Towards Optimal Locality in Mesh-Indexings

Rolf Niedermeier[1], Klaus Reinhardt[1], and Peter Sanders[2]

[1] Wilhelm-Schickard-Institut für Informatik, Universität Tübingen, Sand 13, D-72076 Tübingen, {niedermr,reinhard}@informatik.uni-tuebingen.de
partially supported by the DFG Project La 618/3-1 KOMET
[2] Fakultät für Informatik, Universität Karlsruhe, D-76128 Karlsruhe, sanders@ira.uka.de

Abstract. The efficiency of many algorithms in parallel processing, computational geometry, image processing, and several other fields relies on "locality-preserving" indexing schemes for meshes. We concentrate on the case where the maximum distance between two mesh nodes indexed i and j shall be a slow-growing function of $|i-j|$ (using the Euclidean, the maximum, and the Manhattan metric). In this respect, space-filling, self-similar curves like the Hilbert curve are superior to simple indexing schemes like "row-major." We present new tight results on 2-D and 3-D Hilbert indexings which are easy to generalize to a quite large class of curves. We then present a new indexing scheme we call *H-indexing*, which has superior locality. For example, with respect to the Euclidean metric the H-indexing provides locality approximately 50% better than the usually used Hilbert indexing. This answers an open question of Gotsman and Lindenbaum. In addition, H-indexings have the useful property to form a Hamiltonian cycle and they are optimally locality-preserving among all cyclic indexings.

1 Introduction

For many algorithms, indexing schemes for (cubic[1]) meshes, that is, bijective mappings $\{0,\ldots,n-1\}^r \rightarrow \{0,\ldots,n^r-1\}$, play a crucial role. For example, in computational geometry one often has to map an r-dimensional raster to a one-dimensional traversal order or storage order. In this case it is advantageous if close-by raster points have close-by indices [1]. Analogous problems also arise in evaluating differential operators or even in a biological setting [7]. On the other hand, it is also important to consider "locality the other way round." For example, in parallel processing on mesh-connected computers one often has to map one-dimensional data structures to the processor-mesh. If the communication requirements within this data structure is predominantly between close-by indices, it is advantageous to map them to close-by processors in order to decrease network contention and latency.

Several mesh-indexing schemes are well-known. Most of these have been developed for the two-dimensional case, but they usually have generalizations for

[1] Generalizations are straightforward.

multiple dimensions, for example, row-major or snakelike row-major. However, these kinds of indexings do not preserve locality of computation and communication very well. So, e.g., for an r-dimensional mesh with side length n and row-major indexing, processors 1 and n are at distance $n-1$ from each other. Hence, a communication between these two processors ties up $n-1$ communication links and has a high latency. This is large compared to the distance of about $r\sqrt[r]{n}$ achievable if the first n processors could be arranged in a cube. A locality-preserving indexing should yield a distance $f(n) \in O(\sqrt[r]{n})$. This should generalize to all pairs of processors within the mesh, that is, processors indexed i and j should be at distance $f(|i-j|)$ from each other.

For example, a simple parallel variant of quicksort can be shown to run in average time $\Theta\left((n+\log m)\frac{m}{n^r}\right)$ for $m \geq n^r$ elements on n^r processors if a locality preserving indexing scheme is used. This is asymptotically optimal and compared to other asymptotically optimal algorithms only $\Theta(\log n)$ rather than $\Theta(n)$ messages are sent on the critical path [10]. Quicksort using row-major indexing and related schemes needs time $\Theta\left((n\log n+\log m)\frac{m}{n^r}\right)$. Various other applications in parallel processing are e.g. discussed in [6,5]. Further applications of this kind of locality can be found in image processing and related fields (see [4] and the references cited there). In this case, the Euclidean metric is sometimes preferred [4].

In this paper, we improve previous work on locality in mesh-indexings using (discrete) space-filling curves. First, we outline a *simplified* and *complete* proof of the result that the Manhattan distance $d(i,j)$ of two arbitrary indices i and j in the Hilbert indexing is bounded by $3\sqrt{|i-j|}-2$, a tight result previously given by Chochia, Cole, and Heywood [3]. We generalize the proof technique here and give a general "recipe" that makes it applicable for other indexing schemes and meshes of higher dimensions. In particular, we get almost tight results for the Manhattan distance of three-dimensional Hilbert indexings, showing, for example, $d(i,j) \leq 4.62\sqrt[3]{|i-j|}-3$ for $|i-j|>41$. Perhaps the most important contribution of this paper is the introduction of so-called *H-indexings* for two-dimensional meshes. We study H-indexings with respect to the three metrics Euclidean, maximum, and Manhattan. H-indexings show better locality than Hilbert indexings. Indeed, we conjecture that they are optimally locality-preserving among all space-filling curves. At least we can show that this holds for the class of cyclic space-filling curves. For H-indexings we prove, for example, that with respect to the Euclidean metric for arbitrary indices i and j it holds that $d(i,j) \leq \sqrt{4|i-j|-2}$, which is tight up to small additive constants. This answers an open question of Gotsman and Lindenbaum [4] for the existence of a family of space-filling curves with locality properties better than those of Hilbert curves in the two-dimensional case. To put it in quantitative terms, H-curves possess 50% better locality than Hilbert curves with respect to the Euclidean metric. Finally, we give improved lower bounds for the locality achievable by space-filling curves with respect to the three metrics mentioned above. Due to space constraints, many details and most proofs had to be moved to the full paper [8].

2 Preliminaries

In this paper we deal with 2-D and 3-D *grids* (or *meshes*, equivalently). We focus attention on quadratic and cubic grids, where, for example, in the two-dimensional case we have n^2 points arranged in an $n \times n$-array. Grids occur in various settings such as parallel computer architectures or image processing and many other fields of computer science. In what follows, we restrict the description of some basic concepts to the two-dimensional case. The generalization to the three-dimensional setting is straightforward.

We deal with *indexing schemes* for grids. An indexing scheme simply is a bijective mapping from $\{0, \dots, n^2-1\}$ onto $\{0, \dots, n-1\} \times \{0, \dots, n-1\}$, thus providing a total ordering of the grid points. We consider discrete space-filling curves as special kinds of indexing schemes, which turn out to have the desired property of preserving locality. To define locality, we need a metric. We will use the *Euclidean*, the *Manhattan*, and the *maximum metric*, which are defined as follows. Assume that $x(i)$ and $y(i)$ denote the position of a point i within the grid with respect to Cartesian coordinates. Then the Euclidean distance of two grid points i and j is defined as $d(i,j) := \sqrt{(x(i)-x(j))^2 + (y(i)-y(j))^2}$, for the Manhattan distance we have $d(i,j) := |x(i)-x(j)|+|y(i)-y(j)|$, and the distance according to the maximum metric is $d(i,j) := \max\{|x(i)-x(j)|, |y(i)-y(j)|\}$.

For a *discrete space-filling curve* $C : \{0, \dots, n^2-1\} \to \{0, \dots, n-1\} \times \{0, \dots, n-1\}$ it holds that $d(C(i), C(i+1)) = 1$, where d shall be e.g. the Euclidean distance. Thus one might say that space-filling curves provide *continuous* indexings. A space-filling curve traverses the grid making unit steps and turning only at right angles. To simplify presentation, we will often write simply i when in fact referring to $C(i)$. The meaning always will be clear from the context. Another feature of space-filling curves besides being continuous usually is their *self-similarity*. Self-similarity here simply means that the curve can be generated by putting together identical (basic construction) units, only applying rotation and reflection to these units. This will become clearer when considering the construction principles of Hilbert and H-curves in subsequent sections. A *segment* $\overline{(i,j)}$ of a space-filling curve is the set $\{C(i), \dots, C(j)\}$ of grid nodes. We deal with the following measure of *locality*. The basic requirement is that if according to the indexing scheme it holds that $|i-j|$ is small, then also $d(i,j)$ shall be small (applying one of the above metrics). We call an indexing *cyclic* if $(\{0, \dots, n^2-1\}, +)$ is the additive group modulo n^2. In this case $|i|$ shall denote the distance from i to 0, thus $|i| \leq n^2/2$.

3 The 2-D Hilbert Indexing

The main intent of this and the following section is to introduce a technique which makes it possible to derive locality properties of self-similar indexings by mechanical inspection. For conciseness, we focus on the well-known Hilbert indexing and the Manhattan metric which is particularly important in the context of parallel processing. Many generalizations to other indexings and other metrics will be possible however.

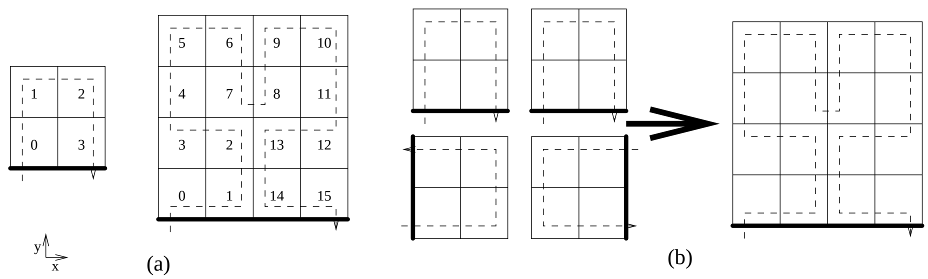


Fig. 1. Hilbert indexings of size 4 and 16 and the general construction principle.

Fig. 1-(a) shows the two smallest Hilbert indexings for meshes of size 4 and 16. Figure 1-(b) shows the general construction principle. For any $k \geq 1$ four Hilbert indexings of size 4^k are combined into an indexing of size 4^{k+1} by rotating and reflecting them in such a way that concatenating the indexings yields a Hamiltonian path through the mesh. Note that the left and the right side of the curve are symmetric to each other. So we only need to keep track of the orientation of the edge which contains the start and end of the curve (drawn with bold lines here).[2] We start with a lower bound for the locality:

Theorem 1. *For every $k \geq 1$, there are indices i and j on the Hilbert indexing such that $|i-j| = 4^{k-1}$ and the Manhattan-distance of i and j is exactly $3\sqrt{|i-j|} - 2 = 3 \cdot 2^{k-1} - 2$.*

Before we come to the matching upper bound, we need a technical lemma which shows how we can bound $\max_{|i-j|=k} d(i,j)$ for a fixed m by inspecting a finite number of segments. Namely the segments of length m which either lie within a single indexing of size $4^{\lceil \log_4 m \rceil}$ or within two such sub-grids which can only have 4 different relative orientations.

Lemma 2. *Let $x(i)$ and $y(i)$ denote the x-coordinate and y-coordinate of the ith point in the Hilbert indexing. Let*

$$d_{\text{int}}(m) := \max\left\{ d(i,j) \;:\; |i-j| = m \wedge 0 \leq i < j < 4^{\lceil \log_4 m \rceil} \right\} \text{ and}$$

$$d_{\text{ext}}(m) := 1 + \max_{0 \leq n < m} \max \begin{pmatrix} y(n)+x(m-n-1)+|x(n)-y(m-n-1)| \\ y(n)+y(m-n-1)+|x(n)-x(m-n-1)| \\ x(n)+y(m-n-1)+|y(n)-x(m-n-1)| \\ x(n)+x(m-n-1)+|y(n)-y(m-n-1)|. \end{pmatrix} .$$

Then $\forall i,j : d(i,j) \leq \max(d_{\text{int}}(|i-j|), d_{\text{ext}}(|i-j|))$.

---

[2] We note without proof that the above rule uniquely defines the Hilbert indexing up to global rotation and reflection. In a sense, the Hilbert curve is the “simplest” self-similar, recursive, locality-preserving indexing scheme for square meshes of size $2^k \times 2^k$.

Theorem 3. *For the Manhattan-distance $d(i,j)$ of two arbitrary indices i and j on the Hilbert indexing with $i \neq j$ we have $d(i,j) \leq 3\sqrt{|i-j|} - 2$.*

Proof. (Outline) By induction over $|i-j|$ the stronger statement "$d(i,j) \leq 3\sqrt{|i-j|} - 2.5$ or i and j are arranged as in Theorem 1" is proved (for $|i-j| \geq 16$)). The induction step exploits the self-similarity of the indexing. The estimations involved only work for $|i-j| > 80$ so there is rather large base case which is proved by a computerized application of Lemma 2. This automatization saves us the elaborate manual case distinction needed for a complete version of the proof in [3]. □

4 A Generalized Technique and its Applications

There are few places where the proof of Theorem 3 makes explicit use of the properties of the Hilbert indexing. We now present a generalized technique which can be applied to a wide spectrum of self-similar indexings in r-dimensional meshes made up of building blocks of size $q_1, \ldots, q_r$. For simplicity, however, we restrict the presentation to cubic building blocks with side-length q and only show how slightly looser upper bounds than that of Theorem 3 can be proved. The latter relaxation allows us to avoid the special treatment of the worst case segments which is necessary in the proof of Theorem 3.

Theorem 4. *Given any indexing scheme for r-dimensional meshes with the property that combining each elementary cube of size q^r in a mesh of size $q^k r$ into a single meta-node yields the indexing for a mesh of size $q^{(k-1)r}$:*
If $\forall q^{(k-1)r} \leq |i-j| \leq q^{kr} : d(i,j) \leq \alpha(\sqrt[r]{|i-j|} - \delta) - r$
where $\delta \geq \dfrac{\sqrt[r]{q^{kr} + q^r - 1} - q^k}{q-1}$
then $\forall |i-j| \geq q^{(k-1)r} : d(i,j) \leq \alpha(\sqrt[r]{|i-j|} - \delta) - r$.

Theorem 4 can be applied as follows to yield upper bounds of the form $d(i,j) \leq \alpha\sqrt[r]{|i-j|} - r$ for the Manhattan-distance of self-similar indexings:

- Determine q and r from the definition of the indexing.
- Fix a k.
- Set $\delta = \dfrac{\sqrt[r]{q^{kr} + q^r - 1} - q^k}{q-1}$.
- Exploit the self-similarity of the indexing to find an analogon to Lemma 2 which makes it possible to bound $d(i,j)$ for indices with $|i-j| = m$ using some mechanizable method.
- Use the above lemma to find a constant α_1 such that $d(i,j) \leq \alpha_1\sqrt[r]{|i-j|} - r$ for $|i-j| \leq q^{(k-1)r}$.
- Similarly, find a constant α_2 such that $d(i,j) \leq \alpha_2(\sqrt[r]{|i-j|} - \delta) - r$ for $q^{(k-1)r} \leq |i-j| \leq q^{kr}$. Applying Theorem 4 we can infer that the same is true for $|i-j| \geq q^{kr}$, i.e. $\forall |i-j| \geq q^{(k-1)r} : d(i,j) \leq \alpha_2(\sqrt[r]{|i-j|} - \delta) - r \leq \alpha_2\sqrt[r]{|i-j|} - r$.

– Set $\alpha = \max(\alpha_1, \alpha_2)$. We can now conclude from the two above points that *for all* i, j, $d(i,j) \le \alpha \sqrt[r]{|i-j|} - r$.

For example, Theorem 4 yields $d(i,j) \le 3.07\sqrt{|i-j|} - 2$ for the 2-D Hilbert indexing ($q = r = 2$) if we are willing to check all $|i-j| \le 256$. By checking larger meshes we can get closer to the tight bound from Theorem 3.

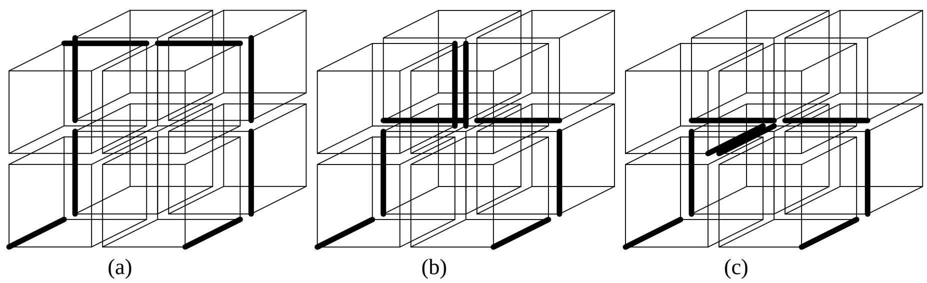


Fig. 2. Rule for building 3-D Hilbert indexings of order k from indexings of order $k-1$. The bottom front edge of the new cube is distinguished by the fact that the indexing starts and ends there. The corresponding edges of the component cubes are drawn with fat lines. The order $k-1$ cubes have to be rotated accordingly.

We have also applied the above technique to the three variants of a 3-D Hilbert indexing shown in Fig. 2. Up to rotation and reflections these are the only variants which are symmetric with respect to an axis. The maximum segment distances can be checked in complete analogy to Lemma 2: Now nine relative orientations are to be checked.[3]

Applying the "recipe" for variants (b) and (c) with $k = 5$ yields $d(i,j) \le 4.820661\sqrt[3]{|i-j|} - 3$ and the systematic search discovers indices with $d(i,j) \ge 4.820248\sqrt[3]{|i-j|} - 3$. Variant (a) has slightly better locality: $d(i,j) \le 4.678428\sqrt[3]{|i-j|} - 3$ and, for example, $d(0,5) = 5 \approx 4.678428\sqrt[3]{|i-j|} - 3$. If we neglect small segments we get $d(i,j) \le 4.640079\sqrt[3]{|i-j|} - 3$ for $|i-j| > 5$, and $d(i,j) \le 4.616161\sqrt[3]{|i-j|} - 3$ for $|i-j| > 333$.

The method could also be applied to the asymmetric variants of the Hilbert indexing described in [2]. We only need to change the procedure for checking maximum segment sizes to take segments starting at both ends of a cube indexing into account. Even generalizations to more complicated schemes like the H^* indexing described in [2] seem possible. (This scheme appears to have a better locality than simple Hilbert indexings.) H^* uses two nonisomorphic building blocks to define larger indexings. But it still has the crucial property, that replacing a $2 \times 2 \times 2$ cube by a unit cube yields an instance of the indexing.

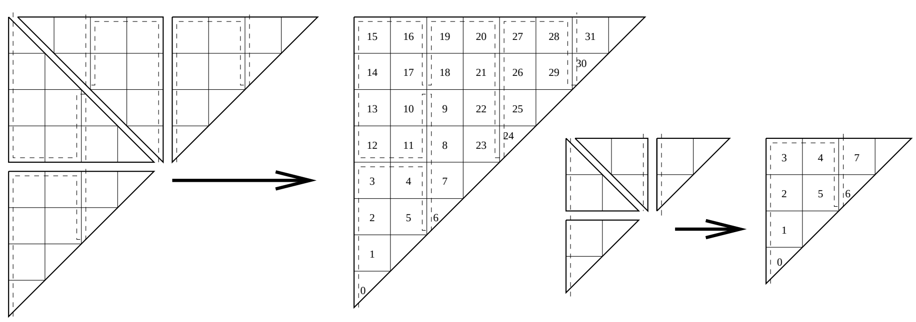
[3] A C-program doing the necessary checks is available under `http://liinwww.ira.uka.de/~sanders/hilbert/`.

5 The H-Indexing

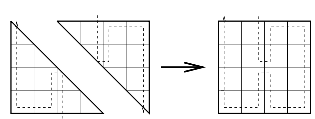
In this section we introduce H-indexings and show that they beat 2D-Hilbert indexings with respect to locality.

5.1 Construction scheme

H-indexings are related to 2D-Sierpiński curves [9]. As the naming already indicates, H-indexings have an “H-shaped” form. In analogy to Hilbert indexings, we obtain indexings for $2^k \times 2^k$-meshes[4], by an inductive method. There is, however, a decisive difference. Whereas in the case of Hilbert indexings the building blocks were four smaller squares, the construction of H-indexings is easier to describe using triangles. As for Hilbert indexings we only have one building block to which we apply rotation or reflection. To build the final mesh indexing, we put together two triangles. The following figure shows the construction of a triangle from 4 smaller triangles.



Observe that the triangles are constructed in such a way that exactly each *other* mesh node along the diagonal belongs to the nodes of the triangle. Thus an indexing scheme for a square mesh can be obtained as follows.



5.2 Euclidean metric

This subsection analyzes H-indexings with respect to the Euclidean metric, which is motivated by applications in image processing [4]. We show that H-indexings provide an improvement in locality compared to Hilbert-curves by roughly 50%, answering an open question of Gotsman and Lindenbaum [4]. They proved that for Hilbert curves C with respect to their locality measure

[4] A Java program for the general case can be found at `http://www-fs.informatik.uni-tuebingen.de/~reinhard/hcurve.html`.

$L_1(C) := \max_{i,j\in\{1,\ldots,n^2\},i<j} \frac{d(i,j)^2}{|i-j|}$ it holds $6 \cdot (1 - O(2^{-k})) \leq L_1(C) \leq 6\frac{2}{3}$, where $n = 2^k$. Now we show that for H-indexings C we have $L_1(C) = 4$. In Section 6 we indicate that this might be even optimal among all discrete space-filling curves. To present our result, we prefer to make a more concrete and more precise statement (which even includes additive constants) than the "$L_1(C)$-notation" allows.

Theorem 5. *For the Euclidean distance $d(i,j)$ of two arbitrary indices i and j on the H-indexing we have $d(i,j) \leq \sqrt{4|i-j|-2}$. This bound is tight up to small additive constants.*

Proof. (Sketch) We consider the triangles of size 2^{2l-1}, $l \geq 1$ according to the construction scheme of H-curves. Note that in order to get common points for the connection of the triangles, we assume that the points r, u, and v (see Fig. 3) also belong to the triangle. To get a $2^k \times 2^k$ mesh, we simply restrict the triangle for $l = k+1$ to $B \cup C$ in Fig. 3 (Because of symmetry, we can use the connection at y). To prove our result we employ an induction on the size of the triangles. As Fig. 3 shows, we extend the scope by three additional points: r, v and for technical reasons we assume a further artificial connection to the point u from a point diagonally under u. (This means u replaces the point left of u.)

We start with the base of the induction for $l = 1$ (see Fig. 3), where we have $|i-j| \leq 3$. In this case $d(i,j) \leq \sqrt{4|i-j|-2} \leq \sqrt{10}$ is always fulfilled.

Now we come to the induction step, assuming that $l \geq 2$. We distinguish between subcases $X - Y$, which means that i lies somewhere in X and j lies somewhere in Y. Clearly, cases $A - A$, $B - B$, $C - C$, and $D - D$ follow easily from the induction hypothesis.

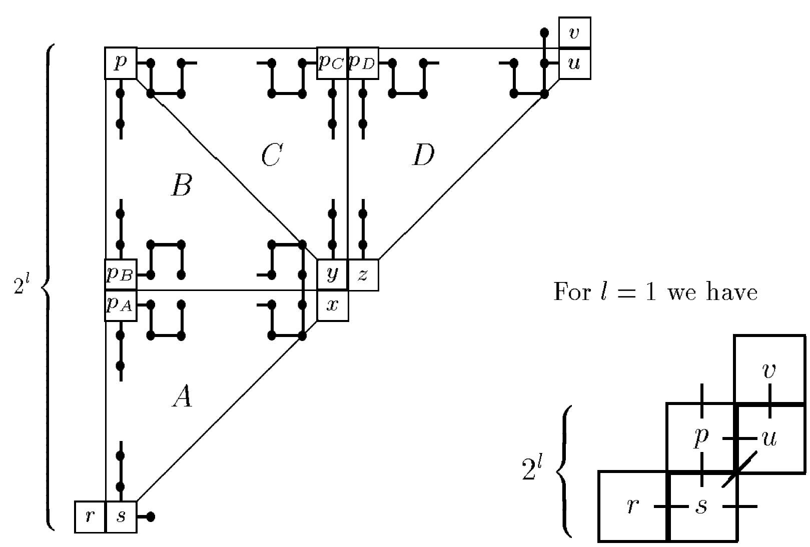


Fig. 3. A triangle of size $2^{2l-1} + 3$, including the additional points r, u and v.

In the case $C-D$ we distinguish between two further subcases. First assume that not both i and j (where w.l.o.g. i shall be in C and j in D) are located exactly at the corresponding diagonal lines. This means that either the angle between the lines from i to y and from y to j (cf. Fig. 3) or the angle between the lines from i to z and from z to j is at most 90^o. Thus we may use the estimation $d(i,j) \leq \sqrt{d^2(i,y)+d^2(y,j)}$ and applying the induction hypothesis for $d(i,y)$ and $d(y,j)$ we obtain $d(i,j) \leq \sqrt{4|i-y|-2+4|y-j|-2} = \sqrt{4|i-j|-4}$ Thus our claim is verified (analogously for z).

If both i and j are lying on the two diagonal lines as drawn in Fig. 4, then we get the following. W.l.o.g. assume that $y_j \geq y_i$. Then

$$d(i,j) = \sqrt{(y_i+y_j-1)^2+(y_j-y_i)^2} = \sqrt{2y_i^2+2y_j^2-2y_i-2y_j+1}.$$

On the other hand,

$$|i-j| \geq \frac{1}{2}(y_i+1)^2-1+\frac{1}{2}(y_j+1)^2-1-1 = \frac{1}{2}y_i^2+\frac{1}{2}y_j^2+y_i+y_j-2.5$$

due to the construction principle of H-curves and some straightforward considerations, yielding the validity of our claim in the case $C-D$.

The case $A-B$ is a bit more tricky. In order to employ a similar argument as for $C-D$, we artificially regard x and y as connection of A and B (see Fig. 3). It is obvious that the Euclidean distance from any point within the triangle to x or y is at least as big as the Euclidean distance to the real points of connection. Thus in a sense we only make things "worse" and we are finished if we can upper bound bound $d(i,j)$ using x or y. Now we can again distinguish two subcases in the same way as in case $C-D$: If not both i and j are located at diagonal lines, then completely the same argumentation as in $C-D$ applies. If both i and j are on diagonal lines, then analogous considerations as for $C-D$ show that $d(i,j) \leq \sqrt{2y_i^2+2y_j^2-2y_i-2y_j+1}$ and $|i-j| \geq \frac{1}{2}y_i^2+\frac{1}{2}y_j-1$, which implies our claim for $y_i, y_j \geq 2$, which always holds.

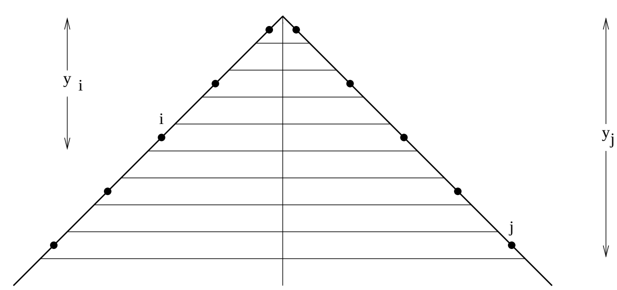


Fig. 4. The case when i and j are lying both on opposite diagonal lines. Note that every second row contains such a point. The number of rows till i and j, respectively, counting from the top are denoted y_i and y_j, respectively.

It remains to handle cases $(A \cup B) - (C \cup D)$. For $i \in A \cup B$ we distinguish the case $i \neq r$, where we use the induction hypothesis for $d(i,p)$ from the case $i = r$ where we use $|r-p| = 4^{l-1}$. For $j \in C \cup D$ we distinguish the case $j \neq v$ and u is reached by the normal way (i.e., not using the artificial diagonal connection), where we use the induction hypothesis for $d(p,j)$, the case $j = v$, where we use $|p-v| = 4^{l-1}+1$ and the case $j = u$ being reached on the artificial way, where we use $|p-u| = 4^{l-1}$. In each of those 6 combined cases we make an estimation using an angle of at most 90^o at p or a point beside p. Simple calculations yield

$$\begin{aligned}
d(i,j) &\leq \sqrt{d(i,p)^2 + d(p,j)^2} &&\leq \sqrt{4|i-j|-4} \\
d(r,j) &\leq \sqrt{(2^l-1)^2 + (d(p,j)+1)^2} &&\leq \sqrt{4|r-j| - 2(2^l - d(p,j))} \\
d(i,v) &\leq \sqrt{(d(p,i)+1)^2 + (2^l-1)^2} &&\leq \sqrt{4|i-v| - 2(2 + 2^l - d(p,i))} \\
d(r,v) &\leq \sqrt{(2^l)^2 + (2^l)^2} &&\leq \sqrt{4|r-v|-4} \\
d(i,u) &\leq \sqrt{(d(p,i)^2 + (2^l-1)^2} &&\leq \sqrt{4|p-u| - 1 - 2\cdot 2^l} \\
d(r,u) &\leq \sqrt{(2^l)^2 + (2^l-1)^2} &&\leq \sqrt{4|r-u| - 2\cdot 2^l + 1}
\end{aligned}$$

and the claim follows for $l \geq 1$ in each case. □

We just claim without proof that the statement of Theorem 5 in fact can be made tight for $|i-j| > 3$. Then we have $d(i,j) \leq \sqrt{4|i-j|-6}$ and a corresponding lower bound is reached for each k by pairs of points $i = 3 \cdot 2^{2k-5} - 1$ and $j = 2^{2k-3} + 1$.

5.3 Maximum and Manhattan metrics

Both maximum metric and the subsequently investigated Manhattan metric are of particular relevance in parallel processing.

Theorem 6. *For the maximum distance $d_1(i,j)$ and the Manhattan distance $d_2(i,j)$ of two arbitrary indices i and j on the H-indexing we have $d_1(i,j) \leq 2\sqrt{|i-j|+1} - 1$ and $d_2(i,j) \leq \sqrt{8(|i-j|-2)}$. These bounds are tight.*

H-indexings provide better locality than Hilbert curves also with respect to the maximum and the Manhattan metrics. A particular advantage of H-indexings over Hilbert indexings is that they not just describe a Hamiltonian path but a Hamiltonian cycle through the mesh. This is useful for parallel algorithms which employ communication along a virtual ring network. Interestingly, H-indexings are optimally locality-preserving among all Hamiltonian cycle through a square mesh, as the next section shows.

6 Lower Bounds for Discrete Space-Filling Curves

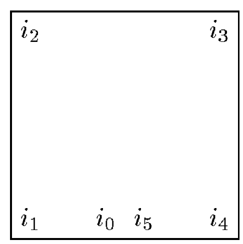
This section indicates that H-indexings might be optimally locality-preserving among all discrete space-filling curves. Indeed, we conjecture that they are optimal for the Euclidean, the maximum and the Manhattan metric. Because the difficulty for a general proof lies in "coming to grips with the loose ends," we advocate this conjecture by showing the optimality among the cyclic indexings.

Theorem 1 of Gotsman and Lindenbaum [4] for $m = 2$ (meaning "two-dimensional") says that for any discrete 2-dimensional space-filling curve on an $n \times n$-mesh it holds $d(i,j) > \sqrt{3(1-1/n)^2|i-j|}$ in Euclidean metric. They also report that by a computerized exhaustive search they have improved the constant factor 3 to 3.25. We improve this to 3.5 by a direct proof. We conjecture that this can be lifted to 4, which would imply the optimality of H-curves among all locality-preserving mesh-indexings (cf. Theorem 5).

Theorem 7. *For each indexing of an $n \times n$-mesh, $n \geq 2$, there must be indices i and j with $d(i,j) > n/4$ such that $d(i,j)$ is $\geq$ the general lower bound in the following table. For the case of a cyclic indexing, $d(i,j)$ is $\geq$ the cyclic lower bound; in particular, this cyclic lower bound holds for two corners i and j of the mesh.*

$d(i,j)$	*Euclidean*	*maximum*	*Manhattan*
Upper bound	$\sqrt{4\|i-j\|} - 2$	$2\sqrt{\|i-j\|+1} - 1$	$\sqrt{8(\|i-j\|-2)}$
General lower bound	$\sqrt{3.5\|i-j\|} - 1$	$\sqrt{3.5\|i-j\|} - 1$	$\sqrt{6\|i-j\|}$
cyclic lower bound	$2\sqrt{\|i-j\|} - 1$	$2\sqrt{\|i-j\|} - 1$	$\sqrt{8\|i-j\|} - 2$

Example: For the cyclic lower bounds, the proof is easily done by considering the distance of corners of the mesh. For the general case some more case distinctions have to be made. We give here only one example case for the Euclidean metric: Let $i_1 < i_2 < i_3 < i_4$ be the indices of the 4 corner points of the $n \times n$-mesh the indexing passes through in the given order. In our example case i_2 and i_3 are on the same side (cf. [4]) and there is no i_0 with $i_1 < i_0 < i_4$ on the opposite line of i_2 and i_3. Let i_0 be the rightmost point on the opposite line of i_2 and i_3 with $i_0 < i_2$ with distance $m - 1 > n/4$ from i_1. The next point i_5 with $i_3 < i_5$ therefore has distance $n - m - 1 > n/4$ from i_4.



Assume to the contrary that for all i and j with $d(i,j) > n/4$ we have $d(i,j) < \sqrt{3.5|i-j|} - 1$, that means $|i-j| > ((d(i,j)+1)^2)/3.5$. Thus we get to a contradiction by $n^2 \geq |i_0 - i_1| + |i_1 - i_2| + |i_2 - i_3| + |i_3 - i_4| + |i_4 - i_5| >$

$$\frac{m^2 + n^2 + n^2 + n^2 + (n-m)^2}{3.5} = \frac{2m^2 + 3n^2 - 2nm + n^2}{3.5} = \frac{3.5n^2 + 2(\frac{n}{2} - m)^2}{3.5}.$$

7 Conclusion

Locality-preserving indexing schemes are increasingly becoming a standard technique for devising simple and efficient algorithms for mesh-connected computers, for processing geometric data, for image processing and several other fields. For the most important, two-dimensional case, the newly presented H-indexing is superior with respect to locality compared to the previously used Hilbert indexing. We conjecture that H-indexings actually are optimal among all possible indexing schemes although we could only prove this for cyclic indexings yet.

With the advent of 3D-mesh-connected computers like the Cray T3E and the increasing interest in processing 3D-geometrical data, locality-preserving 3D-mesh indexings will become more important.[5] We were able to prove an almost tight bound of $d(i,j) \leq 4.62\sqrt[3]{|i-j|} - 3$ for large $|i-j|$ in a 3-D variant of the Hilbert indexing. This improves the upper bound $d(i,j) \leq 5.04\sqrt[3]{|i-j|}$ proven in [2]. For other, "non-Hilbert" indexings, experiments in [2] suggest that $d(i,j) \approx 4.1\sqrt[3]{|i-j|}$ might be achievable as a maximum segment size. We expect that the method from Section 4 can also be used for these numbering schemes as well as for deriving bounds on the Euclid metric and maximum metric. Applying the scheme to the H-curve and other schemes based on composing triangles may be more difficult.

Acknowledgment We are grateful to Henning Fernau for valuable pointers to the literature.

References

1. T. Asano, D. Ranjan, T. Roos, E. Welzl, and P. Widmayer. Space filling curves and their use in the design of geometric data structures. In *LATIN '95: theoretical informatics: second Latin American Symposium*, number 911 in LNCS, page 36ff, Valparaiso, Chile, April 1995.
2. G. Chochia and M. Cole. Recursive 3D mesh indexings with improved locality. Technical report, University of Edinburgh, 1997. short version to be published in HPCN'97.
3. G. Chochia, M. Cole, and T. Heywood. Implementing the hierarchical PRAM on the 2D mesh: Analyses and experiments. In *Symposium on Parallel and Distributed Processing*, pages 587–595, Los Alamitos, Ca., USA, Oct. 1995. IEEE Computer Society Press.
4. C. Gotsman and M. Lindenbaum. On the metric properties of discrete space-filling curves. *IEEE Transactions on Image Processing*, 5(5):794–797, May 1996.
5. C. Kaklamanis and G. Persiano. Branch-and-bound and backtrack search on mesh-connected arrays of processors. *Mathematical Systems Theory*, 27:471–489, 1994.
6. R. Miller and Q. F. Stout. Mesh computer algorithms for computational geometry. *IEEE Transactions on Computers*, 38(3):321–340, March 1989.
7. G. Mitchison and R. Durbin. Optimal numberings of an $N \times N$ array. *SIAM J. Alg. Disc. Meth.*, 7(4):571–582, October 1986.
8. R. Niedermeier, K. Reinhard, and P. Sanders. Towards optimal locality in mesh-indexings. Technical Report IB 12/97, University of Karlsruhe, 1997.
9. H. Sagan. *Space-Filling Curves*. Universitext. Springer-Verlag, 1994.
10. P. Sanders and T. Hansch. On the efficient implementation of massively parallel quicksort. In *4th International Symposium on Solving Irregularly Structured Problems in Parallel*, LNCS. Springer, 1997. to appear.

[5] On modern parallel machines, good locality has mainly the indirect effect to increase the usable bandwidth whereas the latency due to the distance in the network is negligible compared to other overheads.