Embedding multidimensional grids into optimal hypercubes

Zevi Miller * Dan Pritikin † I.H. Sudborough ‡

March 13, 2014

Abstract

Let $G$ and $H$ be graphs, with $|V(H)| \geq |V(G)|$, and $f : V(G) \to V(H)$ a one to one map of their vertices. Let \( \text{dilation}(f) = \max \{ \text{dist}_H(f(x), f(y)) : xy \in E(G) \} \), where $\text{dist}_H(v, w)$ is the distance between vertices $v$ and $w$ of $H$. Now let $B(G, H) = \min \{ \text{dilation}(f) \}$, over all such maps $f$.

The parameter $B(G, H)$ is a generalization of the classic and well studied “bandwidth” of $G$, defined as $B(G, P(n))$, where $P(n)$ is the path on $n$ points and $n = |V(G)|$. Let $[a_1 \times a_2 \times \cdots \times a_k]$ be the $k$-dimensional grid graph with integer values $1$ through $a_i$ in the $i$'th coordinate. In this paper, we study $B(G, H)$ in the case when $G = [a_1 \times a_2 \times \cdots \times a_k]$ and $H$ is the hypercube $Q_n$ of dimension $n = \lceil \log_2(|V(G)|) \rceil$, the hypercube of smallest dimension having at least as many points as $G$. Our main result is that

$$B([a_1 \times a_2 \times \cdots \times a_k], Q_n) \leq 3k,$$

provided $a_i \geq 2^{22}$ for each $1 \leq i \leq k$. For such $G$, the bound $3k$ improves on the previous best upper bound $4k + O(1)$. Our methods include an application of Knuth’s result on two-way rounding and of the existence of spanning regular cyclic caterpillars in the hypercube.

1 Introduction

In this paper we will usually follow standard graph theoretic terminology, as may be found for example in [28]. We let $P(t)$ stand for the path on $t$ vertices. The cartesian product $G \times H$ of two graphs $G$ and $H$ is the graph with vertex set $V = \{(v, w) : v \in V(G), w \in V(H)\}$ and edge set $E = \{(v, w)(v', w') : \text{either } v = v' \text{ and } ww' \in E(H), \text{ or } vv' \in E(G) \text{ and } w = w'\}$. All logarithms are taken base 2.

---

*Dept. of Mathematics, Miami University, Oxford, OH 45056, USA, millerz@muohio.edu
†Dept. of Mathematics, Miami University, Oxford, OH 45056, USA, pritikd@muohio.edu.
‡Program in Computer Science, University of Texas-Dallas, Richardson, Texas
1.1 Background and main result

The analysis of how effectively one network can simulate another, and the resulting implications for optimal design of parallel computation networks, are important topics in graph theoretic aspects of computer science. One of the measures of the effectiveness of a simulation is the dilation of the corresponding map (or “embedding”) of networks, defined as follows. Let \( G \) and \( H \) be two graphs and \( f : V(G) \rightarrow V(H) \) a map from the vertices of \( G \) to those of \( H \). As a convenience we typically write such a map as \( f : G \rightarrow H \), with the meaning that it is a map from vertices to vertices. Similarly we sometimes write \(|G|\) for \(|V(G)|\). Apart from an exception indicated below in a review of previous research on our topic, we will sup-

Let \( G \) and \( H \) be two graphs and \( f : V(G) \rightarrow V(H) \) a map from the vertices of \( G \) to those of \( H \). As a convenience we typically write such a map as \( f : G \rightarrow H \), with the meaning that it is a map from vertices to vertices. Similarly we sometimes write \(|G|\) for \(|V(G)|\). Apart from an exception indicated below in a review of previous research on our topic, we will sup-

Let \( G \) and \( H \) be two graphs and \( f : V(G) \rightarrow V(H) \) a map from the vertices of \( G \) to those of \( H \). As a convenience we typically write such a map as \( f : G \rightarrow H \), with the meaning that it is a map from vertices to vertices. Similarly we sometimes write \(|G|\) for \(|V(G)|\). Apart from an exception indicated below in a review of previous research on our topic, we will sup-

Let \( G \) and \( H \) be two graphs and \( f : V(G) \rightarrow V(H) \) a map from the vertices of \( G \) to those of \( H \). As a convenience we typically write such a map as \( f : G \rightarrow H \), with the meaning that it is a map from vertices to vertices. Similarly we sometimes write \(|G|\) for \(|V(G)|\). Apart from an exception indicated below in a review of previous research on our topic, we will sup-

As a convenience we typically write such a map as \( f : G \rightarrow H \), with the meaning that it is a map from vertices to vertices. Similarly we sometimes write \(|G|\) for \(|V(G)|\). Apart from an exception indicated below in a review of previous research on our topic, we will sup-

We follow the traditional view whereby \( V(Q_n) \) is the set of all strings of length \( n \) over the alphabet \( \{0,1\} \), where two such strings are joined by an edge if they disagree in exactly one coordinate. This departs in a trivial way from our notation above, where we would have required \( 1 \leq x_i \leq 2 \). Clearly \(|V(Q_n)| = 2^n\) and we let Opt\((G)\) be the smallest hypercube containing at least \(|V(G)|\) vertices, so Opt\((G) = Q_t\) where \( t = \lceil \log_2(|V(G)|) \rceil \).

There is a substantial literature on the simulation of various networks by hypercubes and their related networks; the butterfly, shuffle exchange and DeBruijn graphs. See the books [21] and [23] for excellent expositions on these topics, both emphasizing bounds on dilation and
congestion in graph embeddings; where the first also includes routing and implementation of various algorithms while the second gives a unified approach to applying separator theorems for deriving such bounds. An early survey on embedding graphs into hypercubes [22] mentions necessary and sufficient conditions (originating in [16]) for a graph to be a subgraph of some hypercube, and also the fact that for the complete binary tree $T_n$ on $2^n - 1$ vertices there is an embedding $f: T_n \rightarrow Q_n$ such that for every edge $xy \in E(T_n)$ we have $\text{dist}_{Q_n}(f(x), f(y)) = 1$ with the exception of a single edge where this distance is 2 [15]. In [6] it is shown how to embed any $2^n$ node bounded degree tree into $Q_n$ with $O(1)$ dilation and $O(1)$ edge congestion, as $n$ grows. In the same paper these results are extended to embedding bounded degree graphs with $O(1)$ separators. In [21] many-to-one maps of binary trees into hypercubes are considered, letting the load be the maximum number of tree nodes mapped onto a hypercube node. Using probabilistic methods and error correcting codes it is shown how to embed an $M$ node binary tree in an $N$ node hypercube with dilation 1 and load $O(M/N + \log(N))$, and how to perform the same embedding with dilation $O(1)$ and load $O(M/N + 1)$.

Concerning the embedding of multidimensional grids into hypercubes, observe first that if $p_1, p_2, \ldots, p_r$ are positive integers summing to $n$, and $G = [P(2^{p_1}) \times P(2^{p_2}) \times \cdots \times P(2^{p_r})]$, then $Q_n = \text{Opt}(G)$ and $Q_n$ contains $G$ as a spanning subgraph. Thus $B(G, \text{Opt}(G)) = 1$ in this case. In fact one can show that $[a_1 \times a_2 \times \cdots \times a_k]$ is a subgraph of $Q_n$ if and only if $n \geq \lceil \log(a_1) \rceil + \lceil \log(a_2) \rceil + \cdots + \lceil \log(a_k) \rceil$; see Problem 3.20 in [21]. Answering a question posed in [22] about 2-dimensional grids $G = [a_1 \times a_2]$, it is shown in [9] and in [8] that $B(G, \text{Opt}(G)) \leq 2$. In [9] it is also shown for arbitrary multidimensional grids $G = [a_1 \times a_2 \times \cdots \times a_k]$ that $B(G, \text{Opt}(G)) \leq 4k + 1$. Independently it was shown in [20] that $B(G, \text{Opt}(G)) \leq 4k - 1$ for such $G$, this upper bound being realized by a parallel algorithm on the hypercube. Still for such $G$, it was shown in [5] that $B(G, \text{Opt}(G)) \leq k$, assuming quite involved and restrictive inequality constraints on the $a_i$. Returning to dimension 2, it was shown in [18] that determining whether a given graph $G$ can be embedded in $\text{Opt}(G)$ with edge congestion 1 is NP-complete. Subsequently it was shown in [26] that any $G = [a_1 \times a_2]$ can be embedded in $\text{Opt}(G)$ with edge congestion at most 2 and dilation at most 3. Following up on a question posed in [22], the issue of many-to-one embeddings of 2 and 3 dimensional grids $G$ into hypercubes was explored in [24]. For these results, let $\text{Opt}(G)/2^t$ denote the hypercube of dimension $\lceil \log(|G|) \rceil - t$. If $f: G \rightarrow \text{Opt}(G)/2^t$ is a many-to-one map, then as above let the load of $f$ be $\max\{|f^{-1}(z)| : z \in \text{Opt}(G)/2^t\}$. It was shown in [24] that for a 2-dimensional grid $G$ there is a many-to-one map $f: G \rightarrow \text{Opt}(G)/2^t$ of dilation 1 and load at most $1 + 2^t$, and when $G$ is 3-dimensional there is a map $f: G \rightarrow \text{Opt}(G)/2$ of dilation at most 2 and load at most 3, and a map $f: G \rightarrow \text{Opt}(G)/4$ of dilation at most 3 and load at most 5.

The main result of the present paper is that $B([a_1 \times a_2 \times \cdots \times a_k], Q_n) \leq 3k$, provided $a_i \geq 2^{22}$ for each $1 \leq i \leq k$. This improves on the $4k - 1$ bound above under this condition on the $a_i$. We construct a one to one map $H^k: G \rightarrow \text{Opt}(G)$ realizing dilation $(H^k) \leq 3k$ and having congestion $O(k)$. Our construction uses the technique of two way rounding and the existence of regular spanning cyclic caterpillars in the hypercube.
1.2 Some notation

We will need to consider multidimensional grids for which each factor (in the cartesian product) with one possible exception is a path $P(m)$, where $m$ is a power of 2 and varies with the factor, as these will play the role of successive approximations to $Opt(G)$. Let $e_i = \lfloor \log_2(a_1 a_2 \cdots a_j) \rfloor$ for $1 \leq i \leq k$, with $e_0 = 0$. Letting $p_i = e_i - e_{i-1}$ for $1 \leq i \leq k$, we let $Opt'(G) = P(2^{p_1}) \times P(2^{p_2}) \times \cdots \times P(2^{p_k})$. So $Opt'(G)$ is a spanning subgraph of $Opt(G)$. For any $1 \leq t \leq k$ let $(Y_i') = P(2^{p_1}) \times P(2^{p_2}) \times \cdots \times P(2^{p_t})$, with these $p_i$. We let $Y_t = (Y_{t-1}) \times P(l)$, where $l$ is large enough. Thus $Y_t$ is the $t$-dimensional grid $P(2^{p_t}) \times P(2^{p_2-e_1}) \times \cdots \times P(2^{e_1-e_t-1}) \times P(l)$. The grids $Y_2, Y_3, \ldots$ will be the aforementioned successive approximations to $Opt(G)$. We will construct one to one maps $f_i : G \rightarrow Y_i$, $2 \leq i \leq k$. The final map $f_k$ will satisfy $f_k(G) \subseteq Opt'(G) \subseteq Opt(G)$.

For any point $x$ in a multidimensional grid, we let $x_i$ be its $i$’th coordinate (as suggested above), and when $i \leq j$ we let $x_{i\to j}$ be the $(j-i+1)$-tuple $(x_i, x_{i+1}, \ldots, x_j)$. So for example, let $f : G \rightarrow H$ be a map where $G$ and $H$ are both multidimensional grids and $H$ is of dimension $r$. Let $f(x) = (b_1, b_2, \ldots, b_r) \in V(H)$ for some $x \in V(G)$. Then by our notation $f(x)_2 = b_2$, while $f(x)_1 = (b_1, b_2, \ldots, b_r)$. We can also express dilation($f$) as dilation($f$) = $\max \{ \sum_{i=1}^{r} |f(x)_i - f(y)_i| : xy \in E(G) \}$.

For $1 \leq t < i$, a $t$-level of $Y_i$ is any $t$-dimensional subgrid of $Y_i$ obtained by fixing the last $i-t$ coordinates of points in $Y_i$. Recalling that $Y_i = \langle Y_{i-1} \rangle \times P(l)$, note that there are $l$ pairwise disjoint $(i-1)$-levels of $Y_i$, each isomorphic to $\langle Y_{i-1} \rangle$, and we denote by $Y_i^c$ the $(i-1)$-level all of whose points have last (that is, $i$’th) coordinate $c, 1 \leq c \leq l$. We also let $Y_i^{(r)} = \bigcup_{c=1}^{l} Y_i^c$. For fixed $i$ and $j, 2 \leq i \leq k$ and $j \geq 1$, we denote by $S_{ij}$ the subgraph of $Y_i$ induced by the vertices $\{(x_1, x_2, \ldots, x_{i-1}, y) \in Y_i : 1 \leq x_q \leq 2^{e_q-e_{q-1}} \text{ for } 1 \leq q \leq i-1, 1+(j-1)2^{e_{i-1}+\ldots+e_1} \leq y \leq j2^{e_{i-1}+\ldots+e_1} \}$, and call $S_{ij}$ an $i$-section of $Y_i$, more precisely, the $j$’th $i$-section. Thus we have $S_{ij} \cong \langle Y_{i-1} \rangle \times P(2^{e_{i-1}+\ldots+e_1}) \cong \langle Y_i \rangle$ for each $j$, and we let $S_{i}^{(r)} = \bigcup_{j=1}^{l} S_{ij}$.

As an example, let $G = [3 \times 7 \times 5 \times 9]$. Then $e_1 = 2, e_2 = 5, e_3 = 7, \langle Y_2 \rangle = [4 \times 8]$, and $Y_2 = 4 \times l$ for large $l$. Further, $Y_3^3 = \{(x_1, 3) : 1 \leq x_1 \leq 4 \}$, so $Y_3^3$ can be thought of as the third column of $Y_2$, and $Y_3^{(3)} = \bigcup_{j=1}^{l} Y_3^j$ is the graph induced by the union of the first three columns of $Y_2$. We have $S_{2}^j \cong \langle Y_2 \rangle = [4 \times 8]$ for each $j \geq 1$, and $S_2^j$ is the $[4 \times 8]$ subgrid of $Y_2$ induced by columns $8(j-1) + 1$ through $8j$. We have $\langle Y_3 \rangle = [4 \times 8 \times 4], Y_3^3 = [4 \times 8 \times l]$ for large $l$, and $Y_3^3 = \{(x_1, x_2, c) : 1 \leq x_1 \leq 4, 1 \leq x_2 \leq 8 \}$ for any fixed integer $c \geq 1$. So $Y_3^c$ is the $c$’th 2-level of $Y_3$ (ordered by altitude, or third coordinate), and $Y_3^{(j)} = \bigcup_{j=1}^{l} Y_3^j$ is the 3-dimensional subgrid of $Y_2$ induced by the first $j$ many 2-levels of $Y_2$. Further $S_{3}^{(3)} \cong [4 \times 8 \times 4]$ is the 3-dimensional subgrid of $Y_3$ induced by the set of 2-levels $\{Y_3^c : 4(t-1) + 1 \leq c \leq 4t \}$.

So $S_3^{(j)} = \bigcup_{j=1}^{l} S_3^j$ is the 3-dimensional subgrid $[4 \times 8 \times 4j]$ of $Y_3$ induced by the first $4j$ many 2-levels of $Y_3$.

We let $u_i(G) = \lfloor \left[ \frac{|G|}{|Y_{i-1}|} \right] \rfloor = \left[ \frac{|G|}{2^{e_{i-1}}} \right]$, which we abbreviate with just $u_i$ when $G$ is understood by context or is an arbitrary grid. Since each $(i-1)$-level of $Y_i$ has size $|Y_{i-1}| = 2^{e_{i-1}}, u_i$ is the minimum number of $(i-1)$-levels of $Y_i$ whose union could contain the image $f_i(G)$. Our maps $f_i$ will satisfy $f_i(G) \subseteq Y_i^{u_i}$ for each $2 \leq i \leq k$.

We will need the analogue of a $t$-level for $G = [a_1 \times a_2 \times \cdots \times a_k]$. For $2 \leq i \leq k-1$, an $i$-page of $G$ is any $i$-dimensional subgrid of $G$ induced by all vertices of $G$ having the same
last \( k - i \) coordinate values. Let \( P_i = a_{i+1}a_{i+2} \cdots a_k \), which is the number distinct \( i \)-pages in \( G \). We define a linear ordering \( \prec_i \) on these \( i \)-pages as follows. Let \( D_i \) and \( D_i' \) be two \( i \)-pages, with fixed last \( k - i \) coordinate values \( c_{i+1}, c_{i+2}, \ldots, c_k \) and \( c'_{i+1}, c'_{i+2}, \ldots, c'_k \) respectively. Then \( D_i \prec_i D_i' \) in this ordering if at the maximum index \( r \), \( i + 1 \leq r \leq k \), where \( c_r \neq c'_r \) we have \( c_r < c'_r \). Now index the \( i \)-pages of \( G \) relative to this ordering by \( D_i^j \), \( 1 \leq j \leq P_i \), where \( r < s \) if and only if \( D_i^r \prec_i D_i^s \). Also let \( D_i^{(0)} = \bigcup_{j=1}^{P_i} D_i^j \).

As an example, consider the 4-dimensional grid \( H = [3 \times 7 \times 3 \times 2] \), containing 6 many 2-pages each isomorphic to \([3 \times 7]\). For fixed \( i \) and \( j \), \( 1 \leq i \leq 3 \) and \( 1 \leq j \leq 2 \), denote by \( D_2(i,j) \) the 2-page of \( G \) given by \( D_2(i,j) = \{(x_1, x_2, i, j) \in H : 1 \leq x_1 \leq 3, 1 \leq x_2 \leq 7\} \). Then the above ordering of 2-pages of \( H \) is given by \( D_2^{1} = D_2(1,1) \prec_2 D_2^{2} = D_2(2,1) \prec_2 D_2^{3} = D_2(3,1) \prec_2 D_2^{4} = D_2(1,2) \prec_2 D_2^{5} = D_2(2,2) \prec_2 D_2^{6} = D_2(3,2) \). There are two 3-pages in \( H \), given by \( D_3^{1} = \{D_1^{1}, D_1^{2}, D_1^{3}\} \) and \( D_3^{2} = \{D_1^{4}, D_1^{5}, D_1^{6}\} \), and we have \( D_3^{1} \prec_3 D_3^{2} \). As this example illustrates, for \( 3 \leq i \leq k \) the ordering \( \prec_{i-1} \) is a refinement of the ordering \( \prec_i \) in that if \( D_{i-1}^{p'} \subseteq D_{i}^{q'} \) and \( D_{i-1}^{p} \subseteq D_{i}^{q} \) with \( p' < q' \), then \( p < q \).

From here on we fix \( G = [a_1 \times a_2 \times \cdots \times a_k] \) to be a \( k \)-dimensional grid. We summarize the above notation, together with selected notation items to be introduced later, in Appendix 2 for convenient reference.

2 The 2-dimensional mapping

Recall the 2-dimensional grid \( Y_2 = P(2^{e_1}) \times P(l) \), with \( e_1 = \lfloor \log_2(a_1) \rfloor \) and \( l \) sufficiently large. In this section we construct a map \( f_2 : G \rightarrow Y_2^{(u_2)} \), and we abbreviate \( m = u_2 \) throughout this section, so \( m = \lfloor \log_2(a_1) \rfloor \). So \( f_2 \) will satisfy \( f_2(G) \subseteq P(2^{e_1}) \times P(2^{\lfloor \log_2(|G|)\rfloor - e_1}) \subseteq \text{Opt}(G) \). Additional work will show that for any edge \( vw \in E(G) \) we have that \( |f_2(v) - f_2(w)| + |f_2(v) - f_2(w)_2| \) is small. This map \( f_2 \), resembling a map constructed in \([9]\) and \([26]\), will be the first step in an inductive construction leading to a low dilation embedding \( f_k : G \rightarrow \text{Opt}'(G) \subseteq \text{Opt}(G) \).

We use the following notation. Let \( G(r) \) denote the infinite 2-dimensional grid having \( r \) rows, so the vertex and edge sets of \( G(r) \) are \( V(G(r)) = \{(x,y) \in \mathbb{Z}^2 : 1 \leq x \leq r, 1 \leq y < \infty\} \), and \( E(G(r)) = \{(x_1,y_1)(x_2,y_2) : |x_1 - x_2| + |y_1 - y_2| = 1\} \). We let \( C_i \) denote the set of vertices \((x,y)\) of \( G(r) \) with \( x = i \), and refer to this set as “chain \( i \)”, or the “\( i \)th chain” of \( G(r) \).

We can view \( V(G) \) as a subset of \( V(G(a_1)) \) by a natural correspondence \( \kappa : V(G) \rightarrow V(G(a_1)) \) defined as follows. Let \( W_i = \prod_{t=2}^{k} a_t \) for \( 2 \leq i \leq k \). For any vertex \( x = (x_1, x_2, \ldots, x_k) \) of \( G \), let \( \kappa(x) = (x_1, y) \), where \( y = (x_k - 1)W_{k-1} + (x_{k-1} - 1)W_{k-2} + \cdots + (x_3 - 1)W_2 + x_2 \). To see the action of \( \kappa \), let \( \rho_t \) be the subset of \( V(G) \) consisting of vertices \( x \in V(G) \) with \( x_t = 1, 1 \leq t \leq a_t \). Then \( \kappa \) maps the points of \( \rho_t \) to the first \( W_k \) points of \( C_t \) in lexicographic order; that is, if \( z = (t, x_2, x_3, \ldots, x_k) \) and \( z' = (t', x'_2, x'_3, \ldots, x'_k) \) are two points of \( \rho_t \), then \( \kappa(z)_1 = \kappa(z')_1 = t \) and \( \kappa(z)_2 < \kappa(z')_2 \) if and only if at the largest index \( r \) where \( x_r \neq x'_r \) we have \( x_r < x'_r \). Recall now the ordering \( D_i^j \), \( 1 \leq j \leq P_i \), of \( i \)-pages defined at the end of the last section. Then for fixed \( 2 \leq i \leq k \) and \( 1 \leq j \leq P_i \), \( \kappa(D_i^j) \) is the subset of points in \( G(a_1) \) given by \( \{(t,y) : 1 \leq t \leq a_1, (j-1)a_2a_3a_4 \cdots a_i + 1 \leq y \leq ja_2a_3a_4 \cdots a_i\} \).

Our method in this section is to first construct a low dilation map \( f : G(a_1) \rightarrow Y_2 \). We then obtain the desired \( f_2 \) as the composition \( f_2 = f \circ \kappa : G \rightarrow Y_2^{(m)} \). For the rest of this section we literally identify any point \( x \in V(G) \) with \( \kappa(x) \), dropping further references to \( \kappa \).
itself. Thus, once \( f : G(a_1) \rightarrow Y_2 \) is constructed, our map \( f_2 : G \rightarrow Y_2^{(m)} \) will henceforth be viewed as the restriction of \( f \) to \( G \subset G(a_1) \) (under the identification \( x \leftrightarrow \kappa(x) \)).

The map \( f : G(5) \rightarrow Y_2 \) is shown in Figure 1(b), and the map \( f_2 : [3 \times 7 \times 4 \times a_4] \rightarrow Y_2 \), obtained by restriction from \( f : G(3) \rightarrow Y_2 \), is shown in Figure 2.

To begin the description of \( f \), let positive integers \( 1 \leq i \leq a_1 \) and \( j \geq 1 \) be given. Then \( f \) will map either 1 or 2 points of chain \( C_i \) to \( Y_2^j \) (which recall is column \( j \) of \( Y_2 \)). We encode this information by defining a \((0,1)\) matrix \( R \) having \( a_1 \) rows and infinitely many columns indexed by the positive integers, where \( R_{ij} = 1 \) (resp. \( R_{ij} = 0 \)) means that \( C_i \) has 2 (resp. 1) image points in \( Y_2^j \) under the map \( f \). In the first case these 2 image points are successive in \( Y_2^j \).

Define the first column of \( R \) by

\[
R_{i1} = \lfloor \left( \frac{2^{e_1} - a_1}{a_1} \right)i \rfloor - \left\lfloor \left( \frac{2^{e_1} - a_1}{a_1} \right)(i-1) \right\rfloor, \quad 1 \leq i \leq a_1.
\]

For \( j > 1 \), let \( R_{i,j} = R_{i-1,j-1} \) where the row index is viewed modulo \( a_1 \). Thus \( R \) is just a circulant matrix whose columns are obtained by successive downward shifts of the first column with wraparound, and is illustrated for \( a_1 = 5 \) in Figure 1(a). The following Lemma shows that the set of 1’s in any set of consecutive entries of some row or column of \( R \) depends only on the number of such entries, and is one of two successive integers depending on that number.

**Lemma 2.1** The matrix \( R \) has the following properties.

(a) The sum of entries in any column of \( R \) is \( 2^{e_1} - a_1 \).

(b) The sum of any \( t \) consecutive entries in any row or column of \( R \) is either \( S_t \) or \( S_t + 1 \), where \( S_t = \lfloor (\frac{2^{e_1} - a_1}{a_1})t \rfloor \).

**Proof.** For (a), it suffices to prove the claim for column 1 of \( R \), since any other column of \( R \) is just a circular shift of column 1. This column sum is \( \sum_{i=1}^{a_1} \lfloor (\frac{2^{e_1} - a_1}{a_1})i \rfloor - \lfloor (\frac{2^{e_1} - a_1}{a_1})(i-1) \rfloor \), which telescopes to \( 2^{e_1} - a_1 \).

Consider (b). From the circulant property of \( R \) and the constant column sum property from part (a), it suffices to prove this claim for any sum of \( t \) consecutive entries (mod \( a_1 \)) in column 1, say a telescoping sum of the form

\[
\sum_{i=r}^{r+t-1} R_{i1} = \lfloor (\frac{2^{e_1} - a_1}{a_1})(r + t - 1) \rfloor - \lfloor (\frac{2^{e_1} - a_1}{a_1})(r - 1) \rfloor.
\]

Now letting \( A = (\frac{2^{e_1} - a_1}{a_1})t \) and \( B = (\frac{2^{e_1} - a_1}{a_1})(r - 1) \), we see that this sum is \( \lfloor A + B \rfloor - \lfloor B \rfloor \). But any such difference is either \( |A| \) or \( |A| + 1 \), so our sum is \( S_t \) or \( S_t + 1 \) as claimed. \( \square \)

The basic idea in constructing \( f \) is to fill out \( Y_2^{(m)} \) with images under \( f \) column by column of \( Y_2^{(m)} \) in increasing column index. For any fixed column of \( Y_2^{(m)} \), say the \( j \)’th column \( Y_2^j \), we proceed down that column mapping points of \( G(a_1) \) in nondecreasing chain index. The set \( f(C_i) \cap Y_2^j \) will consist of either 1 point or 2 successive points of \( Y_2^j \), depending on whether \( R_{ij} = 0 \) or 1 respectively.

The construction of \( f \) follows. While reading the construction below, the reader may wish to consult Figure 1 illustrating the matrix \( R \) and corresponding map \( f : G(5) \rightarrow Y_2 \). For \( 1 \leq r \leq a_1 \), we let \( C^{(r)} = \bigcup_{i=1}^r C_i \).
Construction of the map \( f : G(a_1) \to Y_2 \)

1. The images of \( f \) in \( Y_2^1 \) are given as follows.
   (1a) Initial points of \( C_1 \) are mapped to \( Y_2^1 \) as follows.
   If \( R_{11} = 0 \), then \( f(1, 1) = (1, 1) \) and \( c_1 \leftarrow 1 \).
   If \( R_{11} = 1 \), then \( f(1, 1) = (1, 1) \), \( f(1, 2) = (2, 1) \), and \( c_1 \leftarrow 2 \).
   (1b) For \( i = 1 \) to \( a_1 - 1 \)

\[
\begin{align*}
\text{begin} \\
\text{Assume inductively that } f(C(i)) \cap Y_2^1 = \{(x, 1) : 1 \leq x \leq c_i \} \text{ for an integer } c_i. \text{ Then initial points of } C_{i+1} \text{ are mapped to } Y_2^1 \text{ as follows.} \\
\text{If } R_{11} = 0, \text{ then } f(i + 1, 1) = (c_i + 1, 1), \text{ and } c_{i+1} \leftarrow c_i + 1. \\
\text{If } R_{11} = 1, \text{ then } f(i + 1, 1) = (c_i + 1, 1), f_2(i + 1, 2) = (c_i + 2, 1), \text{ and } c_{i+1} \leftarrow c_i + 2. \\
\text{If } i = a_1 - 1, \text{ then go to step 2.} \\
\text{end}
\end{align*}
\]

2. For all \( 1 \leq t \leq N \) (with \( N \) sufficiently large), the images of \( f \) in \( Y_2^j \) are given inductively as follows.

For \( j = 1 \) to \( N \)

\[
\begin{align*}
\text{begin} \\
\text{Suppose inductively that every point of } Y_2^{(j)} \text{ has been defined as an image under } f. \text{ Suppose also that for each } 1 \leq i \leq a_1, \text{ we have } C_i \cap f^{-1}(Y_2^{(j)}) = \{(i, t) : 1 \leq t \leq n_{ij} \} \text{ for suitable integers } n_{ij}, \text{ with } n_{0j} = n_{a0} = 0. \\
\text{[Comment : So the first } n_{ij} \text{ points of } C_i \text{ have been mapped to } Y_2^{(j)}.] \\
\text{Now define } f(G(a_1)) \cap Y_2^{j+1} \text{ as follows.} \\
\text{For } r = 0 \text{ to } a_1 - 1 \\
\text{begin} \\
\text{Suppose inductively that } f(C(r)) \cap Y_2^{j+1} \text{ is an initial segment of } Y_2^{j+1}, \text{ say given by} \\
f(C(r)) \cap Y_2^{j+1} = \{(x, j + 1) : 1 \leq x \leq c_{r,j+1} \}, \text{ where } c_{0,j} = 0 \text{ for all } j \geq 1. \\
\text{Define } f(C_{r+1}) \cap Y_2^{j+1} \text{ as follows.} \\
\text{(2a) If } R_{r+1,j+1} = 0, \text{ then } f(r + 1, 1 + n_{r+1,j}) = (c_{r,j+1} + 1, j + 1), \text{ and } c_{r+1,j+1} \leftarrow c_{r,j+1} + 1. \\
\text{(2b) If } R_{r+1,j+1} = 1, \text{ and } j + 1 \text{ is even then } f(r + 1, 1 + n_{r+1,j}) = (c_{r,j+1} + 2, j + 1), f(r + 1, 2 + n_{r+1,j}) = (c_{r,j+1} + 1, j + 1), \text{ and } c_{r+1,j+1} \leftarrow c_{r,j+1} + 2. \\
\text{(2c) If } R_{r+1,j+1} = 1 \text{ and } j + 1 \text{ is odd, then } f(r + 1, 1 + n_{r+1,j}) = (c_{r,j+1} + 1, j + 1), f(r + 1, 2 + n_{r+1,j}) = (c_{r,j+1} + 2, j + 1), \text{ and } c_{r+1,j+1} \leftarrow c_{r,j+1} + 2. \\
\text{end}
\end{align*}
\]

Toward analyzing this construction, recall that \( m = \lceil \frac{|G|}{2a_1} \rceil \). With the goal of showing (in the theorem which follows) that \( f(G) \subseteq Y_2^{(m)} \), we analyze \( f^{-1}(Y_2^{(m)}) \). Let \( C_i(t) = \{(i, y) : 1 \leq y \leq t \} \) be the set of the first \( t \) points of chain \( C_i \). Now by steps 2a and 2b, \( f(C_i) \) contributes either one point or two successive points to any column \( Y_2^j \), depending on whether \( R_{ij} = 0 \) or 1 respectively. So \( f(C_i) \) contributes exactly \( j + \sum_{t=1}^{j} R_{it} \) points to \( Y_2^{(j)} \). Thus letting \( N_{ij} = j + \sum_{t=1}^{j} R_{it} \), we have \( f^{-1}(Y_2^{(m)}) \cap C_i = C_i(N_{im}) \). So let \( G(a_1, N_{im}) \) be the subgraph
of $G(a_1)$ induced by $\bigcup_{i=1}^{a_1} C_i(N_{im})$. The next theorem gives various properties of $f$, including $Y_2^{(m-1)} \subset f(G) \subseteq f(G(a_1, N_{im})) = Y_2^{(m)}$.

Finally, we define $f_2 : G \rightarrow Y_2^{(m)}$ as the restriction of $f$ to the subgraph $G$ of $G(a_1)$. In Figure 2 we illustrate part of $f_2(G)$ for $G = [3 \times 7 \times 4 \times a_4]$ for some $a_4 > 1$. Each 2-page $D^i_2$ of $G$ is isomorphic to $[3 \times 7]$, and the images $f_2(D^i_2)$, $1 \leq i \leq 4$, are shown in detail with dividers separating the images $f_2(D^i_2)$ and $f_2(D^{i+1}_2)$ of successive 2-pages. Near these dividers, for each chain and each 2-page we have placed a box around the image of the chain’s first point in that 2-page. For example, under the letter $C$ there are three boxed points, representing the images of the first points of each of the three chains in the third 2-page $D^3_2$. We have also labeled three points in the figure by their preimages in $G$. For example $(2, 4, 4, ...)$ indicates the preimage $(x_1, x_2, x_3, \cdots, x_k) \in D^4_2 \subset G$ with $x_1 = 2, x_2 = 4, x_3 = 4$, and so on. There are four 2-pages of $G$ in each 3-page of $G$, and Figure 2 includes the image $f_2(D^1_2)$ of the first 3-page, consisting of $\bigcup_{i=1}^{4} f_2(D^i_2)$. The beginnings of $f_2(D^5_2)$, this being the first 2-page of the 3-page $D^3_2$, and of $f_2(D^3_2) = \bigcup_{i=5}^{8} f_2(D^i_2)$ are also illustrated at the far right in the same figure.

**Theorem 2.2** The map $f : G(a_1) \rightarrow Y_2$ constructed above has the following properties. Let $m = \lceil \frac{|G|}{a_1} \rceil$, $S_i = \lfloor \frac{a_i}{a_1} \rfloor t$ (as in Lemma 2.1), $N_{ij} = j + \sum_{t=1}^{j} R_{it}$, and let $G(a_1, N_{im})$ be the subgraph of $G(a_1)$ induced by $\bigcup_{i=1}^{a_1} C_i(N_{im})$.

(a) For each $i$ and $j$, $|f(C_i) \cap Y^j_2| = 1$ or 2, depending on whether $R_{ij} = 0$ or 1 respectively. Further, if $R_{ij} = 1$, then $f(C_i) \cap Y^j_2$ consists of two successive points of $Y^j_2$. Also, $f$ is monotone in the sense that $f(i, j, 2) \leq f(i, j + 1, 2) \leq f(i, j + 1, 1)$ for each $i, j$.

(b) Let $L_r(j) = f(C^{(r)}) \cap Y^j_2$. Then $L_r(j)$ is an initial segment, say $\{(d, j) : 1 \leq d \leq |L_r(j)|\}$, of $Y^j_2$, with $|L_r(j)| = r + \sum_{i=1}^{r} R_{ij}$.

(c) For $i, h, j \geq 1$ with $h \leq j$, let $\pi(i, h \rightarrow j)$ be the number of points of $C_i$ mapped to columns $h$ through $j$ of $Y_2^{(m)}$. Then $\pi(i, 1, j) = N_{ij}$. Further, for any $r, s \geq 1$ we have $\pi(i, r \rightarrow r + j) = j + 1 + S_{j+1}$ or $j + 2 + S_{j+1}$, and $|\pi(i, r \rightarrow r + j) - \pi(i, s \rightarrow s + j)| \leq 1$.

(d) For any $1 \leq r_1 < r_2 \leq a_1$ we have $N_{r_1, j'} - N_{r_2, j'} \leq 1$. Also with $L_r(j)$ as in (b), for any $1 \leq j_1 < j_2 \leq m$ we have $|L_r(j_1)| - |L_r(j_2)| \leq 1$, and $|L_{r+1}(j_1)| - |L_r(j_2)| \leq 2$.

(e) $f(G(a_1, N_{im})) = Y_2^{(m)}$.

(f) $G \subset f(G(a_1, N_{im}))$, and $Y_2^{(m-1)} \subset f(G)$.

(g) For any $i, j, r$ we have $|f(i, r) - f(j, r)| \leq 1$.

(h) For any $i, j, r$ we have $|f(i, j) - f(i, r)| \leq 2$.

(i) Suppose that $|f(C_r) \cap Y^j_2| = 2$. Then $|L_r(j)| \geq |L_r(j + 1)|$ and $N_{r, j} \geq N_{r+1, j}$.

We omit the involved but straightforward proof of this theorem here, and give it in Appendix 1. The properties listed above can be easily verified in the examples illustrated in Figures 1 and 2.

The following corollary will be used later in proving our dilation bound and the containment $f_2(G) \subseteq Opt(G)$. Its proof also appears in Appendix 1.

**Corollary 2.3** Let $v$ and $w$ be adjacent points of $G$. The map $f_2 : G \rightarrow Y_2^{(m)}$ has the following properties.
(a) \(|f_2(v)_1 - f_2(w)_1| \leq 3, \text{ and } |f_2(v)_2 - f_2(w)_2| \leq 1.\)
(b) \(|f_2(v)_1 - f_2(w)_1 + |f_2(v)_2 - f_2(w)_2| \leq 3.\)
(c) \(f_2(G) \subseteq Opt(G).\)
(d) Let \(T\) and \(T'\) be segments of \(p\) consecutive points on chains \(C_i\) and \(C_j\) respectively, \(1 \leq i, j \leq a_1,\) where possibly \(i = j.\) Let \(c\) and \(c'\) be the number of columns of \(Y_2\) spanned by \(f_2(T)\) and \(f_2(T')\) respectively. Then \(|c - c'| \leq 1.\)
(e) For \(1 \leq r \leq P_2,\) let \(r' = \min\{c: f_2(D_2^{(r)}) \subseteq Y_2^{(c)}\}.\) Then \(Y_2^{(r') - 1} \subset f_2(D_2^{(r)})\) and \(|Y_2^{(r')} - f_2(D_2^{(r')})| < 2^{e_1}.\)

3 The idea of the general construction, with examples

In this section we give the idea behind our general construction, saving complete details and proofs of validity for later sections. We continue with the notation of Section 1.2.

The overall plan is to construct a sequence of maps \(f_i: G \rightarrow [(Y_{i-1}) \times P(u_i)] = Y_i^{(u_i)},\)

\(u_i = \lceil \frac{|G|}{|V_{i-1}|} \rceil = \lceil \frac{|G|}{2^{e_i-1}} \rceil, \) \(2 \leq i \leq k,\) the first of which is \(f_2\) from section 2. Since \(u_i \leq 2^{|log_2(G)|-e_i-1,\) we will have \(f_i(G) \subseteq Y_i^{(u_i)} \subseteq (Y_{i-1}) \times P(2^{|log_2(G)|-e_i-1}).\) The last graph is a spanning subgraph of \(Opt(G),\) so \(f_i(G) \subseteq Opt(G)\) for \(2 \leq i \leq k.\) In particular, for \(i = k\) we get \(f_k(G) \subseteq Opt(G).\) The maps \(f_i\) will be successive approximations to \(f_k\) in that for any \(x \in V(G)\) and \(2 \leq i \leq k\) we will have \(f_i(x)_{1-|i-1|} = f_k(x)_{1-|i-1|}.\)

The final map \(f_k\) gives the basic geometry of our construction. We then apply a labeling \(L\) of the points of \(Opt(G)\) with hypercube addresses from \(Opt(G)\) to obtain the final embedding \(H^k: G \rightarrow Opt(G),\) where \(H^k = L \circ f_k.\) We construct the maps \(f_i\) inductively, letting \(f_{i+1}\) be the composition \(f_{i+1} = \sigma_i \circ f_i \circ f_i,\) using maps \(I_i\) and \(\sigma_i\) described below. Recall that \(P_i = a_{i+1}a_{i+2} \cdots a_k\) is the number of \(i\)-pages in \(G.\)

So suppose \(f_i: G \rightarrow Y_i^{(u_i)}\) has been constructed, and we outline the construction of \(f_{i+1}: G \rightarrow Y_i^{(u_i+1)}\). Now \(I_i: Y_i^{(u_i)} \rightarrow S_i^{(P_i)} \subset Y_i\) is a one to one “inflation” map which spreads out the image \(f_i(G)\) “evenly” in \(S_i^{(P_i)}\) by successively “skipping over” certain carefully chosen \((i-1)\)-levels of \(Y_i\) that are designated “blank”. See Figure 3 for an example where \(i = 2,\) and where \(f_2\) is the map of Figure 2, and blank 1-levels (columns) are shaded.

We let \(S_i^{(P_i)}(G)\) be the set of points in \(S_i^{(P_i)}\) lying in nonblank \((i-1)\)-levels (i.e. levels not designated “blank”) of \(S_i^{(P_i)}\). We stipulate that \((I_i \circ f_i)(G) \subset S_i^{(P_i)}(G).\) The number of successive \(i\)-sections \(S_i^{(P_i)}\) in the range of \(I_i\) is the same as the number of \(i\)-pages \(D_i^r\) in \(G,\) each equaling \(P_i,\) and for each \(r\) we associate \(S_i^{(P_i)}\) with \(D_i^r\) in a sense to be made clear below.

Define \(I_i: Y_i^{(u_i)} \rightarrow S_i^{(P_i)}\) as follows, for now assuming that certain \((i-1)\)-levels of \(Y_i\) (all lying within \(S_i^{(P_i)}\) ) have been designated blank. For \(x = (x_1, x_2, \ldots, x_i) \in Y_i^{(u_i)},\) we let \(I_i(x) = (x_1, x_2, \ldots, x_{i-1}, x'_i),\) where \(x'_i\) is the common \(i\)-coordinate in the \(x_i\)th nonblank \((i-1)\)-level of \(Y_i\) (in order of increasing \(i\)-coordinate). Now for any \(S \subseteq Y_i^{(u_i)},\) let \(I_i(S) = \{I_i(s) : s \in S\}.\) Observe that by its definition, \(I_i\) preserves \((i-1)\)-levels; that is \(I_i(Y_i^s) = Y_i^{s'},\) is an \((i-1)\)-level of \(Y_i.\) Also for \(1 \leq s < t \leq u_i,\) where \(I_i(Y_i^s) = Y_i^{s'}\) and \(I_i(Y_i^t) = Y_i^{t'},\) we have \(s' < t';\) that is, \(I_i\) preserves the order (by increasing \(i\)-coordinate) of \((i-1)\)-levels. We can picture \(I_i\) as an order preserving spreading out of the \(u_i\) many \((i-1)\)-levels of \(Y_i^{(u_i)}\) containing \(f_i(G)\) among
the $P_2e_i-e_{i-1}$ many $(i - 1)$-levels of $S_i^{(r)}$. The image $I_i(Y_i^{(u_i)})$ becomes the set of nonblank $(i - 1)$-levels $(u_i$ of them) in $S_i^{(r)}$. The remaining $P_2e_i-e_{i-1}-u_i$ many $(i - 1)$-levels of $S_i^{(r)}$ are blank, and distributed among the nonblank $(i - 1)$-levels so that certain balance properties outlined below are satisfied.

Consider the example $G' = [3 \times 7 \times 4 \times 3]$. Start with $f_2 : G' \rightarrow Y_2^{(u_2(G'))}$ given in the previous section, where $u_2(G') = \lceil \frac{529}{4} \rceil = 63$. Figure 2 gives the initial part of $f_2(G')$, while Figures 3(a) and b), illustrate the map $I_2 : Y_2^{(u_2(G'))} \rightarrow S_2^{(r)}$, $P_2 = 12$. Each 2-section $S_2^{(r)}$, $1 \leq j \leq 12$, satisfies $S_2^j \cong (Y_i) \times P(2e_i-e_{i-1}) = P(4) \times P(8)$. So $S_2^{(r)}$ has $P_22e_i-e_{i-1} = 12 \cdot 8 = 96$ columns (i.e. $(i - 1)$-levels where $i = 2$), of which the 63 nonblank ones comprising $I_2(Y_2^{(63)})$ contain $(I_2 \circ f_2)(G')$, while the remaining 33 blank ones (shaded) contain no points of $(I_2 \circ f_2)(G')$. Note in the figure how $I_2$ preserves the order in $S_2^{(12)}$ of the 63 columns in the domain $Y_2^{(63)}$ of $I_2$, and the 33 blank columns (shaded) of $S_2^{(12)}$ are distributed throughout $S_2^{(12)}$.

Turning to arbitrary $G$, the quantity and distribution of blank $(i - 1)$-levels within $S_i^{(r)}$ will be such that for each $1 \leq r \leq P_i$, the subgraph $S_i^{(r)}$ of $Y_i$ has barely enough nonblank $(i - 1)$-levels to host $(I_i \circ f_i)(D_i^{(r)})$. In effect, we want $S_i^{(r)}$ to have at least as much nonblank volume as $|D_i^{(r)}|$ for each $r \geq 1$, but barely so in increments the size of an $(i - 1)$-level in $Y_i$. We formulate this condition precisely as follows. Let $s_i(j)$ be the number of blank $(i - 1)$-levels in section $S_i^{(j)}$. Observing that $|D_i^{(r)}| = a_1a_2 \cdots a_i$ for any $j$, that each $(i - 1)$-level of $Y_i$ has $2e_i-e_{i-1}$ points, and that each section $S_i^{(j)}$ consists of $2e_i-e_{i-1}$ many pairwise disjoint $(i - 1)$-levels, we require that

$$\left\lfloor \frac{r a_1a_2 \cdots a_i}{2e_i-e_{i-1}} \right\rfloor + \sum_{j=1}^{r} s_i(j) = r 2e_i-e_{i-1} \quad (1)$$

for each $1 \leq r \leq P_i$.

Next, the map $\sigma_i : (I_i \circ f_i)(G) \rightarrow (Y_i) \times P(u_{i+1}) = Y_i^{(u_{i+1})}$ “stacks” the sets $S_i^{(j)} \cap (I_i \circ f_i)(G)$, $1 \leq j \leq P_i$, over the single section $S_i^{(j)} \cong (Y_i)$ as follows. Let $x = (x_1, x_2, \ldots, x_i) \in S_i^{(j)} \cap (I_i \circ f_i)(G)$, $1 \leq j \leq P_i$. Let $\bar{x}_i \equiv x_i \pmod{2e_i-e_{i-1}}$, $1 \leq \bar{x}_i \leq 2e_i-e_{i-1}$, be the congruence class of $x_i$ mod $2e_i-e_{i-1}$. Define the first $i$ coordinates of $\sigma_i(x)$ by $\sigma_i(x)_{1 \rightarrow i} = (x_1, x_2, \ldots, x_{i-1}, \bar{x}_i)$. Observe that since $x \in S_i^{(j)}$, we have $x_i = (j-1)2e_i-e_{i-1}+\bar{x}_i$, and that $\sigma_i(x)_{1 \rightarrow i} \in S_i^{(1)}$. To get the $i+1$st (and last) coordinate, let $c$ be the number of points $y = (y_1, y_2, \ldots, y_i) \in S_i^{(1)} \cap (I_i \circ f_i)(G)$, $1 \leq t \leq j$, satisfying $\sigma_i(y)_{1 \rightarrow i} = \sigma_i(x)_{1 \rightarrow i}$. Then define $\sigma_i(x) = (x_1, x_2, \ldots, x_{i-1}, \bar{x}_i, c)$.

Finally define $f_{i+1}$ as the composition $f_{i+1} = \sigma_i \circ I_i \circ f_i$.

For a fixed point $(x_1, x_2, \ldots, x_{i-1}, \bar{x}_i) \in S_i^{(1)}$, we view the set of images $(x_1, x_2, \ldots, x_{i-1}, \bar{x}_i, c)$ under $\sigma_i$ as a stack addressed by $(x_1, x_2, \ldots, x_{i-1}, \bar{x}_i)$ extending into the $(i + 1)$st dimension. Thus each point $(x_1, x_2, \ldots, x_{i-1}, \bar{x}_i)$ of $S_i^{(1)}$ becomes the address of such a stack, and the image $\sigma_i(x) = (x_1, x_2, \ldots, x_{i-1}, \bar{x}_i, c)$ is the $c$’th point “up” in this stack. See Figure 4 to which we return later with a full explanation, for an initial look at $\sigma_2$. Since the domain of $\sigma_i$ is $(I_i \circ f_i)(G)$, which is a set contained in the collection of nonblank $(i - 1)$-levels of $S_i^{(r)}$, it follows that the points in blank $(i - 1)$-levels of $S_i^{(r)}$ make no contribution under the map $\sigma_i$ to the aforementioned stacks. We can picture the images $\sigma_i(x)$, $x \in S_i^{(j)} \cap (I_i \circ f_i)(G)$, as “falling through” blank $(i - 1)$-levels $Y_i^{d} \subset S_i^{(r)}$, $t < j$, with $d \equiv x_i$ (mod $2e_i-e_{i-1}$). But if such a $Y_i^{d}$ is nonblank, and if $z = (x_1, x_2, \cdots, x_{i-1}, d) \in Y_i^{d} \cap (I_i \circ f_i)(G)$, then $\sigma_i(z)_{1 \rightarrow i} = \sigma_i(x)_{1 \rightarrow i}$. Thus $z$ does contribute (under $\sigma_i$) to the same stack as $x$; that
is, both $\sigma_i(z)$ and $\sigma_i(x)$ belong to the stack addressed by $(x_1, x_2, \ldots, x_{i-1}, x_{i}) \in S_i^2$. Further, $\sigma_i(z_{i+1}) < \sigma_i(x_{i+1})$ by definition of $c$ and since $t < j$; that is, $\sigma_i(z)$ appears “below” $\sigma_i(x)$ in this stack. So we see that each section $S_i^2$, $1 \leq j < P_i$, contributes (via $\sigma_i$) either 0 or 1 point to the stack $(x_1, x_2, \ldots, x_{i-1}, \bar{x}_{i}) \in S_i^2$. It contributes 1 point precisely when its unique $(i - 1)$-level $Y_i^d \subset S_i^2$ satisfying $d \equiv \bar{x}_{i} \pmod{2^{e_{i-1}-e_{i-1}}}$ is not blank and the point $z = (x_1, x_2, \ldots, x_{i-1}, d) \in V_i^d$ of $S_i^2$ lies in $(I_i \circ f_1)(G)$.

As an example, consider $G'' = [3 \times 7 \times 4] \subset G'$. We construct $f_2(G'') = (\sigma_2 \circ I_2 \circ f_2)(G'')$. Note that $e_1 = 2$, $e_2 = 5$, $P_2 = 4$, $u_2(G'') = [\frac{84}{7}] = 21$, and $u_3(G'') = [\frac{84}{32}] = 3$. Again each 2-section $S_i^2$, $1 \leq j \leq P_2 = 4$, is isomorphic to $P(2^{e_1}) \times P(2^{e_2-e_1}) = P(4) \times P(8)$, and $S_2^{(P_2)}$ has $P_2 \cdot 2^{e_2-e_1} = 4 \cdot 8 = 32$ columns.

Step 1: Start with $f_2 : G'' \to Y_2^{(u_2(G''))} = Y_2^{(21)}$ given in section 2 and illustrated as $f_2(D_2^1)$ in Figure 2.

Step 2: Perform the map $I_2 : Y_2^{(21)} \to S_2^{(P_2)} = S_2^{(4)}$ by letting $I_2((a, b)) = (a', b')$, where $b'$ is the $b$th nonblank column of $S_2^{(P_2)}$, ordered by increasing second coordinate. The result is shown in Figure 3a, where $I_2(f_2(G''))$ is unshaded and blank columns are shaded. The choice of blank columns will be discussed later.

On comparing with Figure 2, we see that $I_2(f_2(G''))$ is indeed obtained from $f_2(G'')$ by skipping over the designated blank columns. Each 2-section $S_i^2 \subset S_2^{(4)}$, $1 \leq i \leq 4$, is the host for the 2-page $D_i^2$ of $G''$, with the possible exception of the images of first points of chains in $D_i^2$ (these being boxed in Figure 3) which may appear in the last nonblank column of $S_i^{2^{-1}}$ (instead of in $S_i^2$).

Step 3: Perform the stacking map $\sigma_2 : (I_2 \circ f_2)(G'') \to Y_3^{(u_3(G''))} = Y_3^{(3)}$. Thus for each $(a, b) \in S_i^2 \cap (I_2 \circ f_2)(G'')$, we let $\sigma_2((a, b)) = (a, \bar{b}, c)$, where $\bar{b} \equiv y \pmod{8}$, $1 \leq \bar{b} \leq 8$, and $c = \{(a, b') \in S_i^2 \cap (I_2 \circ f_2)(G'') : t \leq j, b' \equiv b \pmod{8}\}$. As above, we view $(a, \bar{b})$ as the address in $S_i^2$ of the stack on which $(a, b)$ has been placed by $\sigma_2$, and $\sigma_2((a, b))$ is the $c$th point “up” on this stack.

In Figure 4 we illustrate $f_3(G'') = \sigma_2(I_2 \circ f_2)(G'')$; that is, how $\sigma_2$ stacks the four sets $S_i^2 \cap (I_2 \circ f_2)(G'')$, $1 \leq j \leq 4$, over $S_i^2 = Y_2$ to yield $f_3(G'')$. At top center we begin with $I_2(Y_2^{(21)}) \subset S_2^{(P_2)} = S_2^{(4)}$ where the four 2-sections $S_2^j$, $1 \leq j \leq 4$ comprising $S_2^{(4)}$, are placed vertically in succession for convenience. We now perform the stacking, with the result shown at lower left in the figure. Here we regard the bottom $4 \times 8$ layer (of the three layers in the result) as a copy of $S_2^1$, each of whose 32 points is the address of a stack. Consider the 4 points lying in column 2 of this $S_2^1$. As shown in the figure, each of these points is the address of a stack of height 3. For brevity let us write $\sigma_2(S_2^i)$ for $\sigma_2(S_2^i \cap (I_2 \circ f_2)(G''))$. The contributions to any one of these 4 stacks come from $\sigma_2(S_2^1)$, $\sigma_2(S_2^3)$, and $\sigma_2(S_2^4)$ (indicated respectively by the labels 1, 3, 4 in these stacks). To see how this happens, refer back to the top center of this figure and look at the points in column 2 in the 4 sections $S_2^j$, $1 \leq j \leq 4$, being stacked. Since column 2 of the second of these, $S_2^2$, is blank, the points in it make no contribution to the stacks whose addresses lie in column 2 of $S_2^1$ at lower left. But the points in column 2 of $S_2^3$ and $S_2^4$ (at top center) lie in nonblank columns and are contained in $(I_2 \circ f_2)(G'')$. Hence they do contribute to the column 2 stacks of $S_2^1$ at lower left. We can view such points in column 2 of $S_2^3$ and $S_2^4$ in the top center as “falling through” blank column 2 in $S_2^2$ under the action of $\sigma_2$. Similarly every point in column 1 of $S_2^1 = Y_2$ at lower left is the address of a stack of height 2. The contributions to these stacks are from $\sigma_2(S_2^2)$, $\sigma_2(S_2^3)$, indicated...
by labels 2 and 3 in these stacks. These contributions “fall through” blank column 1 of $S_j^1$ (at top center) under the action of $\sigma_2$. The lower right of the figure shows how individual image points are affected by this stacking. For example, the images under $\sigma_2$ of paths in $S_j^2 \cap (I_2 \circ f_2) (G'')$ (in bold at top center) jump between levels of the final result at lower right. Note also that the maximum stack height is indeed $u_3(G'') = 3$, achieved at stacks addressed by points in columns 2, 3, 5, 6, and 8 of $S_j^2$. So $f_3(G'') \subseteq Y_3^3$.

In the rest of this section we show how to assign blank columns in $S_j^{(P_2)}$ (used in constructing $I_2$ and then $f_3$), and generally how to assign blank $(i - 1)$-levels in $S_i^{(P_1)}$ (used in constructing $I_i$ and then $f_{i+1}$) for an arbitrary multidimensional grid $G$. We will see that the ability to make this assignment is implied by the existence of a certain class of $(0, 1)$ matrices, whose construction we give in the next section. For now we describe conditions which our assignment will satisfy that are sufficient for making our embedding $f_k$ (and later $H^k$) have the required dilation and containment properties. To start, consider the assignment of blank columns in $S_j^{(P_2)}$.

(a) The number $s_2(j)$ of blank columns in each 2-section $S_j^2$, $1 \leq j \leq P_2$, is chosen so that after $(I_2 \circ f_2)(G)$ “skips over” these columns, each subgraph $S_2^{(r)}$ (the union of the first $r$ many 2-sections), $1 \leq r \leq P_2$, has barely enough nonblank columns to host the image $(I_2 \circ f_2) (D_2^{(r)})$ of the first $r$ many 2-pages of $G$. This is expressed in equation (1) in the case $i = 2$, which says that \[ \left\lceil \frac{r |e|}{2} \right\rceil + \sum_{j=1}^r s_2(j) = r 2^{e-\epsilon_1} \] for each $1 \leq r \leq P_2$. Thus each image $(I_2 \circ f_2) (D_2^r)$ lies in its own section $S_2^r$, apart from images of first points of chains in $D_2^r$ as explained previously and shown in Figure 3. In the example $G'' = [3 \times 7 \times 4]$, the sequence $s_2(j)$ can be found recursively from equation (1) (with $i = 2$) to be $s_2(1) = 2$, $s_2(2) = 3$, $s_2(3) = 3$, and $s_2(4) = 3$, and these numbers of blank columns are shown in the four sections respectively in Figure 3a).

(b) The blank columns are distributed over the various 2-sections $S_j^2$ so that for $xy \in E(G)$ the contribution to $\text{dist}_{Y_k}(f_k(x), f_k(y))$ from $|f_k(x)_2 - f_k(y)_2|$ is small. As background, note that $\text{dist}_{Y_k}(f_k(x), f_k(y)) = \sum_{i=1}^k |f_k(x)_i - f_k(y)_i|$. It turns out that $f_k$ will satisfy $|f_k(x)_i - f_k(x)_i| = |f_i(x)_i - f_{i+1}(x)_i| = |(I_2 \circ f_1)(x)_i - (I_2 \circ f_1)(y)_i|$ for each $1 \leq i \leq k - 1$, where the last difference is taken mod $2^{e-\epsilon_1}$. The last equality for $i = 2$ can be verified in the construction of $f_3(G'')$ and figures above.

Focusing on the case $i = 2$, suppose that $x$ and $y$ agree in their first two coordinates. So $x$ and $y$ are corresponding points in their respective 2-pages, say $x \in D_2^s$ and $y \in D_2^t$. Now we wish to keep $|(I_2 \circ f_2)(x)_2 - (I_2 \circ f_2)(y)_2|$ small mod $2^{e_2-\epsilon_1}$. Thus we want $(I_2 \circ f_2)(x)$ to be skipping over roughly the same number of blank columns in $S_j^2$ as does $(I_2 \circ f_2)(y)$ in $S_j^2$. To accomplish this, given that $s$ and $t$ are arbitrary as are $x$ and $y$ as corresponding points, we will require a strong balance, over all 2-sections $S_j^2$, in the frequency of blank columns in any initial segment of columns in $S_j^2$.

A similar balance will be required in the frequency of blank $(i - 1)$-levels in any initial segment of $(i - 1)$-levels of any $i$-section $S_i^j$. A precise formulation of this requirement with additional detail will be described below.

(c) The assignment of blank columns across all 2-sections $S_j^2$, $1 \leq j \leq P_2$, is uniformly distributed mod $2^{e_2-\epsilon_1}$. This condition allows us to stack the sections $S_j^2$ on top of each other, over a single 2-section $S_1^2 = \langle Y_2 \rangle$, so that the stacks addressed in $S_2^2$ have roughly equal stack
heights. In the example $G'' = [3 \times 7 \times 4]$ illustrated in Figure 3a, note that the 32 columns in 2-sections $S'_2$, $1 \leq j \leq 4$, can be partitioned into congruence classes mod 8 by column number, and each congruence class has exactly 1 or 2 of its columns designated blank. The result is that stack heights differing by at most 1.

The goals described above in (a)-(c) can be formulated as combinatorial conditions to be satisfied by the designation of blank columns, and generally of blank $(i - 1)$-levels. Starting with the designation of blank columns, it will be convenient to define a $P_2 \times 2^{e_2 - e_1}$, $(0, 1)$ matrix $F(2) = (f_{cd}(2))$. The rows of $F(2)$ correspond to the 2-sections $S'_2$, $1 \leq c \leq P_2$, and the columns of $F(2)$ to the columns $(2^{e_2 - e_1}$ of them) within each 2-section. We let $f_{cd}(2) = 1$ if column $d$ in section $S'_2$ (which recall is column $Y_2^{(c-1)(2^{e_2 - e_1})+d}$ in $Y_2$) is blank, and $f_{cd}(2) = 0$ if that column is nonblank. Since $s_2(c)$ is the number of blank columns in $S'_2$, the sum of entries in row $c$ of $F(2)$ is

$$f_{cd}(2) = s_2(c).$$

for $1 \leq c \leq P_2$.

Toward formulating the goal expressed in (c), consider now the contribution, through the stacking map $\sigma_2$, from $S'_2$ to any stack addressed by a point in $S'_2$. For any stack address $(x, y) \in S'_2$, let $Stack_3((x, y), r) = \{\sigma_2(z) : \sigma_2(z)_{1..2} = (x, y), z \in S^{(r)}_2\}$, which is the set of points in the stack addressed by $(x, y)$ whose preimages under the map $\sigma_2$ come from $S^{(r)}_2$. Now for any $\sigma_2(z) \in Stack_3((x, y), r)$, we have $z = (x, d)$, with $d \equiv y \pmod{2^{e_2 - e_1}}$, and column $Y_2^d$ is nonblank. Then for $r < P_2$ we see that the stack height $|Stack_3((x, y), r)|$ is the number of zeros of the matrix $F(2)$ lying in column $y$ and within rows 1 through $r$. So to keep stack heights nearly equal over all stack addresses $(x, y) \in S'_2$, we require that this number of zeros is nearly the same over all columns $y$ in $F(2)$. For this, it suffices to have the number of 1’s in rows 1 through $r$ of any column nearly the same; that is, to have nearly equal initial column sums. This becomes the condition

$$|\sum_{c=1}^{r} f_{cy}(2) - \sum_{c=1}^{r} f_{cy'}(2)| \leq 1.$$ (3)

for any $1 \leq y, y' \leq 2^{e_2 - e_1}$ and $1 \leq r \leq P_2$. It says that the blank columns are uniformly distributed mod $2^{e_2 - e_1}$.

To formulate (b), for integers $a$ and $b$ let $|a - b|$ be the difference $a - b$ taken mod $2^{e_2 - e_1}$. Let $xy \in E(G)$, say with $x \in D_2$ and $y \in D_2^2$, $s \neq t$. Since $x$ and $y$ agree in their first two coordinates, $x$ and $y$ are each the $p'$th points in $D_2$ and $D_2^2$ of their respective chains, for some $1 \leq p \leq a_2$. By equation (1) for $i = 2$, $S^{(r)}_2$ has barely enough nonblank columns to contain $(I_2 \circ f_2)(D_2^{(r)})$ for any $1 \leq r \leq P_i$. Thus by the monotonicity property of $f_2$ in Theorem 2.2.4$^*$, for any $j \geq 1$, the image under $I_2 \circ f_2$ of the first point of any chain in $D_2^2$ must lie either in the last nonblank column of $S^{(r)}_2$ or the first nonblank column of $S^{(r)}_2$. Now let $T$ (resp. $(T'))$ be the set of the first $p$ points in $D_2^2$ (resp. $D_2$) in the chain containing $x$ (resp. $y$). Also let $c$ (resp. $c'$) be the number of columns of $Y_2$ spanned by $f_2(T)$ (resp. $f_2(T')$), and thus the number of nonblank columns of $Y_2$ spanned by $(I_2 \circ f_2)(T)$ (resp. $(I_2 \circ f_2)(T')$). By Corollary 2.3.1 we have $|c - c'| \leq 1$. Thus the contribution to $||(I_2 \circ f_2)(x) - (I_2 \circ f_2)(y)||$
from the difference between the number of nonblank columns in $S_2^* \ (\text{resp. } S_2^j)$ preceding $(I_2 \circ f_2)(x) \ (\text{resp. } (I_2 \circ f_2)(y))$ is small (in fact $\leq 1$). The same contribution due to the difference in starting columns of $(I_2 \circ f_2)(T)$ and $(I_2 \circ f_2)(T')$ is also small ($\leq 1$) by the above. Thus $\| (I_2 \circ f_2)(x)_2 - (I_2 \circ f_2)(y)_2 \|$ depends primarily on the number $N_1 \ (\text{resp. } N_2)$ of blank columns in $S_2^*$ (resp. $S_2^j$) preceding the column containing $(I_2 \circ f_2)(x) \ (\text{resp. } (I_2 \circ f_2)(y))$. Each column counted by $N_1 \ (\text{resp. } N_2)$ pushes the image $(I_2 \circ f_2)(x) \ (\text{resp. } (I_2 \circ f_2)(y))$ one more column to the right in $S_2^* \ (\text{resp. } S_2^j)$. So we want to keep $|N_1 - N_2|$ small. Since each blank column corresponds to a 1 in $F(2)$, we see that each of $N_1$ and $N_2$ is just an initial row sum in $F(2)$ (row $s$ for $N_1$ and row $t$ for $N_2$). These considerations motivate the goal of keeping the difference between corresponding initial row sums in $F(2)$ small. It will suffice for our purposes to have

$$|\sum_{j=1}^{b} f_{sj}(2) - \sum_{j=1}^{b} f_{tj}(2)| \leq 2.$$  \hspace{1cm} (4)

for any, $1 \leq s, t \leq a_3a_4 \cdots a_k, 1 \leq b \leq 2^{e_2-e_1}$.

For fixed $i$, we will see that the sequence $\{s_i(j)\}, 1 \leq j \leq P_i$, recursively defined by \[1\] satisfies $|s_i(j_1) - s_i(j_2)| \leq 1$ for all $1 \leq j_1, j_2 \leq P_i$. Applying this to the case $i = 2$, we can view the satisfying of conditions (a)-(c), as relying on the construction of a $P_2 \times 2^{e_2-e_1}$, $(0,1)$ matrix $F(2)$ with prescribed row sums $s_2(c), 1 \leq c \leq P_i$, these sums differing by at most 1 (from the preceding sentence and \[2\]). Further, $F(2)$ will have balanced initial column sums and balanced initial row sums (from \[3\] and \[4\]).

Such an $F(2)$ for the embedding $f_3 : G'' = [3 \times 7 \times 4] \rightarrow Y_3^{(3)}$ discussed above, where $F(2)$ has $P_2 = 4$ rows, is illustrated by the $(0,1)$ matrix in the right of Table 1a). The fractional matrices at left from which this and the other $(0,1)$ matrices in this table are derived will be explained later. This $F(2)$ encodes which columns are designated ‘blank’ in performing the inflation step $I_2 : Y_2^{(u_2(G''))} \rightarrow S_2^{(4)}$. The stacking map $\sigma_2$ is then applied to yield the final embedding $f_3 = \sigma_2 \circ I_2 \circ f_2 : [3 \times 7 \times 4] \rightarrow Y_3^{(3)}$.

The corresponding requirements for arbitrary dimension $i \geq 2$ are analogous. We return to the construction of $f_{i+1}$ from $f_i$ as the composition $f_{i+1} = \sigma_i \circ I_i \circ f_i$. Recall that there are $P_i = a_{i+1}a_{i+2} \cdots a_k$ many $i$-sections $S_i^j, 1 \leq j \leq P_i$, and each such $i$-section has $2^{e_i-e_{i-1}}$ many $(i-1)$-levels. So let $F(i) = (f_{cd}(i))$ be the $P_i \times 2^{e_i-e_{i-1}}$, $(0,1)$ matrix (analogous to $F(2)$) defined by $f_{cd}(i) = 1$ if the $d$’th $(i-1)$-level of $i$-section $S_i^j$ (that is, $(i-1)$-level $Y_i^{(c-1)2^{e_i-e_{i-1}}+d}$ of $Y_i$) is blank, and $f_{cd}(i) = 0$ otherwise. So we require the analogues of the relations \[2\]-\[4\]:

$$\sum_{j=1}^{2^{e_i-e_{i-1}}} f_{cj}(i) = s_i(c),$$  \hspace{1cm} (5)

for $1 \leq c \leq P_i$,

$$|\sum_{c=1}^{r} f_{cy}(i) - \sum_{c=1}^{r} f_{cy'}(i)| \leq 1,$$  \hspace{1cm} (6)

for $1 \leq y, y' \leq 2^{e_i-e_{i-1}}, 1 \leq r \leq P_i$. 

14
for $1 \leq s, t \leq P_i$, $1 \leq r \leq 2^{e_i - e_i - 1}$.

Relation (5) says that there are $s_i(c)$ many blank $(i - 1)$-levels in $S_i^c$. Relation (6) gives balanced initial column sums in $F(i)$. This implies that blank $(i - 1)$-levels are uniformly distributed mod $2^{e_i - e_i - 1}$. This ensures balanced stack heights, under the map $\sigma_i$, for stacks addressed by $S_i^t \cong (Y_i)$. Finally (7) will imply (after some work) that $||((I_i \circ f_i)(x), (I_i \circ f_i)(y))||$ is small for any $xy \in E(G)$, with $x \in D_i^s$ and $x \in D_i^t$, $1 \leq s, t \leq P_i$. This will make the contribution to $\text{dist}_{Y_i}(f_k(x), f_k(y))$ from $|f_k(x) - f_k(y)|$ small for corresponding points $(x$ and $y)$ in distinct $i$-pages. Applying this requirement for all $i$ will keep $\text{dilation}(f_k)$ small.

We are thus reduced to the construction of a $(0, 1)$ matrix $F(i)$ for each $2 \leq i \leq k - 1$, with prescribed row sums $s_i(j)$ (with values recursively computed using (1)), $1 \leq j \leq P_i$, differing by at most 1, and balanced initial column and row sums as required in (6) and (7).

We construct such matrices in the next section.

Returning to $G' = [3 \times 7 \times 4 \times 3]$ we illustrate the construction of $f_4 : G' \rightarrow Y_4^{(u_4(G'))} = Y_4^{(2)}$. We have $u_2(G') = \left\lceil \frac{|G'|}{2^4} \right\rceil = \left\lceil \frac{252}{4} \right\rceil = 63$, $u_3(G') = \left\lceil \frac{|G'|}{2^3} \right\rceil = \left\lceil \frac{252}{32} \right\rceil = 8$, and $u_4(G') = \left\lceil \frac{|G'|}{2^4} \right\rceil = \left\lceil \frac{252}{128} \right\rceil = 2$.

The maps $f_2, f_3,$ and $f_4$ are constructed in succession. The map $f_2 : G' \rightarrow Y_2^{(63)}$ is given in the previous section. To build $f_3 = \sigma_2 \circ I_2 \circ f_2$ the next step is to define the inflation map $I_2$. For this, note that $P_2 = 12$, so there will be 12 many 2-sections $S_2^j$, $1 \leq j \leq 12$, each isomorphic to $P(4) \times P(8)$. Thus $I_2$ has the form $I_2 : Y_2^{(63)} \rightarrow S_2^{(12)}$. The sequence $\{s_2(j)\}$, $1 \leq j \leq 12$, is calculated inductively using (1) for $i = 2$, with $e_2 = 2$ and $e_2 = 5$, to give $2, 3, 3, 3, 2, 3, 3, 2, 3, 3, 3$. A balanced distribution of blank columns among the sections $S_2^j$, $1 \leq j \leq 12$ satisfying (1) - (4) is given by the $12 \times 8, (0, 1)$ matrix $F(2)$ shown at right in Table 1b). So $I_2$ will make $f_2$ skip over the blank columns in $S_2^{(12)}$ as in the previous example. The blank columns (shaded) among the first 4 sections $S_2^1, S_2^2, S_2^3, S_2^4$, as encoded by the first 4 rows of the matrix in the right of Table 1b), are illustrated in Figure 3b). The blank columns among the remaining 8 sections $S_2^j$, $5 \leq j \leq 12$, (encoded by the last 8 rows of the same matrix) are illustrated in Figure 3b), where these 8 sections are surrounded by a box. We put small boxes around images of first points of chains within each 2-page.
In Figure 6. Here we identify the points addresses in columns 4 and 5 of S sections contributing to such a stack are height, and the resulting stack height is (σ is column 2 of S).

You see that points in column 2 of S “fall through” blank columns 2 in S, points in column 2 of S “fall through” blank columns 2 in S and S, points in column 2 of S “fall through” blank columns 2 in S, S, and S, and so on. The resulting stack height is |Stack3((x, 2), 12)| = 7. By contrast, for any stack address (x, 5) ∈ S, 1 ≤ x ≤ 4, the sections contributing to such a stack are j = 2, 3, 4, 5, 7, 9, 10, 12 in order of increasing stack height, and the resulting stack height is |Stack3((x, 5), 12)| = 8. In fact the maximum stack height over all 32 stacks addressed by points of S is 8.

A more detailed look at the members of various stacks addressed by points of S is given in Figure 6. Here we identify the points x ∈ G such that (σ ∘ I ∘ f3)(x) belongs to stack addresses in columns 4 and 5 of S. For example, consider the members of Stack3((3, 4), 7) (bolded in the figure), consisting by definition of those images (σ ∘ I ∘ f3)(x) satisfying (σ ∘ I ∘ f3)(x),1−2 = (3, 4) ∈ S and (I ∘ f3)(x) ∈ S for some 1 ≤ j ≤ 7. Here we must have (I ∘ f3)(x) = (3, d), where d ≡ 4(mod 8) by the definition of σ and 8(j − 1) + 1 ≤ d ≤ 8j.

\[
\begin{bmatrix}
1/4 & 1/4 & \cdots & 1/4 \\
3/8 & 3/8 & \cdots & 3/8 \\
3/8 & 3/8 & \cdots & 3/8 \\
3/8 & 3/8 & \cdots & 3/8 \\
1/4 & 1/4 & \cdots & 1/4 \\
3/8 & 3/8 & \cdots & 3/8 \\
3/8 & 3/8 & \cdots & 3/8 \\
3/8 & 3/8 & \cdots & 3/8 \\
\end{bmatrix}
\begin{bmatrix}
1/4 & 1/4 & 1/4 & 1/4 \\
1/4 & 1/4 & 1/4 & 1/4 \\
1/2 & 1/2 & 1/2 & 1/2 \\
\end{bmatrix}
\rightarrow
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
\end{bmatrix}
\]

b) Matrix F\(2\) encoding blank columns for the embedding \(f_3 : [3 × 7 × 4 × 3] \rightarrow Y_3\)

\[
\begin{bmatrix}
1/4 & 1/4 & \cdots & 1/4 \\
3/8 & 3/8 & \cdots & 3/8 \\
3/8 & 3/8 & \cdots & 3/8 \\
3/8 & 3/8 & \cdots & 3/8 \\
1/4 & 1/4 & \cdots & 1/4 \\
3/8 & 3/8 & \cdots & 3/8 \\
3/8 & 3/8 & \cdots & 3/8 \\
3/8 & 3/8 & \cdots & 3/8 \\
\end{bmatrix}
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
\end{bmatrix}
\rightarrow
\begin{bmatrix}
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
\end{bmatrix}
\]

c) Matrix F\(3\) encoding blank 2-levels for the embedding \(f_4 : [3 × 7 × 4 × 3] \rightarrow Y_4\)

Table 1: Matrices encoding blank levels

Next we apply the stacking map \(\sigma_2 : (I_2 ∘ f_2)(G) \rightarrow Y_3^{(u_3(G'))} = Y_3^{(8)}\) as defined above, thereby yielding the map \(f_3 = \sigma_2 ∘ I_2 ∘ f_2 : G' \rightarrow Y_3^{(8)}\). Note that \(\sigma_2\) stacks the 12 many 2-sections of Figure 3 onto \(S_2\), with the result shown in Figure 5. The left subfigure shows the stacking of the first 7 sections, while the right subfigure all 12 sections.

Again for each stack address \((x, y) \in S_2 \equiv (Y_2)\) we indicate by label \(j\), given to various cross sections of this stack, which image sets \(\sigma_2(S_2^j), 1 ≤ j ≤ 12\), contribute points to this stack. For example one can check in the right subfigure that for any stack address \((x, 2) \in S_2^1, 1 ≤ x ≤ 4\), (so \((x, 2)\) lies in column 2 of \(S_2^1\)) the image sets \(\sigma_2(S_2^j)\) contributing to Stack3((x, 2), 12) satisfy \(j = 1, 3, 4, 6, 7, 9, 11\) in order of increasing stack height. To see why, refer to Figure 8. There you see that points in column 2 of \(S_2^3\) and \(S_2^2\) “fall through” blank column 2 of \(S_2^2\), points in column 2 of \(S_2^6\) and \(S_2^7\) “fall through” blank columns 2 in \(S_2^3\) and \(S_2^2\), points in column 2 of \(S_2^9\) “fall through” blank columns 2 in \(S_2^3, S_2^5, S_2^6, S_2^2\), and \(S_2^2\), and so on. The resulting stack height is |Stack3((x, 2), 12)| = 7. By contrast, for any stack address \((x, 5) \in S_2^1, 1 ≤ x ≤ 4\), the sections contributing to such a stack are \(j = 2, 3, 4, 5, 7, 9, 10, 12\) in order of increasing stack height, and the resulting stack height is |Stack3((x, 5), 12)| = 8. In fact the maximum stack height over all 32 stacks addressed by points of \(S_2^1\) is 8.
since \((I_2 \circ f_2)(x) \in S^2_3\). We find such an \(x\) when \(j = 1\), with \((I_2 \circ f_2)(x) = (3, 4) \in S^1_2\), and Figure 3 shows that \(x = (2, 4, 1, 1)\). Here \(x_{1 \rightarrow 3} = (2, 4, 1)\) since \(x\) is the 4’th point on chain 2 of \(D^3_2\), while \(x_4 = 1\) since \(x \in D^1_3\). There is no such \(x\) when \(j = 2\) since column 4 of \(S^2_3\) is blank. We find such an \(x\) when \(j = 3\), thus satisfying \((I_2 \circ f_2)(x) = (3, 20) \in S^2_2\), and the figure shows that \(x = (2, 4, 3, 1)\) since \(x\) is point \((2, 4)\) in \(D^3_2\). The remaining two such points \(x\) are found similarly; \(x = (2, 3, 1, 2) \in D^5_2\) with \((I_2 \circ f_2)(x) = (3, 36) \in S^2_2\), and \(x = (2, 4, 2, 2) \in D^6_2\) with \((I_2 \circ f_2)(x) = (3, 44) \in S^2_2\). For the last \(x\) note that the first point of \(D^6_2\) in the same chain as \(x\) is \((2, 1, 2, 2)\) and \((I_2 \circ f_2)((2, 1, 2, 2)) \in S^2_3\).

To construct \(f_1 = \sigma_3 \circ I_3 \circ f_3\), we begin with the parameters needed to define \(I_3\). Observe that \(P_3 = 3\) and \(e_3 = 7\), so \(e_3 - e_2 = 2\) and \(S^2_3 = P(4) \times P(8) \times P(4)\) for each \(1 \leq j \leq 3\). So \(I_3\) has the form \(I_3 : Y^{(u_3(G'))}_3 \rightarrow S^{(P_3)}_3 = S^{(3)}_3\). Since \(|D^3_3| = 84\), we get by \([4]\) the sequence \(s_3(1) = 1\), \(s_3(2) = 1\), and \(s_3(3) = 2\). A balanced distribution of blank 2-levels (among the sections \(S^2_3\)) satisfying \([4]\) together with \([5]\) - \([7]\) is given by the \(3 \times 4\), \((0, 1)\) matrix \(F(3)\) given in Table 1c. We then apply the map \(I_3 : Y^{(8)}_3 \rightarrow S^{(3)}_3\), which distributes the eight 2-levels in \(Y^{(8)}_3\) among the twelve 2-levels of \(S^{(3)}_3\), leaving four of the latter twelve levels blank.

The result is represented in the left side of Figure 7 as the set of sections \(S^2_3\), \(1 \leq j \leq 3\), each having four 2-levels, with blank 2-levels shaded. For example, \(S^2_3\) has 2-level 1 as blank (as specified by row 1 of \(F(3)\)), \(S^2_3\) has 2-level 3 as blank (as specified in row 2), and \(S^3_3\) has 2-levels 2 and 4 as blank (as specified in row 3).

Now we apply the stacking map \(\sigma_3 : S^{(3)}_3 \cap (I_3 \circ f_3)(G') \rightarrow Y^{(u_4(G'))}_4 = Y^{(2)}_4\) (whose image recall we abbreviate as \(\sigma_3(S^{(3)}_3)\)), which stacks each of the sets \(S^2_3 \cap (I_3 \circ f_3)(G')\), \(1 \leq j \leq 3\), in succession onto \(S^1_3 \cong \langle Y^3_3 \rangle\) as defined previously. One can check that the maximum stack height \(|Stack_4(z, 3)|\) over all 128 stack addresses \(z \in \langle Y^3_3 \rangle\) is 2, again either by checking that each column of \(F(3)\) has at most (in fact exactly) 2 zeros, or by seeing in Figure 7 that each of the 128 stack addresses in \(\langle Y^3_3 \rangle\) receives at most 2 points under the map \(\sigma_3\). In the remainder of Figure 7, we illustrate the stacking map \(\sigma_3\) in stages. First \(S^2_3\) is stacked on top of \(S^1_3 \cong \langle Y^3_3 \rangle\), with the result that 64 of the stack addresses \(z\) (those lying in levels 2 and 4 of \(\langle Y^3_3 \rangle\)) so far have 4-dimensional stack height 2 (i.e. \(|Stack_4(z, 2)| = 2\) for these \(z\)), while 64 of the stack addresses \(z\) (those lying in levels 1 and 3 of \(\langle Y^3_3 \rangle\)) so far have 4-dimensional stack height 1 (i.e. \(|Stack_4(z, 2)| = 1\) for these \(z\)). Next, \(S^3_3\) is stacked on top of the previous stacking, so that now every stack address has 4-dimensional stack height 2 (i.e. \(|Stack_4(z, 3)| = 2\) for all \(z \in \langle Y^3_3 \rangle\)).

Consider the final result \(\sigma_3(S^{(3)}_3)\), yielding the map \(f_4\), illustrated in the upper right of Figure 7. There, focus on the stack addresses \((x, y, z) \in \langle Y^3_3 \rangle\) with \(z = 2\), \(1 \leq x \leq 4\), \(1 \leq y \leq 8\), these lying in the second 2-level of \(\langle Y^3_3 \rangle\) (since \(z = 2\)). The sets \(\sigma_3(S^1_3)\), \(1 \leq j \leq 3\), contributing points to any such stack \(Stack_4((x, y, 2), 3)\) satisfy \(j = 1\) or \(2\), where in Figure 7 the contribution of \(\sigma_3(S^1_3)\) is represented by \(1b\), and of \(\sigma_3(S^2_3)\) by \(2b\).

In Figure 8 we look in detail at the individual stacks \(Stack_4((3, 1, 2), 3)\), \(Stack_4((3, 4, 2), 3)\), and \(Stack_4((1, 4, 2), 3)\). For example, the members of \(Stack_4((3, 1, 2), 3)\) are the two images \((\sigma_3 \circ I_3 \circ f_3)(\alpha)\), indicated by their preimages \(\alpha \in G’\) which are \(\alpha = (3, 2, 2, 1)\) and \(\alpha = (2, 1, 4, 2)\) in increasing order of stack height.
4 Tools for the general construction

In this section we develop two tools used in our general construction:
(1) the designation of blank \((i - 1)\)-levels in \(i\)-sections \(S^i_j\), \(1 \leq j \leq P_i\), and
(2) the construction of an ordering (by consecutive integer labels) of the vertices of any hypercube such that for any reasonably long interval of successive label values, any two vertices whose labels lie in that interval are at fairly small hypercube distance.

4.1 The Construction of Blank Levels

In this subsection we describe the sequence \(\{s_i(j)\}\), \(1 \leq i \leq k, 1 \leq j \leq P_i\), where \(s_i(j)\) is the number of \((i - 1)\)-levels of \(S^i_j\) that are designated blank under the map \(I_i \circ f_i\). We show that this sequence satisfies equation (1). That is, for each \(1 \leq r \leq P_i\) we show that \(S^i_r\) has just enough nonblank \((i - 1)\)-levels to contain \((I_i \circ f_i)(D^i_r)\). Finally, we construct for each \(j\) the actual set of \(s_i(j)\) many \((i - 1)\)-levels in \(S^i_j\) that are designated blank, and show that the required properties (5)-(7) are satisfied. This construction is based on a theorem of Knuth on simultaneous roundings of sequences.

Given \(G = [a_1 \times a_2 \times \cdots \times a_k]\) and \(1 \leq i \leq k - 1\), define the sequence \(\{s_i(j)\}\), with \(1 \leq j \leq P_i\), by

\[s_i(j) = 2^{e_i-e_{i-1}} - \left[ \frac{a_1 a_2 \cdots a_j}{2^{e_i-1}} \right] + \lfloor j \phi_i \rfloor - \lfloor (j-1) \phi_i \rfloor \tag{8}\]

where \(\phi_i = \left[ \frac{a_1 a_2 \cdots a_i}{2^{e_i-1}} \right] - \frac{a_1 a_2 \cdots a_i}{2^{e_i-1}}.

Lemma 4.1 Let \(P_i = a_{i+1} a_{i+2} \cdots a_k\), \(1 \leq r \leq P_i\), and \(1 \leq i \leq k\).
(a) The sequence \(\{s_i(j)\}\), \(1 \leq j \leq P_i\), defined above satisfies \(\sum_{j=1}^{r} s_i(j) = r 2^{e_i-e_{i-1}}\).

In particular, taking \(s_i(j)\) to be the number of blank \((i - 1)\)-levels in \(S^i_j\), then the number of nonblank \((i - 1)\)-levels in \(S^i_r\) is \(\sum_{j=1}^{r} s_i(j) = \left[ \frac{a_1 a_2 \cdots a_i}{2^{e_i-1}} \right] \). Part (a) follows.

(b) \(s_i(j) = p\) or \(p + 1\), where \(p = 2^{e_i-e_{i-1}} - \left[ \frac{a_1 a_2 \cdots a_i}{2^{e_i-1}} \right]\). Also \(\frac{s_i(j)}{2^{e_i-e_{i-1}}} \leq \frac{1}{2}\).

Proof. For (a), observe that the sum \(\sum_{j=1}^{r} s_i(j)\) telescopes. Note also that \(r \left[ \frac{a_1 a_2 \cdots a_i}{2^{e_i-1}} \right]\) is an integer. Thus \(\sum_{j=1}^{r} s_i(j) = r 2^{e_i-e_{i-1}} - r \left[ \frac{a_1 a_2 \cdots a_i}{2^{e_i-1}} \right] + r \left[ \frac{a_1 a_2 \cdots a_i}{2^{e_i-1}} \right] = r 2^{e_i-e_{i-1}} - \left[ r \frac{a_1 a_2 \cdots a_i}{2^{e_i-1}} \right]\). Part (a) follows.

Consider (b). Since \(\phi_i < 1\) we have \(0 \leq \lfloor j \phi_i \rfloor - \lfloor (j-1) \phi_i \rfloor \leq 1\), giving the first statement. Since \(a_1 a_2 \cdots a_i \geq 2^{e_i-1} + 1\), we get \(s_i(j) \leq 2^{e_i-e_{i-1}} - \left[ \frac{2^{e_i-1}+1}{2^{e_i-1}} \right] + 1 \leq 2^{e_i-e_{i-1}-1}\). The second statement follows.

We now specify, for each \(2 \leq i \leq k - 1\) and \(1 \leq j \leq P_i\), which \(s_i(j)\) out of the \(2^{e_i-e_{i-1}}\) many \((i - 1)\)-levels in \(S^i_j\) will be designated blank. Keep in mind that this designation must satisfy the balance properties (5)-(7) we required in the overview. For this, we need the following theorem of Knuth [19].

Let \(x_1, \ldots, x_n\) be a sequence of reals, and \(\gamma\) a permutation of \(\{1, 2, \ldots, n\}\). Let \(S_k = x_1 + \cdots + x_k\) and \(\Sigma_k = x_{\gamma(1)} + \cdots + x_{\gamma(k)}\) be the partial sums for these two independent orderings of the \(x_i\)’s. Consider a rounding of the \(x_i\)’s; that is, a designation of integers \(\bar{x}_i\) satisfying \(\lfloor x_i \rfloor \leq \bar{x}_i \leq \lfloor x_i \rfloor + 1\) for \(1 \leq i \leq n\). Now let the corresponding partial sums be
\( \bar{S}_k = \bar{x}_1 + \cdots + \bar{x}_k \) and \( \Sigma_k = \bar{x}_{\gamma(1)} + \cdots + \bar{x}_{\gamma(k)} \). We say that a rounding of the \( x_i \)'s is consistent with the original sequence \( \{x_i\} \) (resp. with the permuted sequence under \( \gamma \)) if \( \lfloor \bar{S}_k \rfloor \leq \bar{S}_k \leq \lfloor S_k \rfloor \) (resp. \( \lfloor \Sigma_k \rfloor \leq \Sigma_k \leq \lfloor \Sigma_k \rfloor \)) for each \( 1 \leq k \leq n \). We also say that the rounding is a two-way rounding if it is simultaneously consistent with both the original sequence and the permuted sequence under \( \gamma \).

**Theorem 4.2** [12] For any finite sequence \( x_1, \ldots, x_n \) of reals and any permutation \( \gamma \) of \( \{1, 2, \ldots, n\} \), there is a two-way rounding of the \( x_i \).

The existence of two-way roundings was shown earlier by Spencer [27] by probabilistic methods, as a corollary to more general results on the discrepancy of set systems [23]. Knuth’s network flow based proof of Theorem 1.2, omitted here, is constructive and yields improved error bounds. The two-way rounding produced is not necessarily unique.

As a consequence of Knuth’s theorem, we obtain the following theorem (from [11] and [12]) on roundings of matrices which are consistent with respect to all initial row and column sums. We give its short proof for completeness. This extends the rounding lemma of Baranyai [4] giving such consistency with respect to all row sums, to all column sums, and to the sum of all matrix entries. Additional results on roundings of matrices, including extensions of previous work, applications to digital halftoning, and improved running time and error bounds in implementation can be found in the work of Doerr ([10], [11], [12] and others), Asano ([1], [2]), and Wright [29] among others.

**Theorem 4.3** Let \( T = (t_{ij}) \) be an \( m \times n \) matrix with \( 0 \leq t_{ij} \leq 1 \) for all \( i \) and \( j \). Then there exists an \( m \times n \), (0,1) “rounding” matrix \( F = (f_{ij}) \) of \( T \); that is, where \( f_{ij} = \lfloor t_{ij} \rfloor \) or \( \lceil t_{ij} \rceil \), satisfying the following properties.

(a) For each \( b \) and \( i \), \( 1 \leq b \leq n \), \( 1 \leq i \leq m \), we have \( |\sum_{j=1}^{b}(t_{ij} - f_{ij})| < 1 \).

(b) For each \( b \) and \( j \), \( 1 \leq b \leq m \), \( 1 \leq j \leq n \), we have \( |\sum_{i=1}^{b}(t_{ij} - f_{ij})| < 1 \).

(c) \( |\sum_{i=1}^{m}\sum_{j=1}^{n}(t_{ij} - f_{ij})| < 1 \).

**Proof.** For the most part we paraphrase proofs in [11] and [12]. First consider parts (a) and (b). We construct an \((m+1) \times (n+1)\) matrix \( Y \) from \( T \) which has integral row and column sums by appending to each row and each column a last entry in \([0,1)\) just large enough to make that row or column have an integer sum. Specifically, let \( y_{ij} = t_{ij} \) for \( 1 \leq i \leq m \) and \( 1 \leq j \leq n \), and then \( y_{m+1,j} = \lfloor \sum_{i=1}^{m}t_{ij} \rfloor - \sum_{i=1}^{m}t_{ij} \) for \( 1 \leq j \leq n \), and \( y_{i,n+1} = \lfloor \sum_{j=1}^{n}t_{ij} \rfloor - \sum_{j=1}^{n}t_{ij} \) for \( 1 \leq i \leq m \), and \( y_{m+1,n+1} = \sum_{i=1}^{m}\sum_{j=1}^{n}t_{ij} \). Then \( Y \) has integral row and column sums.

Consider now the following two orderings of the entries of \( Y \). First we order these entries by “row-major” order; that is, first the entries of row 1, then those of row 2, etc., until row \( m+1 \), and within any row \( i \) place \( y_{ij} \) ahead of \( y_{ir} \) iff \( j < r \). Similarly consider the “column-major” order where we place \( y_{ij} \) before \( y_{rs} \) iff either \( j < s \), or \( j = s \) and \( i < r \).

Applying Knuth’s theorem, there is a (not necessarily unique) two-way rounding matrix \( \bar{Y} = \bar{y}_{ij} \) relative to these two orders. Since every row and column sum of \( Y \) is already an integer, we get \( \sum_{i=1}^{p}\sum_{j=1}^{n+1}(y_{ij} - \bar{y}_{ij}) = 0 \) for any \( 1 \leq p \leq m+1 \), and \( \sum_{j=1}^{q}\sum_{i=1}^{m+1}(y_{ij} - \bar{y}_{ij}) = 0 \) for any \( 1 \leq q \leq n+1 \). Thus taking suitable partial sums in the row major order we get initial row sum estimates \( |\sum_{j=1}^{b}(y_{ij} - \bar{y}_{ij})| < 1 \) for any fixed \( i \) and \( 1 \leq b \leq n+1 \). Similarly, suitable partial
sums in column major order yield \(|\sum_{i=1}^{b}(y_{ij} - \bar{y}_{ij})| < 1\) for any fixed \(j\) and \(1 \leq b \leq m + 1\). Finally, let \(F\) be the upper left \(m \times n\) submatrix of \(Y\); that is, \(F = (f_{ij})\) where \(f_{ij} = \bar{y}_{ij}\) for \(1 \leq i \leq m\) and \(1 \leq j \leq n\). Then recalling that \(t_{ij} = y_{ij}\) for such \(i\) and \(j\), we get parts (a) and (b) of the theorem.

For (c), let \(R = \sum_{i=1}^{m+1} \sum_{j=1}^{n+1} (y_{ij} - \bar{y}_{ij})\), \(S = \sum_{i=1}^{m+1} (y_{i,n+1} - \bar{y}_{i,n+1})\), and \(S' = \sum_{j=1}^{n+1} (y_{m+1,j} - \bar{y}_{m+1,j})\), noting that these three quantities are all 0 since each row or column sum in \(Y\) is integral. Then \(|\sum_{i=1}^{m} \sum_{j=1}^{n} (t_{ij} - \bar{y}_{ij})| = |R - S - S' + (y_{m+1,n+1} - \bar{y}_{m+1,n+1})| \leq |y_{m+1,n+1} - \bar{y}_{m+1,n+1}| < 1.\)

The setting in which we apply Theorem 4.3 is as follows. Let \(X = \{s_1, s_2, \ldots, s_m\}\) be an integer sequence such that for all \(1 \leq i \leq m\) we have \(s_i = k\) or \(k + 1\) for some fixed integer \(k\) independent of \(i\). Also let \(n\) be a positive integer satisfying \(k + 1 \leq n\). Define the \(m \times n\) matrix \(T^X = (t_{ij})\) as follows. For each fixed row index \(i, 1 \leq i \leq m\), let \(t_{ij} = \frac{s_i}{n}\) for all \(j, 1 \leq j \leq n\). Thus all entries in a given row of \(T^X\) have the same constant value. For some rows this constant is \(\frac{k}{n}\) while for the rest it is \(\frac{k+1}{n}\), and row \(i\) of \(T^X\) has row sum \(s_i\). Now let \(F^X = (f_{ij})\) be an \(m \times n\), \((0,1)\) rounding matrix of \(T^X\) as guaranteed to exist by Theorem 4.3. That is, with \(T^X\) and \(F^X\) playing the roles of \(T\) and \(F\) respectively in that theorem. For suitably chosen integers \(m, n\) and integer sequence \(X\) determined by the construction in the next section, this \(F^X\) will be, for fixed \(i\), the matrix \(F(i)\) introduced in the previous section that encodes which \((i-1)\)-levels of \(Y_i\) will be designated blank. This \(F^X\) has certain balance properties described in the following theorem.

**Theorem 4.4** Let \(k\) and \(n\) be positive integers with \(1 \leq k + 1 \leq n\), and \(X = \{s_1, s_2, \ldots, s_m\}\) a sequence with \(s_i = k\) or \(k + 1\) for all \(i\). Let \(T^X = (t_{ij})\) and \(F^X = (f_{ij})\), a rounding of \(T^X\), be the \(m \times n\) matrices as defined above. Then \(F^X\) has the following properties.

(a) \(\sum_{j=1}^{n} f_{ij} = s_i\) for \(1 \leq i \leq m\).

(b) \(|\sum_{j=1}^{b} f_{jr} - \sum_{j=1}^{b} f_{js}| \leq 1\) for \(1 \leq r, s \leq n, 1 \leq b \leq m\).

(c) \(|\sum_{j=1}^{b} f_{rj} - \sum_{j=1}^{b} f_{sj}| \leq 2\) for \(1 \leq r, s \leq m, 1 \leq b \leq n\).

**Proof.** For (a), note that row \(i\) of \(T^X\) has the integer \(s_i\) as its sum. Then (a) follows from Theorem 4.3b on taking \(b = n\).

For (b), observe that for any \(b, 1 \leq b \leq m\), the sequences formed by the first \(b\) entries in any two columns \(r\) and \(s\) of \(T^X\) are identical. So \(\sum_{j=1}^{b} t_{jr} = \sum_{j=1}^{b} t_{js}\) for any \(r\) and \(s\). Thus by Theorem 4.3b, \(\sum_{j=1}^{b} f_{rj}\) and \(\sum_{j=1}^{b} f_{sj}\) are roundings of the same quantity, and hence can differ by at most 1, proving (b).

For (c), recall that any two rows \(r\) and \(s\) of \(T^X\) are either identical, or one has constant entries \(\frac{k}{n}\) and the other constant entries \(\frac{k+1}{n}\). So the maximum difference between any two initial row sums of \(T^X\) occurs in the second possibility, and in that case the difference we get is \(|\sum_{j=1}^{b} t_{rj} - \sum_{j=1}^{b} t_{sj}| \leq \frac{b}{n} \leq 1\). By Theorem 4.3c, the initial row sums \(\sum_{j=1}^{b} f_{rj}\) and \(\sum_{j=1}^{b} f_{sj}\) of \(F^X\) are roundings of the corresponding initial row sums in \(T^X\). Since the latter sums differ by at most 1, part (c) follows, where the upper bound of 2 can be achieved only if there is an integer strictly between the corresponding initial row sums in \(T^X\).

For any two integers \(r\) and \(s\), \(1 \leq r, s \leq m\), we will need to bound the column difference between the \(p\)th zero from the left in row \(r\) of \(F^X\) and the \(q\)th zero from the left in row \(s\) of
$F_X$ as a function of $|p - q|$, independent of $r$ and $s$. For this purpose, let $N_i(d)$ be the column
index of the $d$'th zero from the left in row $i$ of $F_X$; that is, $N_i(d) = \min \{ b : b = d + \sum_{j=1}^b f_{ij} \}$. We
omit mention of $X$ in the notation $N_i(d)$, since $X$ will be clear by context. We will call
the entry $f_{ih}$ forward if $\sum_{j=1}^h f_{ij} = \lceil \sum_{j=1}^h t_{ij} \rceil$. Otherwise we call $f_{ih}$ backward, so in that case
$\sum_{j=1}^h f_{ij} = \lceil \sum_{j=1}^h t_{ij} \rceil - 1$. Recall that every entry of $F_X$ is either forward or backward by
Theorem 4.3b.

It will be convenient to interpret the entries $f_{ij}$ of the $m \times n$ matrix $F_X$ and the function
$N_i(\ast)$ using "wraparound". For example we let $f_{i,n+3} = f_{i+1,3}$. Similarly if row $i$ of $F_X$ has $q$
many 0's then we let $N_i(q + 3) = N_i(1,3)$, while if row $i + 1$ has $q'$ many 0's then let
$N_i(q + q' + 3) = N_{i+2}(3)$, and so on. Also for $t \le 0$, let $N_{i+1}(t) = N_i(q + t)$.

**Corollary 4.5** Let $X = \{s_1, s_2, \ldots, s_m\}$, $T_X = (t_{ij})$, $F_X = (f_{uw})$, and $k$ be as in the preceding
theorem, and assume now that $k + 1 \le \frac{n}{2}$. Let $c = \frac{k+1}{n}$. Then the $m \times n$, $(0,1)$-matrix $F_X$
has the following properties.

(a) If $f_{ih}$ is forward, then $\sum_{j=h+1}^{h+2e} f_{ij} \le e$. In particular, if $f_{i,N_i(d)}$ is forward,
then $N_i(d + e) - N_i(d) \le 2e$.

(b) If $f_{ih}$ is backward, then $\sum_{j=h+1}^{h+2e} f_{ij} \le e + 1$. In particular, if $f_{i,N_i(d)}$ is backward,
then $N_i(d + e) - N_i(d) \le 2e + 2$.

(c) For integers $r, s, d, e$ with $1 \le r, s \le m$, we have $N_r(d + e) - N_s(d) \le 2e + 4$.

(d) Let $2 \le i \le k - 1$. Set $m = P_i = a_{i+1}a_{i+2} \ldots a_k$, $n = 2^{e_i-e_{i-1}} - 1$, $k = \min\{s_i(j) : 1 \le j \le P_i\}$,
$s_j = s_i(j)$, $1 \le j \le m = P_i$, and $f_{uw} = f_{uw}(i)$. With these settings, $F_X$ satisfies
$\frac{i}{n} (3)$ (which include the conditions $21$ as the special case $i = 2$).

**Proof.** For part (a), begin by observing that for any positive integer $p$ we have $\sum_{j=1}^{N_i(d)+p} t_{ij} \le pc \le \frac{p}{n}$. Now take $p = 2e$ to be an even integer. Then since $f_{ih}$ is forward, the number of 1's
among the entries $f_{ij}$, $h + 1 \le j \le h + 2e$, is at most $e$, so $\sum_{j=h+1}^{h+2e} f_{ij} \le e$ as claimed. For the
second statement, since the number of these entries is $2e$ and $\sum_{j=h+1}^{h+2e} f_{ij} \le e$, it follows that
the number of 0's among these entries is at least $e$, and hence $N_i(d + e) - N_i(d) \le 2e$.

For part (b), we proceed as in part (a), taking $p = 2e$ to be an even integer. The difference
is that since $f_{i,N_i(d)}$ is backward, the number of 1's among the entries $f_{ij}$, $h + 1 \le j \le h + 2e$
can be at most $1 + e$ (since that number now includes the 1 corresponding to $\lceil \sum_{j=1}^{N_i(d)} t_{ij} \rceil$).
It follows that $\sum_{j=h+1}^{h+2e} f_{ij} \le e + 1$. For the second statement, since $\sum_{j=h+1}^{h+2e} f_{ij} \le e + 1$ it follows
that the number of 0's among these entries is at least $e - 1$. Now replacing $2e$ by $2e + 2$ in the above reasoning, we find that the number of 0's among the $2e + 2$ entries $f_{ij}$,
h + 1 \le j \le h + 2e + 2, is at least $e$. It follows that $N_i(d + e) - N_i(d) \le 2e + 2$.

Consider now part (c). By Theorem 4.4b we have $N_r(d - 2) \le N_s(d)$; that is, there are at least
d − 2 many 0's among the entries $f_{rj}$, $1 \le j \le N_s(d)$. Suppose first that $N_r(d - 1) > N_s(d)$, so
that there are exactly $d - 2$ many 0's among the entries $f_{rj}$, $1 \le j \le N_s(d)$. It follows that
$\sum_{j=1}^{N_r(d)} f_{rj} - \sum_{j=1}^{N_r(d)} f_{sj} = 2$. Thus $f_{r,N_r(d)}$ is forward. So by part (a) we have $\sum_{j=1}^{N_r(d)+2e+4} f_{rj} \le e+2$. So there are at least $e + 2$ many 0's among the entries $f_{rj}$, $N_s(d)+1 \le j \le N_s(d)+2e+4$.
Combining these 0's with the $d - 2$ many 0's among the entries $f_{rj}$, $1 \le j \le N_s(d)$, obtain
$N_r(d + e) \le N_s(d) + 2e + 4$, as required. Now assume $N_r(d - 1) \le N_s(d)$. Then we have
$N_r(d + e) - N_s(d) \leq N_r(d + e) - N_r(d - 1) \leq 2(e + 1) + 2 = 2e + 4$, where we have used the more generous bound from part (b) in bounding $N_r(d + e) - N_r(d - 1)$. This completes (c).

Finally consider part (d). By Lemma 4.1b we have $s_i(j) = k$ or $k + 1$ for each term $s_i(j)$, $1 \leq j \leq P_i$, of our sequence $X$. Hence we may apply Theorem 4.4a,b,c with the given settings for $m, n, X$, and for the entries $f_{uv}$ of $F^X$, to obtain (5), (6), and (7) respectively. ■

We now apply Theorems 4.3 and 4.4 to construct the matrices $F(i)$ of the previous section. Recall that $F(i) = (f_{cd}(i))$ is the $P_i \times 2^{e_i - e_{i-1}}$, (0,1)-matrix such that for $1 \leq d \leq 2^{e_i - e_{i-1}}$ and $1 \leq c \leq P_i$, the $(i - 1)$-level $Y_i^{(c-1)2^{e_i - e_{i-1}} + d}$ of $i$-section $S_i^c$ (i.e. the $d$'th $(i - 1)$-level of $S_i^c$) is blank precisely when $f_{cd}(i) = 1$.

**Construction of the matrix $F(i)$, $1 \leq i \leq k - 1$**

1. Given $G$ and fixed $i$, compute the sequence $X = \{s_i(j)\}, 1 \leq j \leq P_i$, using formula (8).
2. Form the $P_i \times 2^{e_i - e_{i-1}}$ matrix $T^X$ and construct rounding $F^X$ of $T^X$ as in Theorem 4.4 that is, letting $m = P_i$, $n = 2^{e_i - e_{i-1}}$, and $s_j = s_i(j)$ in the definition of $T^X$ (preceding Theorem 4.4). We note that the hypotheses of Theorem 4.4 hold with this $X$ by Lemma 4.1b.
3. Let $F(i) = F^X$.

To illustrate, recall $G'' = [3 \times 7 \times 4]$ from the previous section. Toward constructing $F(2)$, apply formula (5) to obtain the sequence $X = \{s_2(1) = 2, s_2(2) = 3, s_2(3) = 3, s_2(4) = 3\}$. So the $P_2 \times 2^{e_2 - e_1} = 4 \times 8$ matrix $T^X$ has constant row entries $\frac{s_2(r)}{8} = \frac{1}{4}$ or constant row entries $\frac{s_2(r)}{8} = \frac{3}{8}$ in any given row $r$, $1 \leq r \leq 4$, as shown at left in Table 1a. The rounding of $T^X$ by Knuth's network flow method or Doerr's approach (Theorems 4.2 and 4.3) yields, as one possibility, the $(0,1)$ matrix $F(2) := F^X$ at right in Table 1a. Next recall $G' = [3 \times 7 \times 4 \times 3]$. To construct $F(2)$, apply formula (5) to obtain the sequence $X = \{s_2(j)\}, 1 \leq j \leq 12$, given by $X = \{(2,3,3,3,2,3,3,2,3,3,3,3,3)\}$. So the $P_2 \times 2^{e_2 - e_1} = 12 \times 8$ matrix $T^X$ has constant row entries $\frac{1}{4}$ or $\frac{3}{4}$ in any given row, and is shown at left in Table 1b. Again applying Theorems 4.2 and 4.3 gives a possible rounding of $T^X$ given by $F(2) = F^X$ at right in Table 1b. Toward constructing $F(3)$, still for $G'$, we apply formula (8) to obtain the sequence $X = \{s_3(1) = 1, s_3(2) = 1, s_3(3) = 2\}$. The $P_3 \times 2^{e_3 - e_2} = 3 \times 4$ matrix $T^X$ therefore has constant row entries $\frac{1}{4}$ or $\frac{1}{2}$, as shown at left in Table 1c, while a possible rounding $F(3) = F^X$ by these theorems is shown at right.

**4.2 An Integer Labeling of Hypercubes**

In this subsection we construct an ordering $L_t$ of the vertices of $Q_t$ such that for any interval of at most $O(\log(t))$ consecutive points under this ordering, any two points $x, y$ in that interval satisfy $\text{dist}_{Q_t}(x, y) \leq 3$. This labeling is based on the existence of certain spanning subgraphs of hypercubes of suitable dimension, as follows. Define a cyclic caterpillar as a connected graph $H$ such that removal of all leaves of $H$ results in a cycle graph $C_e$ for some $e \geq 3$. A cyclic caterpillar is $r$-regular if each vertex of its cycle subgraph $C_e$ has exactly $r$ neighboring leaves not on the cycle. Denote such an $r$-regular cyclic caterpillar by $\text{Cat}(e, r)$.

We are interested in finding spanning subgraphs $\text{Cat}(e, r)$ of hypercubes. Clearly if $\text{Cat}(e, r)$ spans $Q_n$, then $e = \frac{2^n}{r + 1}$, so $r + 1$ must be a power of 2. Papers on this subject include [7],
In this section we inductively construct a series of embeddings $f_i : G \to ([Y_{i-1}] \times P(u_i)) = Y_i^{(u_i)} \subset Y_i$, $3 \leq i \leq k$, where $u_i = \lfloor \frac{G}{[Y_{i-1}]} \rfloor = \lfloor \frac{G}{2^{i-1}} \rfloor$. At the end we relabel the points of $\langle Y_{k-1} \rangle \times P(u_k)$ with hypercube addresses coming from $Opt(G)$, using the inverse of the labeling of Corollary 4.8. The composition of this relabeling with the map $f_k$ yields the final embedding $H^k : G \to Opt(G)$. We follow the general plan outlined in section 3.

### Theorem 4.6 [Corollary 5.8 in 7] There exists a spanning cyclic regular caterpillar $Cat(e, 2r + 1)$ of $Q_n$ provided that $r + 1 = 2^i$ and $n = 2^{i+1} + 2i$ for some integer $i \geq 0$.

From this we easily obtain the following.

**Corollary 4.7** Suppose $r + 1 = 2^i$ and $t \geq 2^{i+1} + 2i$ for an integer $i \geq 0$. Then $Q_t$ contains a spanning cyclic $(2r + 1)$-regular caterpillar $Cat(e, 2r + 1)$ for suitable $e$.

**Proof.** Recall that $Q_{t+1}$ is the cartesian product $Q_{t+1} = Q_t \times K_2$. Hence if $Q_t$ contains a spanning subgraph $Cat(e, 2r + 1)$, then $Q_{t+1}$ contains a spanning subgraph $Cat(2e, 2r + 1)$. The corollary follows by induction on $t$. $\blacksquare$

We can now construct our desired labeling of hypercubes of sufficiently large dimension.

**Corollary 4.8** (a) Let $r, i, t$ be positive integers satisfying $r + 1 = 2^i$ and $t \geq 2^{i+1} + 2i$. Then there exists a one to one integer labeling $L_t : V(Q_t) \to \{1, 2, \ldots, 2^t\}$ such that for any $x, y \in V(Q_t)$ we have $|L_t(x) - L_t(y)| \leq 2r + 3 \Rightarrow \text{dist}_{Q_t}(x, y) \leq 3$, where the indicated difference is taken modulo $2^t$.

(b) Let $t \geq 22$. Then there exists a one to one integer labeling $L_t : V(Q_t) \to \{1, 2, \ldots, 2^t\}$ such that for any $x, y \in V(Q_t)$ we have $|L_t(x) - L_t(y)| \leq 17 \Rightarrow \text{dist}_{Q_t}(x, y) \leq 3$, where the indicated difference is taken modulo $2^t$.

**Proof.** For (a), consider the spanning subgraph $Cat(e, 2r + 1)$ of $Q_t$, $t \geq 2^{i+1} + 2i$, from Corollary 4.7. For $1 \leq i \leq e$, let $x_i$ be the vertices of the cycle $C_e$ in $Cat(e, 2r + 1)$ indexed consecutively around this cycle. Also let $x_{i+1}, x_{i+2}, \ldots, x_{i+2r+1}$ be the leaf neighbors of $x_i$. Now define $L_t$ by letting $L_t(x_i) = (2r + 2)i$ for $1 \leq i \leq e$, and $L_t(x_{i+j}) = (2r + 2)(i + j)$ for fixed $i$ and $1 \leq j \leq 2r + 1$. For the claim, the critical case to check is when $x$ or $y = x_{i+2r+1}$ for any $1 \leq i \leq e$. The $2r + 3$ points which follow this $x$ in this ordering are (in order) $x_i$, $\{x_{i+1}, x_{i+2}, \ldots, x_{i+2r+1}\}$, and $x_{i+1}$, all at distance $3$ from $x$. Further $2r + 3$ is best possible since the $(2r + 4)$th point following $x$ is $x_{i+2,1}$, and $\text{dist}_{Q_t}(x, x_{i+2,1}) = 4$.

For (b), simply apply part (a) with $r = 7$, yielding $i = 3$, and $t \geq 22$. $\blacksquare$

We will apply the labeling $L_t$ when $t = e_i - e_{i-1}$, for each $1 \leq i \leq k$. To apply Corollary 4.8b we need the condition $t = e_i - e_{i-1} \geq 22$. We therefore assume $a_i \geq 2^{22}$, $1 \leq i \leq k$, from now on, which ensures that this condition holds.

### 5 The general construction

In this section we inductively construct a series of embeddings $f_i : G \to ([Y_{i-1}] \times P(u_i)) = Y_i^{(u_i)} \subset Y_i$, $3 \leq i \leq k$, where $u_i = \lfloor \frac{G}{[Y_{i-1}]} \rfloor = \lfloor \frac{G}{2^{i-1}} \rfloor$. At the end we relabel the points of $\langle Y_{k-1} \rangle \times P(u_k)$ with hypercube addresses coming from $Opt(G)$, using the inverse of the labeling of Corollary 4.8. The composition of this relabeling with the map $f_k$ yields the final embedding $H^k : G \to Opt(G)$. We follow the general plan outlined in section 3.
Recall $S_i(G)$, the set of points of $S_i^{(P)}$ lying in nonblank columns of $S_i^{(P)}$. We let $S_i^{r}(G) = S_i(G) \cap S_i^{r}$, the set of points lying in nonblank columns of $S_i^{r}$, and $S_i^{(r)}(G) = \bigcup_{j=1}^{r} S_i^{j}(G)$. Recall we stipulated that $(I_i \circ f_i)(G) \subseteq S_i(G)$. For $r \leq P_i$ we let $S_i^{r}(G) = S_i^{r}(G) \cap (I_i \circ f_i)(G)$ and $S_i^{(r)}(G)' = S_i^{(r)}(G) \cap (I_i \circ f_i)(G)$. Thus $S_i^{(P)}(G)'$ is the domain of the stacking map $\sigma_i$, and $S_i^{(r)}(G)'$ (resp. $S_i^{(r)}(G)'$) is the subset of that domain lying in section $S_i^{r}$ (resp. $S_i^{(r)}$). We will see later that $S_i^{(P-1)}(G) = S_i^{(P-1)}(G)'$.

### 5.1 The three dimensional embedding

In this subsection we construct the map $f_3 : G \rightarrow [Y_2^{(u_2)}] \subset Y_3$ for a 3-dimensional grid $G = [a_1 \times a_2 \times a_3]$. Since $f_3(G)$ is a 3-dimensional grid, we can visualize its construction along lines of Section 3\(\text{\textsection}3\) using Figures 3, 4, and 5 to which we refer the reader toward following the construction which follows. This 3-dimensional case will hopefully help in understanding the generalization to higher dimensions in the next subsection.

**Construction of the map $f_3 : [a_1 \times a_2 \times a_3 \times \cdots \times a_k] \rightarrow Y_3$**

1. Set $G = [a_1 \times a_2 \times a_3 \times \cdots \times a_k]$. Construct the embedding $f_2 : G \rightarrow Y_2^{(u_2)}$ of section 2\(\text{\textsection}2\).
2. [Designation of Blank Columns]
   a) Set $P_2 = \Pi_{i=3}^{k} a_i$.
   b) Construct the matrix $F(2) = (f_{ij}(2))$ by the procedure following Corollary 4.5 for the case $i = 2$. (Comment: The matrix entries $f_{ij}(2)$ now satisfy (2)-(4) by Corollary 4.5d.)
   c) Now designate ‘blank’ columns in the subgraph $S_2^{(P_2)}$ of $Y_2$ as follows.

   Column $j$ of $S_2^{r}$ is ‘blank’ if $f_{ij}(2) = 1$, and is ‘nonblank’ if $f_{ij}(2) = 0$.

   (Comment: Recall that column $j$ of $S_2^{r}$ is column $(i-1)2^{e_2-e_1} + j$ of $Y_2$.)

   d) For $1 \leq r \leq P_2$, let $S_2^{r}(G)$ (resp. $S_2^{(r)}(G)$) be the set of points lying in nonblank columns of $S_2^{r}$ (resp. $S_2^{(r)}$).

3. [The Inflation Step]
   a) “Inflate” the image $f_2(G)$ by the map $I_2 : Y_2^{(u_2(G))} \rightarrow S_2^{(P_2)}(G)$ given as follows. For any $(a,b) \in Y_2^{(u_2(G))}$, let

   $$I_2(a,b) = (a,b'),$$

   where $Y_2^{(u_2)}$ is the $b$th nonblank column in $S_2^{(P_2)}$ in increasing order of $b$.

   b) Let $S_2^{r}(G)' = S_2^{r}(G) \cap (I_2 \circ f_2)(G)$ and $S_2^{(r)}(G)' = S_2^{(r)}(G) \cap (I_2 \circ f_2)(G)$.

4. [The Stacking Step]
   “Stack” the sets $S_2^{r}(G)'$, $1 \leq r \leq P_2$, successively “over” $S_1^{(Y_2)}$ by the stacking map $\sigma_2 : S_2^{(P_2)}(G)' \rightarrow S_2^{\times} \times P(u_3) = Y_3^{(u_3)}$ defined as follows. Take $y = (a,b) \in S_2^{(P_2)}(G)'$, say with $y \in S_2^{r}(G)'$.

   a) Let $n_y = |\{z \in S_2^{(r)}(G)' : z_1 = a \text{ and } z_2 \equiv b \text{ (mod $2^{e_2-e_1}$)}\}|$.
   b) Let $b$ satisfy $b \equiv b \text{ (mod $2^{e_2-e_1}$)}$, with $1 \leq b \leq 2^{e_2-e_1}$.
   c) Define $\sigma_2(a,b) = (a,b,n_y)$. 

24
(Comment: Take any \((a,b) \in S_2(G)\). Then \(1 \leq \sigma_2(a,b)_2 = \bar{b} \leq 2^{e_2-e_1}\), and \(1 \leq \sigma_2(a,b)_1 = a \leq 2^{e_1}\). Thus \(\sigma_2(a,b)_{1,2} \in S_2\) so we can view \(\sigma_2\) as “stacking” the sets \(S_2^r(G)'\), \(1 \leq r \leq P_2\), in succession by increasing \(r\) “over” \(S_2^1 \cong \langle Y_2 \rangle\). So we have \((\sigma_2 \circ I_2 \circ f_2)(G) \subseteq \langle Y_2 \rangle \times P(u) \subseteq Y_3\), where \(u\) is the maximum of \(n_y\) over all \(y\) belonging to the last set \(S_2^{P_2}(G)\). We will prove later that \(u = u_3(G)\).)

5. [The Composition Step]

Finally, define \(f_3 : G \to Y_3\) as the composition \(f_3 = \sigma_2 \circ I_2 \circ f_2\).

5.2 Embeddings of grids of higher dimension

Again set \(G = [a_1 \times \cdots \times a_k]\). In this subsection we inductively construct embeddings \(f_i : G \to Y_i^{(u_i)}\) for \(3 \leq j \leq k\). Assume then that we have constructed the required maps \(f_2, \ldots, f_i\), \(3 \leq i < k\), and we construct \(f_{i+1} : G \to Y_i^{(u_{i+1})}\).

Starting from \(f_i(G) \subseteq Y_i^{(u_i)}\), we will use direct analogues \(I_i\) and \(\sigma_i\) of the inflation map \(I_2\) and the stacking map \(\sigma_2\) used in constructing \(f_3\) from \(f_2\). In particular, \(I_i\) will inflate \(f_i(G)\) by introducing blank \((i-1)\)-levels \(Y_i^j\) of \(Y_i\) using a matrix \(F(i) = F^X\) (constructed by the procedure following Corollary 4.5) that encodes which \((i-1)\)-levels of \(Y_i\) will be introduced as blank. Then a stacking map \(\sigma_i : S_i^{(P_i)}(G)' \to S_i^1 \times (u_{i+1}) = Y_i^{(u_{i+1})}\) stacks the sets \(S_i^r(G)'\), \(1 \leq r \leq P_i\), in succession by increasing \(r\) over \(S_i^1 \cong \langle Y_i \rangle\) to yield the final image \(f_{i+1}(G)\). Thus we may write \(f_{i+1}\) as the composition \(f_{i+1} = \sigma_i \circ I_i \circ f_i\), and inductively

\[
f_{i+1} = \sigma_i \circ I_i \circ \sigma_{i-1} \circ I_{i-1} \circ \cdots \circ \sigma_2 \circ I_2 \circ f_2.
\]

**Construction of the map** \(f_{i+1} : [a_1 \times a_2 \times a_3 \times \cdots \times a_k] \to Y_{i+1}^{(u_{i+1})}, \ 2 \leq i < k\)

1. Set \(G = [a_1 \times a_2 \times a_3 \times \cdots \times a_k]\). Assume inductively that the map \(f_i : [a_1 \times a_2 \times a_3 \times \cdots \times a_k] \to Y_i^{(u_i)}\) has been constructed. We now operate on \(f_i(G)\) to obtain our image \(f_{i+1}(G) \subseteq Y_{i+1}^{(u_{i+1})}\).

2. [Designation of Blank \((i-1)\)-Levels]

a) Set \(P_i = \Pi_{t=1}^{i+1} a_t\).

b) Construct the \(P_i \times 2^{e_i-e_{i-1}}, (0,1)\)-matrix \(F(i) = (f_{cd}(i))\) by the procedure following Corollary 4.5.

[Comment: Thus by Corollary 4.5, the matrix entries \(f_{cd}(i)\) satisfy relations (5)-(7).]

c) For \(1 \leq c \leq P_i\), define “level \(j\) of \(S_i^c\)”, or the “\(j\)'th level of \(S_i^c\)”, to be the \((i-1)\)-level \(Y_i^{(c-1)2^{e_i-e_{i-1}+j}} \subseteq Y_i\). Designate level \(j\) of \(S_i^c\) as being either ‘blank’ or ‘nonblank’ as follows.

Level \(j\) of \(S_i^c\) is ‘blank’ if \(f_{cj}(i) = 1\), and is ‘nonblank’ if \(f_{cj}(i) = 0\).

d) Linearily order all the nonblank \((i-1)\)-levels in \(S_i^{(P_i)}\) by increasing \(i\)'th coordinate; that is, for any two nonblank levels \((i-1)\)-levels \(Y_i^t\) and \(Y_i^{t'}\) we have \(Y_i^t < Y_i^{t'}\) if and only if \(t < t'\).

e) Let \(S_i^c(G)\) (resp. \(S_i^{(c)}(G)\)) be the set points of \(S_i^c\) lying in nonblank \((i-1)\)-levels of \(S_i^c\) (resp. \(S_i^{(c)}\)).

3. [The Inflation Step]

a) “Inflate” the image \(f_i(G)\) by the map \(I_i : Y_i^{(u_i(G))} \to S_i^{(P_i)}(G)\) as follows. Take any \(z = (z_1, z_2, \ldots, z_i) \in Y_i^{(u_i(G))}\). Then let

\[
I_i(z) = (z_1, z_2, \ldots, z_{i-1}, z_i')
\]
where $Y_i^{z_i}$ is the $z_i$'th nonblank $(i - 1)$-level in $S_i^{(P_i)}$ (in the ordering of nonblank $(i - 1)$-levels in $S_i^{(P_i)}$ from step 2(d)).

[Comment: Let $b$ and $r$ be the integers such that $Y_i^{z_i}$ is the $b$'th nonblank $(i - 1)$-level of $S_i^r(G)$. Recall that $N_r(b)$ is the column index in matrix $F(i)$ of the $b$'th zero in row $r$ of $F(i)$. Then by step 2(c), we have $z'_i = (r - 1)2^{e_i-1} + N_r(b), 1 \leq N_r(b) \leq 2^{e_i-1}$, so

$$I_i(z) = (z_1, z_2, \ldots, z_{i-1}, (r - 1)2^{e_i-1} + N_r(b)).]$$

b) Let $S_i^r(G)' = S_i^r(G) \cap (I_i \circ f_i)(G)$ and $S_i^{(r)}(G)' = S_i^{(r)}(G) \cap (I_i \circ f_i)(G).

4. [The Stacking Step]

“Stack” the sets $S_i^r(G)'$, $1 \leq r \leq P_i$, from step 3 (The Inflation Step) on top of each other “over” $S_i^1$ in succession as $r$ increases. We do this by the stacking map $\sigma_i : S_i^{(P_i)}(G)' \to \bigcup_{u_{i+1}} \sigma_i : S_i^{(P_i)}(G)' \to S_i^1 \times P(u_{i+1}) \subseteq Y_{i+1}^{(u_{i+1})}$ defined as follows. Let $y = (y_1, y_2, \ldots, y_{i-1}, y_i) \in S_i^{(P_i)}(G)'$, say with $y \in S_i^r(G)'$, $1 \leq r \leq P_i$.

a) Let $n_y = |\{z \in S_i^r(G)' : z_{i-1} = y_{i-1} \text{ and } z_i \equiv y_i \pmod {2^{e_i-1}}\}|$.

b) Let $\bar{y}_i$ be such that $\bar{y}_i \equiv y_i \pmod {2^{e_i-1}}$, with $1 \leq \bar{y}_i \leq 2^{e_i-1}$. So $\bar{y}_i = N_r(b)$ as in the comment to step 3 (The Inflation Step).

c) Define $\sigma_i(y) = (y_1, y_2, \ldots, y_{i-1}, \bar{y}_i, n_y) = (y_1, y_2, \ldots, y_{i-1}, N_r(b), n_y)$. [Comment: Since $\sigma_i(y_1, y_2, \ldots, y_{i-1}, \bar{y}_i) \in S_i^1$, we see that $\sigma_i(S_i^{(P_i)}(G)') \subseteq S_i^1 \times P(m)$, where $m$ is the maximum of $n_y$ over all $y$ belonging to the last set $S_i^{(r)}(G)'$. We will see later that $m = u_{i+1}(G)$.]

5. [The Composition Step]

Finally define $f_{i+1} : G \to Y_{i+1}^{(u_{i+1})}$ as the composition $f_{i+1} = \sigma_i \circ I_i \circ f_i$. In particular, for any $z = f_i(v) \in f_i(G)$ we have $f_{i+1}(v) = \sigma_i(I_i(z))$.

Consider now the construction of $f_i$ from $f_{i-1}$ by the above construction. The stacking step 4 suggests that we can regard $f_i(G) \subseteq S_i^1 \times P(m)$ (as in the comment to step 4) as a collection of stacks addressed by the points of $S_i^1 \subseteq (Y_i-1)$. The height of each such stack extends into the $i$'th dimension of $Y_i$. Specifically, for any $x \in (Y_i-1)$, we let $Stack_i(x, r) = \{\sigma_i-1(y) : (\sigma_i-1(y))_{i-1} = x \text{ and } y \in S_i^{(r)}(G)\}'$, the stack addressed by $x$, for a given integer $r \leq P_i$. So $Stack_i(x, r)$ consists of images $\sigma_i-1(y)$ which project onto $x$ in their first $i-1$ coordinates, and such that $y$ comes from the first $r$ many sets $S_i^j(G)'$, $1 \leq j \leq r$. These sets $S_i^j(G)'$, $1 \leq j \leq r$, are stacked on top of $S_i^1 \subseteq (Y_i-1)$ in order of increasing $j$. We view the “height” of $\sigma_i-1(y)$ in $Stack_i(x, r)$ as its $i$’th coordinate $\sigma_i-1(y)_i = n_y$, and the height of $Stack_i(x, r)$ as $|Stack_i(x, r)|$. Now define $[r]_i = \max\{|Stack_i(x, r)| : x \in (Y_i-1)\}$, the maximum height of any of these stacks addressed by points of $(Y_i-1)$. Thus by definition $\sigma_i-1(S_i^{(r)}(G)')$ is contained in the first $[r]_i$ many $(i-1)$-levels of $Y_i$; thus that, $\sigma_i-1(S_i^{(r)}(G)') \subseteq Y_i^{[r]_i}$. As a convenience, for a stack address $x \in S_i^1 \subseteq (Y_i-1)$, we will refer to $x$ either as a member of $S_i^1$ or of $(Y_i-1)$, in most cases of $(Y_i-1)$.

The parameter $[r]_i$ plays a role in our containment result. We will see in Lemma 5.1 that $I_{i-1}(f_{i-1}(D_i^{(r)})) \subseteq S_i^{(r)}(G)$, as suggested in section 3. Thus $[r]_i \geq \lceil \frac{|D_i^{(r)}|}{|Y_i-1|} \rceil = \lceil \frac{|D_i^{(r)}|}{2r-1} \rceil$. We will also see ( Lemma 5.1.1 ) that in fact equality holds in the first inequality for each relevant $r$ and $i$; that is, $[r]_i$ is as small as it could possibly be. Thus taking $r = P_i$ and recalling that $D_i^{(P_i-1)} = G$, we obtain $f_i(G) \subseteq Y_i^{(u_i(G))} \subseteq Opt(G)$ for each $i \geq 2$, our containment result.
The stacking step 4 of the construction will yield the following monotonic properties of Stack\(_i(x,r)\). First, if \(w', w'' \in \text{Stack}\(_i(x,r)\)\) with \(\sigma_{i-1}(w') \in S_{i-1}^c\) and \(\sigma_{i-1}(w'') \in S_{i-1}^d\), \(1 \leq c,d \leq r\), then \(w'_1 > w''_1\) implies that \(c > d\). That is, the originating \((i-1)\)-section number under \(\sigma_{i-1}\) is strictly increasing as we move up any fixed stack. For the second property, let \(\text{Stack}\(_i(x,r) = f_1(D_{i-1}^{(r)}) \cap \text{Stack}\(_i(x,r)\)\). We will see that \(\text{Stack}\(_i(x,r)\) is an initial substack of \(\text{Stack}\(_i(x,r)\)\). Specifically, if \(w', w'' \in \text{Stack}\(_i(x,r)\)\) with \(f_1^{-1}(w') \in D_{i-1}^c\) and \(f_1^{-1}(w'') \in D_{i-1}^d\), then \(w'_1 > w''_1\) implies that \(c \geq d\). Thus the originating \((i-1)\)-page number under \(f_1^{-1}\) is nondecreasing as we move up any fixed stack. The first monotonicity property (which we call stack monotonicity) is immediate from our construction, and will be noted for the record in the Lemma which follows (part g). The second monotonicity property (which we call page monotonicity) will be proved later in Theorem 6.1(a).

**Lemma 5.1** (a) For \(x \in V(G)\). For \(2 \leq j \leq k\) the maps \(f_j\) satisfy the following.

(a1) For \(1 \leq i \leq k - 1\) and \(i + 1 \leq j \leq k\) we have \(f_j(x)_{1-i} = f_{i+1}(x)_{1-i}\).

(a2) In particular, suppose \(2 \leq i \leq k - 1\) and take \(z = f_j(x)\). Let \(I_i(z) = (z_1, z_2, \ldots, z'_i)\) as in step 3 (the Inflation Step) of the above construction, and express \(z'_i = z_i = (r - 1)2^{e - e_i - 1} + N_i(b)\) as in the comment to step 3.

Then \(f_k(x)_i = N_i(b)\), and \(1 \leq f_k(x)_i \leq 2^{e - e_i - 1}\).

(b) For any \(x, y \in V(G)\), if \(|f_i(x)_i - f_i(y)_i| \leq e\) then \(|f_k(x)_i - f_k(y)_i| \leq 2e + 2\), where the last difference is interpreted mod \(2^{e - e_i - 1}\).

(c) Consider the map \(\sigma_i : y \to \sigma_i(y)_{1-i}, \) with \(y \in S_i^{(P)}(G')\), obtained from \(\sigma_i\) by projecting onto the first \(i\) coordinates. Then \(\sigma_i\) is one to one when restricted to any one \(i\)-section; that is, to the set \(\{y \in S_i^{r}(G')\}\) for a given \(1 \leq r \leq P_i\). Hence \(\sigma_i\) is one to one, and \(f_i\) is one to one for all \(2 \leq i \leq k\).

(d) Let \(1 \leq r < P_1\) and \(1 \leq t < P_{i-1}\) be integers. Then

(e1) \(f_i(D_{i-1}^{(r)}) = S_i^{(r)}(G')\)

(e2) \(f_i(D_{i-1}^{(r)}) = S_i^{(r)}(G')\)

Then for \(i \geq 3\)

(e1) for \(1 \leq r \leq P_{i-1}\) we have \([r]_i = l(i, r), \) and \(|Y_i(l(i, r))| - |f_i(D_{i-1}^{(r)})| < 2^{e_i - 1}\), and

(e2) for \(1 \leq r \leq P_{i-1}\) we have \(f_i(D_{i-1}^{(r)}) \subseteq Y_i^{(l(i, r))}, \) and for \(1 \leq r \leq P_i, f_i(D_{i-1}^{(r)}) \subseteq Y_i^{(l(i, r))}\).

(e3) \(f_i(G) \subseteq Y_i^{(u_i)} \subseteq \text{Opt}(G)\). Moreover, we have \(f_k(G) \subseteq \text{Opt}'(G)\) and \(H_k(G) \subseteq \text{Opt}(G)\).

(f) For \(i \geq 2\) we have \(I_i(f_i(D_{i-1}^{(r)})) \subseteq S_i^{(r)}(G)\) for \(1 \leq r \leq P_i\), and \(S_i^{(r)}(G) \subseteq I_i(f_i(D_{i-1}^{(r+1)}))\) for \(1 \leq r \leq P_i - 1\).

(g) (Stack monotonicity) Suppose \(w', w'' \in \text{Stack}\(_i(x,r)\)\), \(1 \leq r \leq P_{i-1}\), with \(\sigma_{i-1}(w') \in S_{i-1}^c\) and \(\sigma_{i-1}(w'') \in S_{i-1}^d\), \(1 \leq c,d \leq r\). Then \(w''_1 > w'_1 \iff d > c\).

**Proof.** Consider first part (a1). We proceed by induction on \(j\). The base case \(j = i + 1\) is trivial. So suppose inductively that \(f_j(x)_{1-i} = f_{i+1}(x)_{1-i}\) for some \(j \geq i + 1\). By the definition of \(I_j\) and \(\sigma_j\) from the inflation and stacking steps respectively, we have \(I_j(z)_{1-j-1} = z_{1-j-1}\) and \(\sigma_j(y)_{1-j-1} = y_{1-j-1}\) for \(z\) and \(y\) in the domain of \(I_j\) and \(\sigma_j\) respectively. Thus since \(j \geq i + 1\)
we have \( I_j(f_j(x))_{1 \to i} = f_j(x)_{1 \to i} \). Therefore \( f_{j+1}(x)_{1 \to i} = \sigma_j(I_j(f_j(x)))_{1 \to i} = I_j(f_j(x))_{1 \to i} = f_j(x)_{1 \to i} = f_{i+1}(x)_{1 \to i} \). This completes the inductive step, so (a1) is proved.

Next consider (a2). For \( i \geq 2 \) we have \( f_{i+1}(x)_i = N_i(b) \) by step 4c (The Stacking Step), and \( 1 \leq N_i(b) \leq 2^e - e_{i-1} \) by the comment to step 3 and the definition of \( N_i(b) \), proving (a2).

Consider (b). We know that \( I_i(f_i(x)) \) is in some nonblank \((i-1)\)-level, say the \( b'th \) one, of some \( i \)-section, say \( S_i(r'_1) \), of \( S_i(P_1) \). So we can write \( I_i(f_i(x))_i = (r-1)2^e - e_{i-1} + N_i(b) \). Since \( |f_i(x)_i - f_i(y)_i| \leq e \) and \( I_i \) preserves the order of \((i-1)\)-levels, it follows that \( |I_i(f_i(x))_i - I_i(f_i(y))_i| \leq e \). Assuming without loss that \( I_i(f_i(y))_i > I_i(f_i(x))_i \) (otherwise interchange the roles of \( x \) and \( y \)), we can write \( I_i(f_i(y))_i = (r-1)2^e - e_{i-1} + N_i(b + t) \) for some \( 0 \leq t \leq e \), where we interpret the function \( N_r(*) \) in row major order as in the discussion just preceding Corollary 4.5. So we have \( |f_k(x)_i - f_k(y)_i| \leq |N_i(b + e) - N_i(b)| \leq 2e + 2 \), where the first inequality follows from part (a2), and the second by using the more generous of the bounds (a) and (b) in Corollary 4.5 and interpreting the differences \( \mod 2^e - e_{i-1} \).

Consider (c). For the first statement, let \( y, z \in S_i'(G)' \) for some \( 1 \leq r \leq P_i \). We must show that \( (\sigma_i(y))_{1 \to i} \neq (\sigma_i(z))_{1 \to i} \). By the comment to step 3 (The Inflation Step) we may write \( y = (y_1, y_2, \ldots, y_{i-1}, (r-1)2^e - e_{i-1} + N_i(d)) \) and \( z = (z_1, z_2, \ldots, z_{i-1}, (r-1)2^e - e_{i-1} + N_i(c)) \) for suitable integers \( d \) and \( c \) and \( 1 \leq N_i(d), N_i(c) \leq 2^e - e_{i-1} \). By step 4 (The Stacking Step), we have \( (\sigma_i(z))_{1 \to i} = (z_1, z_2, \ldots, z_{i-1}, N_i(r)) \), and \( (\sigma_i(y))_{1 \to i} = (y_1, y_2, \ldots, y_{i-1}, N_i(c)) \). If \( y \) and \( z \) disagree at one of their first \( i - 1 \) coordinates, then obviously \( (\sigma_i(y))_{1 \to i} \neq (\sigma_i(z))_{1 \to i} \) by these formulas. So we can suppose that \( y_{1,i-1} = z_{1,i-1} \). But since \( y \) and \( z \) are distinct, the formulas above for \( y \) and \( z \) force them to disagree in their \( i'th \) coordinates, so \( N_i(d) \neq N_i(c) \). Thus \( y \neq z \) \( \mod 2^e - e_{i-1} \). So by definition of \( \sigma_i \) we get \( (\sigma_i(y))_{1 \to i} \neq (\sigma_i(z))_{1 \to i} \), as desired. For one-to-oneness of \( \sigma_i \) itself, it only remains to show that if \( y, z \) come from distinct sections, then \( \sigma_i(y) \neq \sigma_i(z) \). So let \( y \in S_i'(G)' \) and \( z \in S_i'(G)' \), say with \( s < t \). If \( (\sigma_i(y))_{1 \to i} \neq (\sigma_i(z))_{1 \to i} \), then the claim follows obviously, so assume \( (\sigma_i(y))_{1 \to i} = (\sigma_i(z))_{1 \to i} \). Then \( \sigma_i(y) \) and \( \sigma_i(z) \) both belong to the stack addressed by \( (\sigma_i(y))_{1 \to i} \), namely, \( \text{Stack}_{i+1}((\sigma_i(y))_{1 \to i}, t) \). But then since \( s < t \), \( \sigma_i(y) \) is lower in this stack than \( \sigma_i(z) \) by definition of \( \sigma_i \). Then \( (\sigma_i(y))_{i+1} < (\sigma_i(z))_{i+1} \), proving one-to-oneness of \( \sigma_i \). As for the claim about the \( f_i \), observe first that \( f_2 \) is one-to-one. So from \( f_{i+1} = \sigma_i \circ I_i \circ f_i \) and the one-to-oneness of \( \sigma_i \) and \( I_i \), it follows by induction on \( i \) that \( f_i \) is one-to-one for all \( 2 \leq i \leq k \).

For (d), we start by proving (d1) for \( i = 2 \). By Theorem 2.2.1, \( Y_2^{(m-1)} \subset f_2(G) \subset Y_2^m \), where \( m = \left\lceil \frac{|G|}{2^e} \right\rceil \) (noting that \( m = u_2(G) \)). That is, all but at most the last column \( Y_2^m \) of \( Y_2^{(m)} \) is contained in \( f_2(G) \). Applying \( I_2 \) to both sides of the last containment and noting that \( (I_2 \circ f_2)(G) = S_2(P_2)(G)' \), we see that every nonblank column of \( S_2(P_2)(G)' \) lies in \( S_2(P_2)(G)' \), except possibly the last one \( I_2(Y_2^d) \), call it \( Y_2^d \). Since every 2-section must contain at least one nonblank column we have \( Y_2^d \subset S_2^P (\text{the last 2-section}) \), so \( S_2^{(P_r-1)}(G) = S_2^{(P_r-1)}(G)' \). Taking subsets, we get \( S_i'^{(r)}(G) = S_i'(r')(G)' \) for \( 1 \leq r < P_2 \).

Next we show that (d1) for \( i \) implies (d2) and (d3) for \( i+1 \). Take \( x = (x_1, x_2, \ldots, x_i) \in (Y_i) \) to be a stack address. By definition of matrix \( F(i) \), its entry \( f_{t,x_i}(i) \) satisfies \( f_{t,x_i}(i) = 0 \) if and only if the \((i-1)\)-level \( Y = Y^{(i-1)2^e - e_{i-1} + x_i} \) (the \( i'th \) \((i-1)\)-level of \( i\)-section \( S_i \)) is nonblank. Assume first that \( f_{t,x_i}(i) = 0 \). Thus \( Y \) is nonblank. Since \( t < P_i \) by the hypothesis for (d2) with \( i+1 \) replacing \( i \), we have \( Y \subset S_i'(P_{i-1})(G) = S_i'(P_{i-1})(G)' \), the last equality holding by the conclusion of (d1) for \( i \). Thus every point of \( Y \) lies in the domain of
\( \sigma_i \). So \( \text{Stack}_{i+1}(x, t) \) receives the point \( \sigma_i(x_1, x_2, \ldots, (t-1)2^{e_i-e_i-1}+x_i) \) under the stacking map \( \sigma_i \). If \( f_{t,x}(i) = 1 \), then \( \text{Stack}_{i+1}(x, t) \) receives no point \( \sigma_i(y) \) with \( y \in S_i^t \). It follows that \( |\text{Stack}_{i+1}(x, t)| = t - \sum_{j=1}^{t} f_{j,x}(i) \). Since by Lemma 4.4b the sum on the right must be one of two successive integers depending on \( t \), it follows that \( |\text{Stack}_{i+1}(x, t)| \) is one of two successive integers depending on \( t \) but independent of \( x \). Thus by definition of \([t]_{i+1}\), we get \( |\text{Stack}_{i+1}(x, t)| = [t]_{i+1} \) or \([t]_{i+1} - 1 \), proving (d2) for \( i+1 \). Consider now (d3) for \( i+1 \), and let \( y \in \langle Y \rangle \) be a stack address with \( \sigma_i(y) = x_i \). Then clearly \( f_{j,y}(i) = f_{j,x}(i) \) for all \( 1 \leq j \leq P_i \) by the definition of matrix \( F(i) \). Thus \( \sum_{j=1}^{t} f_{j,x}(i) = \sum_{j=1}^{t} f_{j,y}(i) \), and hence \( |\text{Stack}_{i+1}(x, t)| = |\text{Stack}_{i+1}(y, t)| \), as required.

Finally we show that (d2) for \( i+1 \) in place of \( i \) implies (d1) for \( i+1 \) in place of \( i \). By assumption we have \( |\text{Stack}_{i+1}(x, t)| = [t]_{i+1} \) or \([t]_{i+1} - 1 \) for \( t \leq P_i - 1 \) and any stack address \( x \in \langle Y \rangle \). Observe that \( |P_i|_{i+1} = |P_i - 1|_{i+1} + 1 \) since the projection \( \sigma_i \) onto the first \( i \) coordinates is one-to-one restricted to any single section, in this case \( S_i^t \), by part (c). Further, \( |\text{Stack}_{i+1}(x, P_i-1)| \geq |P_i - 1|_{i+1} - 1 \) for all stack addresses \( x \) by assumption. It follows that \( Y_i^{[P_i-1]_{i+1}} \subseteq f_{i+1}(G) \subseteq Y_i^{[P_i]_{i+1}} \subseteq Y_i^{[P_i-1]_{i+1}} \).

So at most two \( i \)-levels of \( Y_i^{[P_i]_{i+1}} \) are possibly not entirely contained in \( f_{i+1}(G) \), these being the top two \( Y_i^{[P_i+1]} \) and \( Y_i^{[P_i]_{i+1} - 1} \), and if this possibility occurs then \( |P_i|_{i+1} = |P_i - 1|_{i+1} + 1 \) (so also \( |P_i|_{i+1} - 1 = |P_i - 1|_{i+1} \)). But still \( Y_i^{[P_i]_{i+1}} \cap f_{i+1}(G) \neq \emptyset \) if and only if \( i \leq j \leq |P_i|_{i+1} \) by definition of \( |P_i|_{i+1} \). Thus every \( i \)-level \( I_{i+1}(Y_i^{[P_i]_{i+1}}) \), \( 1 \leq j \leq |P_i|_{i+1} \), is nonblank, and by the previous sentence all but at most two of these nonblank \( i \)-levels belong to \( S_i^{(P_i+1)}(G)' \), the two possible exceptions being \( I_{i+1}(Y_i^{[P_i+1]}) \) and \( I_{i+1}(Y_i^{[P_i]_{i+1} - 1}) \).

Since any \((i+1)\)-section \( S_{i+1}^j \), \( 1 \leq j \leq P_{i+1} \), contains at least two nonblank \( i \)-levels (since \( a_i > 4 \) for all \( i \)) it follows that the top two nonblank \( i \)-levels of \( S_i^{(P_i+1)} \) belong to the top \((i+1)\)-section \( S_{i+1}^{P_{i+1}} \). So the remaining \((i+1)\)-sections \( S_{i+1}^j \), \( j < P_{i+1} \), satisfy \( S_{i+1}^j(G) = S_{i+1}^j(G)' \). In particular we have \( S_i^{(P_i+1)-1}(G) = S_i^{(P_i+1)-1}(G)' \), and taking subsets we obtain (d1) for \( i+1 \).

Consider now (e1). The case \( i = 2 \) follows immediately from Theorem 2.2e,f, and we assume \( i \geq 3 \).

For the lower bound \([r]_i \geq \left\lfloor \frac{|P_i|_1}{2^{i-2}} \right\rfloor = l(i, r) \) , suppose first that \( r < P_{i-1} \). By definition we have \([r]_i \geq \left\lfloor \frac{|P_i|_1}{2^{i-2}} \right\rfloor = \left\lfloor \frac{|S_{i-1}^r(G)|}{2^{i-1}} \right\rfloor \). So since \( \sigma_i \) is one to one, it suffices to show that \( |S_{i-1}^r(G)'| \geq \left\lfloor \frac{|P_i|_1}{2^{i-2}} \right\rfloor \cdot 2^{e_i-2} \). For then we get \([r]_i \geq \left\lfloor \frac{|P_i|_1}{2^{i-2}} \right\rfloor \cdot 2^{e_i-2} - 1 \geq \left\lfloor \frac{|P_i|_1}{2^{i-2}} \right\rfloor \), and since \([r]_i \) is an integer \([r]_i \geq \left\lfloor \frac{|P_i|_1}{2^{i-2}} \right\rfloor = l(i, r) \). By part (d1) and Lemma 4.1 we have \( S_{i-1}^r(G) = \sum_{i=1}^{r} S_{i-1}^r(G)' \), and \( S_{i-1}^r(G)' \) consists of \( \left\lfloor \frac{|P_i|_1}{2^{i-2}} \right\rfloor \) many \((i-2)\)-levels in \( Y_{i-1} \). Since each such \((i-2)\)-level has \( 2^{e_i-2} \) points, we get \( |S_{i-1}^r(G)'| \geq \left\lfloor \frac{|P_i|_1}{2^{i-2}} \right\rfloor \cdot 2^{e_i-2} \) as required. Now suppose that \( r = P_{i-1} \). We have \( |S_{i-1}^{P_{i-1}}(G)'| = |G| = |P_{i-1} a_{i-2} a_{i-3} \cdots a_1| \). So again since \( \sigma_i \) is one to one we get \( |P_{i-1}| \geq \left\lfloor \frac{|S_{i-1}^{P_{i-1}}(G)'|}{|Y_{i-1}|} \right\rfloor \geq \frac{|P_{i-1} a_{i-2} a_{i-3} \cdots a_1|}{2^{i-1}} = \left\lfloor \frac{|P_{i-1}|_1}{2^{i-1}} \right\rfloor \), as required.

It remains to show that \([r]_i \geq l(i, r) \). Recall that for any stack address \( x \in \langle Y_{i-1} \rangle \), \( |\text{Stack}_i(x, r) \) is at most the number of 0's in column \( x_i \) among the first \( r \) rows in matrix \( F(i-1) \). By Theorem 4.4b this number of 0's is either the same for all \( x \) or is one of two successive integers, call them \( \alpha_r \) or \( \alpha_{r-1} \), depending only on \( r \) (and \( i \), which we fix in this argument). Since \([r]_i \) is the maximum of \( |\text{Stack}_i(x, r)| \) over all \( x \in \langle Y_{i-1} \rangle \), it suffices to prove
that $\alpha_r \leq l(i, r)$. By our construction, the total number of 0’s in the first $r$ rows of $F(i - 1)$ is the number of of nonblank $(i - 2)$-levels in $S_i^{(r)}$. So by Lemma 4.1a this number of 0’s is $\lceil \frac{r\alpha_r - 1}{2^{k+1}} \rceil$. Let $\frac{p}{2^{k+1}} = \lceil \frac{r\alpha_r - 1}{2^{k+1}} \rceil - \frac{r\alpha_r - 1}{2^{k+1}}$. Since the number of columns of $F(i - 1)$ is $2^{e_i} - e_i - 2$, we have $\alpha_r = \lfloor \frac{r\alpha_r - 1}{2^{k+1}} \rfloor + \frac{p}{2^{k+1}} = \lceil \frac{r\alpha_r - 1}{2^{k+1}} \rceil$, as required.

For the second claim of (e1), again note that each $(i - 1)$-level of $Y_i$ has size $2^{e_i-1}$. Thus $|Y_i^{(l(r))}| - |f_i(D_i^{(r)})| = 2^{e_i-1} - \lceil |D_i^{(r)}| \rceil = 2^{e_i-1}$, completing the proof of (e1).

For (e2) we prove both $f_i(D_i^{(r)}) \subseteq Y_i^{(l(i,r))}$ and $f_i(D_i^{(r)}) \subseteq Y_i^{(l'(i,r))}$ by induction. If $i = 2$, then both statements hold by Theorem 2.2d,e.f. So let $i > 2$ be given, and assume inductively that both of these containments hold for $i - 1$ in place of $i$. By this assumption we have $f_{i-1}(D_i^{(r)}) \subseteq Y_i^{(l'(i-1,r))}$. Applying the inflation map $I_{i-1}$ to both sides, we obtain $(I_{i-1} \circ f_{i-1})(D_i^{(r)}) \subseteq S_{i-1}^{(r)}(G)$, since by Lemma 4.1a we see that $S_{i-1}^{(r)}$ has the same number $l'(i-1,r)$ of nonblank $(i - 2)$-levels as the number of $(i - 2)$-levels in $Y_i^{(l'(i-1,r))}$, and $I_{i-1}$ preserves the order of the latter set of $(i - 2)$-levels. Thus $(I_{i-1} \circ f_{i-1})(D_i^{(r)}) \subseteq S_{i-1}^{(r)}(G')$ by definition of $S_{i-1}^{(r)}(G')$. Now applying $\sigma_{i-1}$ to the last containment, and recalling that $f_i = \sigma_{i-1} \circ I_{i-1} \circ f_{i-1}$, we get $f_i(D_i^{(r)}) = (\sigma_{i-1} \circ I_{i-1})(f_{i-1}(D_i^{(r)})) \subseteq \sigma_{i-1}(S_{i-1}^{(r)}(G')) \subseteq Y_i^{(l'(i,r))} = Y_i^{(l(i,r))}$, where the last containment holds by definition of $[r]$. By (e1) we have $[r] = l(i, r)$, thereby proving $f_i(D_i^{(r)}) \subseteq Y_i^{(l(i,r))}$.

To show $f_i(D_i^{(r)}) \subseteq Y_i^{(l'(i,r))}$, apply $f_i(D_i^{(r)}) \subseteq Y_i^{(l(i,r))}$ using $a_i r$ in place of $r$ and observing that $D_i^{(r)} = D_i^{(a_i r)}$. This completes the inductive step, and hence the proof of (e2).

Consider now (e3). Recall that $P_{i-1} = \prod_{t=1}^{k} a_t$, so that $V(G) = V(D_i^{(P_{i-1})})$. Noting that $l(i, P_{i-1}) = \lceil \frac{|D_i^{(P_{i-1})}|}{2^{k+1}} \rceil = \lceil \frac{|Y_i^{(l'(i,r))}|}{2^{k+1}} \rceil = u_i(G)$, by part (e2) we have $f_i(G) \subseteq Y_i^{(l(i,P_{i-1}))} = Y_i^{(u_i(G))}$, proving the first containment. For the rest, note that $u_i(G) \leq 2^{[\log_2(|G|)]-e_i-1}$, so $f_i(G) \subseteq Y_i^{(u_i(G))} \subseteq \langle Y_{i-1} \rangle \times P(2^{[\log_2(|G|)]-e_i-1})$. The last graph is a spanning subgraph of $Opt(G)$ for each $i$, yielding the second containment, and for $i = k$ yielding the final two containments after performing the relabeling with hypercubes addresses by the map $H^k$.

For (f), start with $f_i(D_i^{(r)}) \subseteq Y_i^{(l'(i,r))}$ from (e2). Then we get $I_i(f_i(D_i^{(r)})) \subseteq I_i(Y_i^{(l'(i,r))}) = S_i^{(r)}(G)$. The equality holds since the right side consists of $l'(i, r)$ many (nonblank) $(i - 1)$-levels, while $I_i$ preserves the order of $(i - 1)$-levels by increasing $i$-coordinate value.

Next we show that $S_i^{(r)}(G) \subseteq I_i(f_i(D_i^{(r+1)}))$. For the case $i = 2$ we apply Theorem 2.2e,f and Corollary 2.3c to $G = D_i^{(r+1)}$, together with Lemma 4.1 to obtain $S_i^{(r)}(G) = I_i(Y_i^{(l'(i,r))}) \subseteq I_i(\langle Y_i^{(l'(2,r+1)-1)} \rangle) = I_i(f_i(D_i^{(r+1)})))$.

Proceeding by induction on $i$, let $i \leq 3$ and assume the statement true for $i - 1$ and $1 \leq r \leq P_{i-1} - 1$, and we prove it for $i$ and $1 \leq r \leq P_i - 1$. It suffices to show that $Y_i^{(l'(i,r))} \subseteq f_i(D_i^{(r+1)})$, since then we can just apply $I_i$ to each side and use $I_i(Y_i^{(l'(i,r))}) = S_i^{(r)}(G)$ (as observed above) to get the result directly.

By (e1) we have $[a_i r+2] = l(i, a_i r+2) \geq l'(i, r)+1$, using $2a_i-a_i-2 \ldots a_1 > 2^{e_i-1}$. By (d), for any $x \in (Y_{i-1})$ we have $[Stack_i(x,a_i r+2)] = [a_i r+2]$, or $[a_i r+2] - 1$, so $[Stack_i(x,a_i r+2)] \geq l'(i, r)$. It follows that $Y_i^{(l'(i,r))} \subseteq \sigma_{i-1}(S_i^{(a_i r+2)}(G')) \subseteq \sigma_{i-1}(I_i-1(f_{i-1}(D_i^{(a_i r+3)})))$, where the last containment is by the inductive hypothesis. Now since $a_i \geq 2^{22} > 3$, we have $a_i r+3 < a_i r+1$, so $D_i^{(a_i r+3)} \subseteq D_i^{(a_i r+1)} = D_i^{(r+1)}$. So we have $Y_i^{(l'(i,r))} \subseteq \sigma_{i-1} \circ I_{i-1} \circ f_{i-1}(D_i^{(r+1)}) =$
For (g) observe that \( \sigma_{i-1} \) stacks the sets \( S_{i-1}^r(G) \) onto \( S_{i-1}^l(G) \) in succession by increasing \( r \); that is, \( S_{i-1}^c(G) \) is stacked before \( S_{i-1}^d(G) \) precisely when \( c < d \). Part (g) follows.

We can now give the final embedding \( H^k \) of \( G \) into \( \text{Opt}'(G) \). It is based on the fact that \( f_k(G) \subseteq \text{Opt}'(G) \) from Lemma 5.1(e3). So by definition we have \( 1 \leq f_k(x)_j \leq 2^{e_j-e_j-1} \) for \( 1 \leq j \leq k \) and all \( x \in V(G) \). Given these facts, we obtain \( H^k \) from \( f_k \) as follows. For each \( 1 \leq j \leq k \) we interpret \( f_k(x)_j \) as a hypercube point in \( Q_{e_j-e_{j-1}} \) using the inverse image of the labeling \( L_{e_j-e_{j-1}} \) of Corollary 4.8. We then concatenate these hypercube points (now strings over \( \{0,1\} \) left to right in order of increasing \( j \) to obtain \( H^k(x) \). The details are as follows.

**Construction of the map** \( H^k : G = [a_1 \times a_2 \times a_3 \times \ldots \times a_k] \rightarrow \text{Opt}(G) \)

1. **Initialization and the case** \( k = 2 \).
   a) Start with the map \( f_2 : G \rightarrow Y_{2^2} \subseteq P(2^{e_1}) \times P(2^{e_2-e_1}) \).
   b) Define \( H^2 : G \rightarrow \text{Opt}(G) \) by \( H^2(x) = (L_{e_1}^{-1}(f_2(x)_1), L_{e_2-e_1}(f_2(x)_2)) \).
2. **For** \( k \geq 3 \), **construct the maps** \( f_3, f_4, \ldots, f_k \) inductively as follows.
   For \( i = 3 \) to \( k - 1 \), construct the maps \( f_{i+1} : G \rightarrow Y_{i+1} \) from \( f_i \) using the procedure given at the beginning of this subsection.
3. **Having obtained the map** \( f_k : G \rightarrow \text{Opt}'(G) \) from the preceding step, **define the map** \( H^k : G \rightarrow \text{Opt}(G) \) by

\[
H^k(x) = (H^k(x)_1, H^k(x)_2, \ldots, H^k(x)_k),
\]

where \( H^k(x)_j = L_{e_j-e_{j-1}}^{-1}(f_k(x)_j) \) for \( 1 \leq j \leq k \), taking \( e_0 = 0 \), and where \( L_{e_j-e_{j-1}} \) is the labeling from Corollary 4.8.

6 **The dilation bound**

From Lemma 5.1(e3) we have the containment result \( H^k(G) \subseteq \text{Opt}(G) \). The goal in this section is to complete the proof of our main result by showing that \( \text{dilation}(H^k) \leq 3k \) when every \( a_i \) is larger than some fixed constant.

We recall some notation. For \( x \in \{Y_{i-1}\} \), recall that \( \text{Stack}_i(x,r) = \{ z = \sigma_{i-1}(y) : z_{1-i-1} = x \text{ and } y \in S_{i-1}^r(G) \} \), that \( \text{Stack}_i'(x,r) = f_i(D_{i-1}^r) \cap \text{Stack}_i(x,r) \), and that \( [r]_i = \max\{|\text{Stack}_i(x,r)| : x \in \{Y_{i-1}\}\} \). So by definition we have \( \text{Stack}_i'(x,r) \subseteq \text{Stack}_i(x,r) \).

To set the context for the next theorem, note that by Lemma 5.1(e1,e2), \( f_i(D_{i-1}^r) \subseteq Y_i^{[r]} \), so that \( |\text{Stack}_i'(x,r)| \leq [r]_i \). In the next theorem we will see that \( \text{Stack}_i'(x,r) \) is always an initial substack of \( \text{Stack}_i(x,r) \); equivalently, that \( \text{Stack}_i'(x,P_{i-1}) \) is page monotone. Also we will see that the “page stack” heights \( |\text{Stack}_i'(x,r)|, x \in \{Y_{i-1}\}, \) fall within a narrow range; for a given \( r \) (and fixed \( i \)) any two such heights differ by at most \( 2 \) independent of \( x \).

For a subset \( S \subseteq Y_i^{[r]} \), let \( v_{i,r}(S) = |S \cap f_i(D_{i-1}^r)| \).

**Theorem 6.1** Let \( i \geq 3 \) and \( 1 \leq r \leq P_{i-1} \).

(a) \textit{(page monotonicity)} Take \( x \in \{Y_{i-1}\} \). Let \( z', z'' \in \text{Stack}_i(x,P_{i-1}) \) with \( f_i^{-1}(z') \in D_{i-1}^r \) and \( f_i^{-1}(z') \in D_{i-1}^r \). If \( z' > z'' \), then \( s \geq t \).

(b) For every stack address \( x \in \{Y_{i-1}\} \), we have \( [r]_i - 2 \leq |\text{Stack}_i'(x,r)| \leq [r]_i \). Moreover, all points of \( Y_i^{[r]} - f_i(D_{i-1}^r) \) lie in the union \( Y_i^{[r]} \cup Y_i^{[r]-1} \) of the top two \((i-1)\)-levels of \( Y_i^{[r]} \).

(c) \( v_{i,r}(Y_i^{[r]-1}) + v_{i,r}(Y_i^{[r]}) > 2^{e_i-1} \).
Proof. Consider part (a). It suffices to show that for any \(x \in \langle Y_{i-1}\rangle\) and \(1 \leq r \leq P_{i-1}\), we have that \(\operatorname{Stack}_i(x, r)\) is page monotone. We prove this by induction on \(r\), for any fixed \(i \geq 3\) and \(x \in \langle Y_{i-1}\rangle\). For the base case \(r = 1\), the claim is trivial since \(\operatorname{Stack}_i(x, 1)\) contains a single entry by one to oneness of \(\bar{\sigma}_{i-1}\) (Lemma 5.1) on any one section (in this case, on \(S_{i-1}^1\)). So suppose inductively that \(\operatorname{Stack}_i(x, r)\) is page monotone for some \(1 \leq r < P_{i-1}\). We use Lemma 5.1, with \(i - 1\) in place of \(i\). Applying \(\sigma_{i-1}\) to the second containment \(S_{i-1}^{(r)}(G) \subseteq I_{i-1} \circ f_i \circ f_{i-1}(D_{i-1}^{(r+1)})\) stated there and noting that \(f_i = \sigma_{i-1} \circ I_{i-1} \circ f_{i-1}\), we see that every entry \(z\) of \(\operatorname{Stack}_i(x, r)\) satisfies \(f_i^{-1}(z) \in D_{i-1}^{(r+1)}\). If \(\operatorname{Stack}_i(x, r + 1) = \operatorname{Stack}_i(x, r)\), then trivially \(\operatorname{Stack}_i(x, r + 1)\) is page monotone by induction. So assume \(\operatorname{Stack}_i(x, r + 1) \not= \operatorname{Stack}_i(x, r)\), and let \(y\) be the unique element (by one to oneness of \(\bar{\sigma}_{i-1}\) on any section) of \(\operatorname{Stack}_i(x, r + 1) - \operatorname{Stack}_i(x, r)\). By the first containment in Lemma 5.1 (again with \(i - 1\) replacing \(i\)) we have, on applying \(\sigma_{i-1}\) to each side again, \(f_i^{-1}(y) \in D_{i-1}^j\) for some \(j \geq r + 1\). Now \(\operatorname{Stack}_i(x, r)\) is page monotone by induction and as just noted \(f_i^{-1}(z) \in D_{i-1}^{(r+1)}\) for all \(z \in \operatorname{Stack}_i(x, r)\). Since \(\operatorname{Stack}_i(x, r + 1)\) is obtained by placing \(y\) at the top of \(\operatorname{Stack}_i(x, r)\) and \(y \in f_i(D_{i-1}^j)\) for some \(j \geq r + 1\), it follows that \(\operatorname{Stack}_i(x, r + 1)\) is page monotone, completing the inductive step.

For part (b), the upper bound \(|\operatorname{Stack}_i'(x, r)\| \leq |[r]|\) follows from \(\operatorname{Stack}_i'(x, r) \subseteq \operatorname{Stack}_i(x, r)\) and the definition of \([r]\).

Consider the lower bound on \(|\operatorname{Stack}_i'(x, r)\|\) in part (b). The statement holds vacuously for \(r = 1\) and all \(i \geq 3\), since \(\operatorname{Stack}_i'(x, 1)\) is 0 or 1 depending on \(x\), and \([1]_i = 1\), by one to oneness of the projection map \(\bar{\sigma}_{i-1}\) on any section from Lemma 5.1. So suppose \(r \geq 2\).

Fixing \(i\), for any stack address \(x \in \langle Y_{i-1}\rangle\), let \(s(x, r) = |\operatorname{Stack}_i(x, r)\|\) and \(s'(x, r) = |\operatorname{Stack}_i'(x, r)\|\). By Lemma 5.1 we have \(S_{i-1}^{(r)}(G) \subseteq I_{i-1} \circ f_i \circ f_{i-1}(D_{i-1}^{(r)})\). Now applying \(\sigma_{i-1}\) to both sides of this containment and using \(\operatorname{Stack}_i(x, r - 1) \subseteq \operatorname{Stack}_i(x, r)\) and \(f_i = \sigma_{i-1} \circ I_{i-1} \circ f_{i-1}\), we have \(\operatorname{Stack}_i(x, r - 1) = [\sigma_{i-1}(S_{i-1}^{(r)}(G)) \cap \operatorname{Stack}_i(x, r - 1)] \subseteq [f_i(D_{i-1}^{(r)}) \cap \operatorname{Stack}_i(x, r)] = \operatorname{Stack}_i'(x, r)\). Thus \(s'(x, r) \geq s(x, r - 1)\). Since \(r - 1 \leq P_{i-1} - 1\), by Lemma 5.1 we have \(s(x, r - 1) = [r - 1]_i\) or \([r - 1]_i - 1\). Therefore \(s'(x, r) \geq [r - 1]_i - 1\). By Lemma 5.1 we have \([r - 1]_i \leq [r]_i \leq [r - 1]_i + 1\). Thus \(s'(x, r) \geq [r - 1]_i - 1 \geq [r]_i - 2\), proving the first sentence of (b).

The second sentence of (b) follows from the first sentence, together with the page monotonicity property of part (a).

Next consider (c). By Lemma 5.1 we have \(|Y_{i-1}^{(r-i)}_i| - |f_i(D_{i-1}^{(r)})| < 2^{e_i - 1}\). Let \(A = [Y_{i-1}^{(r-i)}_i] \cup Y_{i-1}^{(r-i)-1}_i\) and \(B = v_{i,r}(Y_{i-1}^{(r-i)-1}_i) + v_{i,r}(Y_{i-1}^{(r-i)}_i)\). By (b) we have \(Y_{i-1}^{(r-i)-2} \subseteq f_i(D_{i-1}^{(r)})\), so \(|Y_{i-1}^{(r-i)}_i| - |f_i(D_{i-1}^{(r-i)})| = A - B\). So \(2^{e_i - 1} > |Y_{i-1}^{(r-i)}_i| - |f_i(D_{i-1}^{(r)})| = A - B = 2 \cdot 2^{e_i - 1} - B\), so \(B > 2^{e_i - 1}\), as required.

We introduce notation for identifying particular \((i-1)\)-subpages of a given \(i\)-page in \(G\). For \(1 \leq j \leq a_i\) let \(D_i^j(j) = D_i^{(r-i)+j}\), and we regard \(D_i^j(j)\) as the \(j\)th \((i-1)\)-subpage of \(D_i^r\) under the ordering of \((i-1)\)-subpages of \(D_i^r\) induced by \(<_{i-1}\). Similarly let \(S_i^j(j)\), \(j \geq 1\), be the \(j\)th nonblank \((i-1)\)-level of \(S_i^1(G)\), ordered by increasing \(i\)-coordinate.

Now suppose \(z = (I_i \circ f_i(x))\) for some \(x \in D_i^r\). We let \(\nu_i(z)\) be the integer such that \(z \in S_i^j(\nu_i(z))\) for suitable \(r\); that is, \(z\) belongs to the \(\nu_i(z)\)'th nonblank \((i-1)\)-level, ordered by increasing \(i\)-coordinate, of the \(i\)-section \(S_i^j\) which contains \(z\). To illustrate, recall the examples at the right of Figure 5 and the left of Figure 7. In Figure 5 we have the first 8 many 2-levels of \(Y_3\) containing the image \(f_3(G)\), where \(G = [3 \times 7 \times 4 \times 3]\). In Figure 7 at left we have inserted 4
blank 2-levels among these (as specified by the matrix in Table 1c), and grouped the resulting 12 many 2-levels into the three 3-sections $S^j_3$, $1 \leq j \leq 3$, preserving the order of the nonblank levels. So for example the 6th 2-level in Figure 3 (ordered by 3’rd coordinate, or height in the figure) becomes, after the map $I_3$ inserts these 4 blank 2-levels, the 3’rd nonblank 2-level $S^2_3(3)$ of section $S^2_3$ (again ordered by 3’rd coordinate, or height within $S^2_3$) in Figure 7. So any point $z = (I_3 \circ f_3)(x)$, where $f_3(x)$ was in the 6th 2-level of Figure 5, satisfies $\nu_3(z) = 3$ since $z$ belongs to the third nonblank 2-level of the 3-section (in this case $S^2_3$) containing $z$ (as shown in Figure 7). Similarly take any point $f_3(x)$ lying in the 7th level of Figure 5 Then the corresponding point $z = (I_3 \circ f_3)(x)$ after inflation satisfies $\nu_3(z) = 1$, since $z$ belongs to the first nonblank 2-level $S^3_3(1)$ of the 3-section $S^3_3$ containing $z$ (as shown in Figure 7).

Let $m_r$ be the number of nonblank $(i - 1)$-levels in $S^j_i$ for any $1 \leq r \leq P_i$. We interpret the $(i - 1)$-levels $S^j_i(j)$ for integers $j$ outside the range $[1, m_r]$, and also interpret the differences $\nu_i(z') - \nu_i(z'')$ using “wraparound” as follows. For example, if $j \leq 0$, then $S^j_i(j)$ is understood as $S^j_i(1)$.

**Corollary 6.2** For $2 \leq i \leq k$ we have the following.

(a) For each $i \geq 2$ and $j \geq 1$, we have $(I_i \circ f_i)(D^j_i) \subseteq S^{j-1}_i \cup S^j_i$. Hence for $1 \leq r \leq P_i - 1$ we have the following.

(1) For any stack address $x \in (Y_{i-1})$ and $1 \leq r \leq P_{i-1}$, $|\text{Stack}_i(x, r) \cap f_i(D^j_{i-1})| \leq 2$.

(2) Suppose $|\text{Stack}_i(x, r) \cap f_i(D^j_{i-1})| = 2$, and let $\text{Stack}_i(x, r) \cap f_i(D^j_{i-1}) = \{z', z''\}$. Then $z'$ and $z''$ are at successive heights in Stack$_i(x, r)$; that is, $|z'_r - z''_r| \leq 1$.

(3) Let $z \in D^j_{i-1}$, with $f_i(z) \in \text{Stack}_i(x, r)$ for some $x \in (Y_{i-1})$. Then $[r]_i - 2 \leq f_i(z) \leq [r]_i$.

(b) For $x, y \in G$, suppose $x \in D^j_i(q)$ and $y \in D^j_i(q)$ for some $1 \leq q \leq a_i$ and $1 \leq r, s \leq P_i$. Let $z' = (I_i \circ f_i)(x)$ and $z'' = (I_i \circ f_i)(y)$. Then $|\nu_i(z') - \nu_i(z'')| \leq 3$.

(c) For $x, y \in G$, let $I_i(f_i(x)) \in S^j_i(G)$ and $I_i(f_i(y)) \in S^j_i(G)$, $1 \leq s \leq t \leq P_i$.

(1) If $s = t$, then $|f_{i+1}(x)_{i+1} - f_{i+1}(y)_{i+1}| \leq 1$.

(2) If $|s - t| = 1$, then $|f_{i+1}(x)_{i+1} - f_{i+1}(y)_{i+1}| \leq 2$.

(d) Suppose $x \in D^j_i$, $y \in D^j_i$, $1 \leq s \leq t \leq P_i$.

(1) If $s = t$, then $|f_{i+1}(x)_{i+1} - f_{i+1}(y)_{i+1}| \leq 2$.

(2) If $|s - t| = 1$, then $|f_{i+1}(x)_{i+1} - f_{i+1}(y)_{i+1}| \leq 3$.

**Proof.** Consider (a). Since $(I_i \circ f_i)(D^j_i) \subseteq D^j_i$ by Lemma 5.11, it suffices to show that $(I_i \circ f_i)(D^j_i) \cap S^{j-2}_i = \emptyset$. Recall that $D^{j-1}_i = D^{a_i(j-1)}$. By Lemma 5.11 we have $[a_i(j-1)]_i = \left[\frac{|D^{a_i(j-1)}|}{2^{a_i-1}}\right] = \left[\frac{2^{a_i-1} - 1}{2^{a_i-1}}\right]$. So by Lemma 4.1 we see that $[a_i(j-1)]_i$ is the number of nonblank $(i - 1)$-levels in $S^{j-1}_i$. Since $I_i$ preserves the order (by increasing $i$ coordinate) of $(i - 1)$-levels, it follows that $I_i(Y_{i}^{[a_i(j-1),j]}_i) = S^{j-1}_i(G)$. By Theorem 6.1b, only the top two $(i - 1)$-levels of $Y_{i}^{[a_i(j-1),j]}_i$ may contain points not in $f_i(D^{a_i(j-1)}_i) = f_i(D^{j-1}_i)$. Again since $I_i$ preserves the order of $(i - 1)$-levels, it follows that only the top two nonblank $(i - 1)$-levels of $I_i(Y_{i}^{[a_i(j-1),j]}_i) = S^{j-1}_i$, call them $Y^{c}_i$ and $Y^{d}_i$, may contain points not in $(I_i \circ f_i)(D^{j-1}_i)$. Thus $(I_i \circ f_i)(D^j_i) \cap S^{j-1}_i \subset (Y^{c}_i \cup Y^{d}_i)$. But since $a_i \geq 2^{a_i} > 4$ for all $i$, these top two nonblank
(i−1)-levels \( Y_i^c \), \( Y_i^d \) of \( S_i^{(j−1)} \) lie in the last \( i \)-section \( S_i^{j−1} \) of \( S_i^{(j−1)} \). So \( (I_i \circ f_i)(D_i^r) \cap S_i^{(j−2)} = \emptyset \) as desired.

We now consider the consequences (a1)-(a3) of (a), starting with (a1). By (a) we have \((I_{i−1} \circ f_{i−1})(D^r_{i−1}) \subset S_{i−1}^r \cup S_i^r \). Thus \( f_i(D^r_{i−1}) \subset \sigma_i−1(S_{i−1}^{(r−1)}(G)) \cup \sigma_i−1(S_{i−1}^r(G)) \). Also the projection \( \sigma_i−1 \) is one to one when restricted to any single \((i−1)\)-section, in particular to \( S_{i−1}^r \) and to \( S_{i−1}^r \). Thus \( f_i(D^r_{i−1}) \) can contribute at most two points to any single stack \( \text{Stack}_i(x, r) \); namely, up to one point from each of \( S_{i−1}^{(r−1)}(G) \) and \( S_{i−1}^r(G) \), proving (a1).

Part (a2) follows immediately from page monotonicity of \( \text{Stack}_i^t(x, r) \) (Theorem 6.1a).

Consider (a3). For the upper bound note that \((I_{i−1} \circ f_{i−1})(D^r_{i−1}) \subset S_i^{(r−1)} \) by part (a). Since \( z \in D^r_{i−1} \), we have \( f_i(z) \leq \max\{\sigma_i−1(u) : u \in S_{i−1}^{(r−1)}(G)\} = [r]_i \) by definition of \([r]_i \). Consider now the lower bound in (a3). If \( f_i(z) \) is the highest point in \( \text{Stack}_i^t(x, r) \), then (a3) follows immediately from Theorem 6.1b. Otherwise, by part (a2) we have that \( f_i(z) \) is the second highest point in \( \text{Stack}_i^t(x, r) \). Let the highest such point be \( f_i(y), y \in D^r_{i−1}, \) so that \( f_i(y) − f_i(z) = 1 \) again by part (a2). As in the proof of (a1), we must have \((I_{i−1} \circ f_{i−1})(z) \in S_{i−1}^{(r−1)} \) and \((I_{i−1} \circ f_{i−1})(y) \in S_i^{(r−1)} \). Thus by stack monotonicity (Lemma 5.1g) \( f_i(z) \) is the highest point in \( \text{Stack}_i^t(x, r−1) \), so \( f_i(z) = |\text{Stack}_i^t(x, r−1)| \). Since \( r−1 < P_{i−1}, \) by Lemma 5.1d2 we have \( f_i(z) = |\text{Stack}_i^t(x, r−1)| \geq [r−1]_i−1 \). It remains to check that \([r−1]_i−1 ≥ [r]_i−2 \), which follows directly from the formula \([r]_i = \left[ \frac{a_i−1}{2} \right] \) proved in Lemma 5.1d.

For (b), recall that \( D_i^r(q) = D_i^{(r−1)a_i+q} \), with the same formula for \( D_i^r(q) \) where \( r \) is replaced by \( s \). Let \( d_i = a_ia_{i−1} \ldots a_1 \). By part (a) we have \((I_{i−1} \circ f_{i−1})(D_i^{(r−1)a_i+q}) \subset S_i^{(r−1)a_i+q−1} \cup S_i^{(r−1)a_i+q} \). Hence using Lemma 5.1d and the upper bound in part (a3), we have \( f_i(x) \leq [(r−1)a_i+q]_i = \frac{[(r−1)a_i+q]_i}{2^{r−1}} \). Now by Lemma 4.1a and our construction we know that \( S_i^{(r−1)} \) has \( [\frac{(r−1)a_i+q}{2^{r−1}}]_i \) nonblank \((i−1)\)-levels. Thus the first \( [\frac{(r−1)a_i+q}{2^{r−1}}] \) nonblank blank \((i−1)\)-levels in the image \( I_i(Y_i^c) \) belong to \( S_i^{(r−1)} \). Hence \( \nu_i(z') \leq \frac{[(r−1)a_i+q]_i}{2^{r−1}} \). Similarly, this time using Lemma 5.1d and the lower bound in (a3), we have \( \nu_i(z'') \geq \frac{[(r−1)a_i+q]_i}{2^{r−1}} − 2\cdot \frac{[(r−1)a_i+q]_i}{2^{r−1}} \). It follows that \( \nu_i(z') − \nu_i(z'') \leq 3 \). A symmetric argument interchanging the roles of \( r \) and \( s \) yields \( \nu_i(z') − \nu_i(z'') \geq −3 \), completing (c).

Consider part (c), starting with (c1). Let \( x' \in Y_i^c \) and \( y' \in Y_i^d \) be the stack addresses of the images \( f_{i+1}(x) \) and \( f_{i+1}(y) \): that is \( x' = f_{i+1}(x)_1 \ldots, \) and \( y' = f_{i+1}(y)_1 \ldots \). Since \( s−1 \leq P_{i−1}, \) by Lemma 5.1a, we have \([s−1]_{i+1} = 1 \leq |\text{Stack}_{i+1}(x', s−1)| \). Since \( I_i(f_i(x)) \in S_i^t(G) \), we have \( f_{i+1}(x) = \sigma_i(I_i(f_i(x))) \in \text{Stack}_{i+1}(x', s) \). By Lemma 5.1a \( f_{i+1}(x) \) is the only member of \( \text{Stack}_{i+1}(x', s) \), and thus \( |\text{Stack}_{i+1}(x', s)| \leq |\text{Stack}_{i+1}(x', s)|+1 \). It follows that \([s−1]_{i+1} \leq f_{i+1}(x)_i \leq [s−1]_{i+1} \), with the same inequality holding for \( f_{i+1}(y)_{i+1} \) and \( [s−1]_{i+1} \). Part (c1) follows.

For (c2), suppose without loss of generality that \( s = t−1 \). Again using Lemma 5.1c, we get \([s]_{i+1} \leq [t]_{i+1} \leq [s]_{i+1} \). Since \( I_i(f_i(x)) \in S_i^t(G) \) we have \( f_{i+1}(y)_{i+1} \leq [t]_{i+1} \) by definition of \([t]_{i+1} \). Again since \( I_i(f_i(x)) \in S_i^t(G) \) and using Lemma 5.1d2 we have \( f_{i+1}(y)_{i+1} = 1 + |\text{Stack}_{i+1}(y')| \geq 1 + [s]_{i+1} = 1 + [s]_{i+1} \). Hence we have \([s]_{i+1} \leq f_{i+1}(y)_{i+1} \leq [s]_{i+1} \). Since \( s ≤ P_{i−1} \) and \( f_{i+1}(x) \) is the topmost entry of \( \text{Stack}_{i+1}(x', s) \), it follows by Lemma 5.1d2 that \([s]_{i+1} \leq f_{i+1}(x)_{i+1} \leq [s]_{i+1} \). Part (c2) follows.

For (d), set \( \alpha = I_i(f_i(x)) \) and \( \beta = I_i(f_i(y)) \), and consider first (d1). By part (a), \( \alpha \) and \( \beta \) lie in the same \( i \)-section or in successive \( i \)-sections \( S_i^{s−1} \), \( S_i^s \) of \( Y_i \). Hence (d1) follows directly from part (c). For (d2), suppose \( s = t−1 \). Again by part (a), \( \alpha \in S_i^{s−1} \cup S_i^s = S_i^{t−2} \cup S_i^{t−1} \),
and $\beta \in S_i^{t-1} \cup S_i^t$. If $\alpha$ and $\beta$ belong to successive $i$-sections (e.g. $S_i^{t-2}$ and $S_i^{t-1}$, or $S_i^{t-1}$ and $S_i^t$ respectively), then $|f_{i+1}(x)_{i+1} - f_{i+1}(y)_{i+1}| \leq 2$ by part (c). So we can suppose that $\alpha \in S_i^{t-2}$ and $\beta \in S_i^t$. Since the projection $\sigma_j$ is one to one on any one $i$-section (Lemma 5.1d), we have $[\sigma]_{i+1} \leq [t-2]_{i+1} + 2$. Since $t-1 \leq P_i - 1$ we have $f_{i+1}(x)_{i+1} \geq [t-2]_{i+1} - 1$ by Lemma 5.1i, while $f_{i+1}(y)_{i+1} \leq [t]_{i+1}$ since $\beta \in S_i^t$. Thus $|f_{i+1}(x)_{i+1} - f_{i+1}(y)_{i+1}| \leq 3$, completing (d2).

We now proceed to a bound on the dilation of our embedding. In the proof we frequently apply Corollary 4.5, where (in the language of its statement) we use $N_u(d)$ relative to the matrix $F^X = (f_{uv})$ with the settings in part (d) of that Corollary. By the construction given immediately after that Corollary, we have $F^X = F(i)$. The value of $i$ will change from one application to another. So in the proof which follows, let $N_{i,u}(d)$ denote the $N_u(d)$ of the corollary, when in the application we intend to use $F^X = F(i)$. So for example, in applying Corollary 4.5 with $e = 3$ and $F^X = F(2)$, we would conclude that $|N_{2,u}(d+3) - N_{2,u}(d)| \leq 2 \cdot 3 + 2 = 8$, using the more generous of the bounds (a) and (b).

**Theorem 6.3** Let $G = [a_1 \times a_2 \cdots \times a_k]$ be a $k$-dimensional grid with $a_i \geq 2^{22}$ for each $i$. Then the embedding $H^k : G \to \text{Opt}(G)$ from section 5.2 satisfies $	ext{dilation}(H^k) \leq 3k$.

**Proof.** Let $x = (x_1, x_2, \ldots, x_k)$ and $y = (y_1, y_2, \ldots, y_k)$ be arbitrary adjacent points of $G$, and set $Q = \text{Opt}(G)$ and $f = f_k$ for short. Recall that

$$H^k(x) = (H^k(x_1), H^k(x_2), \ldots, H^k(x_j), \ldots, H^k(x_k)) \in Q,$$

where $H^k(x_j) = L_{e_j-e_{j-1}}^{-1}(f(x_j))$ is the (0,1) string of $Q_{e_j-e_{j-1}}$ equivalent to $f(x_j)$ under the labeling $L_{e_j-e_{j-1}}$ of Corollary 4.8. It suffices to show for each $j$, $1 \leq j \leq k$, that $|f(x_j) - f(y_j)| \leq 17$. Then by the definition of $H^k$ just given and Corollary 4.8b we have $\text{dist}_{Q_{e_j-e_{j-1}}}(H^k(x_j), H^k(y_j)) \leq 3$ for each $1 \leq j \leq k$.

So we obtain $\text{dist}_Q(H^k(x), H^k(y)) = \sum_{j=1}^{k} \text{dist}_{Q_{e_j-e_{j-1}}}(H^k(x_j), H^k(y_j)) \leq 3k$, and our desired dilation bound follows. Thus we are reduced to showing that $|f(x_j) - f(y_j)| \leq 17$ for each $j$.

Let $i_0 := j_0(x,y)$, $1 \leq i_0 \leq k$, be the unique coordinate at which the above adjacent points $x, y \in V(G)$ disagree, so $|x_{i_0} - y_{i_0}| = 1$. Also let $L(j, i_0) = \max\{|f(x_j) - f(y_j)| : xy \in E(G), |x_{i_0} - y_{i_0}| = 1\}$. For any given $j$, the maximum of $L(j, i_0)$ over all $i_0$ serves as an upper bound for $|f(x_j) - f(y_j)|$, over all edges $xy \in E(G)$. So it suffices to show that for each $1 \leq j \leq k$ this maximum is at most 17. Let $x(y)(j)$ be an edge at which this maximum occurs for a given $j$, and refer to this edge simply as $xy$ since $j$ will be understood by context.

We begin with $j = 1$, starting with the case $i_0 = 1$ for a bound on $(1,1)$. Since $i_0 = 1$, $x$ and $y$ are corresponding points on successive chains $C_i$ and $C_{i+1}$ of $G(a_1)$. So we have $L(1,1) = |f(x_1) - f(y_1)| = |f_2(x_1) - f_2(y_1)| \leq 3 < 17$ as required, where the second equality follows from Lemma 5.1a and the inequality following from Corollary 2.3a. Now suppose $i_0 \geq 2$. Thus $x$ and $y$ agree in their first coordinate, so can be considered as lying on the same chain of $G(a_1)$. So similarly by Corollary 5.1b and Corollary 2.3b we have $L(1, i_0) = |f(x_1) - f(y_1)| = |f_2(x_1) - f_2(y_1)| \leq 2 < 17$. So $\max\{L(1, i_0) : 1 \leq i_0 \leq k\} \leq 17$.

Now suppose that $j = 2$. If $i_0 = 1$, then again $x$ and $y$ are corresponding points on successive chains of $G(a_1)$. Thus $|f_2(x_2) - f_2(y_2)| \leq 1$ by Corollary 2.2g. Hence by Lemma 5.1b (with $e = 1$) we get $L(2,1) = |f(x_2) - f(y_2)| \leq 4 < 17$.

If $i_0 = 2$, then $x$ and $y$ are successive points on the same chain of $G(a_1)$, so by Corollary 2.3b we have $|f_2(x_2) - f_2(y_2)| \leq 1$. So again by Lemma 5.1b, $L(2, 2) = |f(x_2) - f(y_2)| \leq 4 < 17$.
Now suppose \( i_0 \geq 3 \), still with \( j = 2 \). Then \( x \) and \( y \) belong to the same chain \( C_i \) of \( G(a_1) \). Since \( x_1 = y_1 \), \( x_2 = y_2 \), and \( i_0 \geq 3 \), we can view \( x \) and \( y \) as corresponding points belonging to a pair of distinct 2-pages of \( G \), say \( x \in D_w^u \) and \( y \in D_w^v \). Consider the segments \( T_1 \) and \( T_2 \) of \( C_i \) given by \( T_1 = \{ (i, (u-1)\alpha_0 + t) : 1 \leq t \leq x_2 \} \) and \( T_2 = \{ (i, (v-1)\alpha_0 + t) : 1 \leq t \leq x_2 \} \). We see that \( f_2(T_1) \) (resp. \( f_2(T_2) \)) is the initial segment of \( x_2 \) points of the chain \( C_i \) in \( D_2^w \) (resp. \( D_2^v \)), and that \( f_2(x) \) and \( f_2(y) \) are the last points of these segments respectively. Let \( c_1 \) and \( c_2 \) be the number of successive columns of \( Y_2 \) spanned by \( f_2(T_1) \) and \( f_2(T_2) \) respectively.

By Corollary 2.3d we have \(|c_1 - c_2| \leq 1\). The columns spanned by \( f_2(T_1) \) (resp. \( f_2(T_2) \)) are transformed under the inflation map \( I_2 \) into a corresponding set of \( c_1 \) (resp. \( c_2 \)) successive nonblank columns in \((I_2 \circ f_2)(G)\). Corollary 2.3d and Lemma 4.1 now imply that \((I_2 \circ f_2)(T_1)\) and \((I_2 \circ f_2)(T_2)\), being transforms of \( f_2(T_1) \) and \( f_2(T_2) \) respectively, each begin in either the first nonblank column of their section \((S_{2}^u \circ f_2)\) for \((I_2 \circ f_2)(T_1)\) and \((S_{2}^v \circ f_2)\) for \((I_2 \circ f_2)(T_2)\) or the last nonblank column of their preceding section \((S_{2}^{u-1} \circ f_2)\) for \((I_2 \circ f_2)(T_1)\) and \((S_{2}^{v-1} \circ f_2)\) for \((I_2 \circ f_2)(T_2)\).

Now let \( z' = (I_2 \circ f_2)(x) \) and \( z'' = (I_2 \circ f_2)(y) \), and assume by symmetry that \( c_2 \geq 1 \). Then \(|c_1 - c_2| \leq 1\) implies that \( \nu_2(z'') - \nu_2(z') = 0, 1 \) or 2. Assume first that \( \nu_2(z'') - \nu_2(z') = 2 \), and write \( d = \nu_2(z') \) and \( d + 2 = \nu_2(z'') \). This case arises when \( c_2 - c_1 = 1 \), and \((I_2 \circ f_2)(T_2)\) begins in the first nonblank column of \( S_{2}^u \) while \((I_2 \circ f_2)(T_1)\) begins in the last nonblank column of \( S_{2}^{u-1} \). Then by Lemma 5.1a and Corollary 4.5c: with \( e = 2 \), we get \(|f(x)_2 - f(y)_2| = |f_3(x)_2 - f_3(y)_2| = |N_{2,u}(\nu_2(z')) - N_{2,v}(\nu_2(z''))| = |N_{2,u}(d) - N_{2,v}(d + 2)| \leq 2 \cdot 2 + 4 = 8 \).

In the case \( \nu_2(z'') - \nu_2(z') = 1 \), we similarly obtain (using \( e = 1 \) in Corollary 4.5c) that \(|f(x)_2 - f(y)_2| \leq 6 \), and in the case \( \nu_2(z'') - \nu_2(z') = 0 \) (using \( e = 0 \) we get \(|f(x)_2 - f(y)_2| \leq 4 \).

So we get \( L(2, i_0) \leq 8 < 17 \) for \( i_0 \geq 3 \), and overall \( \max\{L(2, i_0) : 1 \leq i_0 \leq k\} \leq 17 \).

Next suppose \( j \geq 3 \). We argue according to the order relation between \( j \) and \( i_0 \).

Suppose first that \( j > i_0 \). Then \( x \) and \( y \) belong to the same \((j - 1)\)-page and successive \((i_0 - 1)\)-pages of \( G \). We obtain our bound on \( L(j, i_0) \) independent of \( i_0 \). Since \( x \) and \( y \) belong to the same \((j - 1)\)-page we have that \( I_{j-1}(f_{j-1}(x)) \) and \( I_{j-1}(f_{j-1}(y)) \) belong to the same \((j - 1)\)-section or to successive \((j - 1)\)-sections \( S_{j-1}^u \) and \( S_{j-1}^v \), \(|t - t'| \leq 1\), by Corollary 6.2a.

Applying Corollary 6.2b, in the first case (same \((j - 1)\)-section) we get \(|f_j(x)_j - f_j(y)_j| \leq 1\) by part (c1), while in the second case (successive \((j - 1)\)-sections) we get \(|f_j(x)_j - f_j(y)_j| \leq 2\) by part (c2). Now we bound \( L(j, i_0) \) by applying Lemma 5.1b (with \( e = 2 \)) to obtain \( L(j, i_0) = |f(x)_j - f(y)_j| \leq 6 < 17 \), as required.

Now suppose \( j = i_0 \). Then \( x \) and \( y \) are corresponding points in successive \((j - 1)\)-pages of \( G \). Hence by Corollary 6.2b we get \(|f_j(x)_j - f_j(y)_j| \leq 3\). Thus by Lemma 5.1b with \( e = 3 \) we get \( L(j, j) = |f(x)_j - f(y)_j| \leq 8 < 17 \).

It remains to consider the case \( 3 \leq j < i_0 \). Here \( x \) and \( y \) are corresponding points in successive \((i_0 - 1)\)-pages, say \( x \in D_{i_0-1}^r \) and \( y \in D_{i_0-1}^s \). Consider the \( j \)-pages containing \( x \) and \( y \), say \( x \in D_j^c \subset D_{i_0-1}^r \) and \( y \in D_j^d \subset D_{i_0-1}^s \) for suitable integers \( c \) and \( d \). Since \( x \) and \( y \) agree in their first \( j \) coordinates, they must also be corresponding points in these \( j \)-pages. In particular, \( x \) and \( y \) belong to corresponding \((j - 1)\)-subpages \( D_j^c(q) \) of \( D_j^c \) and \( D_j^d(q) \) of \( D_j^d \) respectively for the same integer \( q \).

As before, let \( z' = (I_j \circ f_j)(x) \) and \( z'' = (I_j \circ f_j)(y) \) for brevity. Then by Corollary 6.2b we get \(|\nu_j(z') - \nu_j(z'')| \leq 3\). Again applying Lemma 5.1a we have \( f(x)_j = N_{j,c}(\nu_j(z')) \) and \( f(y)_j = N_{j,d}(\nu_j(z'')) \). So applying Corollary 4.5c: with \( e = 3 \) we get \( L(j, i_0) = |f(x)_j - f(y)_j| \leq |N_{j,c}(\nu_j(z')) - N_{j,d}(\nu_j(z''))| \leq 2 \cdot 3 + 4 = 10 < 17 \). So for any \( j \geq 3 \) we have \( \max\{L(j, i_0) : 1 \leq i_0 \leq k\} \leq 17 \), completing the proof.
7 Concluding Remarks

1. There is a routing of edge congestion $O(k)$ associated to our embedding $H^k$. We outline the idea here, omitting full details of the proof.

As notation, for any graph $H$ and permutation $\pi : V(H) \to V(H)$, a $\pi$-routing is an assignment $P : V(H) \to \{\text{paths in } H\}$ such that $P(x)$ is a path in $H$ from $x$ to $\pi(x)$. If $H$ is directed, then $P(x)$ is a directed path from $x$ to $\pi(x)$. The congestion of a $\pi$-routing is the maximum of \( \{n(e) : e \in E(H)\} \), where $n(e)$ is the number of paths in the $\pi$-routing which use the edge $e$.

For background, let $Q_n^\rightarrow$ be the directed graph obtained from $Q_n$ by replacing each edge of $Q_n$ by 4 directed edges, two pointing in one direction and two in the opposite direction. Using the classic Benes routing method, one can show [14] that for any permutation $\pi$ of $Q_n^\rightarrow$ there exists a $\pi$-routing such that the paths of the routing are edge disjoint. Consequently, for each permutation $\pi$ of the undirected $Q_n$ there is a $\pi$-routing with congestion $O(1)$.

For an undirected graph $H$, suppose there is a partition of $E(H)$ into $k$ sets, $E(H) = \bigcup_{i=1}^{k} E_i$, such that each $E_i$ is a vertex disjoint union of edges in $H$, where the vertices of each cycle are ordered in one of the two natural ways. Further, let $g : H \to Q_n$ be a one to one map. Now consider the permutation $\pi_i$ on $Q_n$ as follows: if $v \notin g(H)$ then $\pi_i(v) = v$, while if $v \in g(H)$ then $\pi_i(v) = w$, where $g^{-1}(v)g^{-1}(w) \in E_i$ and $g^{-1}(w)$ follows $g^{-1}(v)$ in the ordering of the vertices of the cycle of $E_i$ containing $g^{-1}(v)$ and $g^{-1}(w)$. Then by the above result there is a $\pi_i$ routing on $Q_n$ of congestion $O(1)$ for each $i$, $1 \leq i \leq k$. Since each $e \in E(H)$ lies in some $E_i$, it would follow that the edge congestion of $g$ (as defined in section 1.1) is $O(k)$.

It therefore suffices to find a graph $G' \supset G = [a_1 \times a_2 \times \ldots \times a_k]$ with $V(G') = V(G)$ such that $G'$ has the required cycle partition of edges. For then (letting our map $H^k$ play the role of $g$ in the above paragraph) the map $H^k : G' \to \text{Opt}(G)$ has edge congestion $O(k)$ by the above argument, so the same is true of its restriction $H^k : G \to \text{Opt}(G)$.

The graph $G'$ is obtained from $G$ as follows. For each $i$, $1 \leq i \leq k$, consider a fixed $(k-1)$-tuple $\mathbf{c}(i) = (c_1, c_2, \ldots, c_{i-1}, c_{i+1}, \ldots, c_k)$ of integer entries, with $1 \leq c_j \leq a_j$ for $j \neq i$. Further, for $1 \leq t \leq a_i$ let $v(\mathbf{c}(i), t) = (c_1, c_2, \ldots, c_{i-1}, t, c_{i+1}, \ldots, c_k) \in V(G)$. We let $V(G') = V(G)$ and $E(G') = E(G) \cup E'$, where $E' = \{v(\mathbf{c}(i), a_i)v(\mathbf{c}(i), 1) : 1 \leq i \leq k, \mathbf{c}(i) \text{ any } (k-1)\text{-tuple as above}\}$. Thus $E'$ is just the set of “wraparound” edges not present in $G$ in each of the $k$ dimensions. The required cycle partition of $E(G')$ is given by $E(G') = \bigcup_{i=1}^{k} E'_i$, where $E'_i = \{v(\mathbf{c}(i), a_i)v(\mathbf{c}(i), 1), v(\mathbf{c}(i), t)v(\mathbf{c}(i), t+1) : 1 \leq t \leq a_i - 1, 1 \leq i \leq k, \mathbf{c}(i) \text{ any } (k-1)\text{-tuple as above}\}$.

2. The lower bound requirement $a_i > 2^{22}$ for our result can be relaxed to $a_i > 2^{112}$, provided one can improve the conclusion of Corollary 4.8a only slightly to say $|L_t(x) - L_t(y)| \leq 2r + 4 \Rightarrow \text{dist}_{Q_4}(x, y) \leq 3$. Then using $r = 3$ and $i = 2$ we obtain $|L_t(x) - L_t(y)| \leq 10 \Rightarrow \text{dist}_{Q_4}(x, y) \leq 3$ for $t \geq 12$, so that $a_i > 2^{112}$ suffices for our result. The proof of Theorem 6.3 is then reduced (as in its first paragraph) to proving the inequality $|f(x) - f(y)| \leq 10$ for each $1 \leq j \leq k$ and $xy \in E(G)$. The proof of Theorem 6.3 shows that this inequality does indeed hold. The improvement in Corollary 4.8a may require a detailed study of the regular cyclic caterpillars we used.

3. The question of finding good lower bounds for $B(G, \text{Opt}(G))$, for some class of multidimensional grids $G$ with $|V(G)| \to \infty$, remains open. A nontrivial lower bound for all multidimensional grids is of course not possible, since if the $a_i$ are all powers of 2, then $G$ is
a spanning subgraph of $Opt(G)$, so $B(G, Opt(G)) = 1$ in that case.

4. The first author thanks Stephen Wright for many useful discussions on rounding of matrices. We also acknowledge a useful discussion with Hal Kierstead which inspired our application of regular cyclic caterpillars.

References

[1] T. Asano, T. Matsui, and T. Tokuyama: Optimal Roundings of Sequences and Matrices, *Nordic J. Comput.* 7 (2000) 241-256.

[2] T. Asano, N. Katoh, K. Obokata, and T. Tokuyama: Matrix Rounding Under the $L_p$-Discrepancy Measure and Its Application to Digital Halftoning, *SIAM J. Comput.* 32 No.6 (2003) 1423-1435.

[3] U. Blass: Short Dominating Paths and Cycles in the Binary Hypercube: *Annals of Combinatorics* 5 No. 1 (2001) 51-59.

[4] Zs. Baranyai: On the Factorization of the Complete Uniform Hypergraph, in *Infinite and Finite Sets, Proc. Coll. Keszthely, 1975*, Colloquia Math. Soc. János Bolyai A. Hajnal, R. Rado, V.T. Sós (Eds.), Colloquia Math. Soc. János Bolyai 10 North-Holland (1975) 91-108.

[5] S. Bettayeb, Z. Miller, and I.H. Sudborough: Embedding Grids into Hypercubes, *J. Comput. System Sci.* 45 No. 3 (1992) 340-366.

[6] S. Bhatt, F.R.K. Chung, F.T. Leighton, and A. Rosenberg: Efficient Embeddings of Trees in Hypercubes, *SIAM J. Comput.* 21 No.1 (1992) 151-162.

[7] R. Caha and V. Koubek: Spanning Regular Caterpillars in Hypercubes, *European J. Combinatorics* 18 (1997) 249-266.

[8] Y. Cai, T.H. Lai, and C.C. Lin: Embedding Two-Dimensional Grids into Hypercubes with Dilation 2 and Congestion 5, Proc. Intern. Conf. on Parallel Processing, Oconomowoc, Wisc., III (1995) 105-112.

[9] M.Y. Chan: Embeddings of Grids into Optimal Hypercubes, *SIAM J. Comput.* 20 No. 5 (1991) 834-864.

[10] B. Doerr: Matrix Rounding with Respect to Small Submatrices, *Random Structures & Algorithms* 28 (2006) 107-112.

[11] B. Doerr, T. Friedrich, C. Klein, R. Osbild: Rounding of Sequences and Matrices, with Applications, *Proceedings of the Workshop on Approximation and Online Algorithms (WAOA 2005)* T. Erlebach, G. Persiano (Eds.) Vol. 3879 of Lecture Notes in Computer Science, Springer Verlag (2006) 96-109.
[12] B. Doerr, T. Friedrich, C. Klein, and R. Osbild: Unbiased Matrix Rounding, Proceedings of the 10th Scandinavian Workshop on Algorithm Theory (SWAT), L. Arge, R. Freivalds (Eds.) Vol. 4059 of Lecture Notes in Computer Science, Springer Verlag (2006) 102-112.

[13] T. Dvorak, I. Havel, J-M Laborde, and M. Mollard: Spanning Caterpillars of a Hypercube, Journal of Graph Theory 24 No. 1 (1997) 9-19.

[14] Q.-P. Gu and H. Tamaki: Routing a Permutation in the Hypercube by Two Sets of Edge Disjoint Paths, J. Parallel Distrib. Comput. 44(2) (1997) 147-152.

[15] I. Havel and P. Liebl: Embedding the Polytopic Tree into the n-cube, Casopis Pest. Mat. 98 (1973) 307-314.

[16] I. Havel and J. Moravec: B-Valuations of Graphs, Czech. Math. J. 22 (1972) 338-351.

[17] F. Harary, M. Lewinter, and W. Widulski: On Two-Legged Caterpillars which Span Hypercubes Congress. Numer. (19’th Southeastern Conference on Comb., Graph Theory, and Computing) 66 (1988) 103-108.

[18] Y.M. Kim and T.H. Lai: The Complexity of Congestion-1 Embedding in a Hypercube, J. of Algorithms 12 No. 2 (1991) 246-280.

[19] D.E. Knuth: Two-Way Rounding, SIAM J. Disc. Math 8 No. 2 (1995) 281-290.

[20] J. Lee: On the Efficient Simulation of Networks by Hypercube Machines, Ph.D. Dissertation, Dept. of Electrical Engineering and Computer Science, Northwestern University (1990).

[21] F.T. Leighton: Introduction to Parallel Algorithms and Architectures: Arrays, Trees, Hypercubes, Morgan Kaufman Publishers, San Mateo, CA (1992).

[22] M. Livingston and Q. Stout: Embeddings in Hypercubes, Mathematics and Computational Modelling 11 (1988) 222-227.

[23] L. Lovasz, J. Spencer, and K. Vendegombi: Discrepancy of Set Systems and Matrices, European J. of Combin. 7 (1986) 151-160.

[24] Z. Miller and I.H. Sudborough: Compressing Grids into small Hypercubes, Networks 24 No. 6 (1994) 327-357.

[25] A. L. Rosenberg, L. S. Heath: Graph Separators with Applications, Kluwer Academic/Plenum Publishers, New York, New York (2001).

[26] M. Rottger, U. Schroeder: Embedding 2-Dimensional Grids into Optimal Hypercubes with Edge Congestion 1 or 2, Parallel Processing Letters 8 No. 2 (1998) 231-242.

[27] J. Spencer: Ten Lectures on the Probabilistic Method: CBMS-NSF Regional Conference Series in Applied Mathematics 52 SIAM, Philadelphia (1987) 37-44.

[28] D.B. West: Introduction to Graph Theory (Second Edition) Prentice Hall (2001).
8 Appendix 1: Proofs of Theorem 2.2 and Corollary 2.3

Proof of Theorem 2.2. Part (a) follows directly from steps 2a and 2b of the construction of \( f \). In part (b), the claim that \( L_r(j) \) is an initial segment of \( Y_2' \) follows directly from part (a), step 2 of the construction of \( f \), and induction on \( r \). The claim on \( |L_r(j)| \) follows from step 2 of this construction on noting that \( C_i \) contributes 1 (resp. 2) to this sum if \( R_{ij} = 0 \) (resp. \( R_{ij} = 1 \).

Consider (c). By part (a), \( \sum_{t=1}^{j} R_{it} \) is the number of columns of \( Y_2' \), 1 \( \leq t \leq j \), such that \( |f(C_t) \cap Y_2'| = 2 \), while in the remaining columns \( Y_2' \), 1 \( \leq t \leq j \), we have \( |f(C_t) \cap Y_2'| = 1 \). So \( \pi(i, 1, j) = N_{ij} \) as claimed. Also observe that \( \pi(i, r \rightarrow r + j) = N_{i,r+j} = j + 1 + \sum_{t=r}^{j} R_{it} \in \{j + 1 + S_{j+1}, j + 2 + S_{j+1}\} \) by the formula for \( N_{ij} \) and Theorem 2.1b. The same equation with \( r \) replaced by \( s \) yields the rest of (c).

Part (d) follows from parts (b), (c), Lemma 2.1b, and the fact that \( S_{r+1} - S_r \leq 1 \). For (e), by part (c) (taking \( j = m \)) and the definition of \( G(a_1, N_{im}) \) we see that \( f(G(a_1, N_{im})) \subseteq Y_2'(m) \). So it suffices to show that \( |f(G(a_1, N_{im}))| = |Y_2'(m)| \). We have \( |f(G(a_1, N_{im})))| = \sum_{i=1}^{m} N_{ij} = \sum_{i=1}^{m} (m + \sum_{t=1}^{m} R_{it}) = a_1 m + m(2^i - a_1) = m2^i - a_1 = |Y_2'(m)| \), using Lemma 2.1a.

Consider (f). Recall the notation \( P_i = a_2 a_3 \cdots a_k \) for the number of 1-pages in \( G \). We use the fact that \( G \) can be identified with the the subgraph of \( G(a_1) \) induced by the union of initial segments \( \bigcup_{i=1}^{m} C_i(P_i) \) via the map \( \kappa \) of section 2. For \( j = m - 1 \) or \( m \) let \( M'_j = \min\{N_{ij} : 1 \leq i \leq a_1\} \), and \( M''_m = \max\{N_{ij} : 1 \leq i \leq a_1\} \). We claim that \( M'_m \geq P_1 \). If not, then \( M'_m < P_1 \), so \( M''_m \leq P_1 \) by (d). It follows that \( |Y_2'(m)| = |f(G(a_1, N_{im})))| = \sum_{i=1}^{m} N_{im} < a_1 P_1 = |G| \). But this contradicts \( |Y_2'(m)| = m2^i \geq |G| \), proving that \( M'_m \geq P_1 \). Hence for each \( 1 \leq i \leq a_1 \) we have \( N_{im} \geq P_1 \), implying that \( G \subseteq G(a_1, N_{im}) \) by the identification of \( G \) above. To show \( Y_2'(m-1) \subseteq f(G) \), observe from the definition of \( m \) that \( |Y_2'(m-1)| < |G| \). Thus \( \sum_{i=1}^{m} N_{im-1} = |Y_2'(m-1)| < |G| = a_1 P_1 \). So one of the terms in the last sum is less than \( P_1 \). It follows that \( M''_{m-1} < P_1 \) and hence \( M''_{m-1} \leq P_1 \) by (d). This says that for each \( i, f(i, P_1 + 1) \geq m \), and hence that \( f(i, j) \geq m \) for \( j \geq P_1 + 1 \) by the monotonicity of \( f \) from part (a). It follows that \( Y_2'(m) \subseteq f(\bigcup_{i=1}^{m} C_i(P_i)) = f(G) \), as required.

For (g), we induct on \( r \). The base case is clear since \( f(i, 1) = 2 \) for all \( i \). Assume inductively that \( |f(i, r + 1) - f(i, r)| \leq 1 \) for some \( r > 1 \), and let \( c = f(i, r) \). If \( f(j, r) = f(i, r) = c \), then we are done since \( f(i, r + 1) = c \) or \( c + 1 \) and the same holds for \( f(j, r + 1) \). So suppose \( |f(i, r) - f(j, r)| = 1 \), and without loss of generality that \( f(j, r) = c + 1 \) (the case \( f(j, r) = c - 1 \) being symmetric). Then \( N_{jc} \leq r - 1 \) by (c) and \( \sum_{t=1}^{r} R_{it} = 1 + \sum_{t=1}^{c} R_{jt} \) by Lemma 2.1b. If \( f(j, r + 1) = f(i, r) = c + 1 \), then we are done since \( f(i, r) \leq f(i, r + 1) \leq f(j, r + 1) \). So we may assume that \( f(j, r + 1) = c + 2 \). If now \( f(i, r + 1) = c + 1 \), then we are done again. So we can also assume that \( f(i, r + 1) = c \). Then we get \( N_{jc} \geq r + 1 \geq N_{jc} + 2 \), contradicting \( \sum_{t=1}^{c} R_{it} = 1 + \sum_{t=1}^{c} R_{jt} \).

Consider (h). Set \( c = f(i, y) \) for arbitrary \( y, 1 \leq y \leq N_{im} \). It suffices to show that \( f(i, y) \) must be one of three consecutive integers which depend only on \( i \), but not on \( y \). By (b), the set \( L_{i-1}(c) \) is the initial segment of size \( i - 1 + \sum_{t=1}^{c-1} R_{tc} \) in \( Y_2' \). So by Lemma 2.1b \( |L_{i-1}(c)| \) is
either $i - 1 + S_{i-1}$ or $i + S_{i-1}$. Now $f(C_i) \cap Y_2^{c}$ consists of either one or two successive points of $Y_2^c$ which immediately follow the point $(|L_{i-1}(c)|, c)$. Hence $f(i, y)$ must be one of the three successive integers $i + S_{i-1}, i + 1 + S_{i-1},$ or $i + 2 + S_{i-1}$, proving (h).

Finally consider (i). The proof is based on $R^{j+1}$ being a downward shift of $R^j$ (with wraparound) for any $j$. We refer to this property as the “downward shift” property. The assumption $|f(C_r) \cap Y_2^c| = 2$ says that $R_{|r|} = R_{r-1, j+1} = 1$ (viewing subscripts modulo $a_1$), by the downward shift property. Hence by part (b) and this same property, we have $|L_r(j)| = |L_r(j + 1)| = R_{r-1, j+1} = 0$, since $R_{r-1, j+1} = 1$. For the second claim, observe that $N_{r,j} = j + \sum_{i=1}^j R_{i,j} = j + \sum_{i=r-j+1}^{r} R_{i,j}$ by the downward shift property. Similarly we have $N_{r+1,j} = j + \sum_{i=r-j+2}^{r} R_{i,j}$. Therefore $R_{r,j} - N_{r+1,j} = R_{r-1, j+1} - R_{r, j+1} = 0$, since $R_{r-1, j+1} = 1$.

This completes the proof of the theorem.

**Proof of Corollary 2.3** As a convenience, we prove these properties with $f$ replacing $f_2$, to facilitate direct reference to the construction above. Of course $f_2$ then inherits these properties since it is a restriction of $f$. For brevity let $d_1 = |f(v)_1 - f(w)_1|$ and $d_2 = |f(v)_2 - f(w)_2|$.

Consider part (a). We get $d_2 \leq 1$ by Theorem 2.2 if $v$ and $w$ are successive points on the same chain, and by Theorem 2.2 otherwise (i.e., if $v = (i, j)$ and $w = (i + 1, j)$ for some $i$ and $j$). We also get $d_1 \leq 2$ by Theorem 2.2 if $v$ and $w$ are successive points on the same chain. So we are reduced to showing that $d_1 \leq 3$ when, say, $v = (i, j)$ and $w = (i + 1, j)$ for some $i$ and $j$. Let $f(v)_2 = c$. If also $f(w)_2 = c$, then by Theorem 2.2, we have that $(f(C_i) \cap f(C_{i+1})) \cap Y_2^c$ is a set of at most 4 successive points of $Y_2^c$, it follows in this case that $d_1 \leq 3$.

So let $f(w)_2 = c' \neq c$, noting that then $|c - c'| = 1$ since we have already shown that $d_2 \leq 1$. By Theorem 2.2 we have $f(v)_1 = |L_i(c)|$ or $|L_i(c)| - 1$, and similarly $f(w)_1 = |L_{i+1}(c') + 1|$ or $|L_{i+1}(c')| - 1$. By Theorem 2.2 we have $|L_{i+1}(c') - |L_i(c)| | \leq 2$, so $d_1 \leq 3$ follows, completing part (a).

The proof of (b) is a tedious case analysis showing that the bound on $d_1$ from part (a) can be strengthened to $d_1 \leq 2$, which combined with (a) yields (b). The proof is omitted here for brevity. We include (b) only for its possible interest, and do not use it later.

Consider part (c). Since $m = \lceil \frac{|G|}{2m} \rceil$, we see that $Y_2^{(m)}$ is a subgraph of the spanning subgraph $Y_2^{(G^{k-1})}$ of Opt($G$). Thus by Theorem 2.2, we obtain $f_2(G) \subset f(H) = Y_2^{(m)} \subset Opt(G)$, as required.

Consider next part (d). As notation, for $a \leq b$ let $Y_2^{(a-b)} = \bigcup_{t=a}^{b} Y_2'^t$ be the union of columns $a$ through $b$ of $Y_2$. Let then $Y_2^{(r-r+c-1)}$ and $Y_2^{(s-s+c-1)}$ be the sets of columns of $Y_2$ spanned by $f_2(T)$ and $f_2(T')$ respectively. Suppose to the contrary that $|c - c'| \geq 2$. By symmetry we can assume that $c' \geq c + 2$.

Let $S = f_2(T) \cap Y_2^{(r-r+c-1)}$, and let $S' = f_2(T') \cap Y_2^{(s-s+c-1)}$. Now for each $t, 1 \leq t \leq s + c' - 1$, we have $f_2(T') \cap Y_2'^{c}$ with the possible exception when $t = s$ (resp. $t = s + c' - 1$), where possibly $f_2(T')$ contains just the second (resp. first) of two points of $C_j \cap Y_2'^{c}$. So since $c < c' - 1$ it follows that $S' = f_2(C_j) \cap Y_2^{(s-s+c-1)}$. So $|S'| = \pi(j, s+1 \rightarrow s+c)$. Thus by Theorem 2.2: we have $|S'| - \pi(i, r \rightarrow r + c - 1)| \leq 1$. Now $|S'| \leq p - 2$, since $S'$ omits at least the two endpoints of the path $f_2(T')$. But also $\pi(i, r \rightarrow r + c - 1) \geq |S| \geq p$, a contradiction.

Finally we prove part (e). The containment in (e) follows from (d) (also from Theorem
\[2.24f\). The bound on \(|Y_{2}^{(r'-c)} - f_{2}(D_{2}^{(r)})|\) follows from \(Y_{2}^{(r'-1)} \subset f_{2}(D_{2}^{(r)})\), since by that containment the set \(Y_{2}^{(r')} - f_{2}(D_{2}^{(r)})\) is a proper subset of the column \(Y_{2}^{r'}\).

9 Appendix 2: Glossary of notation

- **G** = \([a_{1} \times a_{2} \times \cdots \times a_{k}]\): the k-dimensional grid, a graph with vertex set \(V(G) = \{x = (x_{1}, x_{2}, \ldots, x_{k}) : x_{i} \text{ an integer}, 1 \leq x_{i} \leq a_{i}\}\) and edge set \(E(G) = \{xy : \sum_{i=1}^{k} |x_{i} - y_{i}| = 1\}\)
- **P(t)**: the path graph on t vertices
- **e_{i} = \{\text{log}_{2}(a_{1}a_{2} \cdots a_{i})\}** for \(1 \leq i \leq k\), with \(e_{0} = 0\)
- **i-page**: a subgraph of \(G\) obtained by fixing the last \(k - i\) coordinates, and letting the first \(i\) coordinates vary over all their possible values
- **\(D_{i}^{r}, D_{i}^{(r)}, D_{i}^{(r)}(j)\)**: \(D_{i}^{r}\) is the \(r\)’th \(i\)-page of \(G\) under the ordering of \(i\)-pages given as follows. Let \(D_{i}\) and \(D_{i}^{'r}\) be two \(i\)-pages, with fixed last \(k - i\) coordinate values \(c_{i+1}, c_{i+2}, \ldots, c_{k}\) and \(c_{i+1}^{'}, c_{i+2}^{'}, \ldots, c_{k}^{'}\) respectively. Then \(D_{i} \prec D_{i}^{'r}\) in this ordering if at the maximum index \(t\), \(i + 1 \leq t \leq k\), where \(c_{t} \neq c_{t}^{'}\) we have \(c_{t} < c_{t}^{'}\). \(D_{i}^{(r)} = \bigcup_{j=1}^{r} D_{i}^{(j)}\). \(D_{i}^{(r)}(j) = D_{i+1}^{(r-1)a_{+}j}\), and we regard \(D_{i}^{(r)}(j)\) as the \(j\)’th \((i - 1)\)-subpage of \(D_{i}^{r}\) under the ordering of \((i - 1)\)-subpages of \(D_{i}^{r}\) induced by \(\prec_{i-1}\).
- **\(P_{i} = a_{i+1}a_{i+2} \cdots a_{k}\)**, the number of \(i\)-pages in \(G\).
- **\(\langle Y_{i} \rangle\)** for \(1 \leq i \leq k\): the \(i\)-dimensional grid given by \(\langle Y_{i} \rangle = P(2^{e_{i}}) \times P(2^{e_{i} - e_{i-1}}) \times P(2^{e_{i} - e_{i-2}}) \times \cdots \times P(2^{e_{i} - e_{1}})\). Note that \(\langle Y_{i} \rangle\) is a spanning subgraph of \(Q_{e_{i}}\), the hypercube of dimension \(e_{i}\).
- **\(Y_{i}\)**: the \(i\)-dimensional grid given by \(Y_{i} = \langle Y_{i-1} \rangle \times P(l_{i})\), where \(l_{i}\) is any integer satisfying \(l_{i} \geq 2^{e_{i} - e_{i-1}}\)
- **\((i - 1)\)-level of \(Y_{i}\)**: the subgraph of \(Y_{i}\) consisting of all points with fixed value \(c\) in the \(i\)’th coordinate, \(1 \leq c \leq l_{i}\).
- **\(Y_{i}^{2}\)**: \(Y_{i}^{2}\) is the \((i - 1)\)-level of \(Y_{i}\), consisting of all points with fixed coordinate value \(j\) in the \(i\)’th coordinate. \((x_{1}, x_{2}, \ldots, x_{i-1}, j)\).
- **\(Y_{i}^{(r)} = \bigcup_{j=1}^{r} Y_{i}^{j}\)**.
- **\(i\)-section of \(Y_{i}, S_{i}^{j}(G), S_{i}^{j}(G)^{'}\)**: \(S_{i}^{j}(G) = \bigcup_{j=a}^{b} Y_{i}^{j}\), where \(a = 1 + (j - 1)2^{e_{i} - e_{i-1}}\) and \(b = j2^{e_{i} - e_{i-1}}\). We call \(S_{i}^{j}\) the \(j\)’th “\(i\)-section” of \(Y_{i}\). \(S_{i}^{j}(G)^{'}\) is the set of points of \(S_{i}^{j}\) lying in nonblank \((i - 1)\)-levels of \(S_{i}^{j}\). \(S_{i}^{j}(G)^{'} = S_{i}^{j}(G) \cap (I_{i} \circ f_{i})(G)\), so \(S_{i}^{j}(G)^{'}\) is the set of points lying in nonblank columns of \(S_{i}^{j}\) which are also in the image of \(I_{i} \circ f_{i}\). \(S_{i}^{j}(c)\) is the \(c\)’th nonblank \((i - 1)\)-level of \(S_{i}^{j}\) in order of increasing \(i\)’th coordinate. We interpret \(S_{i}^{j}(c)\) for certain \(c\) using wraparound (see the comments preceding Theorem 6.2).
- **\(S_{i}^{(r)}, S_{i}^{(r)}(G), S_{i}^{(r)}(G)^{'}\)**: \(S_{i}^{(r)} = \bigcup_{j=1}^{r} S_{i}^{j}\), the union of the first \(r\) many \(i\)-sections of \(Y_{i}\). \(S_{i}^{(r)}(G)\) is the set points in \(S_{i}^{(r)}\) lying in nonblank \((i - 1)\)-levels of \(S_{i}^{(r)}\). \(S_{i}^{(r)}(G)^{'} = S_{i}^{(r)}(G) \cap (I_{i} \circ f_{i})(G)\).
- **\(s_{i}(j)\)**: the number of \((i - 1)\)-levels in \(S_{i}^{j}\) which are designated blank. By construction we have \(s_{i}(j) = 2^{e_{i} - e_{i-1}} - \left[\frac{a_{i+2}^{i}}{2^{e_{i-1}}}\right] + \left[\frac{i}{j^{2e_{i}}}\right] - [(j - 1)\phi_{i}\right]\), where \(\phi_{i} = \left[\frac{a_{i+2}^{i}}{2^{e_{i-1}}}\right] - \frac{a_{i+2}^{i}}{2^{e_{i-1}}}\).
- **\(F(i)\)**: the \(P_{i} \times 2^{e_{i} - e_{i-1}}\) \((0, 1)\) matrix \(F(i) = f_{uv}(i)\) that encodes which \((i - 1)\)-levels of \(S_{i}^{(r)}(P_{i})\) are to be designated blank as follows. We have \(f_{uv}(i) = 1\) if the \(u\)’th \((i - 1)\)-level of \(S_{i}^{u}\) (which is \(Y_{i}^{v+u-1}(2^{e_{i} - e_{i-1}})\)) is blank, and \(f_{uv}(i) = 0\) otherwise. In particular, the sum of entries in the \(u\)’th row of \(F(i)\), being the number of blank \((i - 1)\)-levels in \(S_{i}^{u}\), is \(s_{i}(u)\). Note that \(F(i)\) is
constructed in the procedure following Corollary 4.5.

- \(N_r(d), N_{i,r}(d)\): For a fixed \(i\) and \(1 \leq r \leq P_i\), \(N_r(d)\) is the column index of the \(d\)'th 0 from the left of row \(r\) of matrix \(F(i)\). When \(i\) can vary, we let \(N_{i,r}(d)\) be the \(N_r(d)\) just defined relative to matrix \(F(i)\). Further we interpret the function \(N_r(\ast)\) using "wraparound" (see the comments preceding Corollary 4.5).

- \(\text{Opt}(G), \text{Opt}'(G)\) for \(G = [a_1 \times a_2 \times \cdots \times a_k]\): \(\text{Opt}(G) = Q_n\) with \(n = \lceil \log_2(\lceil |V(G)| \rceil) \rceil\). This is the hypercube of smallest dimension having at least as many vertices as \(G\). \(\text{Opt}'(G) = P(2^{e_1}) \times P(2^{e_2-e_1}) \times P(2^{e_3-e_2}) \times \cdots \times P(2^{e_k-e_{k-1}})\), a spanning subgraph of \(\text{Opt}(G)\).

- \(u_i = u_i(G)\): For \(G\) a \(k\)-dimensional grid, we let \(u_i = \lceil \frac{|G|}{2^i-1} \rceil\). For each \(i, 2 \leq i \leq k\), we constructed a map \(f_i : G \to Y_i^{(u_i)}\), that is, a map of \(G\) into the first \(u_i\) \((i - 1)\)-levels of \(Y_i\). Each such \((i - 1)\)-level has size \(2^{e_i-1}\), so \(u_i\) is the minimum number of \((i - 1)\)-levels required in the image of any one to one map \(G \to Y_i\).

- \(x_{i,j} = (x_i, x_{i+1}, x_{i+2}, \ldots, x_{j-1}, x_j)\), for a \(t\)-tuple \(x = (x_1, x_2, x_3, \ldots, x_t)\), \(1 \leq i, j \leq t\).

- \(I_i\): the “inflation” map \(I_i : f_i(G) \to S_i^{(P_i)}(G)\) defined in step 3 (The Inflation Step) of the construction of \(f_{i+1}\) as follows. For any \(y = (z_1, z_2, \ldots, z_i) \in f_i(G)\), let \(I_i(y) = (z_1, z_2, \ldots, z_{i-1}, z_i')\), where \(Y_i^{z_i'}\) is the \(z_i\)'th nonblank level in \(S_i^{(P_i)}\) in order of increasing \(i\)'th coordinate.

- \(\sigma_i\): the “stacking” map \(\sigma_i : S_i^{(P_i)}(G)' \to S_i^1 \times P(u_{i+1}) \subseteq Y_i^{(u_{i+1})}\) defined in step 4 (The Stacking Step) of the construction of \(f_{i+1}\). Informally, \(\sigma_i\) stacks the sets \(S_i^r(G)'\) “on top of” \(S_i^1 \cong \langle Y_i \rangle\) in order of increasing \(r\) as follows. Suppose \(x = (x_1, x_2, \ldots, x_i)\) lies in \(S_i^j(G)'\). Then let \(\sigma_i(x) = (x_1, x_2, \ldots, x_{i-1}, \bar{x}_i, c)\), where \(\bar{x}_i, 1 \leq \bar{x}_i \leq 2^{e_i-e_i-1}\), is the integer congruent to \(x_i\) mod \(2^{e_i-e_i-1}\), and \(c = \sigma_i(x)_{i+1}\) is the number of images \(\sigma_i(y)\) with \(y \in S_i^j(G), 1 \leq j \leq r\), and \(\sigma_i(x)_{1 \to i} = \sigma_i(y)_{1 \to i}\).

- \(\bar{\sigma}_i\): \(\bar{\sigma}_i\) is the projection of \(\sigma_i\) onto the first \(i\) coordinates. So for \(y \in S_i^{(P_i)}(G)'\), we have \(\bar{\sigma}_i(y) = \sigma_i(z)_{1 \to i}\). By Lemma 5.1, the restriction of \(\bar{\sigma}_i\) to any one \(i\)-section is one to one.

- \(\text{Stack}_i(x, r)\): for \(x \in S_i^{1 \cdot} \cong \langle Y_i \rangle\) and \(1 \leq r \leq P_i \cdot (\text{defined in the paragraph just preceding Lemma 5.1})\): \(\text{Stack}_i(x, r) = \{z = \sigma_{i-1}(y) : z_{i-1 \to 1} = x\} \subseteq S_{i-1}^r(G)'\). We regard \(\text{Stack}_i(x, r)\) as a stack, with “height” extending into the \(i\)'th dimension, addressed by \(x \in \langle Y_{i-1} \rangle\). Its elements are images \(\sigma_{i-1}(y) \in Y_i\), with \(y \in S_{i-1}^r(G)'\), having projection \(\sigma_{i-1}(y)_{1 \to i-1} = x\) onto the first \(i - 1\) coordinates.

- \([\mathbf{r}]_i = \max \{|\text{Stack}_i(x, r)| : x \in \langle Y_{i-1} \rangle\}\). This is the maximum “height” of \(\text{Stack}_i(x, r)\), over all \(x \in \langle Y_{i-1} \rangle\).

- \(\nu_{i,r}(S)\): for \(S \subseteq Y_i^{(r)}\), let \(\nu_{i,r}(S) = |S \cap f_i(D_{i-1}^r)|\).

- \(\nu_{i}(z)\): For \(z = (I_i \circ f_i)(x)\), where \(x \in D_i^r\), \(\nu_{i}(z)\) is the integer such that \(z \in S_i^r(\nu_i(z))\). That is, \(z\) belongs to the \(\nu_i(z)\)'th nonblank \((i - 1)\)-level (ordered by increasing \(i\)-coordinate) of the \(i\)-section \(S_i^r\) containing \(z\). We interpret differences \(\nu_{i}(z') - \nu_{i}(z'')\) using wraparound (see the comments preceding Theorem 6.2).

10 Appendix 3: Figures
Figure 1:  a) The matrix $R$, and b) the corresponding map $f : G(5) \rightarrow Y_2$

Figure 2: The map $f_2 : [3 \times 7 \times 4 \times a_4] \rightarrow Y_2$, with $a_4 > 1$

Figure 3:  a) $I_2 : Y_2^{(21)} \rightarrow S_2^{(4)}$, used for constructing $f_3([3 \times 7 \times 4])$; a) and b) combined $I_2 : Y_2^{(63)} \rightarrow S_2^{(12)}$, used for constructing $f_3([3 \times 7 \times 4 \times 3])$
Figure 4: The stacking map $\sigma_2 : (I_2 \circ f_2)([3 \times 7 \times 4]) \to Y_3^{(3)}$; stacking the four sections $S^j_2$, $1 \leq j \leq 4$.

Figure 5: $\sigma_2$ applied to the first 7 and all 12 sections $S^j_2$, yielding the map $f_3 : [3 \times 7 \times 4 \times 3] \to Y_3^{(8)}$. 
Figure 6: Contents of stacks in columns 4 and 5 after stacking the first 7 sections $S_j^i$, $1 \leq j \leq 7$; points are labeled by their preimages in $G$.

Figure 7: The stacking map $\sigma_3(S_3^{(3)}) \subseteq Y_4^{(2)}$, yielding the map $f_4 : [3 \times 7 \times 4 \times 3] \to Y_4^{(2)}$.
Figure 8: Three individual stacks, each addressed by points of $S^1_3 \cong \langle Y_3 \rangle$, each of height 2, extending into the 4'th dimension