On orders and bounds for multimodular functions

Bruno Gaujal* Arie Hordijk † Dinard van der Laan ‡

April 15, 2002

Abstract

This paper presents several partial orders which can be used to compare the regularity of a stream of customers entering a system of several queues in parallel (cone order, multimodular order, graph order and unbalance). The different notions are compared and are used to derive upper bounds on the waiting times of customers in parallel queues when such sequences are used as routing policies.

1 Introduction

In several papers ([4, 5]), various attemps have been made in order to study the question which of two deterministic periodic admission sequences (periodic sequences of nonnegative integers) gives the smaller average expected waiting time. A partial order, called the cone order is introduced in [4], and it is shown that the average waiting time and more generally any multimodular function is monotone with respect to the cone order. It is natural to define a multimodular order by requiring that any multimodular function is monotone. In contrast to [11], where the convex order for stochastic admission sequences is used, we consider deterministic admission sequences in this paper. Note that deterministic admission sequences can only be ordered for all multimodular functions if they are equal. Therefore we consider multimodular functions with a fixed minimal point as we did with the cone order. In Section 2 we introduce the multimodular order and we show that the cone order is equivalent to the multimodular order. In Subsection 2.3 the shift invariant counterparts of these orders are studied and it is shown that the regular admission sequence is the minimal point.

For the optimal routing problem to N queues a lower bound is obtained by using regular admission sequences (or what is the same bracket sequences) with 'minimizing' routing fractions (densities) as admission sequences to all queues (see [3]). But generally only in the routing to $N = 2$ queues, the routing fractions will be balanceable (see [2]) and only in that case the admission sequences can be glued together and be made to a feasible routing policy.

In Section 5 we derive an upper bound for the difference between the average expected waiting time for customers which are routed to one queue with a given periodic admission sequence which has a fixed routing density, and the lower bound corresponding to the regular admission sequence with the same density. In order to do that, we first introduce a combinatorial notion which we call the unbalance of the admission sequence in Section 3. Roughly speaking it measures its 'distance' to the regular admission sequence. We show that the unbalance is a shift invariant and multimodular function. In Section 6 we derive an upperbound for the average expected waiting time for a routing policy to N queues as a function of its unbalance. For a routing policy the unbalance is the sum of the unbalances of the induced admission sequences to the queues $1, \ldots, N$.

*ENS Lyon, 46 Allée d'Italie, 69007 Lyon, France. Email: Bruno.Gaujal@ens-lyon.fr

†Dept. of Mathematics, Leiden University, P.O.Box 9512, 2300RA Leiden, The Netherlands. E-mail: hordijk@math.leidenuniv.nl

‡Dept. of Mathematics, Leiden University, P.O.Box 9512, 2300RA Leiden, The Netherlands. E-mail: dvdlaan@math.leidenuniv.nl

The difference between the upper and lower bound depends on the average interarrival times but it is insensitive for their distributions and the service time distributions. The upper bound is tight for the heavy traffic situation in which the interarrival and service times are deterministic and the traffic load is 1 for each of the queues. This means that the upper bound is correct for the deterministic heavy traffic case. Clearly, it holds for low traffic that the average expected waiting time is close to the lower bound.

The unbalance is a combinatorial notion and the analysis in this chapter is mainly combinatorial, it turns out that the unbalance is also useful in other models than routing to queues. Indeed, the bounds can be generalized from waiting times to similar bounds for sequences of multimodular functions as in [3] (see the remark at the end of Section 5).

For the definition of the unbalance we introduce in Section 3 the graph order, which gives the distance of a given admission sequence to the bracket sequence and thereby defines its unbalance. The relations between the shift invariant orders and the graph order are studied in Section 5. We show through counterexamples that they are not equivalent. This paper is part of a monograph ([6]), in which the reader can find more details.

2 The multimodular order and the cone order

Let D be a multimodular matrix with rowvectors $d_0, d_1, \ldots, d_n \in \mathbb{R}^n$ verifying $d_0 + \ldots + d_n = 0$ where $d_1, d_2, \ldots, d_n$ are linearly independent over $\mathbb{R}$ and let $M_D \subset \mathbb{R}^n$ be the corresponding mesh, that is the set of all the points $\{a_0 d_0 + a_1 d_1 + \cdots + a_n d_n, \quad a_i \in \mathbb{Z}, \quad i = 0, \cdots, n\}$.

Definition 2.1 *A function $f : M_D \to \mathbb{R}$ is D-multimodular if and only if for all $a \in M_D$, and for all $0 \leq i < j \leq n$,*

$$f(a + d_i) + f(a + d_j) \geq f(a) + f(a + d_i + d_j). \tag{1}$$

For $r \in M_D$ let

$$\mathcal{F}(D, r) = \left\{ f : M_D \to \mathbb{R} \text{ s. t. } f \text{ is } D-\text{multimodular and } f(r) = \min_{x \in M_D} f(x) \right\}$$

be the set of D-multimodular functions with global minimum point in r.

2.1 The cone order

This section is devoted to the definition of the cone order. More on this can be found in [4].

Definition 2.2 *A D-atom is a simplex in $\mathbb{R}^n$, made of the $n+1$ points*

$$\begin{aligned} p_0 &= a, \\ p_1 &= a + d_{\xi(0)}, \\ p_2 &= a + d_{\xi(0)} + d_{\xi(1)} \\ &\vdots \\ p_n &= a + d_{\xi(0)} + d_{\xi(1)} + \cdots + d_{\xi(n-1)} \end{aligned} \tag{2}$$

where $a \in M_D$ (called the root) and ξ is a permutation of $\{0, \cdots, n\}$. This atom will be denoted as $\mathcal{S}(p_0, \cdots, p_n)$.

Theorem 2.3 *The set of all the D-atoms forms a triangulation of $\mathbb{R}^n$, called a multimodular triangulation.*

Proof. Let x be a point in $\mathbb{R}^n$. We can write $x = \alpha_0 d_0 + \cdots + \alpha_n d_n$, with non-negative coordinates, one of which is 0 since $d_0 + \cdots + d_n = 0$. We construct $a = \lfloor \alpha_0 \rfloor d_0 + \cdots + \lfloor \alpha_n \rfloor d_n$ and ξ such that $\alpha_{\xi(i)} - \lfloor \alpha_{\xi(i)} \rfloor \geq \alpha_{\xi(i+1)} - \lfloor \alpha_{\xi(i+1)} \rfloor$. We define $\beta_n = 0$, $\beta_{n-1} = \alpha_{\xi(n-1)} - \lfloor \alpha_{\xi(n-1)} \rfloor$, $\beta_i = \alpha_{\xi(i)} - \lfloor \alpha_{\xi(i)} \rfloor - (\alpha_{\xi(i+1)} - \lfloor \alpha_{\xi(i+1)} \rfloor)$. All of them verify $0 \leq \beta_i \leq 1$ and $\sum_{i=0}^{n-1} \beta_i \leq 1$. We have $x = a + \beta_0 d_{\xi(0)} + \cdots + \beta_{n-1}(d_{\xi(0)} + \cdots + d_{\xi(n-1)})$. Therefore, x belongs to the atom with root a and permutation ξ.

Now, assume that a point x belongs to the interior of two different atoms with respective roots a and b and permutations ξ and τ. Since everything is shift invariant, we may assume that $b = 0$ and τ is the identity. We may write

$$
\begin{aligned}
x &= a + \alpha_0 d_{\xi(0)} + \cdots + \alpha_{n-1}(d_{\xi(0)} + \cdots + d_{\xi(n-1)}) \\
&= \beta_0 d_0 + \cdots + \beta_{n-1}(d_0 \cdots + d_{n-1})
\end{aligned}
$$

with $\sum_{i=0}^{n-1} \beta_i \leq 1$ and $\sum_{i=0}^{n-1} \alpha_i \leq 1$. Since x is in the interior of both atoms, we also have $\beta_i > 0$ and $\alpha_i > 0$ for $i = 0, \cdots, n-1$. Therefore, by uniqueness of the decomposition of x, and writing $a = a_0 d_0 + \cdots + a_n d_n$,

$$
\begin{aligned}
a_0 + \sum_{j=\xi^{-1}(0)}^{n-1} \alpha_{\xi(j)} &= \beta_0 + \cdots + \beta_{n-1} \\
a_1 + \sum_{j=\xi^{-1}(1)}^{n-1} \alpha_{\xi(j)} &= \beta_1 + \cdots + \beta_{n-1} \\
&\vdots \\
a_{n-1} + \sum_{j=\xi^{-1}(n-1)}^{n-1} \alpha_{\xi(j)} &= \beta_{n-1} \\
a_n + \sum_{j=\xi^{-1}(n)}^{n-1} \alpha_{\xi(j)} &= 0.
\end{aligned}
$$

Since the partial sums of the α_i or of the β_i are smaller than one and since a_i are integer numbers, we have $a_i = 0$ for all $i = 0, \cdots, n$. Both atoms have the same root.

Now, the equality of the partial sums taken one by one implies first that $\sum_{j=\xi^{-1}(n)}^{n-1} \alpha_{\xi(j)} = 0$. Since $\alpha_i > 0$ for all i, the only possibility is $\xi^{-1}(n) = n$.

Considering vectors d_k and d_{k+1}, we have:

$$
\sum_{j=\xi^{-1}(k)}^{n-1} \alpha_{\xi(j)} - \sum_{j=\xi^{-1}(k+1)}^{n-1} \alpha_{\xi(j)} = \beta_k > 0.
$$

This implies that $\xi^{-1}(k) > \xi^{-1}(k+1)$. Thus ξ is the identical permutation. Therefore, both atoms are equal. ■

Consider one atom $S = \mathcal{S}(p_0, p_1, \cdots, p_n)$ containing $r = p_0$ as a vertex (the global minimum point of the multimodular functions considered above). Let ξ and τ be the permutations on $\{0, \cdots n\}$ such that $\tau(0) = 0$ and

$$
\begin{aligned}
p_{\tau(1)} &= r + d_{\xi(0)} && (3) \\
p_{\tau(2)} &= p_{\tau(1)} + d_{\xi(1)} && (4) \\
\vdots &= \vdots && (5) \\
p_{\tau(n)} &= p_{\tau(n-1)} + d_{\xi(n-1)} && (6) \\
r &= p_{\tau(n)} + s_{\xi(n)}. && (7)
\end{aligned}
$$

Now, we define for all $1 \leq i \leq n$, $b_i = \sum_{j=1}^{i} d_{\xi(j)}$ and $b = (b_1, \cdots, b_n)$ the n x n matrix for which the i -th column is vector b_i. Note that $p_{\tau(i)} = r + b_i$ for $i = 1, 2, \ldots, n$. The cone associated with such an atom S, denoted by $C(S)$, contains all the points x of M_D such that $x = r + c^t(x)b$ for some non-negative column vector $c = c(x)$ in $\mathbb{N}_0^n$.

Conversely, for any point x in M_D, there exists such a cone containing x and we will have $x = r + c^t(x)b$, where the columns of b corrsponding to the support of $c(x)$ are uniquely defined. We shall denote $d(r, x) = c_1(x) + \cdots + c_n(x)$ and call it the cone-distance from x to r.

It is easilly seen that there are $(n+1)!$ atoms with root in r and thus there are exactly $(n+1)!$ cones corresponding to the given multimodular matrix D and root r. Let $\mathcal{C} = \{C_i\}_{i=1}^{(n+1)!}$ be the set of cones corresponding to D and root $r = (0, 0, \ldots, 0)$ (the origin). Note that from the preceding argument and Theorem 2.3 it follows that $\{\cup C_i : C_i \in \mathcal{C}\} = \mathbb{R}^n$. For $C_i \in \mathcal{C}$ let C_i' be the intersection of C_i and M_D. In the following proposition which follows directly from the definitions and preceding results we elucidate the properties of the objects we have associated to a multimodular matrix D with rowvectors $d_0, d_1, \ldots, d_n \in \mathbb{R}^n$.

Proposition 2.4 *Let $C_i \in \mathcal{C}$. Then there exist $b_1, b_2, \ldots, b_n \in \mathbb{R}^n$ and a bijection $\sigma : \{1, 2, \ldots, n+1\} \to \{0, 1, \ldots, n\}$ such that*

1. *$b_1, b_2, \ldots, b_n$ are linearly independent.*
2. *$b_j = \sum_{i=1}^{j} d_{\sigma(i)}$ for $j = 1, 2, \ldots, n$.*
3. *$C_i = \{\sum_{i=1}^{n} \lambda_i \cdot b_i : \lambda_i \geq 0$ for $i = 1, 2, \ldots, n\}$.*
4. *$M_D = \{\sum_{i=1}^{n} \lambda_i \cdot b_i : \lambda_i \in \mathbb{Z}$ for $i = 1, 2, \ldots, n\}$.*
5. *$C_i' = \{\sum_{i=1}^{n} \lambda_i \cdot b_i : \lambda_i \in \mathbb{Z}_{\geq 0}$ for $i = 1, 2, \ldots, n\}$.*

For a convex cone $C \subset \mathbb{R}^n$ we say that a set of vectors $v_1, v_2, \ldots, v_k$ are generators of C if $v_1, v_2, \ldots, v_k$ are linearly independent and $C = \{\sum_{i=1}^{k} \lambda_i \cdot v_i : \lambda_i \geq 0$ for $i = 1, 2, \ldots, k$ $\}$. In that case we say that C is a k - dimensional cone. In particular we have for $b_1, b_2, \ldots, b_n$ as in Proposition 2.4 that they are generators of the n -dimensional convex cone C_i.

Definition 2.5 *For $x, y \in M_D$ we denote $x \leq_C y$ if there exists some cone $C_i \in \mathcal{C}$, with generators $b_1, b_2, \ldots, b_n$, such that $x, y \in C_i$ and $c(x) \leq c(y)$ component-wise in cone C_i.*

If $x \leq_C y$ then we say that x and y are cone ordered and $\leq_C$ is called the cone order.

2.2 The multimodular order and the cone order

Definition 2.6 *Let the multimodular matrix D and global minimum point $r \in M_D$ be given. Then for $x, y \in M_D$ we say that $x \leq_{mm} y$ if $f(x) \leq f(y)$ for every $f \in \mathcal{F}(D, r)$.*

It is easy to verify that both the multimodular order $\leq_{mm}$ and the cone order $\leq_C$ are reflexive, antisymmetric and transitive and thus they are partial orders on M_D. The following theorem says that these partial orders on M_D are equivalent.

Theorem 2.7 *For given multimodular matrix D and global minimum point $r \in M_D$ we have for $x, y \in M_D$ that $x \leq_{mm} y$ if and only if $x \leq_C y$.*

First we prove that $x \leq_C y$ implies $x \leq_{mm} y$. We have that x and y are in the same cone C and let $b_1, \cdots, b_n$ be the generators of this cone.

Note that we can assume for $x \leq_C y$ that $d(x, y) := \mid d(r, y) - d(r, x) \mid = 1$. If not, then we prove step by step along the path from x to y along the direction of the generators, say $x = x_1 \leq_C \cdots \leq_C x_m = y$ that $f(x) = f(x_1) \leq \cdots \leq f(x_m) = f(y)$ for $f \in \mathcal{F}(D, r)$.
The proof will now proceed by induction on $d(r, x)$. The property is true if $d(x, r) = 0$, since r is

the argmin of f on S. Now, assume $d(x,r) \geq 1$. Pick a point z in cone C such that $x = z + b_i$. (equivalently, we have $c(y) + e_i = c(x)$ and $c(y) \leq 0$). Note that $d(z,r) = d(x,r) - 1$ and $z \leq_C x$. By induction, this means $f(z) \leq f(x)$. Since $d(x,y) = 1$, there exists some j such that $y = x + b_j$. Now we have two cases.

- If $i = j$, then by convexity of f,

$$f(z + b_i) - f(z) \leq f(z + b_i + b_i) - f(z + b_i).$$

  We also know by induction that $f(z + b_i) - f(z) \geq 0$. Thus $f(x) \leq f(y)$.

- If $i \neq j$, then we choose yet another point, w, such that $w = z + b_j$. We may assume that $i > j$ (the case $j < i$ is similar by inverting the role played by b_i and b_j in the following).

  Since b_i is a sum of base vectors, it is also a sum of opposites of base vectors, since all base vectors add up to 0. Note that all these base vectors are distinct from the base vectors involved in b_j.

  We have:

$$\begin{aligned} b_j &= d_{\xi(1)} + \cdots + d_{\xi(j)}, & (8)\\ b_i &= -d_{\xi(i+1)} - d_{\xi(i+2)} - \cdots - d_{\xi(n+1)}. & (9) \end{aligned}$$

  Therefore, by the multimodularity of f,

$$f(w) - f(z) = f(x + d_{\xi(1)} + \cdots + d_{\xi(j)} + d_{\xi(i+1)} + \cdots + d_{\xi(n+1)}) - f(x + d_{\xi(i+1)} + \cdots + d_{\xi(n+1)}) \leq$$

$$f(x + d_{\xi(1)} + \cdots + d_{\xi(j)}) - f(x) = f(y) - f(x).$$

  Since $d(r,w) = d(r,z) + 1$, we have $f(w) - f(z) \geq 0$ by induction. Thus $f(y) - f(x) \geq 0$.

To finish the proof of Theorem 2.7 we have to show that $x \leq_{mm} y$ implies that $x \leq_C y$. By translation we can assume without loss of generality that the global minimum point $r = (0,0,\ldots,0) \in M_D$. Let $b_1, b_2, \ldots, b_n$ be the generators of the cone $C_i \in \mathcal{C}$ as in Proposition 2.4. If $x \in M_D$ then we have by Proposition 2.4 that there exist unique $x_j \in \mathbb{Z}$ such that $x = \sum_{j=1}^n x_j \cdot b_j$. Moreover, for $x = \sum_{j=1}^n x_j \cdot b_j \in \mathbb{R}^n$ we say that $\mathrm{supp}_i(x) = \{k \in \{1, 2, \ldots, n\}$ for which $x_k > 0\}$. Note that we assume for $k \in \mathrm{supp}_i(x)$ not only that $x_k \neq 0$ but that $x_k > 0$ which is stronger than in the standard definition of support. For $k = 1, 2, \ldots, n$ we define the function $f_k : M_D \to \mathbb{Z}_{\geq 0}$ by

$$f_k(x) = \max(x_k, 0). \tag{10}$$

Proposition 2.8 *For every $k \in \{1, 2, \ldots, n\}$ the function $f = f_k$ is D -multimodular with global minimum point in $r = \mathbf{0}$.*

Proof. Since $f(x) \geq 0$ for every $x \in M_D$ and $f(r) = 0$ it follows that r is a global minimum point of f. It remains to prove that f is a D -multimodular function. So, we have to show that for every $x \in M_D$, $i, j \in \{0, 1, \ldots, n\}$ with $i \neq j$ it holds that

$$f(x + d_i) + f(x + d_j) \geq f(x) + f(x + d_i + d_j). \tag{11}$$

According to property 2 of Proposition 2.4 there exists a bijection $\sigma : \{1, 2 \ldots, n+1\} \to \{0, 1, 2, \ldots, n\}$ such that $b_j = \sum_{i=1}^j d_{\sigma(i)}$ for $j = 1, 2, \ldots, n$. From this it follows that $d_{\sigma(1)} = b_1$, $d_{\sigma(j)} = b_j - b_{j-1}$ for $2 \leq j \leq n$ and $d_{\sigma(n+1)} = -\sum_{i=1}^n d_{\sigma(i)} = -b_n$. Put $l = \sigma(k)$, $m = \sigma(k+1)$. If $i \notin \{l, m\}$ and $j \notin \{l, m\}$ then we have for every $x \in M_D$ that $f(x + d_i) = f(x + d_j) = f(x) = f(x + d_i + d_j)$. Hence (11) holds . Suppose that $i \in \{l, m\}$ and $j \notin \{l, m\}$. Then we have for every $x \in M_D$ that $f(x + d_i) = f(x + d_i + d_j)$ and $f(x + d_j) = f(x)$ and thus (11) holds. Suppose that $\{i, j\} = \{l, m\}$ and assume without loss of generality that $i = l$ and $j = m$. If $f(x) > 0$ then it follows that $f(x + d_i) = f(x) + 1$, $f(x + d_j) = f(x) - 1$, $f(x + d_i + d_j) = f(x)$ and thus (11) holds.

If $f(x) = 0$ then it follows that $f(x + d_i) \geq f(x)$, $f(x + d_j) = f(x + d_i + d_j) = f(x) = 0$ and thus (11) holds. So, we can conclude that (11) holds in every case and thus $f \in \mathcal{F}(D, r)$. ■

For $i = 1, 2, \ldots, (n+1)!$ let $\mathcal{G}(C_i)$ be the set of functions $\{f_1, f_2, \ldots, f_n\}$ for cone C_i and let $\mathcal{G} = \cup_{i=1}^{(n+1)!} \mathcal{G}(C_i)$. We have the following corollaries of Proposition 2.8.

Corollary 2.9 *For given multimodular matrix D and global minimum point $r = (0, 0, \ldots, 0) \in \mathbb{R}^n$ we have that $\mathcal{G} \subseteq \mathcal{F}(D, r)$.*

Corollary 2.10 *If $x, y \in M_D$, $x \leq_{mm} y$ and there exists some cone $C_i \in \mathcal{C}$ such that $x, y \in C_i$ then $x \leq_C y$.*

Proof. Let $b_1, b_2, \ldots, b_n$ be the generators of cone C_i as in proposition 2.4 and put $x = \sum_{i=1}^n x_i \cdot b_i$ and $y = \sum_{i=1}^n y_i \cdot b_i$. Then we have by Proposition 2.4 that $x_i, y_i \in \mathbb{Z}_{\geq 0}$ for $i = 1, 2, \ldots, n$. So, for $f_k \in \mathcal{G}(C_i)$ we have that $f_k(x) = x_k$ and $f_k(y) = y_k$ for $k = 1, 2, \ldots, n$. Hence by Proposition 2.8 and $x \leq_{mm} y$ we have that $0 \leq x_k = f_k(x) \leq f_k(y) = y_k$ for $k = 1, 2, \ldots, n$. Thus $x \leq_C y$. ■

Corollary 2.11 *If $x, y \in M_D$, $x \leq_{mm} y$ then $\mathrm{supp}_i(x) \subset \mathrm{supp}_i(y)$ for $i = 1, 2, \ldots, (n+1)!$.*

Corollary 2.11 can be proved in the same way as Corollary 2.10. The following corollary follows immediately from Corollary 2.10 and Corollary 2.11.

Corollary 2.12 *If $x, y \in M_D$, $x \leq_{mm} y$ and there exists some cone $C_i \in \mathcal{C}$ such that x is an interior point of C_i then $x \leq_C y$.*

To prove Theorem 2.7 it suffices by Corollary 2.10 to prove that there exists some cone $C_i \in \mathcal{C}$ such that $x, y \in C_i$ if $x \leq_{mm} y$. If $x = r = \mathbf{0}$, $y = r = \mathbf{0}$ or $y = \lambda \cdot x$ with $\lambda > 0$ then it follows directly that $x, y \in C_i$ for some $C_i \in \mathcal{C}$. Suppose that $y = \lambda \cdot x$ with $\lambda < 0$ and $x \neq r$. Let $C_i \in \mathcal{C}$ be a cone containing x with generators $b_1, b_2, \ldots, b_n$. Then $x = \sum_{i=1}^n x_i \cdot b_i$ and $y = \sum_{i=1}^n \lambda \cdot x_i \cdot b_i$ with $x_i \geq 0$ for $i = 1, 2, \ldots, n$. Moreover we have that $x_k > 0$ for some $k \in \{1, 2, \ldots, n\}$ since $x \neq r$. Then for $f_k \in \mathcal{G}(C_i)$ we have that $f_k(x) = x_k > 0$ and $f_k(y) = 0$, which contradicts $x \leq_{mm} y$. So, we can assume that x and y are linearly independent vectors in $\mathbb{R}^n$ and let $H := \mathrm{span}(x, y)$ the two dimensional subspace that contains x and y. For $C_j \in \mathcal{C}$ let C_j^* be the intersection of C_j and H. Then we have that the C_j^* are convex cones of dimension smaller or equal than two in the subspace H. To prove Theorem 2.7 we have the following lemmas.

Lemma 2.13 *Suppose that $C_i^* = C_i \cap H$ is a two dimensional cone generated by u and v. Then $\mathrm{supp}_i(u)$ is not a subset of $\mathrm{supp}_i(v)$ and $\mathrm{supp}_i(v)$ is not a subset of $\mathrm{supp}_i(u)$.*

Proof. Let $b_1, b_2, \ldots, b_n$ be the generators of C_i as in proposition 2.4. Then $u = \sum_{k \in supp_i(u)} u_k \cdot b_k$ with $u_k > 0$ and $v = \sum_{k \in \mathrm{supp}_i(v)} v_k \cdot b_k$ with $v_k > 0$. For a small positive number ε we have that $w := v - \epsilon \cdot u \in H$ and if $\mathrm{supp}_i(u) \subseteq \mathrm{supp}_i(v)$ then it is easily seen that $w = \sum_{k \in \mathrm{supp}_i(v)} w_k \cdot b_k$ with $w_k > 0$ and thus $w \in C_i^*$. However, $w \in C_i^*$ contradicts the fact that u and v are generators of C_i^*. If $\mathrm{supp}_i(v) \subseteq \mathrm{supp}_i(u)$ we get a contradiction similarly. ■

Lemma 2.14 *Let $u, v, w \in H$ with u, v linearly independent and u, w linearly independent. Let A_1 be the two dimensional cone generated by u and v, A_2 the two dimensional cone generated by u and w and A_3 the one dimensional cone generated by u. Let $w = a_1 \cdot u + a_2 \cdot v$. Then $A_1 \cap A_2 = A_3$ if and only if $a_2 < 0$.*

Proof. It is obvious that $A_3 \subseteq A_1 \cap A_2$. Let $x = \nu \cdot u + w$, where $\nu \geq \max(-a_1, 0)$. Then it is easily seen that $x \in A_2 \setminus A_3$. It follows that $x = (\nu + a_1) \cdot u + a_2 \cdot v$ and thus $x \in A_1 \cap A_2$ if $a_2 \geq 0$. So, $A_1 \cap A_2 = A_3$ implies that $a_2 < 0$. Conversely, suppose that

$$x' = \lambda \cdot u + \lambda' \cdot w \in A_1 \cap A_2.$$

Since $x' \in A_2$ we have that $\lambda \geq 0$ and $\lambda' \geq 0$. Moreover,

$$x' = \lambda \cdot u + \lambda' \cdot (a_1 \cdot u + a_2 \cdot v) = (\lambda + \lambda' \cdot a_1) \cdot u + \lambda' \cdot a_2 \cdot v$$

and since $x' \in A_1$ it follows that $\lambda' \cdot a_2 \geq 0$. Thus if $a_2 < 0$ then $\lambda' = 0$ and $x' \in A_3$. ∎

We are now ready to finish the proof of Theorem 2.7

Proof. We have that $\cup_{i=1}^{(n+1)!} C_i^* = H$ since $\cup_{i=1}^{(n+1)!} C_i = \mathbb{R}^n$. Therefore one of the following two cases holds:

1. There exists some two dimensional cone C_i^* such that in the subspace H we have that x is in the interior of cone C_i^*.

2. There exist two dimensional cones C_i^*, C_j^* such that in the subspace H we have that x is on the boundary of C_i^*, x is on the boundary of C_j^* and x is in the interior of $C_i^* \cup C_j^*$.

Suppose we have case 1 and let a and b be generators of the cone C_i^*. Then $a, b \in C_i$ and $x = \lambda \cdot a + \mu \cdot b$ with $\lambda, \mu > 0$. Hence $x \in C_i$ and $\mathrm{supp}_i(x) = \mathrm{supp}_i(a) \cup \mathrm{supp}_i(b)$. Suppose that $y \notin C_i^*$. Then $y = y_1 \cdot a + y_2 \cdot b$ with $\min(y_1, y_2) < 0$ and suppose without loss of generality that $y_1 < 0$. According to Lemma 2.13 there exist $k \in \mathrm{supp}_i(a) \setminus \mathrm{supp}_i(b)$. Then it follows that $k \notin \mathrm{supp}_i(y)$, $k \in \mathrm{supp}_i(x)$. Thus $\mathrm{supp}_i(x)$ is not a subset of $\mathrm{supp}_i(y)$, but this contradicts Corolary 2.11. Hence $y \in C_i^*$ and thus $x, y \in C_i$. According to Corollary 2.10 it follows that $x \leq_C y$.

Suppose we have case 2. Then there exist $a, b \in H$ such that x and a are generators of cone C_i^* and x and b are generators of cone C_j^*. Let u_1, u_2, v_1, v_2 be such that $y = u_1 \cdot x + u_2 \cdot a = v_1 \cdot x + v_2 \cdot b$. If $u_1 < 0$ then it follows analogously to the proof in case 1 that $\mathrm{supp}_i(x)$ is not a subset of $\mathrm{supp}_i(y)$, which yields a contradiction again. Thus $u_1 \geq 0$ and analogously it follows that $v_1 \geq 0$. We will prove that $y \in C_i^* \cup C_j^*$. Suppose that $y \notin C_i^*$. Then it follows that $u_2 < 0$. Let a_1, a_2 be such that $a = a_1 \cdot x + a_2 \cdot b$. By Lemma 2.14 we have that $a_2 < 0$. Then $y = (u_1 + u_2 \cdot a_1)x + (u_2 \cdot a_2)b$. Hence $v_2 = u_2 \cdot a_2 > 0$. Thus $y \in C_j^*$ if $y \notin C_i^*$ and thus $y \in C_i^* \cup C_j^*$. Since $x \in C_i^* \cap C_j^*$ it follows that $x, y \in C_i$ or $x, y \in C_j$. Hence $x \leq_C y$ by Corollary 2.10. ∎

2.3 Shift invariant counterparts

We consider sequences (of nonnegative integers) of a given length $T \in \mathbb{N}$ and given sum $S \in \mathbb{N}$. The set of such sequences is denoted by $P(T, S)$. So, $P(T, S) = \{(x_1, x_2, \ldots, x_T) : x_i \in \mathbb{Z}_{\geq 0} \forall i, \sum_{i=1}^{T} x_i = S\}$ and this is a submesh of dimension $T - 1$ of $\mathbb{Z}^T$. Let D' be a multimodular matrix of size T x T induced by the submesh $P(T, S)$.

Let $x, x' \in P(T, S)$. Then we say that x and x' are conjugate if they are cyclic permutations of each other. It holds that x and x' are conjugate if there exist finite (possibly empty) sequences v and w such that $x = vw$ and $x' = wv$. It is easily seen that this conjugacy is an equivalence relation on $P(T, S)$ and we write $x \sim x'$ if x and x' are conjugate. We denote by $\widetilde{x}$ the conjugacy class of $x \in P(T, S)$, which is the set of all cyclic permutations of x. By $\widetilde{P}(T, S)$ we denote the set of conjugacy classes of $P(T, S)$. If $x \sim x'$ then we also say that x and x' are shifts of each other. If $\widetilde{x} = \widetilde{y}$ then we say that x is a representative of $\widetilde{y}$.

Let $f : P(T, S) \to \mathbb{R}$ be a function such that $f(x) = f(x')$ if $x \sim x'$. Then we say that f is a shift invariant function. A shift invariant function $f : P(T, S) \to \mathbb{R}$ induces a function $\widetilde{f} : \widetilde{P}(T, S) \to \mathbb{R}$ by $\widetilde{f}(y) = f(x)$ where $x \in P(T, S)$ is a representative of $y \in \widetilde{P}(T, S)$. We denote by $\mathcal{F}^{\mathrm{shift}}(D')$ the set of functions mapping $\widetilde{P}(T, S)$ to $\mathbb{R}$ which are induced by D' multimodular functions that are shift invariant. We use $\mathcal{F}^{\mathrm{shift}}(D')$ to define a partial order $\leq_{mms}$ on $\widetilde{P}(T, S)$ in the same way as $\mathcal{F}(D, r)$ was used to define the partial order $\leq_{mm}$ on M_D. The partial order $\leq_{mms}$ is called the shift invariant multimodular order.

Definition 2.15 *Let a multimodular matrix D' induced by the submesh $P(T, S)$ be given. Then for $x, y \in \widetilde{P}(T, S)$ we say that $x \leq_{mms} y$ if $f(x) \leq f(y)$ for every $f \in \mathcal{F}^{\mathrm{shift}}(D')$.*

Let $x \in P(T, S)$. Then we say that $x = (x_1, x_2, \ldots, x_T)$ is regular if the induced infinite sequence $x^\infty := (x_1, \ldots, x_T, x_1, \ldots, x_T, \ldots)$ is a bracket or regular sequence and in that case the conjugacy class $\widetilde{x} \in \widetilde{P}(T,S)$ is also called regular. We denote by $R(T,S) \subseteq P(T,S)$ the subset of regular sequences. The following lemma can be proved analogously to the proof for sequences of zeros and ones that is given in [12].

Lemma 2.16 *Let $T, S \in \mathbb{N}$ be given. Then there exists exactly one element in $\widetilde{P}(T,S)$ which is regular.*

We denote by $\widetilde{\omega} = \widetilde{\omega}(T,S)$ the unique element of $\widetilde{P}(T,S)$ that is regular. Then we have the following theorem (see [4])

Theorem 2.17 *Let a multimodular matrix D' induced by the submesh $P(T,S)$ be given and let $f \in \mathcal{F}^{\text{shift}}(D')$. Then a global minimum of f is attained in $\widetilde{\omega}(T,S)$.*

Note that the global minimum $\widetilde{\omega}(T,S)$ does not depend on D'.

Corollary 2.18 *For every D' we have that $\widetilde{\omega}(T,S)$ is the smallest element for the partial order $\leq_{mms}$ on $\widetilde{P}(T,S)$.*

We have the multimodular order $\leq_{mm}$ and we have the shift invariant multimodular order $\leq_{mms}$ on $\widetilde{P}(T,S)$. We also define a shift invariant cone order $\leq_{Cs}$ on $\widetilde{P}(T,S)$ which is the counterpart of the cone order $\leq_C$.

Definition 2.19 *Let D' be as in Theorem 2.17. Then for $u, v \in \widetilde{P}(T,S)$ we say that $u \leq_{Cs} v$ if and only if there exist representatives u' of u and v' of v such that $u' \leq_C v'$, where $\leq_C$ is the cone order for multimodular matrix D' and root r with r being a representative of $\widetilde{\omega}(T,S)$.*

According to Theorem 2.7 it holds that $u' \leq_C v'$ if and only if $f(u') \leq f(v')$ for every $f \in \mathcal{F}(D',r)$. Thus the following corollary follows directly from Theorem 2.7 and the definitions.

Corollary 2.20 *Let D' be as in Theorem 2.17. Then for $u, v \in \widetilde{P}(T,S)$ we have that $u \leq_{mms} v$ if $u \leq_{Cs} v$.*

So, according to Corollary 2.20 the shift invariant cone order implies the shift invariant multimodular order.

3 The graph order and the unbalance

Let $u = (u_1, u_2, \dots, u_T) \in P(T,S)$ and for $l = 0, 1, \dots, T-1$ let $u^{(l)} := (u_{l+1}, u_{l+2}, \dots, u_T, u_1, u_2, \dots, u_l) \in P(T,S)$ be the l-th cyclic permutation of u. For a sequence $u \in P(T,S)$ we define a counting function $\kappa_u : \{0, 1, \dots, T\} \to \mathbb{Z}$ by $\kappa_u(n) = \sum_{t=1}^{n} u_t$ and we define a discrepancy function $\phi_u : \{0, 1, \dots, T\} \to \mathbb{Q}$ by $\phi_u(n) = \kappa_u(n) - n \cdot \frac{S}{T}$ for $n = 0, 1, \dots, T$.

Lemma 3.1 *For $u \in P(T,S)$ and $l \in \{0, 1, \dots, T-1\}$ we have that $\phi_{u^{(l)}}(n) \geq 0$ for $n = 0, 1, \dots, T$ if and only if $\phi_u(l) = \min_{n=0,1,\dots,T} \phi_u(n)$. Moreover, $\phi_{u^{(l)}}(n) \leq 0$ for $n = 0, 1, \dots, T$ if and only if $\phi_u(l) = \max_{n=0,1,\dots,T} \phi_u(n)$.*

Proof. Suppose that $\phi_{u^{(l)}}(n) < 0$ for some $n \in \{0, 1, \dots, T-1\}$. Then $\phi_u((l+n) \pmod{T}) = \phi_u(l) + \phi_{u^{(l)}}(n) < \phi_u(l)$. Thus $\phi_u(l) = \min_{n=0,1,\dots,T} \phi_u(n)$ implies that $\phi_{u^{(l)}}(n) \geq 0$ for $n = 0, 1, \dots, T$. Suppose that $\phi_u(l') < \phi_u(l)$ for some $l \neq l' \in \{0, 1, \dots, T-1\}$. Then $\phi_{u^{(l)}}((l'-l) \pmod{T}) = \phi_u(l') - \phi_u(l) < 0$. Thus $\phi_{u^{(l)}}(n) \geq 0$ for $n = 0, 1, \dots, T$ implies that $\phi_u(l) = \min_{n=0,1,\dots,T} \phi_u(n)$. Analogously it follows that $\phi_{u^{(l)}}(n) \leq 0$ for $n = 0, 1, \dots, T$ if and only if $\phi_u(l) = \max_{n=0,1,\dots,T} \phi_u(n)$. ■

We define a partial order $\preceq$ on $P(T,S)$. This induces partial orders on $\widetilde{P}(T,S)$ from which we derive bounds for the expected average waiting time.

Definition 3.2 *For $u, v \in P(T,S)$ we say that $u \preceq v$ if $\kappa_u(n) \leq \kappa_v(n)$ for $n = 1, 2, \ldots, T$.*

We have the following (see [12] where it is proved for sequences of zeros and ones).

Proposition 3.3 *Let $u \in R(T,S)$. Then $u' \in R(T,S)$ if and only if $u \sim u'$. Moreover the partial order $\preceq$ on $P(T,S)$ induces a total order on $R(T,S)$.*

Since $R(T,S)$ is finite it follows from Proposition 3.3 that $R(T,S)$ contains a greatest element for this order. We denote this greatest element by $\omega(T,S)$ or just ω if no confusion is possible. For this element ω the partial sums $\kappa_\omega(n)$ for $n = 1, 2, \ldots, T$ are as great as possible under the restriction that ω is regular. For example $\omega(7,4)$ is the sequence $(1,1,0,1,0,1,0)$. From [12] we have the following lemma which can be used to determine $\omega(T,S)$ quickly.

Lemma 3.4

$$\kappa_{\omega(T,S)}(n) = \lceil n \cdot \frac{S}{T} \rceil \text{ for } n = 0, 1, \ldots, T.$$

We have seen in Proposition 3.3 that the regular sequences $R(T,S)$ form a conjugacy class in $P(T,S)$. It follows that $u \in R(T,S)$ if and only if u is a representative of $\widetilde{\omega}(T,S) \in \widetilde{P}(T,S)$. Combining Lemma 3.1, Proposition 3.3 and Lemma 3.4 we obtain the following theorem in which the partial order $\preceq$ is used to give a characterising property of the conjugacy class $R(T,k) = \widetilde{\omega}(T,S)$ of regular sequences in $\widetilde{P}(T,S)$.

Theorem 3.5 *Every conjugacy class $\widetilde{u}$ of $P(T,S)$ contains an upper bound of $R(T,S)$, i.e for every $u \in P(T,S)$ there exists a $v \in P(T,S)$ such that $v \sim u$ and $v \succeq w$ for every $w \in R(T,S)$.*

Proof. For $u \in P(T,S)$ let $l \in \{0, 1, \ldots, T-1\}$ be such that $\phi_u(l) = \min_{n=0,1,\ldots,T} \phi_u(n)$. Let $v = u^{(l)}$. Then $v \sim u$ and by Lemma 3.1 we have that $\phi_v(n) \geq 0$ for $n = 0, 1, \ldots, T$. Hence $\kappa_v(n) \geq \lceil n \cdot \frac{S}{T} \rceil$ for $n = 0, 1, \ldots, T$. So, by Lemma 3.4 we have that $v \succeq \omega(T,S)$. By Proposition 3.3 we have that $\omega(T,S) \succeq w$ and thus $v \succeq w$ for every $w \in R(T,S)$. ∎

We have the following preorders $\preceq_{\overline{g}}$, $\preceq_{\underline{g}}$ and $\preceq_g$ on $P(T,S)$ called the primal or upper graph order, the dual or lower graph order and the strong graph order respectively.

Definition 3.6 *Let $u, v \in P(T,S)$. Then $u \preceq_{\overline{g}} v$ if there exist $u', v' \in P(T,S)$ such that $u' \in \widetilde{u}$, $v' \in \widetilde{v}$ and $0 \leq \phi_{u'}(n) \leq \phi_{v'}(n)$ for $n = 0, 1, \ldots, T$ and $u \preceq_{\underline{g}} v$ if there exist $u'', v'' \in P(T,S)$ such that $u'' \in \widetilde{u}$, $v'' \in \widetilde{v}$ and $0 \geq \phi_{u''}(n) \geq \phi_{v''}(n)$ for $n = 0, 1, \ldots, T$. Further $u \preceq_g v$ if $u \preceq_{\overline{g}} v$ and $u \preceq_{\underline{g}} v$.*

The preorders $\preceq_{\overline{g}}$, $\preceq_{\underline{g}}$ and $\preceq_g$ on $P(T,S)$ are not partial orders. Namely, if $u, v \in P(T,S)$ are cyclic permutations of each other then it is easily seen that $u \preceq_g v$ and $v \preceq_g u$. Thus these orders are not antisymmetric. However, they induce (see Definition 3.7) partial orders on $\widetilde{P}(T,S)$.

Definition 3.7 *Let $u, v \in \widetilde{P}(T,S)$ and $u' \in P(T,S)$ be a representative of u and $v' \in P(T,S)$ be a representative of v. Then we say that $u \preceq_g v$ if and only if $u' \preceq_g v'$ and similar for the orders $\preceq_{\overline{g}}$ and $\preceq_{\underline{g}}$.*

In [12] it is shown that these induced graph orders on $\widetilde{P}(T,S)$ are partial orders.

For a sequence $u \in P(T,S)$ we have a primal unbalance $\overline{I}(u)$ and a dual unbalance $\underline{I}(u)$.

Definition 3.8 *Let $u \in P(T,S)$. Then the primal unbalance of u is*

$$\overline{I}(u) := \frac{1}{T} \cdot \sum_{n=1}^{T} (\kappa_{u'}(n) - \lceil n \cdot \frac{S}{T} \rceil),$$

where $u' \in \widetilde{u}$ *such that* $\phi_{u'}(n) \geq 0$ *for* $n = 1, 2, \ldots, T$. *The dual unbalance of* u *is*

$$\underline{I}(u) := \frac{1}{T} \cdot \sum_{n=1}^{T} (\lfloor n \cdot \frac{S}{T} \rfloor - \kappa_{u'}(n)),$$

where $u' \in \widetilde{u}$ *such that* $\phi_{u'}(n) \leq 0$ *for* $n = 1, 2, \ldots, T$.

Remark. For $u \in P(T, S)$ the primal unbalance is well defined. Namely, by Lemma 3.1 there exist $u' \in \widetilde{u}$ such that $\phi_{u'}(n) \geq 0$ for $n = 1, 2, \ldots, T$. Moreover, if $u', u'' \in \widetilde{u}$ are such that $\phi_{u'}(n) \geq 0$ and $\phi_{u''}(n) \geq 0$ for $n = 1, 2, \ldots, T$ then $\sum_{n=1}^{T} \kappa_{u'}(n) = \sum_{n=1}^{T} \kappa_{u''}(n)$ (see Theorem 3.9). Analogously it follows that the dual unbalance is well defined. It is easily seen that for every sequence $u \in P(T, S)$ it holds that $\overline{I}(u) \geq 0$ and $\underline{I}(u) \geq 0$. Namely, for $u' \in P(T, S)$ we have that $\phi_{u'}(n) \geq 0$ if and only if $\kappa_{u'}(n) \geq \lceil n \cdot \frac{S}{T} \rceil$ and $\phi_{u'}(n) \leq 0$ if and only if $\kappa_{u'}(n) \leq \lfloor n \cdot \frac{S}{T} \rfloor$. Moreover, a sequence $u \in P(T, S)$ is regular if and only if $\overline{I}(u) = 0$ if and only if $\underline{I}(u) = 0$ (see [12]).

Theorem 3.9 *Let* $u \in P(T, S)$. *Then the following statements are equivalent for* $l \in \{0, 1, \ldots, T-1\}$.

(i) $\overline{I}(u) = \frac{1}{T} \cdot \sum_{n=1}^{T} (\kappa_{u^{(l)}}(n) - \lceil n \cdot \frac{S}{T} \rceil)$.

(ii) $\sum_{n=1}^{T} \kappa_{u^{(l)}}(n) = \max_{i=0,1,\ldots,T-1} \sum_{n=1}^{T} \kappa_{u^{(i)}}(n)$.

(iii) $\min_{n=0,1,\ldots,T-1} \phi_{u^{(l)}}(n) = \phi_{u^{(l)}}(0) = 0$.

(iv) $\phi_u(l) = \min_{i=0,1,\ldots,T-1} \phi_u(i)$.

Proof. By Lemma 3.1 we have that (iii) implies (iv) and vice versa. We now prove that (ii) implies (iii). Suppose there exists some $t \in \{1, 2, \ldots, T-1\}$ such that $\phi_{u^{(l)}}(t) = -\mu$ with $\mu > 0$. Let x be the prefix of length t of $u^{(l)}$ and let y be the suffix of length $T - t$ of $u^{(l)}$. Then $u^{(l)} = xy$. Let $z := yx = u^{((l+t) \pmod{T})} \in \widetilde{u}$. Then $\phi_{u^{(l)}}(n) = \phi_z(n-t) - \mu$ for $n = t, t+1, \ldots, T$ and $\phi_{u^{(l)}}(n) = \phi_z(n+T-t) - \mu$ for $n = 1, 2, \ldots, t-1$. Hence $\sum_{n=1}^{T} \phi_{u^{(l)}}(n) = \sum_{n=1}^{T} \phi_z(n) - T \cdot \mu$ and thus

$$\sum_{n=1}^{T} \kappa_{u^{(l)}}(n) = \sum_{n=1}^{T} \kappa_z(n) - T \cdot \mu < \sum_{n=1}^{T} \kappa_z(n) \leq \max_{i=0,1,\ldots,T-1} \sum_{n=1}^{T} \kappa_{u^{(i)}}(n),$$

which contradicts (ii). From the definition of the primal unbalance it follows directly that (iii) implies (i) and to finish the proof we show that (i) implies (ii). Suppose that $\sum_{n=1}^{T} \kappa_{u^{(l)}}(n) < \max_{i=0,1,\ldots,T-1} \sum_{n=1}^{T} \kappa_{u^{(i)}}(n)$ and let $l \neq l' \in \{0, 1, \ldots, T-1\}$ such that $\sum_{n=1}^{T} \kappa_{u^{(l')}}(n) = \max_{i=0,1,\ldots,T-1} \sum_{n=1}^{T} \kappa_{u^{(i)}}(n)$. Since we have proved that (ii) implies (i) it follows that

$$\overline{I}(u) = \frac{1}{T} \cdot \sum_{n=1}^{T} (\kappa_{u^{(l')}}(n) - \lceil n \cdot \frac{S}{T} \rceil) > \frac{1}{T} \cdot \sum_{n=1}^{T} (\kappa_{u^{(l)}}(n) - \lceil n \cdot \frac{S}{T} \rceil),$$

which contradicts (i). □

The following result follows immediately from the definitions.

Lemma 3.10 *Let* $u, v \in P(T, S)$. *If* $u \preceq_{\overline{g}} v$ *then* $\overline{I}(u) \leq \overline{I}(v)$. *If* $u \preceq_{\underline{g}} v$ *then* $\underline{I}(u) \leq \underline{I}(v)$. *If* $u \preceq_g v$ *then* $\overline{I}(u) \leq \overline{I}(v)$ *and* $\underline{I}(u) \leq \underline{I}(v)$.

We have the following theorem, which shows that the primal and dual unbalance are multimodular with respect to the L - triangulation.

Theorem 3.11 *The primal unbalance and the dual unbalance are shift invariant and multimodular on* $P(T, S)$ *with respect to the base* $d_1, d_2, \ldots, d_T$ *where* $d_i = e_i - e_{i+1}$ *for* $i = 1, 2, \ldots, T-1$, $d_T = e_T - e_1$ *and* e_i *is the* i *-th unit vector of length* T *for* $i = 1, 2, \ldots, T$.

Proof. For $u \in P(T,S)$ let $f(u) = \overline{I}(u)$ and $g(u) = \underline{I}(u)$. Thus $f : P(T,S) \to \mathbb{Q}$ is the primal unbalance function and $g : P(T,S) \to \mathbb{Q}$ is the dual unbalance function. From the definition it follows that if $u', u'' \in P(T,S)$ are cyclic permutations of each other then $f(u') = f(u'')$ and $g(u') = g(u'')$. Thus the primal unbalance and dual unbalance are shift invariant. To prove the multimodularity of the primal unbalance function f we have to show that for every $u \in P(T,S)$ and $i, j \in \{1, 2, \ldots, T\}$, $i \neq j$ it holds that

$$f(u + d_i) + f(u + d_j) \geq f(u) + f(u + d_i + d_j). \tag{12}$$

We first show that for every $x \in P(T,S)$ and $i \in \{1, 2, \ldots, T\}$ it holds that

$$f(x + d_i) \leq f(x) + \frac{1}{T}. \tag{13}$$

Put $y = x + d_i$. Without loss of generality we can assume that $\sum_{n=1}^{T} \kappa_y(n) = \max_{i=0,1,\ldots,T-1} \sum_{n=1}^{T} \kappa_{y^{(i)}}(n)$. Then $f(y) = \frac{1}{T} \cdot \sum_{n=1}^{T} (\kappa_y(n) - \lceil n \cdot \frac{S}{T} \rceil)$ and

$$f(x) = f(y - d_i) \geq \frac{1}{T} \cdot \sum_{n=1}^{T} (\kappa_{y-d_i}(n) - \lceil n \cdot \frac{S}{T} \rceil) \geq \frac{1}{T} \cdot (\sum_{n=1}^{T} (\kappa_y(n) - \lceil n \cdot \frac{S}{T} \rceil) - 1) = f(y) - \frac{1}{T}$$

by Theorem 3.9. So, (13) holds and next we show that (12) holds. Without loss of generality we assume that $\min_{n=0,1,\ldots,T-1} \phi_u(n) = \phi_u(0) = 0$. Then $f(u) = \frac{1}{T} \cdot \sum_{n=1}^{T} (\kappa_u(n) - \lceil n \cdot \frac{S}{T} \rceil)$ by Theorem 3.9. Suppose that $i \neq T$ and $j \neq T$. Then it is easily seen that $\kappa_{u+d_i}(n) \geq \kappa_u(n)$, $\kappa_{u+d_j}(n) \geq \kappa_u(n)$ and $\kappa_{u+d_i+d_j}(n) \geq \kappa_u(n)$ for $n = 0, 1, \ldots, T$. Hence

$$\min_{n=0,1,\ldots,T-1} \phi_{u+d_i}(n) = \min_{n=0,1,\ldots,T-1} \phi_{u+d_j}(n) = \min_{n=0,1,\ldots,T-1} \phi_{u+d_i+d_j}(n) = \min_{n=0,1,\ldots,T-1} \phi_u(n) = 0.$$

From the definition of the primal unbalance it follows that $f(u + d_i) = f(u + d_j) = f(u) + \frac{1}{T}$ and $f(u + d_i + d_j) = f(u) + \frac{2}{T}$ and thus $f(u + d_i) + f(u + d_j) = f(u) + f(u + d_i + d_j)$. It remains to show that (12) holds in case $i = T$ or $j = T$ and we may assume that $j = T$ and $i \neq T$. Then from the preceding argument we have that $f(u + d_i) = f(u) + \frac{1}{T}$. Moreover by (13) we have that $f(u + d_i + d_j) \leq f(u + d_j) + \frac{1}{T}$. Hence

$$f(u + d_i) + f(u + d_j) \geq (f(u) + \frac{1}{T}) + (f(u + d_i + d_j) - \frac{1}{T}) = f(u) + f(u + d_i + d_j).$$

Thus the primal unbalance function f is multimodular and it follows analogously that the dual unbalance function g is multimodular. ■

Note. Since the primal and dual unbalance functions are shift invariant on $P(T,S)$ they induce functions on $\widetilde{P}(T,S)$ by $\overline{I}(v) = \overline{I}(u)$ and $\underline{I}(v) = \underline{I}(u)$ if $u \in P(T,S)$ is a representative of $v \in \widetilde{P}(T,S)$. Then by Theorem 3.11 we have for these induced primal and dual unbalance functions that they are elements of $\mathcal{F}^{\text{shift}}(D)$, where D is the multimodular matrix having the multimodular base $d_1, d_2, \ldots, d_T$ as row vectors.

Corollary 3.12 *For this multimodular base $d_1, d_2, \ldots, d_T$ we have for $u, v \in \widetilde{P}(T,S)$ that $u \leq_{mms} v$ implies that $\overline{I}(u) \leq \overline{I}(v)$ and $\underline{I}(u) \leq \underline{I}(v)$.*

4 Relations and counterexamples

4.1 The shift invariant cone order does not imply the graph order

Consider the following example. Let $u = (1, 3, 2, 2, 3) \in \mathcal{P}(5, 11)$ and $v = (2, 1, 3, 2, 3) \in \mathcal{P}(5, 11)$. Thus we consider sequences of length 5 and sum 11. It is easily seen that u and v are not ordered for the dual lower graph order (but they are ordered for the primal upper graph order). So, u and v are not graph ordered. However, we show that u and v are ordered for shift invariant multimodular functions with respect to the standard base d_0, d_1, d_2, d_3, d_4 where $d_i = e_i - e_{i+1}$ for $i = 1, 2, 3, 4$ and $d_0 = (-1, 0, 0, 0, 1)$. Namely, for a shift invariant multimodular function f we have that $f(u+d_4) = f((1, 3, 2, 3, 2)) = f(v)$ and also $f(u + d_1) = f((2, 2, 2, 2, 3) = f((2, 2, 2, 3, 2)) = f(u + d_1 + d_4)$ by shift invariance. By multimodularity we have that $f(u + d_1) + f(u + d_4) \geq f(u) + f(u + d_1 + d_4)$ and thus $f(u) \leq f(u + d_4) = f(v)$ for every multimodular shift invariant function f.

Conclusion. The shift invariant multimodular order does not imply the graph order. Thus the shift invariant multimodular order does not imply the shift invariant cone order or the shift invariant cone order does not imply the graph order (or both do not hold).

We now investigate whether $u \leq_{Cs} v$. To do this we take the regular sequence $r = (3, 2, 2, 2, 2)$ as root of all the cones we consider and we let u', v' run through all the shifts of u and v respectively. Then we check whether $u' \leq v'$ for some cone. Doing this we find that u and v are indeed ordered for the shift invariant cone order. Namely, for $u' = (3, 1, 3, 2, 2)$ and $v' = (2, 1, 3, 2, 3)$ we have that $u' = r + d_0 + d_1 + d_3 + d_4$ and $v' = r + 2d_0 + d_1 + d_3 + d_4$. Hence, if we consider the cone generated by $b_1 = d_0$, $b_2 = d_0 + d_1$, $b_3 = d_0 + d_1 + d_3$ and $b_4 = d_0 + d_1 + d_3 + d_4$ then $u' = r + b_4$ and $v' = r + b_1 + b_4$. By the way, this u' and v' is the only pair of shifts that are cone ordered for some cone with root in r. Since $u' \leq_C v'$ and thus $u \leq_{Cs} v$ we have the following conclusion.

Conclusion. The shift invariant cone order does not imply the graph order.

Remarks. The explanation for the fact that the ordering of u' and v' in this cone does not imply the lower graph order is the following. For u' you have to start with the second coordinate to get the graph for the lower graph order and for v' you have to start with the first coordinate. This is possible, because for r you have the start at the second coordinate for the lower graph order, while the generating base vectors of this cone "suggest" the lower graph order for starting at the first coordinate. One of the consequences is that this cone does contain other regular sequences than r. Namely, $r + b_1 = (2, 2, 2, 2, 3)$ is for example another regular sequence in this cone. It turns out that the cone order in such cones does not imply the graph order.
If you consider the mirrored sequences of u and v then you get a similar problem with the upper graph order instead of the lower graph order.
Note that u and v are ordered for the unbalance (despite that they are not ordered for the graph order). Namely, for the primal (upper) unbalance $\overline{I}$ we have that $\overline{I}(u) = 1 \leq 2 = \overline{I}(v)$ and for the dual (lower) unbalance $\underline{I}$ we have that $\underline{I}(u) = 1 = \underline{I}(v)$. Of course this order for the unbalance follows immediately from $\widetilde{u} \leq_{mms} \widetilde{v}$ and Corollary 3.12.

4.2 The shift invariant multimodular order does not imply the shift invariant cone order

A counterexample is given here to show that the shift invariant multimodular order does not imply the shift invariant cone order. In fact it is shown that the shift invariant cone order is not even a partial order, because it is not transitive.

Counterexample. Again we consider sequences in $P(5, 11)$ and we have the same multimodular standard base $\{d_i\}_{i=0,1\ldots,5}$ as before. Let $u = (1, 2, 3, 2, 3)$, $v = (2, 2, 3, 3, 1)$ and $w = (2, 1, 0, 4, 4)$ and as regular sequence we have again $r = (3, 2, 2, 2, 2)$. Now we have for the shift invariant cone order that $u \leq_{Cs} v$ and $v \leq_{Cs} w$. However, we do not have that $u \leq_{Cs} w$. Namely, $u' = (3, 2, 3, 1, 2) = r + d_3 = r + b_1$ and $v' = (2, 3, 3, 1, 2) = r + d_2 + 2d_3 + d_4 + d_0 = r + b_1 + b_4$, where $b_1 = d_3$, $b_2 = d_3 + d_0$, $b_3 = d_3 + d_0 + d_2$ and $b_4 = d_3 + d_0 + d_2 + d_4$. Hence $u \leq_{Cs} v$.
Further $v'' = (3, 1, 2, 2, 3) = r + d_0 + d_1 = r + b_2$ and $w'' = (2, 1, 0, 4, 4) = r + 4d_0 + 3d_1 + 2d_2 + 2d_4 = r + b_1 + b_2 + 2b_4$, where $b_1 = d_0$, $b_2 = d_0 + d_1$, $b_3 = d_0 + d_1 + d_2$ and $b_4 = d_0 + d_1 + d_2 + d_4$. Hence $v \leq_{Cs} w$. However, there exist no cone with root in r such that some shift of u is smaller in that cone than some shift of w. Hence, u and w are not shift invariant cone ordered, while from

$u \leq_{Cs} v$ and $v \leq_{Cs} w$ it follows that u and w are ordered for the shift invariant multimodular order.

Remark. So, it turns out that the shift invariant cone order is not a partial order, because it is not transitive. It is natural to consider the smallest preorder which contains the shift invariant cone order and is transitive. It is easily seen that this order implies the shift invariant multimodular order, but we do not know whether the converse is also true.

4.3 The graph order does not imply the shift invariant multimodular order

In this counterexample we show that the graph order does not imply the shift invariant multimodular order. Consider the sequences
$u = (0,0,0,1,0,0,1,0,0,1,0,0,0,1,0,1,0,1,1,0,0,1,1,1,1,1,0,1,0,1,0,1) \in P(32,15)$ and
$v = (0,0,0,0,0,0,0,0,0,0,0,1,1,1,0,1,1,0,1,1,0,1,1,0,1,1,0,1,1,1,0,1) \in P(32,15)$. Then we have that $u \leq_g v$. Thus u and v are ordered for the graph order and we have that $\overline{I}(u) < \overline{I}(v)$ and $\underline{I}(u) < \underline{I}(v)$. However, it does not hold that $f(u) \leq f(v)$ for every shift invariant multimodular function $f : P(32,15) \rightarrow \mathbb{R}$. Namely, we consider u and v as period cycles for the splitting sequences corresponding to the admission of arriving customers to some server. Suppose that the interarrival times are deterministic and equal to 1 and the service times are deterministic and equal to $(1+\epsilon)$, where ϵ is a small positive number. For $x \in P(32,15)$ let $W(x)$ be the average waiting time of customers admitted according to x. The average waiting time is multimodular with respect to the standard base of the L - triangulation (see [4]) and since the traffic intensity for the server is smaller than one we have that the average waiting time function W is also shift invariant (see [12]). Thus W is a shift invariant multimodular function. However, for the given interarrival and service times we have that $W(u) = \frac{11}{15} \cdot \epsilon$ and $W(v) = \frac{10}{15} \cdot \epsilon$. Thus $W(v) < W(u)$ for these service times and thus u and v are not ordered for the shift invariant multimodular order.

5 Bounding the Difference in Expected Average Waiting Time of Routing Sequences for One Queue

In this section we give a bound for the expected average waiting time for customers routed to a single server queue according to a routing sequence of zeros and ones. This bound depends on the unbalance of the routing sequence. To state the results we first give some definitions and notations.

Let u and v be infinite periodic integer sequences. Then we say that u is equivalent to v, $u \sim v$ if there exists a finite sequence w such that w is a period cycle of both u and v.
For the application as routing sequences it is shown in [12] that equivalent infinite periodic sequences have the same performance. Let $\mathcal{P}$ be the set of infinite periodic sequences of zeros and ones (modulo the equivalence relation), $\mathcal{R} \subseteq \mathcal{P}$ the subset of regular periodic sequences and for $d \in \mathbb{Q}$, $0 \leq d \leq 1$ let $\mathcal{P}(d) \subseteq \mathcal{P}$ be the subset of sequences with density d. We denote the regular sequence in $\mathcal{P}(d)$ by $\omega(d)$ or just ω. Note that if $u' \in \mathcal{P}(T,k)$ is a period cycle of $u \in \mathcal{P}$ then $u \in \mathcal{P}(\frac{k}{T})$. We define the unbalance for infinite periodic sequences of zeros and ones as follows.

Definition 5.1 *Let $u \in \mathcal{P}$ and let $u' \in \mathcal{P}(T,k)$ be a period cycle of u. Then we define the (primal) unbalance of u as $\overline{I}(u) := \overline{I}(u')$.*

The unbalance is well defined on $\mathcal{P}$. Namely, if u' and u'' are both period cycles of u then $\overline{I}(u') = \overline{I}(u'')$.

Let $\{T_i\}_{i=1,2,\ldots}$ be a sequence of arrival times of customers, with the convention that $T_1 = 0$. Put $\delta_i := T_{i+1} - T_i$ for $i = 1,2,\ldots$. Then $\{\delta_i\}$ is the sequence of interarrival times. Further these arriving customers are routed to a server according to some routing sequence $u = (u_1, u_2, \ldots)$ of zeros and ones. For such routing sequences we have the counting function $\kappa_u(n) = \sum_{t=1}^{n} u_t$

and we define a partial order $\preceq$ that is similar as the order on $P(T,S)$. Namely, we say that $u \preceq v$ if $\kappa_u(n) \leq \kappa_v(n)$ for $n = 0,1,\ldots$. Further we define the the following related function $\nu_u(i) : \mathbb{Z}_{\geq 0} \to \mathbb{Z}$ with $\nu_u(j) = \min\{n \in \mathbb{Z}_{\geq 0} : \kappa_u(n) = j\}$ and we put $\tau_u(j) = \sum_{i=\max(\nu_u(j-1),1)}^{\nu_u(j)-1} \delta_i$ for $j = 1,2,\ldots$. Then $\tau_u(j)$ is the time elapsed between the routing of the $(j-1)$-th and j-th customer to the server according to routing sequence u. If we put $\lambda_u(j) = \sum_{k=1}^{j} \tau_u(k)$ for $j = 1,2,\ldots$ then $\lambda_u(j)$ is the time at which the j-th customer is routed to the server according to routing sequence u. We have a sequence of service times $\{\sigma_j\}_{j=1,2,\ldots}$, where σ_j is the service time of the j-th customer that is routed to the server according to the routing sequence. Further we define $W_u(j)$ to be the workload for the server at the moment the j-th customer is routed to the server according to routing sequence u. In other words $W_u(j)$ is the waiting time for the j -th customer that is routed to the server. We assume that the server starts empty at $T_1 = 0$ and thus $W_u(1) = 0$. If the interarrival times $\{\delta_i\}$ and service times $\{\sigma_j\}$ are random variables then we say that $\overline{W}(u)$ is the almost sure long-run average waiting time of customers routed to the server according to routing sequence u if $\lim_{m\to\infty} \frac{1}{m} \cdot \sum_{j=1}^{m} W_u(j) = \overline{W}(u)$ with probability one. From ergodic theory we have the following theorem. See [12] for a proof of this theorem.

Theorem 5.2 *Suppose that the interarrival times $\{\delta_i\}$ of customers arriving at the system are independent and identically distributed (i.i.d.) random variables with mean δ and the service times $\{\sigma_j\}$ of the considered server are i.i.d. random variables with mean σ and independent of the interarrival times. Further let u' and u'' be routing sequences of zeros and ones that are both representatives of some $u \in \mathcal{P}(d)$ where $d \in \mathbb{Q}$ (thus they have a common period cycle) and $\frac{\sigma}{\delta} \cdot d < 1$. Then $\overline{W}(u')$ and $\overline{W}(u'')$ exist and are finite. Moreover $\overline{W}(u') = \overline{W}(u'')$.*

Let δ , σ and d be as in Theorem 5.2. Then we say that $\rho := \frac{\sigma}{\delta} \cdot d$ is the traffic intensity for the server. By Theorem 5.2 all the routing sequences which are representative of some $u \in \mathcal{P}(d)$ have the same long-run average waiting time if the stability condition $\rho < 1$ is fulfilled. Therefore in this case we denote this long-run average waiting time simply by $\overline{W}(u)$. The following theorem which is obtained by a sample path argument is the main result of this section.

Theorem 5.3 *Let the interarrival times $\{\delta_i\}$ and the service times $\{\sigma_j\}$ be as in Theorem 5.2. Further let $u \in \mathcal{P}(d)$ for some $d \in \mathbb{Q}$ such that $\rho < 1$ and let $\omega = \omega(d) \in \mathcal{P}(d)$ be the regular sequence of density d. Then*

$$\overline{W}(u) \leq \overline{W}(\omega) + \frac{\delta}{d} \cdot \overline{I}(u). \tag{14}$$

Remark. From the results of [1] and [3] we have that $\overline{W}(u) \geq \overline{W}(\omega)$.

We give an outline of the proof of Theorem 5.3. For a complete proof for this and following results in sections 5 and 6 see [12]. We first assume that the interarrival times $\{\delta_i\}_{i=1,2,\ldots}$ and service times $\{\sigma_j\}_{j=1,2,\ldots}$ are fixed sequences of non-negative real numbers. Then we have the following lemma.

Lemma 5.4 *Let $u = (u_1, u_2, \ldots)$ and $v = (v_1, v_2, \ldots)$ be routing sequences of zeros and ones and suppose that $u \preceq v$. Then*

$$W_v(j) + \lambda_v(j) \leq W_u(j) + \lambda_u(j) \text{ for every } j \in \mathbb{N}. \tag{15}$$

Note that $W_u(j) + \lambda_u(j)$ is the moment that the server starts serving the j-th customer that is routed to the server according to u. So, Lemma 5.4 states that if $u \preceq v$ then the moment when the service of the j -th customer that is routed to the server starts, is for v not later than for u. This can be proved by induction to j.

For every $j \in \mathbb{N}$ we have that $\lambda_u(j) - \lambda_v(j) = \sum_{i=\nu_v(j)}^{\nu_u(j)-1} \delta_i$. This is a sum of interarrival times δ_i, where the number of terms in the sum is $\nu_u(j) - \nu_v(j)$. Therefore, if we put $N_{uv}(m) = \sum_{j=1}^{m}(\nu_u(j) - \nu_v(j))$ for $m = 1,2,\ldots$, then we have the following Corollary by Lemma 5.4.

Corollary 5.5 *Let u and v be routing sequences with $u \preceq v$. For every $m \in \mathbb{N}$ we have that*

$$\sum_{j=1}^{m} W_v(j) \leq \sum_{j=1}^{m} W_u(j) + \sum_{j=1}^{m} \sum_{i=\nu_v(j)}^{\nu_u(j)-1} \delta_i. \tag{16}$$

In the double sum the number of terms is $N_{uv}(m)$.

In [12] it is shown that if $u = (u')^\infty$ and $v = (v')^\infty$ for some $u', v' \in \mathcal{P}(T,k)$ and $\omega(T,k) \preceq u' \preceq v'$ then

$$\lim_{m\to\infty} \frac{N_{uv}(m)}{m} = \frac{T}{k} \cdot (\overline{I}(v) - \overline{I}(u)). \tag{17}$$

Let $u \in \mathcal{P}(d)$ be as in Theorem 5.3. Then by Theorem 3.5 there exist $T, k \in \mathbb{N}$ with $d = \frac{k}{T}$ and $u'' \in \mathcal{P}(T,k)$ such that $u' := (u'')^\infty$ is a representative of u and $\omega' \preceq u'$ where $\omega' = \omega(T,k)^\infty$ is a representative of $\omega(d)$. By Theorem 5.2 we have that both $\overline{W}(u) = \overline{W}(u')$ and $\overline{W}(\omega) = \overline{W}(\omega')$ exist. We make a coupling between the interarrival times for u' and ω' and we also do this for the service times. After the coupling we can apply Corollary 5.5 and (17) and it follows that

$$\overline{W}(u') - \overline{W}(\omega') \leq \frac{\delta \cdot T}{k} \cdot (\overline{I}(u) - \overline{I}(\omega)) = \frac{\delta}{d} \cdot I(u),$$

which proves Theorem 5.3.

Remark. From the preceding arguments it follows that it is possible to refine Theorem 5.3 as follows: Let $u, v \in \mathcal{P}(d)$ be as in Theorem 5.3 and let ω' be a representative of $\omega(d)$ as above. Then, if there exist representatives u' and v' of u and v respectively such that $\omega' \preceq u' \preceq v'$ it holds that

$$\overline{W}(v) - \overline{W}(u) \leq \frac{\delta}{d}(\overline{I}(v) - \overline{I}(u)). \tag{18}$$

The reader can check that the condition for (18) is equivalent to the following: there exist $T, k \in \mathbb{N}$ and period cycles u'' and v'' of u and v respectively such that $d = \frac{k}{T}$, $u'', v'' \in \mathcal{P}(T,k)$ and $u'' \preceq_{\overline{g}} v''$. Moreover, the condition $\rho < 1$ in Theorem 5.3 is just necessary to apply Theorem 5.2. However, if both the interarrival times and the service times are deterministic then Theorem 5.2 also holds if $\rho = 1$ (see [14]). In fact we have for that case the following stronger result from which it follows that the bound of Theorem 5.3 is tight.

Theorem 5.6 *Suppose that the interarrival times and service times are deterministic and equal to δ respectively σ. Let $u \in \mathcal{P}(d)$ for some $d \in \mathbb{Q}$ such that $\delta = d\sigma$, hence $\rho = 1$, and let $\omega = \omega(d)$ be the regular sequence of density d. Then*

$$\overline{W}(u) = \overline{W}(\omega) + \frac{\delta}{d} \cdot \overline{I}(u). \tag{19}$$

Proof. (Outline) Let ω' be a representative of $\omega(d)$ as above and let u' be a representative of u such that $\omega' \preceq u'$. Then it can be shown that $W_{u'}(j) + \lambda_{u'}(j) = W_{\omega'}(j) + \lambda_{\omega'}(j) = (j-1)\sigma$ for every $j \in \mathbb{N}$. Hence we have for every $m \in \mathbb{N}$ that (16) holds with equality for u' and ω' with $\delta_i = \delta$ for $i = 1, 2, \ldots$. From this it follows that (14) also holds with equality. ■

Remark. The upper bound can be generalized to a sequence of multimodular functions satisfying the relations of section 3 of [3]. Indeed, it follows from the results there that the regular or bracket sequence is the minimal admission sequence. If the density is rational then the corresponding regular admission sequence is periodic. In order to obtain upper bounds as in Theorem 5.3 and Theorem 6.3 we need a generalization of Lemma 5.4, especially the relation (15). It is easily seen that for multimodular functions f_k corresponding to $W_u(j)$ it suffices that

$$f_k(a_1, \ldots, a_l + 1, a_{l+1} - 1, \ldots, a_k) \leq f_k(a_1, \ldots, a_l, a_{l+1}, \ldots, a_k) + \delta,$$

for $1 \leq l \leq k-1$ and some constant δ. Clearly, also other multimodular models do satisfy this relation, e.g. the models of [7].

6 Routing to Parallel Queues

In this section we derive upper bounds for the expected average waiting time for routing arriving customers to $N \geq 2$ parallel queues according to some periodic routing policy $U = (U_1, U_2, \ldots)$. The routing sequence $U = (U_1, U_2, \ldots)$ can be seen as an $\mathbb{N}$ - word on the alphabet $\{1, 2, \ldots, N\}$. Further to every letter $i \in \{1, 2, \ldots, N\}$ corresponds a sequence of zeros and ones $u^i = (u^i_1, u^i_2, \ldots)$ via the support of the letter, i.e. $u^i_n = 1$ if and only if $U_n = i$. We call u^i the routing or splitting sequence for queue i ([10], [13], [1], [14] and [12]). Extending the notion of unbalance for periodic sequences of zeros and ones we define a (total) unbalance $O(U)$ for periodic integer sequences and corresponding words on some finite alphabet. Then we extend the result of Theorem 5.3 for one queue to multiple queues by using the unbalance $O(U)$ of U.

Let $d_1, d_2, \ldots, d_N \in \mathbb{Q}_{>0}$ with $\sum_{i=1}^N d_i = 1$. Then we denote by $\mathcal{Q}(d_1, d_2, \ldots, d_N)$ all the infinite periodic $\mathbb{N}$ - words U on the alphabet $\{1, 2, \ldots, N\}$ for which $u^i \in \mathcal{P}(d_i)$ for $i = 1, 2, \ldots, N$. Further for $T, k_1, k_2, \ldots, k_N \in \mathbb{N}$ with $\sum_{i=1}^N k_i = T$ we denote by $\mathcal{Q}(\{T\}, k_1, k_2, \ldots, k_N)$ all the $\mathbb{N}$-words on the alphabet $\{1, 2, \ldots, N\}$ for which every subword of length T contains exactly k_i letters i for $i = 1, 2, \ldots, N$. Note that $\mathcal{Q}(\{T\}, k_1, k_2, \ldots, k_N) \subseteq \mathcal{Q}(\frac{k_1}{T}, \frac{k_2}{T}, \ldots, \frac{k_N}{T})$. For example if $U = (1, 2, 1, 2, 1, 3, 1, 2, 1, 1, 2, 3)^\infty$ then $U \in \mathcal{Q}(\{12\}, 6, 4, 2) \subseteq \mathcal{Q}(\frac{1}{2}, \frac{1}{3}, \frac{1}{6})$. Further $u^1 = (1, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0)^\infty \in \mathcal{P}(\frac{1}{2})$, $u^2 = (0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0)^\infty \in \mathcal{P}(\frac{1}{3})$ and $u^3 = (0, 0, 0, 0, 0, 1)^\infty \in \mathcal{P}(\frac{1}{6})$.

Let ψ be a routing policy and U the corresponding routing sequence. Then for $t \in \mathbb{N}$ we define $W(t) = W_\psi(t)$ as the waiting time of the t -th arriving customer if policy ψ is applied, which is the remaining workload for server U_t at the moment that the t-th customer arrives. We assume that all the servers $i \in \{1, 2, \ldots, N\}$ are empty at $T_1 = 0$ and thus $W(t) = 0$ if $t = \nu_{u^i}(1)$ for some $i \in \{1, 2, \ldots, N\}$. If the interarrival times and service times of the various servers are random variables then we say that $\overline{W}(\psi) = \overline{W}(U)$ is the almost sure long-run average waiting time of the arriving customers routed according to policy ψ if $\lim_{\tau \to \infty} \frac{1}{\tau} \cdot \sum_{t=1}^{\tau} W_\psi(t) = \overline{W}(\psi)$ with probability one. Let $\{\delta_i\}$ be the sequence of interarrival times and let $\{\sigma^i_j\}$ be the sequence of service times of server i for $i = 1, 2, \ldots, N$, i.e σ^i_j is the service time of the j - th customer that is routed to server i. We define $\overline{W}^i(\psi) = \overline{W}^i(u^i)$ as the almost sure long-run average waiting time of customers routed to server i if policy ψ is applied in the same way as we did in the previous section. The only differences are that routing sequence u is replaced by routing sequence u^i and the sequence of service times $\{\sigma_j\}$ is replaced with the sequence of service times $\{\sigma^i_j\}$. From Theorem 5.2 we obtain the following theorem.

Theorem 6.1 *Suppose that the interarrival times $\{\delta_i\}$ of customers arriving at the system are i.i.d. random variables with mean δ and for every $i \in \{1, 2, \ldots, N\}$ the service times $\{\sigma^i_j\}$ are i.i.d random variables with mean σ_i, and independent of the interarrival times. Let ψ be a routing policy that corresponds to some word $U \in \mathcal{Q}(d_1, d_2, \ldots, d_N)$ such that $\rho_i := \frac{\sigma_i}{\delta} \cdot d_i < 1$ for $i = 1, 2, \ldots, N$. Then $\overline{W}(\psi)$ exists and is finite. Moreover $\overline{W}(\psi) = \sum_{i=1}^n d_i \cdot \overline{W}^i(\psi)$.*

Definition 6.2 *Let $U \in \mathcal{Q}(d_1, d_2, \ldots, d_N)$ for some $d_i \in \mathbb{Q}_{>0}$ with $\sum_{i=1}^N d_i = 1$. Then we define the (total) unbalance of U as*

$$O(U) := \sum_{i=1}^{N} \overline{I}(u^i). \tag{20}$$

If it is clear from the context what is meant we just say unbalance and not total unbalance. Let the interarrival times $\{\delta_i\}$ and the service times $\{\sigma^i_j\}$ for $i = 1, 2, \ldots, N$ be as in Theorem 6.1. Further let $d_i \in \mathbb{Q}_{>0}$ for $i = 1, 2, \ldots, N$ with $\sum_{i=1}^N d_i = 1$. Recall that $\omega(d_i)$ is the regular sequence of density d_i. If $\overline{W}^i(\omega(d_i))$ exists and is finite for $i = 1, 2, \ldots, N$ then we put

$$\widetilde{R} = \widetilde{R}(d_1, d_2, \ldots, d_N) := \sum_{i=1}^{N} d_i \cdot \overline{W}^i(\omega(d_i)). \tag{21}$$

Remark. If the interarrival times and service times are random variables then for the existence of $\widetilde{R}$ it suffices that $\rho_i < 1$ for $i = 1, 2, \ldots, N$, while if the interarrival times and service times are deterministic then it suffices that $\rho_i \leq 1$ for $i = 1, 2, \ldots, N$. It is possible to extend the definition of $\widetilde{R}$ and some of the results to the case of irrational d_i. In general $\widetilde{R}$ depends on the distribution of the interarrival times and service times and in some cases it is possible to compute $\widetilde{R}$ explicitly. See [8] and [14] for an exact computation in case of deterministic interarrival and service times and see [13] for computations and bounds in general.

Combining Theorem 5.3 and Theorem 6.1 we obtain the following theorem.

Theorem 6.3 *Let the interarrival times $\{\delta_i\}$ and the service times $\{\sigma_j^i\}$ for $i = 1, 2, \ldots, N$ be as in Theorem 6.1. Let $d_i \in \mathbb{Q}_{>0}$ for $i = 1, 2, \ldots, N$ with $\sum_{i=1}^N d_i = 1$. Suppose that a routing policy ψ is applied with corresponding word $U \in \mathcal{Q}(d_1, d_2, \ldots, d_N)$ and $\rho_i < 1$ for $i = 1, 2, \ldots, N$. Then*

$$\overline{W}(\psi) - \widetilde{R} \leq \delta \cdot O(U).$$

From [3], $\overline{W}(\psi) \geq \widetilde{R}$, which is the lower bound obtained by replacing the routing sequence to queue i by the regular routing sequence with the same density for any queue i. Hence we have the following bounds.

$$\widetilde{R} \leq \overline{W}(\psi) \leq \widetilde{R} + \delta \cdot O(U). \tag{22}$$

Remark. Suppose that we have a queueing system where the arrivals are according to a Poisson process with parameter λ. Suppose that the service times of server i are exponentially distributed with parameter μ_i. If d_i, the fraction of jobs that is routed to server i, equals $\frac{1}{q_i}$ for some $q_i \in \mathbb{N}$, then $\overline{W}^i(\omega(d_i))$ can be calculated in the following way. For the routing sequence $\omega(d_i)$ we have that among every q_i arriving jobs exactly one job is routed to server i. So, the interarrival times at server i consist of q_i Poisson arrivals with parameter λ. Hence the interarrival times for the queue of server i are Erlang distributed, namely according to an $E_\lambda^{q_i}$ distribution. Thus $\overline{W}^i(\omega(d_i))$ is the same as the average waiting time for a $E_\lambda^{q_i}/M/1$ queue, where the parameter of the service times is μ_i. So (see [9]) if x_i is the unique solution in the interval $(0, 1)$ of the equation

$$x = \left(\frac{\lambda}{\lambda + \mu_i - \mu_i \cdot x}\right)^{q_i}, \tag{23}$$

then

$$\overline{W}^i(\omega(d_i)) = \frac{x_i}{\mu_i \cdot (1 - x_i)}. \tag{24}$$

In the following example we have calculated $\overline{W}^i(\omega(d_i))$ for $i = 1, 2, 3, 4$ by applying (24). Further we have explicitly calculated the lower bound $\widetilde{R}$ and upper bound $\widetilde{R} + \delta \cdot O(U)$ of (22) for $\overline{W}(\psi)$, where U is the word corresponding to the applied routing policy ψ.

Example 6.4 *We consider a queueing system with 4 parallel servers where the arrivals are according to a Poisson process with parameter $\lambda = 11$. Hence for the mean interarrival time δ we have that $\delta = \frac{1}{11}$. The arriving jobs are routed to the servers according to the policy ψ that corresponds to the word*

$$U = (1, 2, 3, 1, 4, 2, 1, 3, 1, 2, 4, 3)^\infty \in \mathcal{Q}(\{12\}, 4, 3, 3, 2) \subseteq \mathcal{Q}(\frac{1}{3}, \frac{1}{4}, \frac{1}{4}, \frac{1}{6}).$$

Then we have that $\overline{I}(u^1) = \frac{1}{12}$, $\overline{I}(u^2) = 0$, $\overline{I}(u^3) = \frac{1}{12}$, $\overline{I}(u^4) = 0$ and thus $O(U) = \frac{1}{6}$. For every $i \in \{1, 2, 3, 4\}$ the service times are exponentially distributed with parameter μ_i and $\mu_1 = 4$, $\mu_2 = \mu_3 = 3$ and $\mu_4 = 2$. Then we find by (24) that $\overline{W}^1(\omega(\frac{1}{3})) = 1.7792$ (rounded to 4 decimals) and thus by Theorem 5.3 we have that $\overline{W}^1(u^1) \leq 1.7792 + 3 \cdot \frac{1}{11} \cdot \frac{1}{12} = 1.8019$. Similarly we have that $\overline{W}^2(\omega(\frac{1}{4})) = \overline{W}^2(u^2) = 2.2105$, $\overline{W}^3(\omega(\frac{1}{4})) = 2.2105$, $\overline{W}^3(u^3) \leq 2.2408$ and $\overline{W}^4(\omega(\frac{1}{6})) = \overline{W}^4(u^4) = 3.0732$. Hence by (21) we have that

$$\widetilde{R} = \frac{1}{3} \cdot \overline{W}^1(\omega(\frac{1}{3})) + \frac{1}{4} \cdot \overline{W}^2(\omega(\frac{1}{4})) + \frac{1}{4} \cdot \overline{W}^3(\omega(\frac{1}{4})) + \frac{1}{6} \cdot \overline{W}^4(\omega(\frac{1}{6})) = 2.2105$$

and $\widetilde{R} + \delta \cdot O(U) = 2.2105 + \frac{1}{11} \cdot \frac{1}{6} = 2.2257$. *So, by (22) we have that* $2.2105 \leq \overline{W}(\psi) \leq 2.2257$.

The following theorem provides conditions under which the right-hand side inequality of (22) actually holds with equality.

Theorem 6.5 *Let the interarrival times* $\{\delta_i\}$ *be deterministic equal to* δ *and the service times* $\{\sigma_j^i\}$ *be deterministic equal to* σ_i *for* $i = 1, 2, \ldots, N$. *Let* $d_i \in \mathbb{Q}_{>0}$ *for* $i = 1, 2, \ldots, N$ *with* $\sum_{i=1}^N d_i = 1$. *Let* $p_i, q_i \in \mathbb{N}$ *be such that* $d_i = \frac{p_i}{q_i}$ *with* $\gcd(p_i, q_i) = 1$ *for* $i = 1, 2, \ldots, N$. *Suppose that a routing policy* ψ *is applied with corresponding word* $U \in \mathcal{Q}(d_1, d_2, \ldots, d_N)$ *and* $\rho_i = 1$ *for* $i = 1, 2, \ldots, N$. *Then*

$$\overline{W}(\psi) = \widetilde{R} + \delta \cdot O(U) = \delta \cdot \left(\frac{1}{2} - \sum_{i=1}^{N} \frac{1}{2q_i} + O(U) \right).$$

By combining Theorem 5.6 and Theorem 6.1 we get the first equality of Theorem 6.5. See [14] for the second equality which follows from the computation of $\widetilde{R}$ for this case.

Example 6.6 *We consider a queueing system with 3 parallel servers where the interarrival times are deterministic and equal to* $\delta = 3$. *The arriving jobs are routed to the servers according to the policy* ψ *that corresponds to the word*

$$U = (1, 2, 1, 2, 1, 3, 1, 2, 1, 3)^\infty \in \mathcal{Q}(\{10\}, 5, 3, 2) \subseteq \mathcal{Q}(\frac{1}{2}, \frac{3}{10}, \frac{1}{5}).$$

All the service times are deterministic and for server 1 they are equal to $\sigma_1 = 6$, *for server 2 equal to* $\sigma_2 = 10$ *and for server 3 equal to* $\sigma_3 = 15$. *Hence* $\rho_i = 1$ *for* $i = 1, 2, 3$. *Further* $\overline{I}(u^1) = 0$, $\overline{I}(u^2) = \frac{1}{10}$, $\overline{I}(u^3) = \frac{1}{10}$ *and thus* $O(U) = \frac{1}{5}$. *So, according to Theorem 6.5 we have that*

$$\overline{W}(\psi) = 3 \cdot \left(\frac{1}{2} - (\frac{1}{4} + \frac{1}{20} + \frac{1}{10}) + \frac{1}{5} \right) = \frac{9}{10}$$

which can be checked by direct calculation.

References

[1] E. Altman, B. Gaujal, and A. Hordijk. Admission control in stochastic event graphs. *IEEE Transaction on Automatic Control*, 45(5):854–868, 2000.

[2] E. Altman, B. Gaujal, and A. Hordijk. Balanced sequences and optimal routing. *Journal of the ACM*, 47(4):752–775, 2000.

[3] E. Altman, B. Gaujal, and A. Hordijk. Multimodularity, convexity and optimization properties. *Mathematics of Operations Research*, 25(2):324–347, 2000.

[4] E. Altman, B. Gaujal, and A. Hordijk. Simplex convexity with application to open-loop stochastic control in networks. In *39th Conf. on Decision and Control*. IEEE, 2000.

[5] E. Altman, B. Gaujal, and A. Hordijk. Regular ordering and applications in control policies. *Journal of Discrete Event Dynamic Systems*, 2002. to appear.

[6] E. Altman, B. Gaujal, and A. Hordijk. Discrete-event control of stochastic networks: Multimodularity and Regularity. in preparation.

[7] B. Gaujal, E. Altman and A. Hordijk. Optimal open-loop control of vacations, polling and service assignment. *Queueing Systems*, 36:303–325, 2000.

[8] B. Gaujal and E. Hyon. Optimal routing policy in two deterministic queues. *Calculateurs Parallèles*, 2001. accepted for publication, also available as INRIA RR-3997.

[9] D. Gross and C. Harris. *Fundamentals of queueing theory.* Series in Probabilty and mathematical statistics. Wiley, 1974.

[10] B. Hajek. Extremal splittings of point processes. *Mathematics of Operation Research*, 10(4):543–556, 1985.

[11] A. Hordijk. Comparison of queues with different discrete-time arrival processes. *Probability in Engineering and Information Sciences*, 15:1–14, 2001.

[12] A. Hordijk and D.A. van der Laan. Periodic routing to parallel queues with bounds on the average waiting time. Technical Report MI N. 2000-44, Leiden University, 2000.

[13] H. Shirakawa, M. Mori, and M. Kijima. Evaluation of regular splitting queues. *Comm. Statist. Stochastic Models*, 5:219–234, 1989.

[14] D.A. van der Laan. Routing jobs to servers with deterministic service times. Technical Report MI N. 2000-20, Leiden University, 2000.