, with $\gcd(i_1, i_2, \ldots, i_n, m) = 1$ can be used to generate any element of $\mathbb{Z}/m\mathbb{Z}$ by a linear combination of $i_1, i_2, \ldots, i_n$. This is a well known algebraic result, see e.g. Baldoni et al. [1].
For our proof we need the following useful property. Given a set $S \subset \mathbb{Z}/m\mathbb{Z}$ with $|S| < |\mathbb{Z}/m\mathbb{Z}|$, and I as above, then for the set $S' = \{s + i \mid s \in S, i \in I \cup \{0\}\}$ the inequality $|S| < |S'|$ holds. Otherwise we cannot generate all elements of $\mathbb{Z}/m\mathbb{Z}$ by linear combinations of $i_1, i_2, \ldots, i_n$.

Let $i \in \{0, 1, \dots, m-1\}$ be a fixed row index with $k_i < m$, let $S = \{j \mid a_{ij} > 0\} \subset \mathbb{Z}/m\mathbb{Z}$ and let $S' = \{j \mid c_{ij} > 0\} \subset \mathbb{Z}/m\mathbb{Z}$. Furthermore let $b_{j_1 0}, b_{j_2 0}, \dots, b_{j_l 0}$, $j_1 < j_2 < \dots < j_l$ be the strictly positive entries in the first column of B. Due to the properties of our matrices, we have

$$\begin{aligned}
|S'| &= |\{(s+b) \bmod m \mid s \in S,\ b = -j_i,\ i \in \{1, 2, \dots, l\}\}| \\
&= |\{(s - j_l + b) \bmod m \mid s \in S,\ b = j_l - j_i,\ i \in \{1, 2, \dots, l\}\}| \\
&= |\{(s - j_l + b) \bmod m \mid s \in S,\ b \in \{j_l - j_1, j_l - j_2, \dots, j_l - j_{l-1}\} \cup \{0\}\}| \\
&= |\{(s + b) \bmod m \mid s \in S,\ b \in \{j_l - j_1, j_l - j_2, \dots, j_l - j_{l-1}\} \cup \{0\}\}| \\
&> |S|,
\end{aligned}$$

which concludes the proof for all rows with less than m strictly positive entries.
Now let $i \in \{0, 1, \dots, m-1\}$ be a fixed row index with $k_i = m$. Since each column of B has at least 1 strictly positive entry, the number of strictly positive entries in the i-th row of $C = AB$ is m which concludes the proof for rows with exactly m strictly positive entries. □

Using this lemma we can obtain a sufficient condition for the case that the product of a number of transition matrices contains only strictly positive entries.

Corollary 1. *Let $A_1, A_2, \dots, A_k$ be cyclic $m \times m$ matrices with only non-negative entries and with at least 1 strictly positive entry in each row, and let $m-1$ of these matrices fulfill the greatest common divisor property. Then $B = \prod_{i=1}^{k} A_i$ contains only strictly positive entries.*

Proof. Since the matrices are cyclic and have at least 1 strictly positive entry in each column the number of strictly positive entries in the series of products is never decreasing in any of the rows. Due to lemma 2 a multiplication with one of the matrices, which fulfills the gcd property, increases the number of strictly positive entries in each row by 1 (if it has not been maximum already). Putting this observations together the corollary can be proved easily. □

To generalize the convergence result of theorem 1 we need the following lemma.

Lemma 3. *Let $A = (a_{ij})$ be a $m \times m$ matrix containing only strictly positive entries, with a smallest entry l and a largest entry u and let B be a $m \times m$ column stochastic matrix. Then $u' - l' \leq u - l$, where u' and l' are the largest, respectively the smallest, strictly positive entries in AB.*

Proof. The elements of AB are convex combinations of elements of A and therefore the elements of AB are bounded from below by l and from above by u. □

Now we are able to generalize the convergence result of theorem 1.

Theorem 2. *Let $(A_n)_{n\in\mathbb{N}}$ be a series of doubly stochastic, cyclic $m \times m$ matrices, whose entries are all bounded from below by a constant $L > 0$. Let the greatest common divisor property be fulfilled by infinitely many of these matrices and let $k_1 < k_2 < \dots$ be the indices of the matrices that fulfill the gcd property. If $\max\{k_{i+1} - k_i \mid i \in \mathbb{N}\} \leq k$ for a constant $k \geq k_1$ then $\lim_{n\to\infty} \prod_{i=1}^{n} A_i = \frac{1}{m} J$, where J is a matrix of ones.*

Proof. To proof this theorem we have to combine the previous results. Here we show like in theorem 1 that the differences between the largest and the smallest element in each row of the successive products $\prod_{i=1}^{n} A_i$ are monotonically decreasing and converge to 0.

It is possible to partition the matrices into blocks of at most $k(m-1)$ consecutive matrices, such that the matrices in each block fulfill the conditions of corollary 1. Let $1 = j_1 < j_2 < \ldots$ be the indices at which the first, second, ... block begins. Then we can bound the smallest entry of the products $\prod_{i=j_l}^{j_{l+1}-1} A_i$, $j \in \mathbb{N}$, by $L' = L^{k(m-1)}$.
We can now create a new series $(B_n)_{n\in\mathbb{N}}$, with

$$B_n := \prod_{i=j_n}^{j_{n+1}-1} A_i.$$

Due to corollary 1 those matrices fulfill the conditions of theorem 1 and thus we have $\lim_{n\to\infty} \prod_{i=1}^{n} B_i = \frac{1}{m}J$. Analogue to the proof of theorem 1 we can now bound the difference between the largest and the smallest element in each row of the product $\prod_{i=1}^{n} B_i = \prod_{i=1}^{j_{n+1}-1} A_i$ from above by $(1-2L')^n$. With lemma 3 we can show that the difference between the largest and the smallest element of $\prod_{i=1}^{j} A_i$ can also be bounded from above by $(1-2L')^n$, if $j_{n+1} \leq j < j_{n+2} - 1$. This concludes the proof. □

Before we discuss the impacts of these results on the VRPSD and the VRPSDC in section 6, we show that for any doubly stochastic and cyclic $m \times m$ matrix A, which does not fulfill the gcd property, the series of products $\prod_{i=1}^{n} A$ does not converge to $\frac{1}{m}J$, where J is again a matrix of ones. Of course, that does not mean, that there do not exist series $(A_n)_{n\in\mathbb{N}}$ of doubly stochastic and cyclic $m \times m$ matrices, which do not fulfill the conditions of theorem 2, but which converge to $\frac{1}{m}J$. However, theorem 2 is the strongest general condition we could obtain and is fortunately moderate enough when applied to the VRPSD and the VRPSDC.

Theorem 3. *Let $A = (a_{ij})$ be a $m \times m$ cyclic matrix with only non-negative entries, which does not fulfill the gcd property. Then the series of products $(A^n)_{n\in\mathbb{N}}$ does not converge to $\frac{1}{m}J$, where J is a matrix of ones.*

Proof. Let $a_{i_1 0}, a_{i_2 0}, \ldots, a_{i_l 0}$, $i_1 < i_2 < \ldots < i_l$ be the strictly positive entries in the first column of A. Since A does not fulfill the gcd property, we have $\gcd(i_l - i_1, i_l - i_2, \ldots, i_l - i_{l-1}, m) = k > 1$. Let a_{0j} be one of the strictly positive entries in the first row of A. We show that in the product $A^{(n)} = A^n$, $n \in \mathbb{N}$, the entry $a^{(n)}_{0,(j-(n-1)i_l+1) \bmod m}$ is always 0, and in this case a convergence towards $\frac{1}{m}J$ is not possible.
This clearly holds for the case $n = 1$, otherwise there would be two strictly positive entries next to each other (module m) in the first row and therefore also in the first column. This contradicts $\gcd(i_l - i_1, i_l - i_2, \ldots, i_l - i_{l-1}, m) > 1$. In $A^{(n)}$, $n > 1$, the entry $a^{(n)}_{0,(j-(n-1)i_l) \bmod m}$ is strictly positive. If $a^{(n)}_{0,(j-(n-1)i_l+1) \bmod m}$ would also be strictly positive, then 1 can be written as a linear combination of $i_l - i_1, i_l - i_2, \ldots, i_l - i_{l-1}$ and m. Since the greatest common divisor of $i_l - i_1, i_l - i_2, \ldots, i_l - i_{l-1}$ and m is greater than 1, this is not possible and the entry $a^{(n)}_{0,(j-(n-1)i_l+1) \bmod m}$ equals 0 as claimed. □

5 Convergence speed for binomial demand distributions

In this section we want to examine the convergence speed, if the demands are distributed according to binomial distributions with the same underlying probability of p. For many practical applications binomial demand distributions are a reasonable assumption. Mathematically the sum of multiple binomial distributions with the same probability parameter p can be expressed as a single binomial distribution, which itself can be decomposed into a sum of identical Bernoulli distributions. Using this observation, we can bound the convergence speed with respect to the number of Bernoulli trials and therefore also with respect to the expected demand and the expected number of restockings. In this way we provide results, which are useful for practical applications and for further theoretical investigations.

We model the stochastic process using the Markov chain model and multiple identical Bernoulli distributions. In the following the success probability, which represents a demand of 1, is denoted by p and the failure probability, which represents a demand of 0, is denoted by $q = 1-p$. Omitting entries representing a 0, the transition matrix T associated with one Bernoulli distribution can then be written in the following way.

$$T = \begin{pmatrix} q & p & & & \\ & q & p & & \\ & & \ddots & \ddots & \\ & & & q & p \\ p & & & & q \end{pmatrix}.$$

Now the idea is to focus on powers of the transition matrix T. For that purpose we want to use the eigendecomposition of T into a product $T = VDV^H$ of complex matrices V and D, where $V^H = V^{-1}$ and D is a diagonal matrix. Then we can calculate the product T^n easily as $T^n = (VDV^H)^n = VD^nV^H$. For the eigendecomposition of T, we have to calculate its complex eigenvalues and corresponding eigenvectors. To get the complex eigenvalues we have to calculate the roots of the characteristic polynomial. If T is a $m \times m$ matrix, we can write the characteristic polynomial in the following form.

$$
\begin{aligned}
\det(\lambda I - T) &= \det \begin{pmatrix} \lambda - q & -p & & & \\ & \lambda - q & -p & & \\ & & \ddots & \ddots & \\ & & & \lambda - q & -p \\ -p & & & & \lambda - q \end{pmatrix} \\
&= (\lambda - q) \det \begin{pmatrix} \lambda - q & -p & & \\ & \ddots & \ddots & \\ & & \lambda - q & -p \\ & & & \lambda - q \end{pmatrix} \\
&\quad + (-1)^{m+1}(-p) \det \begin{pmatrix} -p & & & \\ \lambda - q & -p & & \\ & \ddots & \ddots & \\ & & \lambda - q & -p \end{pmatrix} \\
&= (\lambda - q)^m + (-1)^{m+1}(-p)^m \\
&= (\lambda - q)^m - p^m
\end{aligned}
$$

Using this form we can derive the following term for the different eigenvalues. Note that $\sqrt[m]{1}$ represents all m complex roots of 1.

$$
\begin{aligned}
& (\lambda - q)^m - p^m &= 0 \\
\Leftrightarrow \quad & (\lambda - q)^m &= p^m \\
\Leftrightarrow \quad & (\lambda - q) &= p \sqrt[m]{1} \\
\Leftrightarrow \quad & \lambda &= q + p \sqrt[m]{1}
\end{aligned}
$$

So the eigenvalues are distributed around a circle with center q and radius p. Furthermore 1 is always an eigenvalue with a corresponding eigenvector $(1, 1, \ldots, 1)^t$. If we denote the different eigenvalues with $\lambda_1, \lambda_2, \ldots, \lambda_m$ and the corresponding normalized eigenvectors with $v_1, v_2, \ldots, v_m$, we can write V and D as

$$
V = \begin{pmatrix} v_1 & v_2 & \ldots & v_m \end{pmatrix},
$$

$$
D = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{pmatrix}.
$$

Now we can rewrite the product T^n as

$$
\begin{aligned}
T^n &= V D^n V^H \\
&= \sum_{i=1}^{m} (v_i v_i^H \lambda_i^n).
\end{aligned}
$$

With $\lambda_1 = 1$ and the corresponding normalized eigenvector $v_1 = (m^{-1/2}, m^{-1/2}, \ldots, m^{-1/2})$ we have

$$\begin{aligned} T^n &= \sum_{i=1}^{m} (v_i v_i^H \lambda_i^n) \\ &= \frac{1}{m} J + \sum_{i=2}^{m} (v_i v_i^H \lambda_i^n), \end{aligned}$$

where J is a matrix of ones. That means we are now able to bound the convergence speed using the error term $\sum_{i=2}^{m} (v_i v_i^H \lambda_i^n)$.

If we denote the entry in row j and column k of V^n by V_{jk}^n and the j-th entry of v_i by $(v_i)_j$, we can bound the error for each $j, k \in \{1, 2, \ldots, m\}$ by

$$\begin{aligned} \left|1/m - V_{jk}^n\right| &= \left|1/m - \sum_{i=1}^{m} (v_i v_i^H \lambda_i^n)_{jk}\right| \\ &= \left|\sum_{i=2}^{m} (v_i v_i^H \lambda_i^n)_{jk}\right| \\ &= \left|\sum_{i=2}^{m} ((v_i)_j \overline{(v_i)_k} \lambda_i^n)\right| \\ &\leq \sum_{i=2}^{m} \left|(v_i)_j \overline{(v_i)_k} \lambda_i^n\right| \\ &\leq \sum_{i=2}^{m} |\lambda_i|^n \\ &\leq (m-1) \left|\lambda_2\right|^n, \end{aligned}$$

where λ_2 is the second largest eigenvalue regarding the standard norm in $\mathbb{C}$. We can calculate the norms of the eigenvalues in the following way. Here k is representing the different eigenvalues with values $k \in \{0, 1, \ldots, m-1\}$.

$$\begin{aligned} |\lambda|^2 &= \left|q + p \cdot e^{(ik2\pi/m)}\right|^2 \\ &= |q + p\cos(k2\pi/m) + p\sin(k2\pi/m)i|^2 \\ &= (q + p\cos(k2\pi/m))^2 + (p\sin(k2\pi/m))^2 \\ &= q^2 + p^2\cos(k2\pi/m)^2 + 2pq\cos(k2\pi/m) + p^2\sin(k2\pi/m)^2 \\ &= p^2 + q^2 + 2pq\cos(k2\pi/m) \\ &= (p+q)^2 - 2pq(1 - \cos(k2\pi/m)) \\ &= 1 - 2pq(1 - \cos(k2\pi/m)) \end{aligned}$$

The eigenvalue with the largest norm is $\lambda_1 = 1$, the eigenvalues with the second largest norm are $q + p \cdot e^{(i2\pi/m)}$ and $q + p \cdot e^{(-ik2\pi/m)}$. Since the norm of the second largest eigenvalue is

$(1 - 2pq(1 - \cos(2\pi/m)))^{1/2}$ and thus a constant smaller than one, we have finally an exponential convergence speed. For our Vehicle Routing Problems with Stochastic Demands this result means that the convergence speed towards the uniform distribution is exponential in the cumulative expected demand of goods.

6 Discussion

In this section we discuss the impacts of the convergence results, especially regarding practical applications. We also show that with these new theoretical results some observations in literature can be explained.

The first question is, what the results mean for the VRPSD and the VRPSDC with the basic restocking strategy. In section 4 we have indicated that the transition matrices of the VRPSD and the VRPSDC can be identified with the matrices of theorem 2. That means if we would have an instance with infinitely many customers (or we would just loop over a finite number of customers), and if the series of transition matrices fulfills the conditions of theorem 2, the distribution of goods in the vehicle converges to the uniform distribution. Now we have to clarify two different things in order to transfer the theoretical results to real world problems. At first, under which conditions does the series of transition matrices fulfill the conditions of theorem 2 and secondly, what happens if we have only a finite number of customers.

The gcd property looks quite technical at first glance. But for practical instances it is usually fulfilled by the transition matrices, or at least by many of them, which would be sufficient to apply our theoretical results. In the previous section we pointed out that the gcd property is fulfilled for a matrix, if it has two strictly positive entries next to each other in the first column. That means for the problems we investigate, that the demand distribution for a customer contains two demands with strictly positive probability next to each other, and this is not a very strong assumption. It is for example fulfilled if the demand distribution is a binomial distribution, which is a common case for these kinds of problems. On the other hand, the gcd property is not fulfilled, if the differences between the possible demands of a certain customer and the capacity of the vehicle are all a multiple of some integer $k > 1$. This could be due to product specific constraints, e.g. it is only possible to order one package of a product consisting of k units. But in this case the problem can be reduced by examining packages instead of actual units. If it is not due to product specific constraints and if the differences of the possible demands for every customer are multiples of (potentially different) integers greater than 1, the conditions of 2 are not fulfilled, but fortunately this is a really artificial case, usually not occurring in real world problems.

The other point is, that we have not an infinite number of customers in real world instances. In section 5 we have shown that the convergence speed is exponential in the total expected demand up to a specific customer, if the demand distributions are binomial distributions of the same type. For instances with a finite number of customers, we know that in the general case the amount of goods approaches the uniform distribution. Preliminary experiments show that for instances, which require multiple restocking actions, the distribution of goods in the vehicle after performing a few restockings is quite close to the uniform distribution. In this case the probability that a restocking action is necessary at a certain customer depends almost only on the distribution of goods, this customers requires, and only to a small amount on the actual tour. On the other hand for instances, which require only a few restocking actions, the convergence is far slower, but in this case the costs of the restocking actions are small compared to the travel costs between customers. In both cases the results underlay our thesis that the VRPSD and the TSP, and the VRPSDC and the PTSP are quite similar.

The second question is, how these theoretical results can be used for the development of good algorithms for the VRPSD and the VRPSDC. A first idea could be to use existing algorithms for the TSP and the PTSP on instances for the VRPSD and the VRPSDC. This approach is used in the so called Cyclic Heuristic, which will be discussed later in this section. Another idea would be to use the theoretical results for the development of good approximations for the solution costs. For example the beginning of the tour could be evaluated by an exact evaluation approach for the VRPSD and the VRPSDC, and after a certain point, the costs could be evaluated by assuming the uniform distribution for the amount of goods in the vehicle. Further research at this point seems promising.

Another mentionable point is, that our results could also be transferred to several variants of Vehicle Routing Problems with Stochastic Demands, that use multiple vehicles. But for these problems, even for large instances, the routes of the single vehicles are rather small. And in that case convergence results are of limited significance.

Additionally our theoretical results give explanations for some observations in literature. In Gendreau et al. [15] the authors observed that "One interest of these studies is to show that stochastic customers are a far more complicating factor than stochastic demands". This is in accordance with our results in some sense. The introduction of stochastic demands does not change the problem significantly, while the introduction of stochastic customers changes the TSP into the PTSP which usually cannot be tackled in the same way, especially if the probabilities for visiting the customers are rather low.

In Haimovich and Rinnooy Kan [16] the authors introduce the so called cyclic heuristic, which is used for theoretical analyses in Bertsimas [2]. The cyclic heuristic starts with a solution for the TSP or the PTSP. It inserts the depot at all possible locations and takes the best tour with respect to the solution cost of the problem to be solved. Our results give again an explanation for the practical success of this heuristic. Using the cyclic heuristic over a good TSP solution usually yields a good solution for the VRPSD, and using the cyclic heuristic over a good PTSP solution usually yields a good solution for the VRPSDC.

An approximation for the difference of the quality of two solutions for the VRPSD and the VRPSDC is introduced in Bianchi et al. [4]. Here the authors implicitly assume a uniform distribution over the amount of goods in the vehicle and they show that the use of that approximation leads to good results in practice. According to our theoretical results the assumption of the uniform distribution is reasonable and so we can give a theoretical explanation for the success of that method.

7 Conclusions

In this paper we have derived theoretical results which show, that the VRPSD and the VRPSDC are under some moderate conditions quite similar to the TSP and the PTSP, respectively. The introduction of stochastic demand values does not affect the problems all too much. Probably analog results can be obtained for other stochastic combinatorial optimization problems. In this case real world problems can be modeled more realistic using stochasticity in their formulation, while well established algorithms and techniques can be used to solve them. Perhaps it is also possible to transfer or improve approximation results from the underlying non stochastic problems to the variants including stochasticity, or even to develop new optimization algorithms directly based on theoretical results, like those presented in this work.

There are several other interesting possibilities for further research on this topic. One approach would be to investigate the convergence speed from a practical point of view. Another approach would be to develop new algorithms for the VRPSD and the VRPSDC, based on the new theoretical results. In particular it is promising to use algorithms based on existing algorithms for the TSP and the PTSP, or algorithms using the new results for good approximations of the solution costs. Moreover it is very interesting to try to generalize the results to other restocking strategies or to different variants of Stochastic Vehicle Routing Problems.

A Cyclic matrices

Lemma 4. *Let $A = (a_{ij})$ be a $m \times m$ cyclic matrix and let $B = (b_{ij})$ be a $m \times m$ cyclic matrix. Then $C = AB$ is a $m \times m$ cyclic matrix.*

Proof. We have $\forall i \in \{0, 1, \ldots, m-1\}$ $\forall j \in \{0, 1, \ldots, m-1\}$

$$\begin{aligned} c_{ij} &= \sum_{k=1}^{m} a_{ik} b_{kj} \\ &= \sum_{k=1}^{m} a_{i+1,k+1} \; b_{k+1,j+1} \\ &= \sum_{k=1}^{m} a_{i+1,k} \; b_{k,j+1} \\ &= c_{i+1,j+1}, \end{aligned}$$

where the indices are taken modulo m. □

B Invariances of the gcd property

First we show that the gcd property does not depend on the index of the strictly positive entry relative to which the differences are taken.

Theorem 4. *Let $A = (a_{ij})$ be a cyclic $m \times m$ matrix with only non-negative entries and with strictly positive entries $a_{i_1 0}, a_{i_2 0}, \ldots, a_{i_l 0}$, $i_1 < i_2 < \ldots < i_l$ in the first column. Then the following statements are equivalent:*

(i) A fulfills the gcd property.

(ii) There exists a $k \in \{1, 2, \ldots, l\}$, s.t. the greatest common divisor of the differences $i_k - i_j$, $j \in \{1, 2, \ldots, l\} \setminus \{k\}$, and m is 1.

(iii) For all $k \in \{1, 2, \ldots, l\}$ the greatest common divisor of the differences $i_k - i_j$, $j \in \{1, 2, \ldots, l\} \setminus \{k\}$, and m is 1.

Proof. The implications (i) $\Rightarrow$ (ii) and (iii) $\Rightarrow$ (i) are trivial. So it is sufficient to prove the implication (ii) $\Rightarrow$ (iii). Let $k \in \{1, 2, \ldots, l\}$ be a value fulfilling (ii). We show that the property also holds for any other value $k' \in \{1, 2, \ldots, l\} \setminus \{k\}$. There exist values $c_1, c_2, \ldots, c_l, d \in \mathbb{Z}$, s.t. the following equivalences hold.

$$\begin{aligned}
& \gcd(i_k - i_1, \ldots, i_k - i_{k-1}, i_k - i_{k+1}, \ldots, i_k - i_l, m) = 1 \\
\Leftrightarrow \quad & \sum_{t=1}^{l} c_t (i_k - i_t) + dm = 1 \\
\Leftrightarrow \quad & \sum_{t=1}^{l} c_t (i_k - i_{k'} + i_{k'} - i_t) + dm = 1 \\
\Leftrightarrow \quad & \sum_{t=1}^{l} c_t (i_{k'} - i_t) + \left(- \sum_{t=1}^{l} c_t \right) (i_{k'} - i_k) + dm = 1 \\
\Leftrightarrow \quad & \gcd(i_{k'} - i_1, \ldots, i_{k'} - i_{k'-1}, i_{k'} - i_{k'+1}, \ldots, i_{k'} - i_l, m) = 1
\end{aligned}$$

Here we use the fact that $\mathbb{Z}$ is a principal ideal ring. For details about properties of principal ideal rings, see e.g. Baldoni et al. [1]. □

Now we show that the property also does not depend on the column, from which we chose the indices of the strictly positive entries.

Theorem 5. *Let $A = (a_{ij})$ be a cyclic $m \times m$ matrix with only non-negative entries. Then the following statements are equivalent:*

(i) A fulfills the gcd property.

(ii) There exists a $k \in \{0, 1, \ldots, m-1\}$, s.t. the greatest common divisor of the differences $i_l^{(k)} - i_j^{(k)}$, $j \in \{1, 2, \ldots, l-1\}$, and m is 1, where $i_1^{(k)} < i_2^{(k)} < \ldots < i_l^{(k)}$ are the indices with strictly positive entries in column k.

(iii) For all $k \in \{0, 1, \ldots, m-1\}$ the greatest common divisor of the differences $i_l^{(k)} - i_j^{(k)}$, $j \in \{1, 2, \ldots, l-1\}$, and m is 1, where $i_1^{(k)} < i_2^{(k)} < \ldots < i_l^{(k)}$ are the indices with strictly positive entries in column k.

Proof. Again the implications (i) $\Rightarrow$ (ii) and (iii) $\Rightarrow$ (i) are trivial and we just have to show the implication (ii) $\Rightarrow$ (iii). For that purpose we show that if the property in (ii) is fulfilled for any $k > 0$ then it is also fulfilled for $k-1$ and if the property is fulfilled for any $k < m-1$ then it is also fulfilled for $k+1$. In this way we can prove (iii) by using this reasoning repeatedly. We show only the implication from k to $k-1$, since the implication from k to $k+1$ is analogue.
So let (ii) be fulfilled for some $k > 0$ and let $i_1^{(k)} < i_2^{(k)} < \ldots < i_l^{(k)}$ be the indices with strictly positive entries in column k. Now we distinguish the two cases, where the first entry in column k is zero or strictly positive, respectively.
In the first case we have $i_1^{(k)} > 0$. Then the strictly positive entries in column $k-1$ are

$$i_1^{(k-1)} = i_1^{(k)} - 1 < i_2^{(k-1)} = i_2^{(k)} - 1 < \ldots < i_l^{(k-1)} = i_l^{(k)} - 1$$

due to the cyclic structure of A. That means the corresponding differences are the same for both columns, since

$$i_l^{(k-1)} - i_j^{(k-1)} = \left(i_l^{(k)} - 1\right) - \left(i_j^{(k)} - 1\right) = i_l^{(k)} - i_j^{(k)}, \ \ j \in \{1, 2, \ldots, l\},$$

and thus (ii) holds for $k-1$.
Now let $i_1^{(k)} = 0$. In this case we have strictly positive entries in column $k-1$ at the row indices

$$i_1^{(k-1)} = i_2^{(k)} - 1 < i_2^{(k-1)} = i_3^{(k)} - 1 < \ldots < i_{l-1}^{(k-1)} = i_l^{(k)} - 1 < i_l^{(k-1)} = m-1,$$

again due to the cyclic structure of A. The differences with respect to the index $i_{l-1}^{(k-1)}$ (which is not the largest index of a strictly positive entry for column $k-1$) are

$$i_{l-1}^{(k-1)} - i_j^{(k-1)} = \left(i_l^{(k)} - 1\right) - \left(i_{j+1}^{(k)} - 1\right) = i_l^{(k)} - i_{j+1}^{(k)}, \;\; j \in \{1,2,\ldots,l-2\} \text{ and}$$

$$i_{l-1}^{(k-1)} - i_l^{(k-1)} = \left(i_l^{(k)} - 1\right) - (m-1) = i_l^{(k)} - m = i_l^{(k)} - i_1^{(k)} - m.$$

Every common divisor of the differences $i_{l-1}^{(k-1)} - i_j^{(k-1)}$, $j \in \{1,2,\ldots,l-2,l\}$, and m would also be a common divisor of the differences $i_l^{(k)} - i_j^{(k)}$, $j \in \{1,2,\ldots,l-1\}$, and m, and vice versa. Since the greatest common divisor of the latter differences and m is 1, the greatest common divisor of the former differences and m has to be 1, too. With theorem 4 we can conclude that the required property also holds for column $k-1$ in this case, which finishes the proof. □

It is also possible to chose a row instead of a column for the definition of the gcd property and to consider the indices of the strictly positive entries in that row. Due to the cyclic structure, the entries in the first row of the matrix are the same as in the last column, but in inverse order. We omit a formal proof for that statement, but a similar reasoning as in the proof of the last theorem can be used here.

Acknowledgments

The first author's research has been supported by the Swiss National Science Foundation, grant 200021-120039/1. The authors are grateful to Nikos Mutsanas for the fruitful discussions.

References

[1] M.W. Baldoni, C. Ciliberto, and G.M.P. Cattaneo. *Elementary Number Theory, Cryptography and Codes*. Springer, 2008.

[2] D.J. Bertsimas. A vehicle routing problem with stochastic demand. *Operations Research*, pages 574–585, 1992.

[3] D.J. Bertsimas, P. Jaillet, and A.R. Odoni. A priori optimization. *Operations Research*, pages 1019–1033, 1990.

[4] L. Bianchi, M. Birattari, M. Chiarandini, M. Manfrin, M. Mastrolilli, L. Paquete, O. Rossi-Doria, and T. Schiavinotto. Hybrid metaheuristics for the vehicle routing problem with stochastic demands. *Journal of Mathematical Modelling and Algorithms*, 5(1):91–110, 2006.

[5] L. Bianchi, M. Dorigo, L.M. Gambardella, and W.J. Gutjahr. A survey on metaheuristics for stochastic combinatorial optimization. *Natural Computing*, 8(2):239–287, 2009.

[6] L. Bianchi, L.M. Gambardella, and M. Dorigo. Solving the homogeneous probabilistic traveling salesman problem by the ACO metaheuristic. *Lecture notes in computer science*, pages 176–187, 2002.

[7] L. Bianchi, L.M. Gambardella, and M. Dorigo. An ant colony optimization approach to the probabilistic traveling salesman problem. *Lecture Notes in Computer Science*, pages 883–892, 2003.

[8] J. Branke and M. Guntsch. Solving the probabilistic TSP with ant colony optimization. *Journal of Mathematical Modelling and Algorithms*, 3(4):403–425, 2004.

[9] A.M. Campbell. Aggregation for the probabilistic traveling salesman problem. *Computers and Operations Research*, 33(9):2703–2724, 2006.

[10] K. Chepuri and T. Homem-de Mello. Solving the vehicle routing problem with stochastic demands using the cross-entropy method. *Annals of Operations Research*, 134(1):153–181, 2005.

[11] C.H. Christiansen and J. Lysgaard. A branch-and-price algorithm for the capacitated vehicle routing problem with stochastic demands. *Operations Research Letters*, 35(6):773–781, 2007.

[12] M. Dror, G. Laporte, and P. Trudeau. Vehicle routing with stochastic demands: Properties and solution frameworks. *Transportation Science*, 23(3):166, 1989.

[13] M. Gendreau, G. Laporte, and R. Seguin. An exact algorithm for the vehicle routing problem with stochastic demands and customers. *Transportation Science*, 29(2):143, 1995.

[14] M. Gendreau, G. Laporte, and R. Seguin. A tabu search heuristic for the vehicle routing problem with stochastic demands and customers. *Operations Research*, pages 469–477, 1996.

[15] M. Gendreau, G. Laporte, and R. Seguin. Stochastic vehicle routing. *European Journal of Operational Research*, 88(1):3–12, 1996.

[16] M. Haimovich and A.H.G. Rinnooy Kan. Bounds and heuristics for capacitated routing problems. *Mathematics of Operations Research*, pages 527–542, 1985.

[17] P. Jaillet. *Probabilistic traveling salesman problems.* PhD thesis, M. I. T., Dept. of Civil Engineering, 1985.

[18] G. Laporte, F.V. Louveaux, and L. Van Hamme. An integer L-shaped algorithm for the capacitated vehicle routing problem with stochastic demands. *Operations Research*, pages 415–423, 2002.

[19] KL Mak and ZG Guo. A genetic algorithm for vehicle routing problems with stochastic demand and soft time windows. In *Systems and Information Engineering Design Symposium, 2004. Proceedings of the 2004 IEEE*, pages 183–190, 2004.

[20] C. Novoa, R. Berger, J. Linderoth, and R. Storer. A Set-Partitioning-Based Model for the Stochastic Vehicle Routing Problem. 2006.