European Journal of Operational Research 79 (1994) 451–473
North-Holland

Theory and Methodology

The general multi-retailer EOQ problem with vehicle routing costs *

Shoshana Anily

Faculty of Management, Tel Aviv University, Tel Aviv 69978, Israel

Received May 1991; revised September 1992

Abstract: We study the distribution problem of a single commodity from one warehouse to n geographically dispersed retailers by a fleet of capacitated vehicles. Each of the retailers faces a continuous constant and deterministic demand rate over the infinite horizon. In addition, each of the retailers is characterized by *its own* inventory holding cost rate. The objective is to obtain a routing and replenishment strategy which minimizes the long-run average transportation and holding cost. We restrict ourselves to a class of strategies which partitions the overall region into subregions. A retailer can be assigned to several subregions: each subregion is responsible for a certain fraction of the sales of each of its retailers. We first show that the optimal solution can be bounded from below by a special partitioning problem whose solution can be given in a closed form. We then present a simple heuristic which is shown to converge to the lower-bound almost surely under mild probabilistic conditions, when the number of retailers is increased to infinity.

Keywords: Heuristics; Partitioning problems; Routing problems; Inventory; Distribution

1. Introduction

In this paper we study the distribution problem of a single commodity from one depot to n geographically dispersed retailers. Each of the retailers is characterized by its geographic location, its demand rate and by its inventory holding cost rate. We consider an infinite-time horizon in which retailers are facing deterministic retailer-specific constant demand rates. All demands must be met on time (i.e., no backlogging is allowed). The cost structure consists of the routing cost which is proportional to the Euclidean distance driven, plus a fixed cost which is paid each time a tour is initiated (this last cost may include the vehicle rental cost or any other fixed costs which do not depend on the load size or the number of stops on the route). In addition, each of the retailers is charged for holding stock in the same way as in the EOQ model. The objective is to find a routing schedule minimizing the average system-wide costs. Such a schedule should specify a list of routes, the frequency each of the routes should be driven as well as the delivery size for each of the retailers on the route. The vehicles are assumed to be identical and their capacity may be limited or non-limited. We may also impose upper bounds on the

* This research has been partially supported by 1) Natural Science Engineering Research Council Canada (NSERCC) Grant No. A4802; 2) Israel Institute of Business Research.

Correspondence to: Dr. S. Anily, Faculty of Management, Tel Aviv University, Tel Aviv 69978, Israel.

SSDI 0377-2217(93)E0055-3

frequency each of the routes is driven or on the sales volume assigned to each of the routes. Such upper bounds may be used by practitioners to (1) limit the number of trucks assigned to a single route; (2) avoid 'heavy' routes, i.e., routes that serve a large percentage of the total sales volume.

This problem is a version of the Capacitated Vehicle Routing Problem (CVRP). Garey and Johnson (1979) show that even the one-period problem with routing costs only (shortly denoted by the 'VRP') is NP-complete. Several variants of the VRP including solution methods as well as a literature review is provided, for example, in Golden et al. (1977) and Magnanti (1981). Recently much interest arose in the development of heuristics with guaranteed error bounds obtained by either asymptotic, probabilistic, statistical or worst-case analyses.

One of the first papers that considers an inventory replenishment problem with vehicle routing costs is by Burns et al. (1985). The paper considers an infinite-horizon one warehouse multiple retailer system with retailers facing constant demand rates. The transportation cost is proportional to the euclidean distance travelled and all holding cost rates are assumed to be *identical*. The authors develop an analytic method for minimizing total replenishment costs.

In this paper we adopt the asymptotic analysis approach. For an excellent survey on the asymptotic analysis of several versions CVRP and VRP with holding costs see Federgruen and Simchi-Levi (1992). The stimulating paper by Haimovich and Rinnooy Kan (1985) provides a breakthrough in this direction. The paper considers a single-period CVRP where each vehicle can serve at most q retailers (all vehicles are assumed identical). By making use of the geometric setting of the problem they developed extremely simple heuristics, based on regional partitioning schemes, which are shown to be asymptotically optimal (i.e., to converge to the optimal solution when the number of retailers is sufficiently large). Federgruen, Rinnooy Kan and Zipkin (1985) extend the results of the above paper and derive asymptotically accurate lower and upper bounds on the minimal expected total cost for an integrated routing and inventory problem: they assume a single-period time horizon and a set of *identical* retailers. Anily and Federgruen (1990b) consider a class of general routing problems where the cost of driving a route depends both on its length and the number of points visited on the route via some general cost function having two arguments. The paper describes a class of simple heuristics of complexity $O(n \log n)$ which are shown to be asymptotically accurate if the cost function satisfies certain conditions. Anily and Federgruen (1990a, 1993) (and Anily, 1987) consider two deterministic, continuous time, infinite horizon inventory-routing problems: a single depot is assumed to supply stock to a set of retailers each facing its own constant demand rate. The holding cost rate is assumed to be *identical* at all retailers. The former paper considers the case where the depot is an outside supplier (R-systems). The later one analyses the case where the depot is a part of the system (DR-system): in addition to the routing and holding costs due to the retailers, the DR-system is charged: a) by a fixed cost each time the depot places an order and b) for holding inventory at the depot. In both systems the routing and delivery schedule to the retailers are to be determined. In DR-systems the reorder-points of the depot are additional decision variables. Both problems are shown to fall in the class of cost functions analyzed in Anily and Federgruen (1990b). In both cases easily computable replenishment strategies are provided. The heuristics for R (DR)-systems are shown to be tight with (to come within 6% of) the optimal solution when the number of retailers is large. Experimental studies show that the proposed heuristics come close to the optimal solution even for problems of moderate size.

Gallego and Simchi-Levi (1990) consider the infinite horizon continuous time R-system with routing costs and retailer specific holding and ordering costs. The authors obtain a simple lower bound on the average total cost. Moreover they show that if the trucks' capacity is small enough relative to the Economic Order Quantity of each of the retailers separately, then a simple heuristic using only 'direct shipments' (i.e., each route consists of a single retailer), comes within 6% of the lower bound given that the number of retailers is sufficiently large.

In this paper we generalize the results of Anily and Federgruen (1990a) for general holding cost rates. In some settings, the variability among the retailers is small and the assumption of identical holding costs may be reasonable; however in large distribution systems we often encounter the situation where the holding cost rates vary substantially among the retailers. Usually the holding cost is an increasing function of the item-value and this one tends to increase with the distance travelled. Other factors as, for example, the proximity to city-centers may also have an impact on the holding costs. Thus, it is important to obtain good replenishment strategies for general holding cost rates.

Intuitively, one will seek a replenishment policy with the property that two retailers located in close proximity, but one having a higher holding cost than the other, then the stock of the first will be replenished at least as frequently as the stock of the second. It is noteworthy that in Anily and Federgruen (1990a) the assignment of retailers to routes is based solely on their geographical location. The extension of the analysis to the general case, especially with respect to the lower bound derivation, is not direct (for reasons explained below), and a new approach is required. The computation of the lower bound and the heuristic, is as simple as for the identical retailers model (complexity $O(n \log n)$). We also remark that the lower bound and the policy obtained by applying the proposed algorithm in this paper to a set of retailers having *identical* holding cost rates do not necessarily coincide with the ones derived in Anily and Federgruen (1990a). For a set of identical retailers we recommend the user to apply the algorithm in Anily and Federgruen (1990a) which is expected to yield more accurate lower bounds and policies of better quality.

2. Problem description

Consider a distribution system consisting of one warehouse and a set of n retailers. Let μ_j denote the demand rate of retailer j. Similarly to Anily and Federgruen (1990a), we assume that all demand rates are represented as integer multiples of some common quantity μ, i.e.

$$\mu_j = k_j \mu, \quad 1 \le k_j \le K, \; j = 1, \ldots, n,$$

for some given integer K that is independent of n. Without loss of generality we let $\mu = 1$. We define a demand point as a point in the plane facing a demand rate of 1. Thus we can view retailer j as consisting of k_j demand points all located at the j-th retailer's site. We restrict ourselves to policies that partition the set of $N = \sum_{j=1}^{n} k_j$ demand points into groups (= regions), and each time the stock of one demand point is replenished then the stock of all other demand points in the same region is replenished as well. Consequently, a retailer may be served by several routes where each of these routes satisfies a certain fraction of the retailer's total demand. We assume that once the routes and the assignment of the demand points to the routes are determined, then each route is controlled independently of the others: the stock delivered on a specific route to a retailer having q demand points on *that* route is aimed to satisfy the demand of these q demand points only, resulting in a corresponding depletion rate of q. Each route is responsible for not allowing the stock of *its* demand points to be backlogged. We denote this class of policies by Φ. For a further discussion on this restriction see Anily and Federgruen (1990c).

Let:

$c =$ The fixed cost per route driven.

$b =$ The capacity of a vehicle (all vehicles are assumed to be identical). If the vehicles are non-limited in capacity, we let $b = \infty$.

$f =$ The upper bound on the frequency with which a given route may be driven. If no frequency constraints prevail, let $f^* = \infty$.

$\overline{M} =$ The upper bound on the number of demand points assigned to a single route, $\overline{M} \le \infty$.

Without loss of generality, (a) we set the variable transportation cost per mile to one; (b) we assume that $f^* b \ge 1$ in order to ensure feasibility.

Let also $X = \{x_1, \ldots, x_N\}$ be the set of demand points in the Euclidean plane, with r_i the distance of demand point x_i from the warehouse. We choose the warehouse as the origin of the plane. Let $r_i^c = r_i + \frac{1}{2}c$. We use h_i to denote the holding cost rate at x_i and without loss of generality we number the demand points in ascending order of r_i^c / h_i, $i = 1, \ldots, N$. We assume also that the sequences $\{r_i\}$, $\{h_i\}$, $i = 1, \ldots, N$, are uniformly bounded from above by the constants $r_{\max}$ and $h_{\max}$ respectively, and that the sequence $\{h_i\}$, $i=1,\dots,N$, is also bounded from below by the constant $h_{\min}>0$. Moreover, in the case that $f^*=\infty$ and $c=0$, we assume the sequence $\{r_i\}_{i=1,\dots,N}$ is bounded from below by the constant $r_{\min}>0$. Let $r^c_{\max}=r_{\max}+\frac{1}{2}c$ and $r^c_{\min}=r_{\min}+\frac{1}{2}c$. If $f^*b=\infty$, we assume that $\overline{M}$, the upper bound on the sales volume assigned to each of the routes is *finite*. The requirement of having a finite sales volume in each of the routes is essential to the analysis and was assumed in many other asymptotic analyses of vehicle routing heuristics.

A partition of the set of demand points X is denoted by χ. We write $\chi=\{X_1,\dots,X_L\}$ if X_i is a subset of X, $i=1,\dots,L$, $X_i\cap X_j=\emptyset$, $i<j$, and $\bigcup_{i=1}^{L}X_i=X$. We use m_ℓ to denote the cardinality of X_ℓ, i.e. $m_\ell=|X_\ell|$, $\ell=1,\dots,L$. For a *specific partition* χ with L sets the determination of the optimal policy in Φ reduces to L constrained EOQ problems one for each region: the order cost of X_ℓ equals to the sum of the vehicle's rental cost with the total length of the route emanating from the warehouse, visiting all the demand points in X_ℓ, and finally returning back to the warehouse. The holding cost rate of X_ℓ is the average holding cost rate of the demand points in X_ℓ and its demand rate is $|X_\ell|$.

Similarly to the EOQ model, the best policy in the class Φ is such that each of the routes is driven at equi-distant epochs and the quantity delivered to any of the retailers on the route is of constant size. The determination of the route is the well known Traveling Salesman Problem (TSP) defined on the warehouse and the corresponding retailers.

Denote by $\mathrm{TSP}(X^0_\ell)$ the length of an optimal traveling salesman tour via the warehouse and the retailers in X_ℓ, and by Q_ℓ the total delivery size to the ℓ-th region. Our objective is to compute $V^*(X)=$ the minimal long-run average cost among all strategies in the class Φ. We define OPT_Φ to be the problem of finding the optimal strategy in Φ:

(OPT_Φ)

$$V^*(X)=\min\left\{\sum_{\ell=1}^{L}\min_{Q_\ell}\left[\frac{1}{2}\sum_{i\in X_\ell}\frac{h_iQ_\ell}{m_\ell}+\frac{m_\ell}{Q_\ell}\left(\mathrm{TSP}(X^0_\ell)+c\right)\right]\right\}$$

$$\begin{aligned}\text{s.t.}\quad & m_\ell/f^*\le Q_\ell\le b,\quad \ell=1,\dots,L,\\ & m_\ell\le\overline{M},\quad \ell=1,\dots,L,\\ & \chi=\{X_1,\dots,X_L\}\text{ is a partition of }X.\end{aligned}$$

The problem can be simplified by using the following considerations: for a given set of demand points X_ℓ with delivery size Q_ℓ, the average total cost is given by the expression in the square brackets in the objective function of OPT_Φ. Similarly to a constrained EOQ problem, the optimal Q_ℓ subject to the first ℓ constraints is given by

$$Q^*_\ell=\min\left\{b,\max\left[\frac{m_\ell}{f^*};\sqrt{2m^2_\ell\left(\mathrm{TSP}(X^0_\ell)+c\right)/\sum_{i\in X_\ell}h_i}\right]\right\}. \tag{1}$$

By substituting (1) into $V^*(X)$, we obtain an equivalent problem whose objective function is a minimization over all partitions χ, of a cost function that depends on the problem's parameters, subject to constraints on the sets' cardinalities: the ℓ first constraints translate into ℓ constraints of the form $m_\ell\le f^*b$, thus the 2ℓ constraints can be written as $m_\ell\le M^*$, $\ell=1,\dots,L$, where $M^*\overset{\text{def}}{=}\min\{\lfloor f^*b\rfloor,\overline{M}\}<\infty$.

For a given region X_ℓ, define the following parameters:

$$\theta_\ell=\mathrm{TSP}(X^0_\ell)+c,\qquad H_\ell=\sum_{X_\ell}h_i,\qquad R_\ell=\sum_{X_\ell}r_i,\qquad R^c_\ell=\sum_{X_\ell}r^c_i.$$

In Lemma 1 we provide the simplified representation for OPT_Φ that will be used in the sequel.

Lemma 1. (a) *The optimal average cost of a given region* X_ℓ *is* $g(\theta_\ell, H_\ell, m_\ell)$ *where*

$$g(\theta_\ell, H_\ell, m_\ell) = \begin{cases} H_\ell/(2f^*) + f^*\theta_\ell & \text{if } \theta_\ell/H_\ell \le 1/(2f^{*2}), \\ (2H_\ell\theta_\ell)^{1/2} & \text{if } 1/(2f^{*2}) \le \theta_\ell/H_\ell \le b^2/(2m_\ell^2), \\ bH_\ell/(2m_\ell) + m_\ell\theta_\ell/b & \text{otherwise}. \end{cases} \quad (2)$$

(b) $V^*(X) = \min\{\sum_{\ell=1}^{L} g(\theta_\ell, H_\ell, m_\ell) \mid \chi = \{X_1, \ldots, X_L\}$ *is a partition of* X *and* $m_\ell \le M^*$, $\ell = 1, \ldots, L\}$.

Proof. (a): The average cost of a region X_ℓ is given by a constrained EOQ model with the fixed cost θ_ℓ, the variable holding cost H_ℓ/m_ℓ, and demand rate of m_ℓ units.

(b): Directly from the definitions of $V^*(X)$ and the function g. □

In the next section we provide a lower bound on $V^*(X)$. For the reader's convenience we summarize the main notations and definitions of Sections 2 and 3 in Table 1 (see end of Section 3).

3. The lower bound

Theorem 1. $V^*(X) \ge \underline{V}(X)$ *where*

$$\underline{V}(X) = \min\left\{ \sum_{\ell=1}^{L} g(2R_\ell^c/m_\ell, H_\ell, m_\ell) \mid \chi = \{X_1, \ldots, X_L\} \text{ is a partition of X and } m_\ell \le M^*, \ \ell = 1, \ldots, L \right\}. \quad (3)$$

Proof. The function $g(\theta_\ell, H_\ell, m_\ell)$, as given in (2) is strictly increasing in $\theta_\ell > 0$. The well known inequality $\text{TSP}(X_\ell^0) \ge 2R_\ell/m_\ell$ implies that $\theta_\ell \ge 2R_\ell^c/m_\ell$, resulting in (3). □

The lower bound, as given in (3), is a partitioning problem with a separable cost function that depends on the following set's attributes: (1) the average of r_i^c, (2) the sum of holding costs, and (3) the set's cardinality. Chakravarty et al. (1982) show that general partitioning problems are NP-complete. Very few partitioning problem types are known to be polynomially solvable, see Chakravarty et al. (1982, 1985) and Barnes et al. (1989). Unfortunately, the partitioning problem $\underline{V}(X)$ is not known to be solvable: not only it depends on three attributes of the sets but also constraints are imposed on the set's cardinalities.

By a careful investigation of $\underline{V}(X)$ and relaxation of its cardinality constraints, we could derive a solvable lower bound on $\underline{V}(X)$. We first rewrite the set function g as an equivalent set function G having as arguments the sum of r_i^c, the sum of holding cost and the set cardinality: let $\xi_\ell = \frac{1}{2}m_\ell\theta_\ell$ and $G(\xi_\ell, H_\ell, m_\ell) = g(\theta_\ell, H_\ell, m_\ell)$. Alternatively,

$$G(R_\ell^c, H_\ell, m_\ell) = g(2R_\ell^c/m_\ell, H_\ell, m_\ell). \quad (4)$$

Below we write the function G explicitly. (For simplicity we omit the index ℓ.)

$$G(R^c, H, m) = \begin{cases} H/(2f^*) + 2f^*R^c/m & \text{if } R^c/H \le m/(4f^{*2}), \\ (4HR^c/m)^{1/2} & \text{if } m/(4f^{*2}) \le R^c/H \le b^2/(4m), \\ bH/(2m) + 2R^c/b & \text{otherwise} \end{cases} \quad (5)$$

(Observe that the set-function's break points in (5) are well defined since be definition $m \le M^* \le bf^*$ implying that $m/(4f^{*2}) \le b^2/(4m)$.) In Appendix A, Lemma 2, we prove the set function G is non-increasing in m. Thus, the optimal average cost of any feasible region X_ℓ, $m_\ell \le M^*$, is bounded from below by $G(R_\ell^c, H_\ell, M^*)$. Let

$$G_{M^*}(R^c, H) \stackrel{\text{def}}{=} G(R^c, H, M^*). \tag{6}$$

Define also,

$$\underline{V}_{M^*}(X) \stackrel{\text{def}}{=} \min\left\{ \sum_{\ell=l}^{L} G_{M^*}(R_\ell^c, H_\ell) \ \chi = \{X_1, \dots, X_L\} \text{ is a partition of } X \right\}. \tag{7}$$

Theorem 2. $\underline{V}(X) \ge \underline{V}_{M^*}(X)$.

Proof. In view of the fact that the function $G(R^c, H, m)$ is monotone non-increasing in m, the requirement that $m \le M^*$ and (6) we get that

$$G(R_\ell^c, H_\ell, m_\ell) \ge G(R_\ell^c, H_\ell, M^*) = G_{M^*}(R_\ell^c, H_\ell)$$

holds for any feasible set X_ℓ. Thus, (3), (4) and (7) imply that

$$\underline{V}(X) \ge \min\left\{ \sum_{\ell=1}^{L} G_{M^*}(R_\ell^c, H_\ell) \ \chi = \{X_1, \dots, X_L\} \text{ is a partition of } X \text{ and } m_\ell \le M^*,\right.$$
$$\left. \ell = 1, \dots, L \right\} \ge \underline{V}_{M^*}(X).$$

The last inequality follows from the relaxation of the partitioning problem which is obtained by omitting the cardinality constrains. □

Unlike $\underline{V}(X)$, $\underline{V}_{M^*}(X)$ and the corresponding partition can be easily obtained. Moreover, it turns out that the last inequality in the proof of Theorem 2 holds as an equality, i.e. omitting the cardinality constraints $m_\ell \le M^*$ from the partitioning problem defined by the set-function G_{M^*} *does not worsen* the lower bound. We prove below that one of the optimal partitions for $\underline{V}_{M^*}(X)$ consists of N singletons. Nevertheless, we will show that $\underline{V}_{M^*}(X)$ is asymptotically accurate with the optimal average cost over all policies in Φ.

Theorem 3. *The partition* $\chi^* = \{\{1\}, \{2\}, \dots, \{N\}\}$ *is an optimal partition for the lower bound partitioning problem* $\underline{V}_{M^*}(X)$ *as defined by* (5)–(7).

Proof. Observe that for a given set X_ℓ, $G_{M^*}(R_\ell^c, H_\ell)$ may be viewed as the optimal average cost of replenishing the demand of the set X_ℓ having (i) joint setup cost of $2R_\ell^c/M^*$, (ii) joint holding cost rate of H_ℓ, and (iii) total demand rate of M^*, under the restriction that T_ℓ = the order interval, $T_\ell = Q_\ell/M^*$, satisfies the inequalities $f^{*-1} \le T_\ell \le b/M^*$. Therefore,

$$\underline{V}_{M^*}(X) = \min\left\{ \sum_{\ell=1}^{L} \min_{f^{*-1} \le T_\ell \le b/M^*} \left[\sum_{i \in X_\ell} \left(\tfrac{1}{2} h_i T_\ell + \frac{2r_i^c}{M^* T_\ell} \right) \right] : \chi \text{ is a partition of } X \right\}$$
$$= \min_{T_1, \dots, T_N} \left\{ \sum_{i=1}^{N} \left(\tfrac{1}{2} h_i T_i + \frac{2r_i^c}{M^*} \right) T_i^{-1} : f^{*-1} \le T_i \le \frac{b}{M^*} \right\}.$$

The second equality follows from the fact that both the joint setup and holding costs are additive functions in the demand points. Thus, no advantage is gained by replenishing demand points simultaneously, i.e., the separability property of $\underline{V}_{M^*}(X)$ with respect to the demand points ensures that $\underline{V}_{M^*}(X) = \sum_{i=1}^{N} \underline{V}_{M^*}(\{x_i\})$. Therefore, $\chi^* = \{\{1\}, \{2\}, \dots, \{N\}\}$ is an optimal partition for $\underline{V}_{M^*}(X)$. □

Table 1

Notation	Description
c	fixed cost per route driven.
f^*	upper bound on the frequency a route may be driven.
b	vehicle's capacity.
$\overline{M}$ (M^*)	upper bound on the total demand in a region ($M^* = \min\{\lfloor f^*b \rfloor, \overline{M}\}$).
$X = \{x_1, \ldots, x_N\}$	set of demand points.
$\chi = \{X_1, \ldots, X_L\}$	a partition of X into L subsets.
r_i (r_i^c)	radial distance of x_i from the depot ($r_i^c = r_i + \frac{1}{2}c$).
R_ℓ (R_ℓ^c)	$= \Sigma_{X_\ell} r_i$ ($R_\ell^c = \Sigma_{X_\ell} r_i^c$).
h_i (H_ℓ)	holding cost rate at x_i ($H_\ell = \Sigma_{X_\ell} h_i$). Note that $r_1^c/h_1 \le r_2^c/h_2 \le \cdots \le r_N^c/h_N$.
m_ℓ	$= \lvert X_\ell \rvert$ = the number of demand points in X_ℓ.
TSP(X_ℓ^0)	optimal tour through the depot and X_ℓ. Routing cost for X_ℓ is TSP(X_ℓ^0) + c.
$g(\theta_\ell, H_\ell, m_\ell)$	optimal average cost of X_ℓ as a function of the routing cost θ_ℓ, H_ℓ, and m_ℓ.
$V^*(X)$	$= \min\{\Sigma_\ell g(\text{TSP}(X_\ell^0) + c, H_\ell, m_\ell) \mid \chi = \{X_1, \ldots, X_L\}$ is a partition of X and $m_\ell \le M^*\}$: the optimal average cost of policies in Φ.
$\underline{V}(X)$	$= \min\{\Sigma_\ell g(2R_\ell^c/m_\ell, H_\ell, m_\ell) \mid \chi = \{X_1, \ldots, X_L\}$ is a partition of X and $m_\ell \le M^*\}$. $V^*(X) \ge \underline{V}(X)$ (Theorem 1).
$G(R_\ell^c, H_\ell, m_\ell)$	$= g(2R_\ell^c/m_\ell, H_\ell, m_\ell)$
$G_{M^*}(R_\ell^c, H_\ell)$	$= G(R_\ell^c, H_\ell, M^*)$. $G(R_\ell^c, H_\ell, m_\ell) \le G_{M^*}(R_\ell^c, H_\ell)$ for $m_\ell \le M^*$ (Lemma A.2).
$\underline{V}_{M^*}(X)$	$= \min\{\Sigma_\ell G_{M^*}(R_\ell^c, H_\ell) \mid \chi = \{X_1, \ldots, X_L\}$ is a partition of $X\}$. $\underline{V}(X) \ge \underline{V}_{M^*}(X)$ (Theorem 2).

Let T_i^* be the minimizer of $V_{M^*}(\{x_i\})$. According to the proof of Theorem 3, the demand points in X can be classified into three (possibly empty) categories:

$$F = \{x_i \mid r_i^c/h_i < M^*/(4f^{*2})\},$$
$$S = \{x_i \mid M^*/(4f^{*2}) \le r_i^c/h_i \le b^2/(4M^*)\},$$
$$C = \{x_i \mid r_i^c/h_i > b^2/(4M^*)\}.$$

The demands points in F are 'replenished' at the highest frequency allowed ($T_i^* = f^{*-1}$) and the ones in C are 'replenished' by fully loaded trucks, i.e. the capacity constraint is tight ($T_i^* = b/M^*$). Note that $b = \infty$ (no capacity limits prevail) implies that $C = \emptyset$, and $f^* = \infty$ (no frequency constraints prevail) implies that $F = \emptyset$.

In the next section we present a regional partitioning scheme which takes into account the radial and polar coordinates as well as the holding cost rates while grouping the retailers into regions of M^* demand points each (with the possible exception of at most three regions which may contain less than M^* demand points).

4. The upper bound: Regional partitioning scheme

As a direct consequence of the previous section, the lower bound $V_{\overline{M}}^*(X)$ can be rewritten as $\underline{V}_{M^*}(X) = \underline{V}_{M^*}(F) + \underline{V}_{M^*}(S) + \underline{V}_{M^*}(C)$ where

$$\underline{V}_{M^*}(F) = \sum_F \left(\frac{h_i}{2f^*} + \frac{2f^*}{M^*} r_i^c \right), \tag{8a}$$

$$\underline{V}_{M^*}(S) = \sum_S \sqrt{4h_i r_i^c / M^*}, \tag{8b}$$

$$\underline{V}_{M^*}(C) = \sum_C \left(\frac{h_i b}{2M^*} + \frac{2r_i^c}{b} \right). \tag{8c}$$

The partition of X into the categories F, S and C is a consecutive partition, i.e., there exist integers f' and c' such that $0 \le f' \le c' \le N$ and $F = \{x_1, x_2, \ldots, x_{f'}\}$, $S = \{x_{f'+1}, x_{f+2}, \ldots, x_{c'}\}$, $C = \{x_{c'+1}, \ldots, x_N\}$.

In the sequel we describe an algorithm which applies a regional partitioning scheme on each of these categories separately. Several regional partitioning schemes for the classical VRP are presented in the literature, some of them are asymptotically optimal, see Karp (1977) and Haimovich and Rinnooy Kan (1985). However, the cost function of the problem considered here, is much more intricate than in the classical VRP as it depends not only on the routing cost but also on the retailers' holding cost rates.

Apparently none of the existing regional partitioning schemes can be guaranteed to preserve the asymptotic optimality convergence when directly applied to the above problem. However, by making use of the cost function structure we develop an asymptotic optimal heuristic that partitions the plane in accordance to both the retailers' location as well as their holding cost rate.

Our heuristic is based on the CRP scheme proposed by Haimovich and Rinnooy Kan (1985) that partitions the points in the plane according to their location. As will be shown below, this procedure does yield an asymptotically optimal solution when applied on each of the categories F and C separately but for S a new approach is needed. The underlying difference between F and C, on one hand, and S, on the other hand, lies in the lower bound expression $\underline{V}_{M^*}$: the frequency (capacity) constraints are tight for each of the demand points, and also for any subset, in F (C). Therefore, regardless of the grouping of these demand points, all of them share the same order interval f^{*-1} (b/M^*). In addition, it is simple to see that for any subset X_ℓ of F (C), $\underline{V}_{M^*}(X_\ell)$ is a linear function of $\Sigma_{x_\ell} h_i$ and $\Sigma_{x_\ell} r_i^c$.

Suppose now that X_ℓ is a subset of F (C) of cardinality M^*, and all its demand points are geographically close one to the other, i.e., the total route length $\text{TSP}(X_\ell^0) \sim 2R_\ell/M^*$. Thus, substituting $2R_\ell^c/M^*$ be $\text{TSP}(X_\ell^0) + c$ in the cost function almost does not effect the optimal reorder interval f^{*-1} (b/M^*). Therefore, also the average holding cost of X_ℓ will remain almost unchanged, i.e.,

$$\sum_{x_\ell} \underline{V}_{M^*}(x_i) = \underline{V}_{M^*}(X_\ell) \sim g\big(\text{TSP}(X_\ell^0) + c,\ H_\ell,\ M^*\big).$$

In other words, the problem of finding a good approximation for F (C) boils down to applying good *regional* partitioning schemes (as the CRP) to solve the CVRP on these sets.

The holding cost plays a more critical role in S where neither the frequency nor the capacity constraints seem to be tight. There, an aggregation based solely on the points' location may result in an extremely poor policy. For these demand points, we propose a technique which first aggregates demand points with similar r_i^c/h_i ratios into large groups and then the (CRP) scheme is applied on each of these groups, separately. We call this scheme the 'Two-Stage Regional Partitioning' scheme.

We next describe the CRP scheme for solving the CVRP as proposed by Haimovich and Rinnooy Kan (1985). Suppose n points in the plane are to be partitioned into regions containing at most M^* points each. Let $r_1, \ldots, r_n$ be the radial distances of the points, $r_{\max} = \max r_i$ and $\bar{r} = \Sigma_{i=1}^n r_i/n$. The CRP first partitions the plane into $o(\sqrt{n})$ equal sectors and then partitions each of the sectors by radial cut into $o(\sqrt{n})$ regions containing exactly M^* points:

The Circular Regional Partitioning scheme (CRP).

Step 1. Partition the circle into $t = \lceil((4\pi\Sigma_{i=1}^n r_i)/(3M^* r_{\max}))^{1/2}\rceil$ disjoint equal sectors.

Step 2. Partition each sector into regions by circular cuts, such that all of them, except possibly the one closest to the center, contain exactly M^* points.

Step 3. Repartition the group of at most t subregions closest to the center and containing less than M^* points each by radial cuts into at most $t-1$ subregions with M^* points each, and at most one subregion containing less than M^* points. Let X_ℓ, $\ell = 1, \ldots, \lceil n/M^*\rceil$, be the generated subregions.

Step 4. In each subregion X_ℓ find $\text{TSP}(X_\ell^0)$, $\ell = 1, \ldots, \lceil n/M^*\rceil$.

We now review some of the results in Haimovich and Rinnooy Kan (1985) that we use in the sequel. Let $\Pi^{\text{RP}}(X)$ ($\Pi^{\text{CRP}}(X)$) be the total perimeter length over all the subregions generated by any regional (circular) partitioning scheme when applied on X.

Lemma 3 (see Lemma 6 in Haimovich and Rinnooy Kan, 1985). *For any regional partitioning* (RP) *scheme which generates the regions* $X_1, \dots, X_L$, $L = \lceil n/M^* \rceil$, *the following inequality holds:*

$$\sum_{\ell=1}^{\lceil n/M^* \rceil} \mathrm{TSP}(X_\ell^0) \le 2(n/M^*)\bar{r} + 2r_{\max} + \mathrm{TSP}(X) + \tfrac{3}{2}\Pi^{\mathrm{RP}}(X).$$

Lemma 4 (see Lemma 5 in Haimovich and Rinnooy Kan, 1985).

$$\Pi^{\mathrm{CRP}}(X) \le 4\sqrt{3\pi(n/M^*)r_{\max}\bar{r}} + (3 + 2\pi)r_{\max}.$$

Lemma 5 (see Theorem 3 in Haimovich and Rinnooy Kan, 1985).

$$\mathrm{TSP}(X) \le 2(\pi n r_{\max}\bar{r})^{1/2} + (2 + \pi)r_{\max}.$$

In the sequel we explain how the CRP can be adapted for the solution of the above described problem. In the analysis we use the indicator set-up for any set X_0.

$$1_{(X_0)} = \begin{cases} 1 & \text{if } X_0 \text{ is not an empty set,} \\ 0 & \text{otherwise.} \end{cases}$$

The set C

Applying the CRP on C results in a partition of C into clusters of M^* demand points each, except possibly one. Let $C_1, \dots, C_{L_c}$ be the generated clusters such that $|C_\ell| = M^*$, $\ell = 2, \dots, L_c$, and $|C_1| \le M^*$. In the proposed heuristic each of the sets C_ℓ is served by a fully loaded truck. Let $U(C_\ell)$ denote the associated average cost of such a policy, i.e.

$$U(C_\ell) \stackrel{\mathrm{def}}{=} \frac{|C_\ell|}{b}\left(\mathrm{TSP}(C_\ell^0) + c\right) + \frac{b}{2|C_\ell|}\sum_{C_\ell} h_i.$$

The next theorem provides a bound on the difference between the average cost of serving $C_1, \dots, C_{L_c}$ by fully loaded trucks and the respective lower bound on C.

Theorem 4. *Let* $\{C_1, \dots, C_{L_c}\}$, *with* $L_c = \lceil |C|/M^* \rceil$, *be the subregions generated by applying the CRP scheme on C such that* $|C_\ell| = M^*$, $\ell = 2, \dots, L_c$, *and* $|C_1| \le M^*$. *Then*

$$\sum_{\ell=1}^{L_c} U(C_\ell) - \underline{V}_{M^*}(C) \le \frac{M^*}{b}\pi^{1/2}\left(2 + 6\left(\frac{3}{M^*}\right)^{1/2}\right)|C|^{1/2}r_{\max} + \left[\frac{M^*}{b}(8.5 + 4\pi)r_{\max} + \frac{b}{2}h_{\max}\right]1_{(C)}.$$

Proof. See Appendix B.

The set F

We apply the CRP also on the set F, resulting in $L_F = \lceil |F|/M^* \rceil$ subregions $\{F_1, \dots, F_{L_F}\}$ where all of them, except possibly F_1, contain exactly M^* demand points. The proposed heuristic serves all of these subregions at the maximum frequency f^*. We denote the average cost of serving F_ℓ at frequency f^* by $U(F_\ell)$, where

$$U(F_\ell)=\sum_{F_\ell} h_i/(2f^*)+f^*\big(\mathrm{TSP}\big(F_\ell^0\big)+c\big).$$

The following theorem provides a bound on the performance of the heuristic on set F:

Theorem 5. *Let* $\{F_1,\ldots,F_{L_F}\}$, *with* $L_F=\lceil |F|/M^*\rceil$, *be the subregions generated by applying the CRP scheme on F such that* $|F_\ell|=M^*$, $\ell=2,\ldots,L_F$, *and* $|F_1|\le M^*$. *Then*

$$\sum_{\ell=1}^{L_F} U(F_\ell)-\underline{V}_{M^*}(F)\le f^*\pi^{1/2}\left(2+6\left(\frac{3}{M^*}\right)^{1/2}\right)|F|^{1/2}r_{\max}+f^*((8.5+4\pi)r_{\max}+c)1_{(F)}.$$

Proof. See Appendix B.

The set S

The lower bound expression on the set S (see (8b)) is non-linear in the holding costs. Below we show that for a partitioning scheme to perform well relatively to the lower-bound, the demand points in each of the regions should 1) be geographically close one to the other, and 2) have similar r_i^c/h_i ratios. Therefore, pure regional partitioning schemes as the CRP may be inefficient. We propose a two-stage regional partitioning scheme that generates $\lceil |S|/M^*\rceil$ subregions, each (except possibly one) consisting of exactly M^* demand points: in the first stage the demand points in S are partitioned into α consecutive clusters $\{S_1,\ldots,S_\alpha\}$ according to the ratios r_i^c/h_i; in the second stage the CRP procedure is applied on each of these clusters separately. By a careful design of the partitioning scheme we obtain a heuristic that is both asymptotically optimal and have a bounded performance ratio.

Let

$$R\overset{\text{def}}{=}\Big(\max_S r_i^c/h_i\Big)/\Big(\min_S r_i^c/h_i\Big)\ \text{ and }\ \alpha=\Big\lceil |S|^{1/2}\Big\rceil.$$

Our assumptions ensure that R is finite since for any demand point x_i, (a) $r_i^c/h_i\le\min\{b^2/(4M^*), r_{\max}^c/h_{\min}\}<\infty$; and (b) r_i^c/h_i is bounded from below by a positive constant as follows:

$$r_i^c/h_i\ge\begin{cases} r_{\min}/h_{\max} & \text{if } f^*=\infty \text{ and } c=0,\\ \max\{c/(2h_{\max}),\ M^*/(4f^{*2})\} & \text{otherwise.}\end{cases}$$

Therefore, we can write that $R\le\tilde{R}<\infty$ where the explicit form of $\tilde{R}$ is given in Appendix C, see (C.1). In the first stage of the algorithm we partition the demand points of S into α consecutive and disjoint clusters so that the ratio between the maximal and the minimal r_i^c/h_i in each cluster does not exceed $R^{1/\alpha}$. In the second stage, the CRP scheme is applied on each of the clusters separately: L_k regions are generated from the k-th cluster, each containing exactly M^* demand points except possibly the L_k-th region that may contain less than M^* demand points. We then combine all the regions containing *less* than M^* demand points into one set and apply again the CRP scheme on this set. The algorithm is formally stated below:

The Two-Stage Partitioning Algorithm for S.
Step 0. $\alpha=\lceil |S|^{1/2}\rceil$; $\lambda:=R^{1/\alpha}$; $R_0:=\min_s(r_i^c/h_i)$; $R_k:=\lambda^k R_0$, $k=1,\ldots,\alpha$; $(R_\alpha=\max_s(r_i^c/h_i).)$
Step 1. $S_k:=\{x_i\mid x_i\in S$ and $R_{k-1}<r_i^c/h_i\le R_k\}$, $k=1,\ldots,\alpha$; $S_1:=S_1\cup\{x_i\mid x_i\in S$ and $r_i^c/h_i=R_0\}$;

Step 2. Apply the CRP on each cluster S_k, $k = 1, \dots, \alpha$, separately. Denote the subregions generated by $\{S_{k,1}, \dots, S_{k,L_k}\}$ where

$$L_k := \lceil |S_k|/M^* \rceil,\ |S_{k,L_k}| \le M^*, \text{ and } |S_{k,j}| = M^*,\ j = 1, \dots, L_k - 1.$$

Step 3. $S_0 := \bigcup\{S_{k,L_k} | 1 \le k \le \alpha \text{ and } |S_{k,L_k}| < M^*\}$. Apply the CRP on S_0 and let the generated subregions be $\{S_{0,1}, \dots, S_{0,L_0}\}$ where

$$L_0 = \lceil |S_0|/M^* \rceil,\ |S_{0,j}| = M^*,\ j = 2, \dots, L_0, \text{ and } |S_{0,1}| \le M^*.$$

Step 4. For $k = 1, \dots, \alpha$ do: if $|S_{k,L_k}| < M^*$ set $S_k := S_k - S_{k,L_k}$ and $L_k := L_k - 1$.
Step 5. For each of the subregions $S_{k,j}$, $0 \le k \le \alpha$, $1 \le j \le L_k$, compute the optimal policy according to (2).

(In Step 4 we update the set S_k by deleting all demand points that were transferred to S_0.)

The following example demonstrates the regional partitioning scheme described above: consider row 13 from Table 2, with 1000 retailers, a total of 5501 demand points, $M^* = 4$, $b = 6.4$ and $f^* = 1$. According to the output 3825 demand points have ratio $r_i^c/h_i < M^*/(4f^{*2}) = 1$ and thus belong to F. The CRP scheme with $M^* = 4$ is applied on these points resulting in 956 regions consisting of 4 demand points each and a single region consisting of a single demand point. All of these regions are replenished at maximum frequency of once per time unit. No demand point has ratio $r_i^c/h_i > b^2/(4M^*) = 2.56$ thus $C = \emptyset$. 1676 demand points fall in the set S and they are partitioned into $\alpha = \sqrt{1676} = 41$ sets. Assuming that $R = 2.56/1 = 2.56$ we obtain that $\lambda = 2.56^{1/41} = 1.023$. Therefore, the first set consists of all demand points satisfying $1 \le r_i^c/h_i \le 1.023$, the second set consists of all demand points satisfying $1.023 < r_i^c/h_i \le 1.046, \dots,$ and the last set consists of all demand points satisfying $2.502 < r_i^c/h_i < 2.56$. The CRP scheme with $M^* = 4$ is then applied on each of these sets separately; all demand points of the S falling into regions of less than 4 demand points are combined together into a single set and the CRP scheme is applied on it resulting in 419 regions each consisting of exactly 4 demand points. In overall the scheme produces 1375 regions of 4 demand points and one region of a single demand point.

The next two lemmas will be used in the performance analysis of the proposed heuristic:

Lemma 6 (Inmann and Jones, 1987). *Given a sequence of paris of real numbers $\{(a_i, b_i)\}_{i=1}^n$ with $a_i > 0$, $b_i > 0$, $i = 1, \dots, n$, such that $a_1/b_1 \le \cdots \le a_n/b_n$, it follows that*

$$\frac{\left(\sum_{i=1}^n a_i \sum_{i=1}^n b_i\right)^{1/2}}{\sum_{i=1}^n (a_i b_i)^{1/2}} \le B(\delta) \text{ where } \delta \overset{\text{def}}{=} \frac{a_1/b_1}{a_n/b_n} \text{ and } B(\delta) = \sqrt{1 + \frac{(1-\delta^{1/2})^2}{2\delta}}. \tag{9}$$

It is easily verified from Lemma 6 that if all ratios a_i/b_i are identified, i.e., $\delta = 1$, then $B(\delta) = 1$ meaning that $(\Sigma a_i \Sigma b_i)^{1/2} = \Sigma(a_i b_i)^{1/2}$. Moreover, as δ decreases, $B(\delta)$ increases and $\lim_{\delta \downarrow 0} B(\delta) = \infty$. However the function B changes at an extremely slow rate when δ is close to one, for example, $B(0.75) = 1.006$, $B(0.5) = 1.042$ but $B(0.1) = 1.82$.

Recall that $\tilde{R} < \infty$ is a constant bounding $R = \max(r_i^c/h_i)/\min(r_i^c/h_i)$, $i \in S$ (see C.1).

Lemma 7. *Let the regions $\{S_{k,j} | 0 \le k \le \alpha,\ 1 \le j \le L_k\}$ be the subregions generated by the 'Two Stage Regional Partitioning Algorithm' on S. Then:*

(a) *For any k and j, $1 \le k \le \alpha$, $1 \le j \le L_k$, $(\Sigma_{S_{k,j}} r_i^c \Sigma_{S_{k,j}} h_i)^{1/2} / \Sigma_{S_{k,j}} (r_i^c h_i)^{1/2} \le B(\tilde{R}^{-1/\alpha})$.*
(b) *For $1 \le j \le L_0$, $(\Sigma_{S_{0,j}} r_i^c \Sigma_{S_{0,j}} h_i)^{1/2} / \Sigma_{S_{0,j}} (r_i^c h_i)^{1/2} \le B(\tilde{R}^{-1})$.*

Proof. Let $\delta_k^j \stackrel{\text{def}}{=} \min_{S_{k,j}}(r_i^c/h_i)/\max_{S_{k,j}}(r_i^c/h_i)$.

(a): By definition of the Two-Stage Regional Partitioning Algorithm, any subset $S_{k,j}$, $1 \le k \le \alpha$, $1 \le j \le L_k$ is a subset of S_k. Thus $S_{k,j}$ is a subset of $\{x_i \mid R_{k-1} \le r_i^c/h_i \le R_k\}$ which implies that $1 \ge \delta_k^i \ge R_{k-1}/R_k = \lambda^{-1} = R^{-1/\alpha} \ge \tilde{R}^{-1/\alpha}$. Since $B(\delta)$ is a decreasing function of δ for $0 < \delta \le 1$ it follows that $B(\delta_k^j) \le B(\tilde{R}^{-1/\alpha})$. Use Lemma 6 to complete the proof.

(b): The proof of part (b) is similar to that of (a), observing that $\delta_0^i \ge R^{-1} \ge \tilde{R}^{-1}$, $j = 1, \dots, L_0$. Thus $B(\delta_0^j) \le B(\tilde{R}^{-1})$. □

Denote by $U(S_{k,j})$ the average cost of the Two-Stage Partitioning Algorithms when applied on $S_{k,j}$, i.e. $U(S_{k,j}) = g(\text{TSP}(S_{k,j}) + c,\ \Sigma_{S_{k,j}} h_i,\ |S_{k,j}|)$.

Recall that $\underline{V}_{M^*}(S) = \Sigma_s (4 r_i^c h_i / M^*)^{1/2}$. Also note that by definition of S, any demand point in S satisfies $M^*/(4f^2) \le r_i^c/h_i \le b^2/(4M^*)$, thus for any subregion $S_{k,j}$, since $|S_{k,j}| \le M^*$, it holds that $|S_{k,j}|/(4f^2) \le \Sigma_{S_{k,j}} r_i^c / \Sigma_{S_{k,j}} h_i \le b^2/(4|S_{k,j}|)$. Therefore, in view of (5), the optimal average cost of $S_{k,j}$ satisfies the following inequality:

$$U(S_{k,j}) \ge g\left(\frac{2}{|S_{k,j}|} \Sigma_{S_{k,j}} r_i^c,\ \Sigma_{S_{k,j}} h_i, |S_{k,j}|\right) = G\left(\Sigma_{S_{k,j}} r_i^c,\ \Sigma_{S_{kj}} h_i,\ |S_{kj}|\right)$$
$$= \left(\frac{4}{|S_{k,j}|} \Sigma_{S_{k,j}} r_i^c \Sigma_{S_{k,j}} h_i\right)^{1/2}.$$

For $0 \le k \le \alpha$, $1 \le j \le L_k$, define

$$W(S_{k,j}) = G\left(\sum_{S_{k,j}} r_i^c,\ \sum_{S_{k,j}} h_i, |S_{kj}|\right) = \left(\frac{4}{|S_{k,j}|} \sum_{S_{k,j}} r_i^c \sum_{S_{k,j}} h_i\right)^{1/2} \tag{10}$$

In order to bound the gap between the average cost of the heuristic and the respective lower bound on S we present this gap as the sum of two differences, as follows:

$$\sum_{k=0}^{\alpha} \sum_{j=1}^{L_k} U(S_{k,j}) - \underline{V}_{M^*}(S) = \sum_{k=0}^{\alpha} \sum_{j=1}^{L_k} \left(U(S_{k,j}) - W(S_{k,j})\right)$$
$$+ \sum_{k=0}^{\alpha} \sum_{j=1}^{L_k} \left(W(S_{k,j}) - \underline{V}_{M^*}(S_{k,j})\right), \tag{11}$$

and bound each of the differences in the r.h.s. separately. The next theorem bounds the second difference in (11).

Theorem 6. *Let the partition $\{S_{k,j} \mid 0 \le k \le \alpha,\ 1 \le j \le L_k\}$ be obtained by applying the Two-Stage Regional Partitioning Algorithm on S. Then*

$$\sum_{k=0}^{\alpha} \sum_{j=1}^{L_k} W(S_{k,j}) - \underline{V}_{M^*}(S) \le 2\left[|S|\left(B(\tilde{R}^{-1/\alpha}) - 1\right) + \left(|S|^{1/2} + 1\right)\left(B(\tilde{R}^{-1}) - 1\right)\right]\left(\frac{r_{\max}^c h_{\max}}{M^*}\right)^{1/2}$$
$$+ 2(M^* r_{\max} h_{\max})^{1/2}. \tag{12}$$

Proof. See Appendix C.

It is worthnoty that as $|S|$ increases to infinity $B(\tilde{R}^{-1/\alpha}) - 1$ decreases exponentially fast to zero (since $\alpha = \lceil |S|^{1/2} \rceil$). Thus the dominating term in (12) is of order $O(|S|^{1/2})$.

Let $\Pi^{(2)}(S)$ be the total perimeter of the subregions generated by applying the Two-Stage Regional Partitioning Algorithm on S. The following lemma provides an upper bound on the total length of $\Pi^{(2)}(S)$:

Lemma 8.

$$\Pi^{(2)}(S) \le \left\{ 4\left(\frac{3\pi}{M^*}\right)^{1/2} |S|^{3/4} + \left[4\left(\frac{6\pi}{M^*}\right)^{1/2} + (3+2\pi) \right] |S|^{1/2} + (6+4\pi) \right\} r_{\max}.$$

Proof. See Appendix C.

In the next theorem we bound the gap between the average cost of the heuristic and the lower bound. We first need the following lemma. For the proofs of the lemma and the theorem see Appendix C. Recall the assumption that if $c=0$ and $f^*=\infty$, then $r_{\min}>0$.

Lemma 9. *Let D be any subset of S of cardinality m, $m \le M^*$. Then there exists a constant $\beta < \infty$ such that*

$$g\Big(\mathrm{TSP}(D^0)+c,\ \sum_D h_i,\ m\Big) - g\Big(2\sum_D r_i^c/m,\ \sum_D h_i,\ m\Big) \le \beta\Big(\mathrm{TSP}(D^0) - 2\sum_D r_i/m\Big)$$

where

$$\beta = \begin{cases} \frac{1}{2}(M^* h_{\max}/r_{\min})^{1/2} & \text{if } f^*=\infty \text{ and } c=0, \\ \min\left\{ \left(\frac{1}{2}M^* h_{\max}/c\right)^{1/2},\ f^* \right\} & \text{otherwise}. \end{cases}$$

Theorem 7. *Let $\{S_{k,j} \mid 0 \le k \le \alpha,\ 1 \le j \le L_k\}$ be the subregions generated by applying the Two-Stage Regional Partitioning Algorithm on S and $W(S_{k,j})$ be defined in (10). Then*

(a) $$\sum_{k=0}^{\alpha} \sum_{j=1}^{L_k} \left\{ g\Big(\mathrm{TSP}(S_{k,j}^0) + c, \sum_{S_{kj}} h_i, |S_{k,j}|\Big) - W(S_{k,j}) \right\} \le \beta\Big[6(3\pi/M^*)^{1/2}|S|^{3/4} + 43.55|S|^{1/2} + 35\Big] r_{\max}$$

(see Lemma 9 for the definition of β).

(b) $$\sum_{k=0}^{\alpha} \sum_{j=1}^{L_k} g\Big(\mathrm{TSP}(S_{k,j}^0) + c,\ \sum_{S_{k,j}} h_i,\ |S_{k,j}|\Big) - \underline{V}_{M^*}(S) \le \gamma_1 |S|^{3/4} + \gamma_2 |S|^{1/2} + \gamma_3 |S| (1 + \bar{R}^{1/\alpha} - 2\bar{R}^{1/(2\alpha)}) + 1_{(S)}\gamma_4,$$

where $\alpha = \lceil |S|^{1/2} \rceil$,

$$\gamma_1 = 6\beta\left(\frac{3\pi}{M^*}\right)^{1/2} r_{\max}, \qquad \gamma_2 = 43.55\beta r_{\max} + 2\left(\frac{r_{\max}^c h_{\max}}{M^*}\right)^{1/2}\big(B(\bar{R}^{-1}) - 1\big),$$

$$\gamma_3 = 0.5\left(\frac{r_{\max}^c h_{\max}}{M^*}\right)^{1/2}, \qquad \gamma_4 = 35\beta r_{\max} + 4\gamma_3\big(M^* + (B(\bar{R}^{-1}) - 1)\big),$$

and $B(\delta)$ is as defined in (9).

Combining the results of this section we conclude that the proposed algorithm yields an average cost which is bounded from above by $\bar{V}(X)$:

$$\bar{V}(X) = \underline{V}_{M^*}(X) + \frac{M^*}{b}\sqrt{\pi}\left(2 + 6\left(\frac{3}{M^*}\right)^{1/2}\right)|C|^{1/2} r_{\max}$$
$$+ f^*\sqrt{\pi}\left(2 + 6\left(\frac{3}{M^*}\right)^{1/2}\right)|F|^{1/2} r_{\max} + \gamma_1 |S|^{3/4} + \gamma_2 |S|^{1/2}$$
$$+ \gamma_3 |S|(1 + \tilde{R}^{1/\alpha} - 2\tilde{R}^{1/(2\alpha)})$$
$$+ \left[\frac{M^*}{b}(8.5 + 4\pi) r_{\max} + \frac{b}{2} h_{\max}\right] 1_{(c)} + 1_{(F)} f^*((8.5 + 4\pi) r_{\max} + c) + 1_{(S)}\gamma_4 \tag{13}$$

(see Theorems 4, 5 and 7). The next theorem is immediate:

Theorem 8. *For any set X containing N demand points $X = \{x_1, \ldots, x_N\}$, $\underline{V}_{M^*}(X) \le V^*(X) \le \bar{V}(X)$, where $\underline{V}_{M^*}(X) = \underline{V}_{M^*}(F) + \underline{V}_{M^*}(S) + \underline{V}_{M^*}(C)$ as defined in* (8) *and $\bar{V}(X)$ is defined in* (13).

Next we summarize the overall proposed algorithm which we call the 'Multi-Retailer EOQ with Routing Costs'.

The Multi-Retailer EOQ with Routing Costs.

Step 1. Find the largest index f' such that $r_{f'}^c/h_{f'} < M^*/(4f^{*2})$. Set $F := \{x_1, \ldots, x_{f'}\}$;
Find the largest index c' such that $r_{c'}^c/h_{c'} \le b^2/(4M^*)$. Set $S := \{x_{f'+1}, \ldots, x_{c'}\}$, $C := \{x_{c'+1}, \ldots, X_N\}$. Calculate $\underline{V}_{M^*}(X)$.

Step 2. Apply the CRP on C and F separately and the Two-Stage Regional Partitioning Scheme on S. Let $\{Y_1, \ldots, Y_{K_L}\}$, $\lceil N/M^* \rceil \le K_L \le \lceil N/M^* \rceil + 3$, be the collection of the subregions generated.

Step 3. Calculate $\bar{V}(X)$ (see (13)).

Step 4. Find the optimal traveling salesman tour (see Remark below) through the depot and the retailers in each of the subregions generated in Step 2, i.e., calculate $\text{TSP}(Y_k^0)$, $k = 1, \ldots, K_L$.

Step 5. Determine the replenishment interval $T(Y_k)$ and the load quantity $Q(Y_k)$ for each of the subregions as follows:

While $k \le K_L$ do begin $H(Y_k) := \sum_{Y_k} h_i$;
  If $(\text{TSP}(Y_k^0) + c)/H(Y_k) < 1/(2f^{*2})$ then $T(Y_k) := f^{*-1}$;
  If $1/(2f^{*2}) \le (\text{TSP}(Y_k^0) + c)/H(Y_k) \le b^2/(2|Y_k|^2)$ then
    $T(Y_k) := (2(\text{TSP}(Y_k^0) + c)/H(Y_k))^{1/2}$;
  If $(\text{TSP}(Y_k^0) + c)/H(Y_k) > b^2/(2|Y_k|^2)$ then $T(Y_k) := b/|Y_k|$;
$Q(Y_k) = T(Y_k)|Y_k|$
endwhile;

Remark. Computing the *optimal* traveling salesman tours in each of the subregions in Step 4 might be time-consuming if the number of retailers in a subregion exceeds 7 or 8. Instead, one can use any heuristic whose worst-case relative error is bounded (for example, Christofides' algorithm, see Christofides, 1976). This will not affect the asymptotic optimality property of the algorithm.

The asymptotic analysis of the proposed heuristic is carried out on the set of retailers. Observe that the demand points are partitioned to the sets F, S and C according to the ratios r_i^c/h_i. Thus any retailer belongs to exactly one of these sets. Therefore the partition of X into F, S and C can also be viewed as a partition of the set of retailers.

Let y_j be a retailer consisting of μ_j demand points. Let (r_j, h_j) denote radial distance and the holding cost rate of the j-th retailer $\underline{V}_{M^*}(y_j) = \mu_j\ G_{M^*}(r_j^c, h_j)$, $1 \le j \le n$. Thus, $Y = \{y_1, y_2, \dots, y_n\}$ represents the set of retailers and denote by $X(Y)$ the set of demand points associated with Y. Then

$$\underline{V}_{M^*}(X(Y)) = \sum_{j=1}^{n} \mu_j G_{M^*}(r_j^c, h_j).$$

Theorem 9. *Let $Y(n) = \{y_1, y_2, \dots, y_n\}$ be a random set of retailers each characterized by a triplet (r_i, h_i, μ_i) and its polar coordinate. The triplets (r_i, h_i, μ_i) are assumed to be i.i.d. and moreover μ_i is a discrete random variable assuming positive integer values which is independent of (r_i, h_i). Let $\xi = E(G_{M^*}(r_i^c, h_i))$ and $d = E(\mu_i)$. Then*
(a) $\lim_{n\to\infty}(1/n)\underline{V}_{M^*}(X(Y(n))) = d\xi$ a.s.
(b) *If, in addition, $\xi > 0$ and the random variables r_i, h_i, μ_i are uniformly bounded as follows:*
(i) $r_i \le r_{\max} < \infty$ *and if* $c = 0$, $r_i \ge r_{\min} > 0$; (ii) $0 < h_{\min} \le h_i \le h_{\max} < \infty$; (iii) $1 \le \mu_i \le K < \infty$,
then

$$\lim_{n\to\infty} \frac{\bar{V}(X(Y(n))) - \underline{V}_{M^*}(X(Y(n)))}{\underline{V}_{M^*}(X(Y(n)))} = 0 \text{ a.s}$$

Proof. (a): $\underline{V}_{M^*}(X(Y(n))) = \sum_{i=1}^{n} \mu_i G_{M^*}(r_i^c, h_i)$. Part (a) follows from the fact that μ_i is independent of (r_i^c, h_i) and the law of large numbers.

(b): In view of part (a), Theorem 8 and the assumption that $\xi > 0$ and $d > 0$ it suffices to show that

$$\lim_{n\to\infty} \frac{1}{n}\left[\bar{V}(X(Y(n))) - \underline{V}_{M^*}(X(Y(n)))\right] = 0 \text{ a.s.}$$

Using (13) we can write

$$\bar{V}(X(Y(n))) - \underline{V}_{M^*}(X(Y(n))) = \alpha_0 + \alpha_c|C|^{1/2} + \alpha_F|F|^{1/2} + \gamma_1|S|^{3/4} + \gamma_2|S|^{1/2} + \gamma_3|S|(1 + \tilde{R}^{1/\alpha} - 2\tilde{R}^{1/(2\alpha)})$$

where $\alpha = \lceil |S|^{1/2} \rceil$ and the constants γ_1, γ_2, γ_3, γ_4 are defined in Theorem 7 and

$$\alpha_0 = 1_{(S)}\gamma_4 + 1_{(C)}\left(\frac{b}{2}h_{\max} + \frac{M^*}{b}(8.5 + 4\pi)r_{\max}\right) + 1_{(F)}f^*((8.5 + 4\pi)r_{\max} + c),$$

$$\alpha_c = \frac{M^*}{b}\sqrt{\pi}\left(2 + 6\left(\frac{3}{M^*}\right)^{1/2}\right)r_{\max}, \qquad \alpha_F = f^*\sqrt{\pi}\left(2 + 6\left(\frac{3}{M^*}\right)^{1/2}\right)r_{\max}.$$

For any given set Y of n retailers the sets F, C and S define a partition of $X(Y(n))$, i.e., $|F| + |C| + |S| = \sum_{i=1}^{n} \mu_i \ge n$. Since the constants α_0, α_c, α_F, γ_1, γ_2 and γ_3 are all independent of n it remains to be shown that $\lim_{n\to\infty} |S|(1 + \tilde{R}^{1/\alpha} - 2\tilde{R}^{1/(2\alpha)})/n = 0$.

Fix a realization of the sequence $\{y_1, y_2, \dots\}$. Assume to the contrary that

$$\limsup_{n\to\infty} \frac{|S(n)|(1 + \tilde{R}^{1/\alpha(n)} - 2\tilde{R}^{1/(2\alpha(n))})}{n} = \varepsilon > 0. \tag{14}$$

Let $\{n_k\}_{k=1}^{\infty}$ be a sequence of integers such that

(i) $\lim_{k\to\infty}|S(n_k)|$ exists;

(ii) $\lim_{k\to\infty}(1/n_k)\left(|S(n_k)|\left(1+\tilde{R}^{1/\alpha(n_k)}-2\tilde{R}^{1/(2\alpha(n_k))}\right)\right)=\varepsilon.$

If $\lim_{k\to\infty}|S(n_k)|<\infty$ then also $\lim_{k\to\infty}\alpha(n_k)=\lim_{k\to\infty}|S(n_k)|^{1/2}<\infty$. Note that $\tilde{R}$ as defined in (C.1) is an upper bound (independent of n) on $R=\max_s(r_i^c/h_i)/\min_s(r_i^c/h_i)$ thus $\tilde{R}\geq 1$ and

$$\lim_{k\to\infty}\left(1+\tilde{R}^{1/\alpha(n_k)}-2\tilde{R}^{1/(2\alpha(n_k))}\right)\leq 1+\tilde{R},$$

yielding a contradiction to the assumption. In this case the left-hand side of (14) converges to zero as fast as $1/n$.

If $\lim_{k\to\infty}|S(n_k)|=\infty$ then also $\lim_{k\to\infty}\alpha(n_k)=\infty$. Thus,

$$\lim_{k\to\infty}\left(1+\tilde{R}^{1/\alpha(n_k)}-2\tilde{R}^{1/(2\alpha(n_k))}\right)=0.$$

Since $|S(n)|/n\leq K<\infty$ where K is the upper limit on the demand rate of the retailers, we obtain that the l.h.s. of (14) converges to zero exponentially fast leading again to a contradiction. □

In overall we conclude that

$$\frac{\bar{V}(X(Y(n)))-\underline{V}_{M^*}(X(Y(n)))}{\underline{V}_{M^*}(X(Y(n)))}$$

converges to zero as fast as $n^{-1/4}$.

As mentioned in Anily and Federgruen (1990b), the assumption that $\{y_1, y_2,\ldots\}$ are i.i.d. is needed only to assure that the lower bound $\underline{V}_{M^*}$ grows linearly in n almost surely. The assumption is needed to preclude heavy concentration of points with small $G_{M^*}(r_i^c, h_i)$-values.

Similarly, the condition that the radial distances and the holding cost rates are uniformly bounded from above is unnecessarily strong. One merely needs that $r_{\max}$ and $h_{\max}$ do not grow 'too fast' as $n\to\infty$ (a.s.). Much simpler conditions with respect to the joint distribution of (r, h) may be invoked, see David (1970).

5. An experimental study

In this section we report on a numerical study conducted to assess the effectiveness of the proposed heuristic. The algorithm was encoded in Fortran (Tops 20-Version 2) and run on an Amdahl 170V8 computer. We analyze both capacitated and uncapacitated systems and we vary the frequency limit. The upper bound $\bar{M}$ on the sales volume per region was chosen to be either 4 or 7. Note that usually a single vehicle will not serve more than 2–4 locations in a single tour. The retailers' locations were randomly generated according to a uniform distribution in a square of 200×200 with the depot placed in the center. The demand rates of the retailers are generated from a uniform distribution on the integers $1,\ldots,10$. In all runs with 100 (1000) retailers we use the same sequence of locations and demands. The holding cost rates of the retailers were uniformly generated on the interval [50.0, 150.0].

The lower bound value LB reported in Table 2 is $\underline{V}_{M^*}(X)$ where the upper bound UB is the average cost resulting from the proposed algorithm. We also report the cardinalities of the sets F, S and C as indicators for the tightness of the capacity and frequency constraints.

As can be seen from Table 2, the algorithm generates good solutions even for problems of moderate size, e.g. $n=100$. The gap between the upper and lower bounds does not exceed 10% and is usually much lower (see the values UB/LB). Moreover, the computational time is extremely low. As can be observed by comparing problems with $n=1000$ $M^*=4$ and $M^*=7$ the computational time grows quickly with M^*. Indeed it increases by a factor of approximately 1000. This indicates that most of the

Table 2
Experimental study

n	N	M^*	b	f^*	$\|F\|$	$\|S\|$	$\|C\|$	UB/LB	CPU time in milliseconds
100	593	4	∞	1	421	172	0	1.036	0.12694
100	593	4	∞	5	1	592	0	1.075	0.17130
100	593	4	∞	10	1	592	0	1.075	0.17228
100	593	4	6.4	1	421	172	0	1.036	0.12079
100	593	4	6.4	5	1	592	0	1.075	0.17070
100	593	4	6.4	10	1	592	0	1.075	0.17031
100	593	3	3.2	1	281	72	240	1.028	0.08996
100	593	4	3.2	5	1	248	344	1.046	0.12874
100	593	4	3.2	10	1	248	344	1.046	0.12810
1000	5501	4	∞	1	3825	1676	0	1.023	3.15734
1000	5501	4	∞	5	1	5500	0	1.045	4.32912
1000	5501	4	∞	10	1	5500	0	1.045	4.38116
1000	5501	4	6.4	1	3825	1676	0	1.023	3.16105
1000	5501	4	6.4	5	1	5500	0	1.045	4.36226
1000	5501	4	6.4	10	1	5500	0	1.045	4.36018
1000	5501	3	3.2	1	2531	648	2322	1.015	2.43309
1000	5501	4	3.2	5	1	1820	3680	1.025	3.09060
1000	5501	4	3.2	10	1	1820	3680	1.025	3.12366
1000	5501	7	∞	1	5354	147	0	1.013	3328.76
1000	5501	7	∞	5	20	5481	0	1.095	6223.50
1000	5501	7	∞	10	6	5495	0	1.098	6504.50
1000	5501	6	6.4	1	5045	264	192	1.017	3000.70

computational time is due to the computation of the Traveling Salesman tours by full enumeration (see Remark in previous section for alternative methods that use heuristics for calculating the optimal Traveling Salesman tours).

6. Conclusions

The paper considers an infinite-horizon replenishment problem of several retailers from a single warehouse with transportation cost and retailer-specific linear holding costs. The retailers are assumed to face constant demand rates. The stock is distributed via a fleet of identical (possibly capacitated) vehicles. In addition, the model allows for bounds on the frequency at which the routes are driven. The retailers are assumed to be dispersed in the Euclidean plane. We restrict ourselves to a class of policies that specifies a collection of regions that covers all outlets: if an outlet belongs to several regions a specific fraction of its sales/operations is assigned to each of these regions and deliveries to different regions are not coordinated. We propose a regional partitioning procedure which is shown to converge to the optimal solution in the above class of policies, when the number of retailers is sufficiently large. In addition, we present an experimental study which demonstrates the algorithm's efficiency on relatively small systems.

It is interesting to note that the geographic regions generated by the proposed regional partitioning scheme may overlap. The retailers are first partitioned according to the ratios *radial distance / holding cost rate* and then according to their geographic location. In this way we ensure that the retailers in a single region are in close proximity one to the other and that they have similar holding cost rates. As a result, the effect of the joint replenishment on the holding cost is negligible relative to the savings on the routing costs.

Appendix A

Lemma 2.

(a) *The set-function G as defined in (5) can be equivalently represented by one of the following forms:*
If $R^c/H \le b/(4f^*)$ *then*

$$G(R^c, H, m) = \begin{cases} (4HR^c/m)^{1/2} & \text{if } m \le 4f^{*2}R^c/H, \\ H/(2f^*) + 2f^*R^c/m & \text{if } 4f^*2R^c/H < m \le M^*. \end{cases}$$

Otherwise

$$G(R^c, H, m) = \begin{cases} (4HR^c/m)^{1/2} & \text{if } m \le b^2H/(4R^c), \\ Hb/(2m) + 2R^c/b & \text{if } b^2H/(4R^c) < m \le M^*. \end{cases}$$

(b) *The set-function G is non-increasing in m.*

Proof. (a): For fixed (R^c, H) and $m \le \min\{4f^{*2}R^c/H, b^2H/(4R^c)\}$, neither the capacity nor the frequency constraints are binding, see (5). If m is such that

$$\min\left\{\frac{4f^{*2}R^c}{H}, \frac{b^2H}{4R^c}\right\} < m \le \min\left\{M^*, \max\left\{\frac{4f^{*2}R^c}{H}, \frac{b^2H}{4R^c}\right\}\right\}$$

and $b^2H/(4R^c)$ is smaller (greater) than $4f^{*2}R^c/H$, the capacity (frequency) constraint is binding. In order to complete the proof, it is sufficient to show that any m greater than $\max\{4f^{*2}R^c/H, b^2H/(4R^c)\}$ is infeasible: suppose $m > \max\{b^2H/(4R^c), 4f^{*2}R^c/H\}$, then equivalently we can write that $b^2H/(2m^2) < 2R^c/m < H/(2f^{*2})$. As a consequence $b^2H/(2m^2) < H/(2f^{*2})$ which implies that $m > bf^* \ge M^*$, thus m is infeasible.

(b): Relaxing the integrality requirement on m one can verify that G is continuous in m, for $0 < m \le M^*$. Moreover, $(\partial G/\partial m)$ exists, is continuous and is non-positive everywhere on $0 < m < M^*$. Thus, for fixed R^c and H, G is non-increasing in m. □

Appendix B

Proof of Theorem 4. Note that

$$\sum_{\ell=1}^{L_c} U(C_\ell) - \underline{V}_{M^*}(C) = \frac{M^*}{b}\sum_{\ell=2}^{L_c}\left(\mathrm{TSP}(C_\ell^0) - \frac{2}{M^*}\sum_{C_\ell} r_i\right) + \frac{|C_1|}{b}\left(\mathrm{TSP}(C_1^0) - \frac{2}{|C_1|}\sum_{C_1} r_i\right)$$
$$+ \frac{b}{2}\left[\frac{1}{|C_1|} + \frac{1}{M^*}\right]\sum_{C_1} h_i \le \frac{M^*}{b}\sum_{\ell=1}^{L_c}\left(\mathrm{TSP}(C_\ell^0) - \frac{2}{M^*}\sum_{C_\ell} r_i\right)$$
$$+ 1_{(C)}\frac{b}{2}h_{\max}.$$

By applying Lemma 3 on the l.h.s. of the above inequality we obtain that

$$\sum_{\ell=1}^{L_c} U(C_\ell) - \underline{V}_{M^*}(C) \le \left[(M^*/b)\left(2r_{\max} + \mathrm{TSP}(C) + (3/2)\Pi^{\mathrm{CRP}}(C)\right] + \tfrac{1}{2}bh_{\max}1_{(C)}.$$

With the help of Lemmas 4 and 5 and the fact that the average radial distance of any subset of X is no greater than $r_{\max}$ we can bound the r.h.s. of the last inequality by the r.h.s. of the desired inequality in Theorem 4. □

Proof of Theorem 5. By definitions of $\underline{V}_{M^*}(F)$ (see (8a)), $U(F_\ell)$ and Lemmas 3, 4 and 5 we obtain

$$\begin{aligned}\sum_{\ell=1}^{L_F} U(F_\ell) - \underline{V}_{M^*}(F) &= f^*\left(L_F c + \sum_{\ell=1}^{L_F} \mathrm{TSP}(F_\ell^0) - 2\sum_F r_i^c/M^*\right)\\ &= f^*\left((\lceil |F|/M^*\rceil - |F|/M^*)c + \sum_{\ell=1}^{L_F}\mathrm{TSP}(F_\ell^0) - 2|F|/M^*\left(\sum_F r_i/|F|\right)\right)\\ &\le f^*\left(2r_{\max}^c + \mathrm{TSP}(F) + 1.5\Pi^{\mathrm{CRP}}(F)\right)1_{(F)}\\ &\le f^*\Big(2r_{\max}^c + 2(\pi|F|)^{1/2} r_{\max} + (2+\pi)r_{\max} + 1.5\big(4(3\pi|F|/M^*)^{1/2} r_{\max}\\ &\qquad + (3+2\pi)r_{\max}\big)\Big)1_{(F)}\\ &= f^*\left(\pi^{1/2}\left(2+6(3/M^*)^{1/2}\right)|F|^{1/2} r_{\max} + (8.5+4\pi)r_{\max} + c\right)1_{(F)},\end{aligned}$$

which is the desired inequality. □

Appendix C

In this appendix we present some of the technical details regarding the analysis of the Two-Stage Regional Partitioning Scheme applied on the set S. We first present the explicit form $\tilde{R}$, the finite upper bound on $R = (\max r_i^c/h_i)/(\min r_i^c/h_i)$, $i \in S$:

$$\tilde{R} \stackrel{\mathrm{def}}{=} \begin{cases} (bf^*/M^*)^2 & \text{if } b<\infty \text{ and } f^*<\infty,\\ b^2 h_{\max}/(4M^* r_{\min}) & \text{if } b<\infty,\ f^*=\infty \text{ and } c=0,\\ b^2 h_{\max}/(2M^* c) & \text{if } b<\infty,\ f^*=\infty \text{ and } c>0,\\ 4f^{*2} r_{\max}^c/(M^* h_{\min}) & \text{if } b=\infty \text{ and } f^*<\infty,\\ r_{\max}^c h_{\max}/(r_{\min} h_{\min}) & \text{if } b=f^*=\infty \text{ and } c=0,\\ 2h_{\max} r_{\max}^c/(c h_{\min}) & \text{if } b=f^*=\infty \text{ and } c>0. \end{cases} \tag{C.1}$$

Proof of Theorem 6. For any k, $1 \le k \le \alpha$, $|S_{k,j}| = M^*$, $1 \le j \le L_k$. Thus,

$$\begin{aligned} W(S_{k,j}) - \underline{V}_{M^*}(S_{k,j}) &= \left(\frac{4}{M^*}\right)^{1/2}\left[\left(\sum_{S_{k,j}} r_i^c \sum_{S_{k,j}} h_i\right)^{1/2} - \sum_{S_{k,j}} (r_i^c h_i)^{1/2}\right]\\ &\le \left(\frac{4}{M^*}\right)^{1/2}\left(B(\tilde{R}^{-1/\alpha}) - 1\right)\sum_{S_{k,j}} (r_i^c h_i)^{1/2}, \end{aligned}$$

where the inequality follows from Lemma 7. Thus,

$$\begin{aligned} \sum_{k=1}^{\alpha}\sum_{j=1}^{L_k} W(S_{k,j}) - \underline{V}_{M^*}(S-S_0) &\le (4/M^*)^{1/2}\left(B(\tilde{R}^{-1/\alpha}) - 1\right)\sum_{S-S_0} (r_i^c h_i)^{1/2}\\ &\le 2|S|\left(B(\tilde{R}^{-1/\alpha}) - 1\right)\left(r_{\max}^c h_{\max}/M^*\right)^{1/2}. \end{aligned}$$

L_0 regions are generated from S_0, all of them have cardinality M^*, with the possible exception of $S_{0,1}$. Note that $L_0 \le \alpha \le \lceil |S|^{1/2} \rceil$. Therefore, for $2 \le j \le L_0$ we obtain from Lemma 7 that

$$W(S_{0,j}) - \underline{V}_{M^*}(S_{0,j}) \le (4/M^*)^{1/2}\big(B(\bar{R}^{-1}) - 1\big)\sum_{S_{0,j}}(r_i^c h_i)^{1/2}.$$

For $j = 1$,

$$W(S_{0,1}) - \underline{V}_{M^*}(S_{0,1}) = \left(\frac{4}{|S_{0,1}|}\right)^{1/2}\left(\sum_{S_{0,1}} r_i^c \sum_{S_{0,1}} h_i\right)^{1/2} - \left(\frac{4}{M^*}\right)^{1/2}\sum_{S_{0,1}}(r_i^c h_i)^{1/2}$$

$$\le \left(\frac{4}{M^*}\right)^{1/2}\big(B(\bar{R}^{-1}) - 1\big)\sum_{S_{0,1}}(r_i^c h_i)^{1/2}$$

$$+\left[\left(\frac{4}{|S_{0,1}|}\right)^{1/2} - \left(\frac{4}{M^*}\right)^{1/2}\right]\left(\sum_{S_{0,1}} r_i^c \sum_{S_{0,1}} h_i\right)^{1/2}.$$

But,

$$\left(\frac{4}{|S_{0,1}|}\right)^{1/2} - \left(\frac{4}{M^*}\right)^{1/2} \le \frac{2\big(M^{*1/2} - |S_{0,1}|^{1/2}\big)}{(|S_{0,1}|M^*)^{1/2}} \le \frac{2M^{*1/2}}{|S_{0,1}|}$$

and

$$\left(\sum_{S_{0,1}} r_i^c \sum_{S_{0,1}} h_i\right)^{1/2} / |S_{0,1}| \le (r_{\max}^c h_{\max})^{1/2}.$$

Therefore,

$$\sum_{j=1}^{L_0}\big(W(S_{0,j}) - \underline{V}_{M^*}(S_{0,j})\big) \le 2|S_0|(r_{\max}^c h_{\max}/M^*)^{1/2}\big(B(\bar{R}^{-1}) - 1\big) + 2(M^* r_{\max}^c h_{\max})^{1/2}.$$

Summation over the respective inequalities yields the desired inequality (12). □

Proof of Lemma 8. Note that the Two-Stage Regional Partitioning Algorithm applies the CRP scheme on each of the subsets S_j, $0 \le j \le \alpha$ ($\bigcup_{j=0}^{\alpha} S_j = S$), separately. In view of Lemma 4,

$$\Pi^{(2)}(S) \le \sum_{j=0}^{\alpha} \Pi^{\text{CRP}}(S_j) \le \sum_{j=0}^{\alpha}\left\{4\big(3\pi|S_j|/M^*\big)^{1/2} r_{\max} + (3 + 2\pi) r_{\max}\right\}$$

$$= \left\{4(3\pi/M^*)^{1/2}\sum_{j=0}^{\alpha}|S_j|^{1/2} + (\alpha + 1)(3 + 2\pi)\right\} r_{\max}.$$

Note that $\{S_j: 0 \le j \le \alpha\}$ is a partition of S into at most $\alpha + 1$ sets, i.e. $\sum_{j=0}^{\alpha}|S_j| = |S|$ and $0 \le |S_j| \le |S|$. Moreover, it is easy to verify that

$$\max\left\{\sum_{j=0}^{t} z_j^{1/2}: \sum_{j=0}^{t} z_j = z,\ 0 \le z_j \le z,\ 0 \le j \le t\right\} = (t+1)\left(\frac{z}{t+1}\right)^{1/2} = (z(t+1))^{1/2},$$

thus implying that

$$\sum_{j=0}^{\alpha}|S_j|^{1/2} \le (\alpha+1)^{1/2}|S|^{1/2} = \left(\left(\left\lceil |S|^{1/2}\right\rceil + 1\right)|S|\right)^{1/2} < \left(|S|^{3/2}+2|S|\right)^{1/2}$$

$$= |S|^{1/2}\left(|S|^{1/2}+2\right)^{1/2} \le |S|^{1/2}\left(|S|^{1/4}+\sqrt{2}\right) = |S|^{3/4} + (2|S|)^{1/2}.$$

Therefore,

$$\Pi^{(2)}(S) \le \left\{4(3\pi/M^*)^{1/2}\left(|S|^{3/4}+(2|S|)^{1/2}\right) + \left(|S|^{1/2}+2\right)(3+2\pi)\right\} r_{\max}$$

$$\le \left\{4(3\pi/M^*)^{1/2}|S|^{3/4} + \left[4(6\pi/M^*)^{1/2}+(3+2\pi)\right]|S|^{1/2} + (6+4\pi)\right\} r_{\max}. \quad \square$$

Proof of Lemma 9. Note that for any demand point x_i in S and $m \le M^*$,

$$m/\left(4f^{*2}\right) \le M^*/\left(4f^{*2}\right) \le r_i^c/h_i \le b^2/(4M^*) \le b^2/(4m). \tag{C.2}$$

Therefore, for any subset D of S of cardinality m, $m \le M^*$, $m/(4f^{*2}) \le \Sigma_D r_i^c/(\Sigma_D h_i) \le b^2/(4m)$. Thus

$$g\left(2\sum_D r_i^c/m, \sum_D h_i, m\right) = G\left(\sum_D r_i^c, \sum_D h_i, m\right) = \left(4\sum_D r_i^c \sum_D h_i/m\right)^{1/2},$$

see (5). Since $\text{TSP}(D^0) + c \ge 2\Sigma_D r_i^c/m$, the capacity (frequency) constraint may (can never) become tight when accounting for the actual routing cost of D instead of its lower-bound. Let the function $\tilde{g}(\theta, H, m)$ be defined as follows:

$$\tilde{g}(\theta, H, m) = \begin{cases} \sqrt{2H\theta} & \text{if } \theta/H \le b^2/(2m^2), \\ Hb/(2m) + m\theta/b & \text{otherwise.} \end{cases}$$

The function g coincides with $\tilde{g}$ on any subregion $D \subseteq S$. Let $D \subseteq S$, $m = |D| \le M^*$ and $H = \Sigma_D h_i$: for fixed H and m, $\tilde{g}(\theta, H, m)$ is continuously concave and differentiable in θ. Thus,

$$\tilde{g}\left(\text{TSP}(D^0)+c, H, m\right) - \tilde{g}\left(\frac{2\Sigma_D r_i^c}{m}, H, m\right) \le \tfrac{1}{2}\left(\frac{2H}{2\Sigma_D r_i^c/m}\right)^{1/2}\left(\text{TSP}(D^0) - \frac{2\Sigma_D r_i}{m}\right)$$

$$\le \tfrac{1}{2}\left(\frac{M^*H}{\Sigma_D r_i^c}\right)^{1/2}\left(\text{TSP}(D^0) - \frac{2\Sigma_D r_i}{m}\right).$$

The value of β, as given in the lemma, is obtained by the above inequality and the following facts: $H \le m h_{\max}$; by (C.2), $H/\Sigma_D r_i^c \le 4f^{*2}/M^*$; if $f^* = \infty$ and $c = 0$ then $\Sigma_D r_i \ge m r_{\min}$; if $c > 0$ then $\Sigma_D r_i^c \ge \frac{1}{2}mc$. $\square$

Proof of Theorem 7. (a): In view of lemmas 3, 5, 8 and 9,

$$\sum_{k=0}^{\alpha}\sum_{j=1}^{L_k}\left\{g\left(\text{TSP}\left(S_{k,j}^0\right)+c, \sum_{S_{k,j}} h_i, |S_{kj}|\right) - W(S_{k,j})\right\}$$

$$\le \beta \sum_{k=0}^{\alpha}\sum_{j=1}^{L_k}\left(\text{TSP}\left(S_{k,j}^0\right) - 2\sum_{S_{k,j}} r_i/|S_{k,j}|\right)$$

$$\le \beta \sum_{k=0}^{\alpha}\sum_{j=1}^{L_i}\left(\text{TSP}\left(S_{k,j}^0\right) - 2\sum_{S_{k,j}} r_i/M^*\right) \le \beta\left[\sum_{k=0}^{\alpha}\sum_{j=1}^{L_k}\text{TSP}\left(S_{k,j}^0\right) - (2|S|/M^*)\left(\sum_S r_i/|S|\right)\right]$$

$$\leq \beta\big(2r_{\max} + \text{TSP}(S) + \tfrac{3}{2}\Pi^{(2)}(S)\big)$$

$$\leq \beta\Big\{2 + 2(\pi|S|)^{1/2} + (2+\pi) + \tfrac{3}{2}\Big[\big(4(3\pi/M^*)^{1/2}|S|^{3/4} + \big(4(6\pi/M^*)^{1/2} + (3+2\pi)\big)|S|^{1/2} + (6+4\pi)\Big]\Big\}r_{\max}$$

$$= \beta\Big[6(3\pi/M^*)^{1/2}|S|^{3/4} + \big(4.5 + 2\pi^{1/2} + 6(6\pi/M^*)^{1/2} + 3\pi\big)|S|^{1/2} + (13+7\pi)\Big]r_{\max}$$

$$\leq \beta\Big[6(3\pi/M^*)^{1/2}|S|^{3/4} + 43.55|S|^{1/2} + 35\Big]r_{\max}.$$

(b):

$$\sum_{k=0}^{\alpha}\sum_{j=1}^{L_k} g\Big(\text{TSP}\big(S_{k,j}^0\big) + c, \sum_{S_{k,j}} h_{i}, |S_{k,j}|\Big) - \underline{V}_{M^*}(S)$$

$$= \sum_{k=0}^{\alpha}\sum_{j=1}^{L_k}\left\{g\Big(\text{TSP}\big(S_{k,j}^0\big) + c, \sum_{S_{k,j}} h_{i}, |S_{k,j}|\Big) - W(S_{k,j})\right\} + \left\{\sum_{k=0}^{\alpha}\sum_{j=1}^{L_k} W(S_{k,j}) - \underline{V}_{M^*}(S)\right\}$$

$$\leq \beta\Big[6(3\pi/M^*)^{1/2}|S|^{3/4} + 43.55|S|^{1/2} + 35\Big]r_{\max} + 2\Big[|S|\big(B(\tilde{R}^{-1/\alpha}) - 1\big) + \big(|S|^{1/2}+1\big)\big(B(\tilde{R}^{-1}) - 1\big)\Big]\big(r_{\max}^c h_{\max}/M^*\big)^{1/2} + 2(M^* r_{\max} h_{\max})^{1/2}.$$

The inequality follows from Theorem 6 and part (a) of this theorem.

Note that for any $Z > 0$, $Z^{1/2} \leq 1 + \frac{1}{2}(Z-1)$ by using the Taylor expansion of $Z^{1/2}$ around $Z_0 = 1$. Thus

$$B(\delta^{1/\alpha}) = \left(1 + \frac{\big(1 - \delta^{1/(2\alpha)}\big)^2}{2\delta^{1/\alpha}}\right)^{1/2} = \big(1.5 + 0.5\delta^{-1/\alpha} - \delta^{-1(2\alpha)}\big)^{1/2}$$

$$\leq 1 + \big(0.25 + 0.25\delta^{-1/\alpha} - 0.5\delta^{-1/(2\alpha)}\big).$$

Therefore, $B(\tilde{R}^{-1/\alpha}) - 1 \leq 0.25(1 + \tilde{R}^{1/\alpha} - 2\tilde{R}^{1/(2\alpha)})$. By using the notation γ_i, $i = 1, \ldots, 4$, given in part (b) of the theorem we terminate the proof. □

References

Anily, S. (1987), "Integrating inventory control and transportation planning", Ph.D. Dissertation, Columbia University, New York, NY 10027.

Anily, S., and Federgruen, A. (1990a), "One warehouse multiple retailer inventory systems with vehicle routing costs", *Management Science* 36, 92–114.

Anily, S., and Federgruen, A. (1990b), "A class of Euclidean routing problems with general cost functions; applications to combined inventory control and vehicle routing problems", *Mathematics of Operations Research* 15, 268–285.

Anily, S., and Federgruen, A. (1993), "Two-echelon distribution systems with vehicle routing and central inventories", in: M. Dror (ed.), Special Issue on Stochastic and Dynamic Models in Transportation, *Operations Research* 41, 37–47.

Anily, S., and Federgruen, A. (1990c), "Rejoinder to comments on one-warehouse multiple retailer systems with vehicle routing costs", *Management Science* 37, 1497–1499.

Barnes, F.R., Hoffman, A.J. and Rothblum, U.G. (1989), "Optimal partitions having disjoint convex and conic hulls", Working Paper, School of Industrial and System Engineering, Georgia Institute of Technology, Atlanta, GA 30332.

Burns, L.D., Hall, R.W., Blumenfeld, D.E., and Daganzo, C.F. (1985), "Distribution strategies that minimize transportation and inventory costs", *Operations Research* 33, 469–490.

Chakravarty, A.K., Orlin, J.B., and Rothblum, U.G. (1982), "A partitioning problem with additive objective with an application to optimal inventory grouping for joint replenishment", *Operations Research* 30, 1018–1020.

Chakravarty, A.K., Orlin, J.B., and Rothblum, U.G. (1985), "Consecutive optimizers for a partitioning problem with applications to optimal inventory groupings for joint replenishment", *Operations Research* 33, 820–834.

Christofides, N. (1976), "Worst-case analysis of a new heuristic for the Traveling Salesman Problem", Report 388, Graduate School of Industrial Administration, Carnegie–Mellon University, Pittsburgh, PA.

David, H.A. (1970), *Order Statistics*, Wiley, New York.

Federgruen, A., and Simchi-Levi, D. (1992), "Analytical analysis of vehicle routing and inventory-routing problems", to be published in: Ball, Magnanti, Monma and Nemhauser (eds.), *Handbooks in Operations Research and Management Science*, Volume on *Networks and Distribution*.

Fredergruen, A., Rinnooy Kan, A.H.G., and Zipkin, P. (1985), "Combined routing and inventory allocation problems: Bounds and heuristics", Working Paper, Graduate School of Business, Columbia University, New York, NY 10027.

Gallego, G., and Simchi-Levi, D. (1990), "On the effectiveness of direct shipping strategy for the one warehouse multi-retailer R-systems", *Management Science* 36, 240–243.

Garey, M.R., and Johnson, D.S. (1979), *Computers and Intractability*, Freeman, San Francisco, CA.

Golden, B.L., Magnanti, T. and Nguyen, H. (1977), "Implementing vehicle routing algorithms", *Networks* 7, 113–148.

Haimovich, M., and Rinnooy Kan, A.H.G. (1985), "Bounds and heuristics for capacitated routing problems", *Mathematics of Operations Research* 10, 527–542.

Inmann, R.R., and Jones, P.C. (1987), "When is the economic lot scheduling problem easy?", Technical Report 87-02, Center of Manufacturing Engineering, Northwestern University, Evanston, IL 60201.

Karp, R.M. (1977), "Probabilistic analysis of partitioning algorithms for the Traveling Salesman Problem in the plane", *Mathematics of Operations Research* 2, 209–244.

Magnanti, T. (1981), "Combinatorial optimization and vehicle fleet planning: Perspectives and prospects", *Networks* 11, 179–213.