Foundations of Computational Mathematics
https://doi.org/10.1007/s10208-018-9400-5

FOUNDATIONS OF COMPUTATIONAL MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

On the Number of Face-Connected Components of Morton-Type Space-Filling Curves

Carsten Burstedde[1] · **Johannes Holke**[2] · **Tobin Isaac**[3]

Received: 21 April 2017 / Revised: 16 March 2018 / Accepted: 23 August 2018

Abstract
The Morton- or z-curve is one example for a space-filling curve: Given a level of refinement $L \in \mathbb{N}_0$, it maps the interval $[0, 2^{dL}) \cap \mathbb{Z}$ one-to-one to a set of d-dimensional cubes of edge length 2^{-L} that form a subdivision of the unit cube. Similar curves have been proposed for triangular and tetrahedral unit domains. In contrast to the Hilbert curve that is continuous, the Morton-type curves produce jumps between disconnected subdomains. We prove that any contiguous subinterval of the curve divides the domain into a bounded number of face-connected subdomains. For the hypercube case in arbitrary dimension, the subdomains are star-shaped and the bound is indeed two. For the simplicial case in dimension 2, the bound is $2(L-1)$, and in dimension 3 it is $2L+1$, where L is the depth of refinement. We supplement the paper with theoretical and computational studies on the distribution of the number of jumps. For the hypercube curve, we can characterize the distribution by the fraction of segments of a given length that have no jump, and find that the fraction has a lower bound of $1/(2^d-1)$ and an asymptotic upper bound of $1/2$. For the simplicial curve, over 90% of all segments have three components or less.

Keywords Space-filling curve · Adaptive mesh refinement · Morton code

Mathematics Subject Classification 65M50 · 65D18

Communicated by Wolfgang Dahmen.

✉ Carsten Burstedde
burstedde@ins.uni-bonn.de

[1] Institut für Numerische Simulation (INS), Endenicher Allee 19b, 53115 Bonn, Germany

[2] German Aerospace Center (DLR), Linder Höhe, 51147 Köln, Germany

[3] College of Computing, Georgia Institute of Technology, North Avenue, Atlanta, GA 30332, USA

Published online: 05 October 2018

1 Introduction

The Peano curve [18] and the Hilbert curve [14] are continuous maps from the line onto the d-dimensional unit cube. A large number of such space-filling curves (SFC) have been described in the literature; see for example [3,13,19] and the references therein as well as [12]. They are usually defined in terms of a recursive prescription. For numerical applications, the curve is made discrete and finite by bounding the depth of the recursion. The smallest units of space that are traversed may be called elements. The Morton- or z-curve, originally described by Lebesgue [15] and adapted to data storage in 2D [17] and 3D [21], also creates such a map, but it is not continuous. In fact, it contains jumps throughout its length (see Fig. 1). This raises the concern that a subsection of the curve may divide the space covered by its image into a large number of disconnected subdomains, which would increase its surface area for a given volume. This is especially relevant when the curve is used to divide a computational mesh between different processors for parallel computation, see, e.g., [1,2,6,11,22], since an increased surface area increases the amount of data to be communicated. Thus, we aim to find upper bounds on the number of disconnected subdomains.

In this paper we prove that the (classic cubical) Morton curve can lead to no more than two subdomains, where we define a set of elements to be of the same subdomain if they are connected by a finite number of element face connections:

Theorem 1 *A contiguous segment of a Morton curve through a uniform or adaptive tree of maximum refinement level L produces at most two distinct face-connected subdomains. This result is independent of the space dimension.*

We note that a proof for the two-dimensional square has been given in [3, pages 175–177] that proceeds by illustrating and enumerating a finite number of cases. (In fact, we adapt these ideas to dimensions two and three in Sect. 3, and also restate the extension to adaptive meshes in Sect. 4.3.) It is said in [3] that the construction extends to dimensions three and higher. This is entirely plausible, yet we see that the number of cases to discuss grows with the space dimension and would eventually

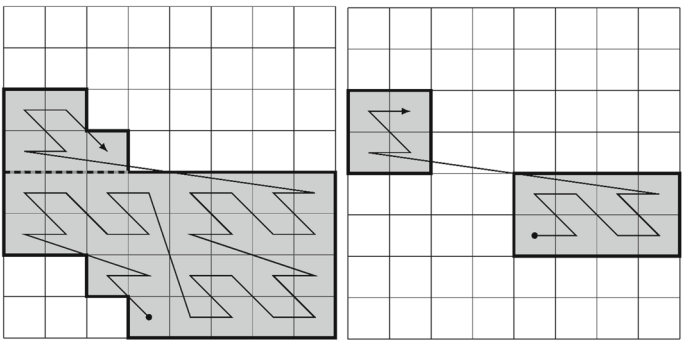


Fig. 1 Contiguous subsections of the Morton curve at refinement level $L = 3$. Observe the jumps when the z-curve runs diagonally. In the right-hand image, this produces two disconnected subdomains. In both pictures shown, the domain decomposes into two star-shaped pieces

require some kind of automation. Thus, we proceed inductively over d to provide dimension-independent results. We also supply a formal non-inductive proof to show that the connected segments are star-shaped.

For the triangular and tetrahedral Morton curves introduced recently [7], we show that the bound is proportional to the depth of refinement L with a leading factor of 2:

Theorem 2 *A contiguous segment of the tetrahedral Morton curve through a uniform or adaptive tree of maximum refinement level $L \geq 2$ produces at most $2(L-1)$ face-connected subdomains in 2D and at most $2L+1$ in 3D. For $L = 1$ there are at most two face-connected subdomains.*

We complete our study with a statement on the lower bound on the fraction of connected segments in the hypercube case and provide an algorithm and numerical results to illustrate the distribution of connected versus disconnected segments. We also compute histograms for the distribution of components in the simplical case. These results suggest that the tetrahedral Morton curve is no worse in practice than the original cubical construction.

2 Concepts and Notation

There is a natural identification between Morton-ordered elements on the one hand and uniform and adaptive quadtrees [10] and octrees [16] on the other. This is true for the tetrahedral Morton curve [7] as well. We will often refer to the elements as (sub)quadrants irrespective of the shape or space dimension d. Different ways exist to formalize the definition of a general space-filling curve; one is to identify a finite set of types of transformations and rules to apply them recursively [13]. In this document we restrict the theory and notation to the minimum required to treat the cubical and the tetrahedral Morton curve.

2.1 The Cubical Morton Curve

The Morton subdivision of a d-dimensional hypercube [17] can be constructed by recursion. When dividing a cube into 2^d half-size subcubes, we enumerate these with the binary index

$$q = (q_d \dots q_1)_2 \in [0, 2^d) \cap \mathbb{Z} \tag{1}$$

comprised of d bits $q_i \in \{0, 1\}$. (We will drop $\cap\mathbb{Z}$ in the following when it is clear that we are referring to integers.) Each of the bits i corresponds to the position of that subcube in the x_i coordinate direction, where 0 denotes the lower and 1 the higher half. In our convention the most significant bit corresponds to the last dimension (z in three dimensions) and the least significant bit to $x \equiv x_1$. When counting through the possible values of q we see that the x_1 coordinate changes its value fastest and the x_d coordinate slowest. Before a bit at position i flips, all numbers in the lower $i-1$ bits have to be counted through first.

We can state one central and well-known fact at this point: The flip of the ith bit amounts to a shift of the corresponding subcube parallel to the coordinate direction i.

If we flip from zero to one, we move up, and else we move down the axis. It is easy to see that flipping one bit transforms the subcube into its neighbor across a face with normal direction $\pm x_i$.

We define a recursion by subdividing each subcube further using the same prescription. The root cube is associated with level $\ell = 0$, with levels increasing with each subdivision. Subcubes exist at any level ℓ and are called level ℓ subquadrants. They are identified with the root of a corresponding level ℓ subtree. A level ℓ subtree has depth $L - \ell$. Level L subtrees are plainly called subquadrants. We count the sequence of level L (sub)quadrants with the index

$$Q = (q^1 \dots q^L)_2 \in [0, 2^{dL}), \tag{2}$$

where each level-wise index q^ℓ is defined as in (1). They designate the choice of subquadrants from the first subdivision $\ell = 1$ to the last at level $\ell = L$. This sequence of choices can be understood as the path from the root to the leaf of a decision tree, where each decision is between 2^d possibilities. The subset of $\mathbb{R}^d$ occupied by the quadrant with index Q is

$$\begin{aligned} \Omega(Q) := {} & [2^{-L}(q_1^1 q_1^2 \dots q_1^L)_2, 2^{-L}((q_1^1 q_1^2 \dots q_1^L)_2 + 1)] \times \\ & [2^{-L}(q_2^1 q_2^2 \dots q_2^L)_2, 2^{-L}((q_2^1 q_2^2 \dots q_2^L)_2 + 1)] \times \\ & \vdots \\ & [2^{-L}(q_d^1 q_d^2 \dots q_d^L)_2, 2^{-L}((q_d^1 q_d^2 \dots q_d^L)_2 + 1)]. \end{aligned} \tag{3}$$

We define a full or complete subtree by the set of all its descendant quadrants. A subtree is incomplete if the quadrants form a strict subset of descendants that are contiguous with respect to the indexing (2). We call such a subset a segment of a Morton curve in the following (two examples are depicted in Fig. 1).

We will make use of the following symmetry property of the Morton curve: It can be traversed forward or in reverse. The reversal amounts to go through the indexing (2) by counting backwards. A quadrant is transformed into the reverse ordering by taking the bitwise negation (the one-complement) of its index,

$$R(Q) = 2^{dL} - 1 - Q. \tag{4}$$

Geometrically, this operation mirrors the quadrant around the center point of the root cube.

2.2 The Simplicial Morton Curve

The tetrahedral Morton (TM) SFC applies to triangular and tetrahedral red-refinement of a mesh (and, conceptually, to higher-dimensional simplices) [7]. We encounter 1:4 refinement in 2D and 1:8 refinement in 3D [5], which means that the quad-/octree interpretation is still valid. We compute the SFC in a bitwise fashion that is an extension

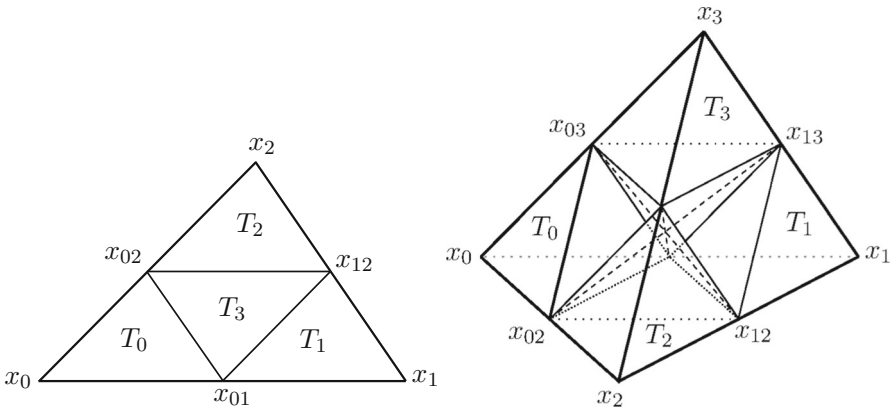


Fig. 2 Left: the refinement scheme for triangles in two dimensions. A triangle $T = [x_0, x_1, x_2] \subset \mathbb{R}^2$ is refined by dividing each face at the midpoints x_{ij}. We obtain four smaller triangles, all similar to T. Right: the situation in three dimensions. If we divide the edges of the tetrahedron $T = [x_0, x_1, x_2, x_3] \subset \mathbb{R}^3$ in half, we get four smaller tetrahedra (similar to T) and one octahedron. By dividing the inner octahedron along any of its three diagonals (shown dashed) we finally end up with a partition of T into eight smaller tetrahedra, all having the same volume. The refinement rule of Bey is obtained by always choosing the diagonal from x_{02} to x_{13} and numbering the corners of the children according to (5)

of the traditional Morton curve. In order to define the TM-SFC we introduce the concept of the type of a simplex.

Definition 3 We describe a d-dimensional simplex $T \subset \mathbb{R}^d$ by $d+1$ ordered vertices $x_0, \ldots, x_d \in \mathbb{R}^d$ and write $T = [x_0, \ldots, x_d]$. By x_{ij} we denote the midpoint between x_i and x_j.

Bey's *red-refinement* rule [5] for a triangle ($d = 2$) or tetrahedron ($d = 3$) amounts to dividing the parent simplex $T = [x_0, \ldots, x_d]$ into 2^d subsimplices that are defined and enumerated as follows (see also Fig. 2):

$$d = 2: \quad \begin{array}{ll} T_0 := [x_0, x_{01}, x_{02}], & T_1 := [x_{01}, x_1, x_{12}], \\ T_2 := [x_{02}, x_{12}, x_2], & T_3 := [x_{01}, x_{02}, x_{12}], \end{array} \tag{5a}$$

$$d = 3: \quad \begin{array}{ll} T_0 := [x_0, x_{01}, x_{02}, x_{03}], & T_4 := [x_{01}, x_{02}, x_{03}, x_{13}], \\ T_1 := [x_{01}, x_1, x_{12}, x_{13}], & T_5 := [x_{01}, x_{02}, x_{12}, x_{13}], \\ T_2 := [x_{02}, x_{12}, x_2, x_{23}], & T_6 := [x_{02}, x_{03}, x_{13}, x_{23}], \\ T_3 := [x_{03}, x_{13}, x_{23}, x_3], & T_7 := [x_{02}, x_{12}, x_{13}, x_{23}]. \end{array} \tag{5b}$$

Definition 4 (Type of a simplex) We start with a unit square/cube divided as in Fig. 3 and pick any of the triangles/tetrahedra as root simplex for refinement. Each subsimplex in a uniform level L refinement of this root simplex is contained in a subsquare/subcube of level L and is exactly one of the two (2D) or six (3D) simplices from Fig. 3. It thus has a unique number, which we define as the *type* of the simplex.

We start on level 0 with the root simplex T_d^0, which can have any of the possible types. In our implementation we pick 0 as the type of the root simplex. The TM code

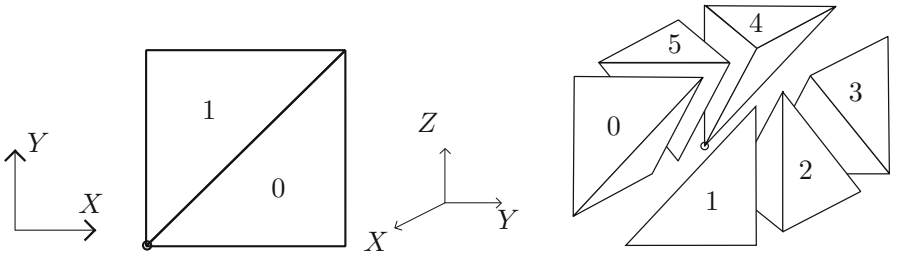


Fig. 3 We can separate an axis aligned cube into subsimplices by dividing it along one diagonal. We enumerate the resulting subsimplices and call the number of a subsimplex its type. Left: We divide a 2D square into two triangles, the lower right one has type zero and the upper left one has type 1. Right: We divide a 3D cube into six tetrahedra and enumerate them counterclockwise from zero to five (exploded view)

$m(T)$ for a descendant T of the root simplex is the interleaving of its anchor (lower left) node coordinates with the types of all of T's ancestor simplices [7]. It creates a total order between all simplices of a given level and thus establishes the SFC. Here we give a second, recursive definition of the SFC that is more suitable for our purposes.

By Proposition 18 in [7] we obtain one permutation $\sigma_b \in \Sigma_{2^d}$ for each possible simplex type b. It relates the ordering of its children to the SFC such that for any d-simplex T with $\text{type}(T) = b$

$$m(T_{\sigma_b(0)}) < m(T_{\sigma_b(1)}) < \cdots < m(T_{\sigma_b(2^d-1)}). \tag{6}$$

It places the child T_i in Bey's order at SFC position $\sigma_b(i)$.

Definition 5 Let T be a level L descendant of T_d^0 such that $T's$ parent P has type b and T is the i-th child of P according to Bey's order (5), $0 \leq i < 2^d$. We call the number $\sigma_b(i)$ the *local index* of the d-simplex T and use the notation

$$I_{\text{loc}}(T) := \sigma_b(i). \tag{7}$$

By definition, the local index of the root simplex is zero, $I_{\text{loc}}(T_d^0) := 0$. Table 1 lists the local indices for each parent type.

Thus, we know for each type $0 \leq b < d!$ how the children of a tetrahedron of type b are traversed. This gives us an approach for describing the SFC arising from the TM-index in a recursive fashion [13]. By specifying for each possible type b the order and types of the children of a type b simplex, we can build up the SFC. In Fig. 4 we describe the SFC for triangles in this way. In three dimensions it is not convenient to draw the six pictures for the different types, yet the SFC can be derived similarly from (5) and Table 1.

Remark 6 In 2D, we will make use of a symmetry property similar to (4): Reversing the TM curve in a uniform refinement of a type 0 triangle results in the (forward) TM curve for a type 1 triangle, and vice versa.

Table 1 The local index of a d-simplex T

I_{loc} 2D		Child T_0	T_1	T_2	T_3	I_{loc} 3D		Child T_0	T_1	T_2	T_3	T_4	T_5	T_6	T_7
b	0	0	1	3	2	b	0	0	1	4	7	2	3	6	5
	1	0	2	3	1		1	0	1	5	7	2	3	6	4
							2	0	3	4	7	1	2	6	5
							3	0	1	6	7	2	3	4	5
							4	0	3	5	7	1	2	4	6
							5	0	3	6	7	2	1	4	5

For each $b = \text{type}(T)$, the 2^d children $T_0, \ldots, T_{2^d-1}$ of T can be ordered according to their TM-indices. The position of the i-th child according to this order is the local index $I_{\text{loc}}(T_i)$

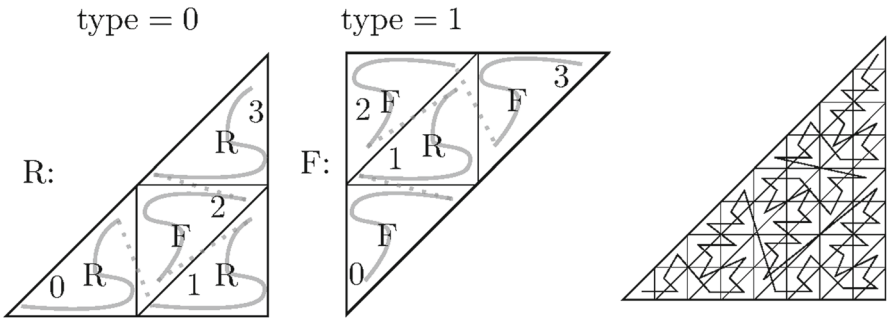


Fig. 4 Left: Using the notation from [13] we recursively describe the space-filling curve arising from the TM-index for triangles. The numbers inside the child triangles T_i are their local indices $I_{\text{loc}}(T_i)$. We write R for the refinement scheme of type 0 triangles and F for type 1 triangles. This pattern can be obtained from (5) and Table 1. Right: the SFC for a uniform level 3 refinement of the root triangle

3 Illustrated Proofs for $d \leq 3$

This section is devoted to proofs that use geometric intuition in two and three dimensions. For the cubical Morton curve, the idea is not new (although the execution in 3D seems to be). For the tetrahedral Morton curve, this is the first such study as far as we know. For abstract proofs for cubes of arbitrary dimension d we refer the reader to Sect. 4.

3.1 The Cubical Case

In this section, we prove a set of statements for cubes up to three dimensions by providing selected illustrations and covering all possible cases. A similar argument has been explored before in two dimensions [3], while an abstract proof for two dimensions can be found in [9].

We begin with statements that assume a curve that either begins with the first subquadrant of the unit cube or ends with its last subquadrant. In a second step, we use these statements to prove the final result. All statements are stated for arbitrary

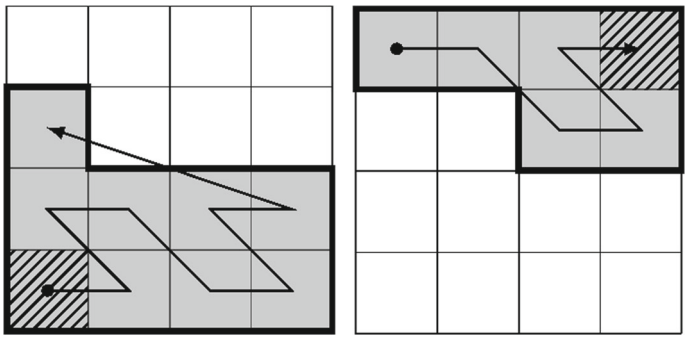


Fig. 5 We show two $L = 2$ examples of Morton curve segments that begin with the first subquadrant of a tree (left) and end with its last subquadrant (right), respectively. In both cases the segment covers one face-connected subdomain

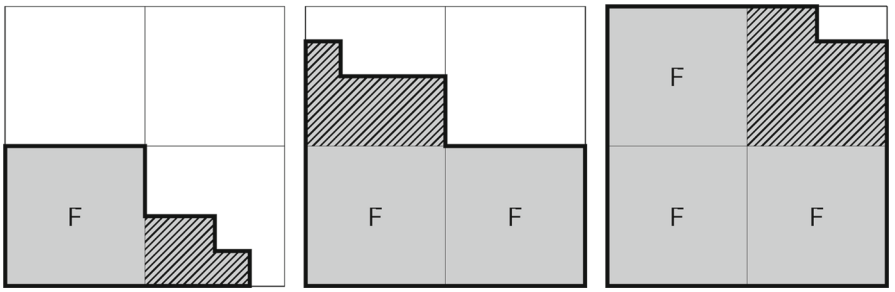


Fig. 6 Proof of Proposition 7: The three non-trivial cases that occur in two dimensions (we choose $L = 3$). The letter F designates a fully covered subtree of level 1. It is crucial that the lower left corner of the hatched area touches at least one of the full subtrees across a face

levels of refinement $L \geq 0$. In fact, all statements are trivially true for one dimension $d = 1$ (with no jumps at all); in this section we cover $d = 2$ and $d = 3$.

Proposition 7 *In a quadtree (or octree) T that is uniformly refined to level L, a contiguous segment of a Morton curve that begins with the first subquadrant in T creates exactly one subdomain of face-connected quadrants, no matter where it ends.*

Corollary 8 *In the situation of Proposition 7, a contiguous segment that ends with the last subquadrant in T creates exactly one face-connected subdomain, no matter where it begins (see Fig. 5 for an illustration).*

Proof Assuming that Proposition 7 is true, we can use the symmetry of the z-curve with respect to reversal to transform the present problem into the setting covered in Proposition 7. □

Proof (Proof of Proposition 7) We proceed by induction over L. Starting with $L = 0$, we only have one element and the statement is true. Supposing $L > 0$, we can identify the number $j \in [0, 2^d)$ that designates in which level 1 subquadrant of the tree the last level L subquadrant of the segment lies. If $j = 0$ then the whole segment is contained

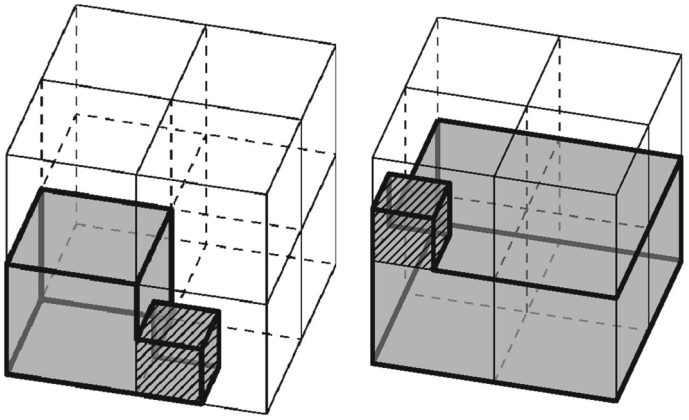


Fig. 7 Proof of Proposition 7: Selected cases in three dimensions ($j = 1, 4$ out of the seven non-trivial ones). The full subtrees are shaded lightly. Again we exploit the fact that the lower left front corner of the last non-empty subtree (hatched) connects to at least one full subtree with a lower subtree index across a face

in a level 1 subtree, which has depth $L-1$, and we can apply the induction assumption. Each of the remaining cases produces j full subtrees and one possibly incomplete one. That last subtree necessarily contains its first level L subquadrant q. Since this subtree produces one subdomain by induction, we are done by arguing that the full subtrees are face-connected to each other and to q, directly or indirectly. For two dimensions we show the three possible cases in Fig. 6, all of which satisfy the statement. For three dimensions we proceed by enumeration as well; we show selected situations in Fig. 7 to conclude the proof. □

Now that we have identified situations that produce one subdomain only; we can prove the main statement for arbitrary segments by a divide-and-conquer approach.

Proposition 9 *In a quadtree or octree that is uniformly refined to level L, a contiguous segment of the Morton curve creates no more than two distinct face-connected subdomains.*

Proof We proceed by induction over L. Again, the case $L = 0$ leaves nothing to prove. If the segment of the curve is contained in one level 1 subtree, the proof is finished by induction. Else we have one subtree in which the segment begins, zero to $2^d - 2$ fully covered subtrees, and one subtree in which the segment ends. To the first non-empty subtree we can apply Corollary 8, while Proposition 7 applies to the last one. Thus we know that the possibly incomplete subtrees lead to one connected piece each. The case of two non-empty subtrees is thus completed, and it remains to consider three or more.

Now, whenever any two adjacent non-empty subtrees have even-odd numbers, they are face-connected since at least one of them must be full. This covers the remaining three- and four-subtree cases in two dimensions. In three dimensions, this clears all situations with three non-empty subtrees. Since we can further reduce the number of

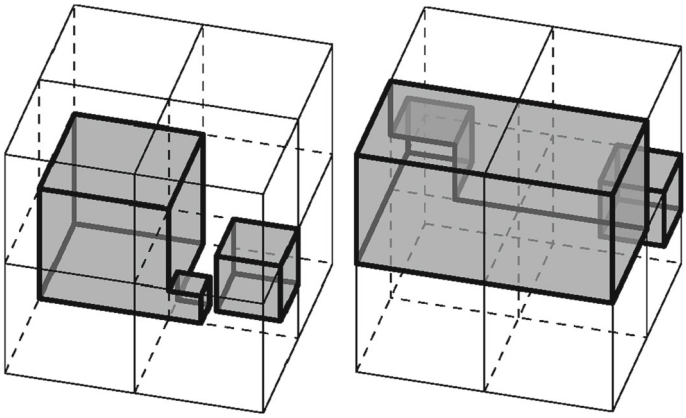


Fig. 8 The cases 1–3 and 3–6 in the proof of Proposition 9 for $d = 3$ dimensions. Each of these examples produces two distinct face-connected subdomains

remaining cases by symmetry, it remains to examine the subtree ranges $(i, \ldots, i+3)$ through $(i, \ldots, 7)$ for $i = 0, \ldots, 3$. All of these cases satisfy our claim; we illustrate a few in Fig. 8. □

We have completed the necessary proofs for a uniform space division into cubes in $d \leq 3$. In Sect. 4.1 we extend the proof to arbitrary dimension d. The case of adaptive space divisions is considered in Sect. 4.3.

3.2 The Simplicial Case

In this subsection we examine the number of face-connected components of a segment of the tetrahedral Morton SFC, $d = 2$ or 3. As we show in Fig. 9, there exist cases where the number of face-connected components in a uniform 2D level L refinement can be as high as $2(L-1)$. We show that this is in fact a sharp upper bound. We also show that in three dimensions the number of face-connected components does not exceed $2L+1$. There exists an example with $2L$ face-connected components and we conjecture that $2L$ is in fact the sharp estimate. The proof of these bounds is fairly analogous to the results for cubes and relies and a divide-and-conquer approach by splitting the segment into subsegments of which we know the number of face-connected components. The main difference to the cubical Morton curve is that we do not have a strong symmetry property like (4), and thus an analogue to Corollary 8 only exists in a weaker form.

Lemma 10 *The following two properties hold for the TM-index in 2D, where we consider a uniform level L refinement of an initial type 0 triangle T.*

- *Each type 1 subsimplex is face-connected to a type 0 subsimplex with a greater TM-index.*
- *Each type 0 subsimplex that is also a descendant of the level 1, type 1 subtriangle T_3 is face-connected to a type 1 subsimplex with a greater TM-index.*

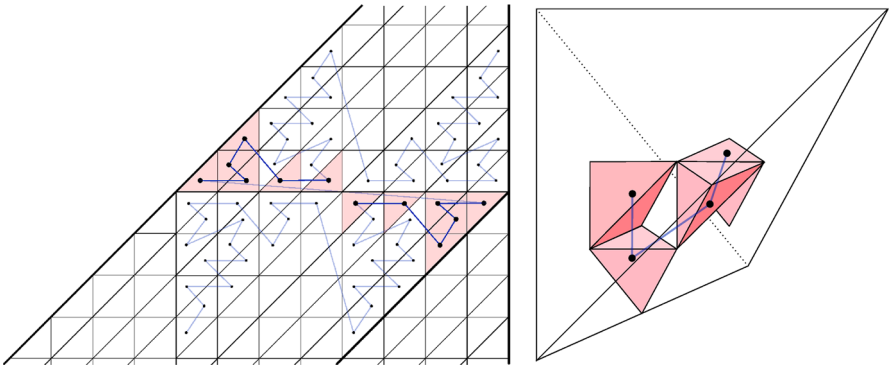


Fig. 9 Left: A segment of the 2D SFC on a level 4 refinement of T_2^0 with six face-connected components (shaded pink). The number of face-connected components in 2D can be as high as $2(L-1)$; this estimate is sharp. Right: a 3D level 2 refinement of T_3^0 with four ($=2L$) face-connected components. We prove that an upper bound on the number of face-connected components is $2L+1$ and conjecture that $2L$ is sharp

Proof The respective face neighbor is the top face neighbor for the type 1 subsimplex and the face neighbor along the diagonal face for the type 0 subsimplex; see Fig. 10. For type 0 we additionally require that the subsimplex is a descendant of T_3, since this ensures that the face neighbor along the diagonal face is inside the root triangle. Despite this detail, the proofs for both items are identical, and we only present one for the first.

Let S denote an arbitrary type 1 subsimplex of level L and let S' be its neighbor across the top face. If S and S' share the same parent P then there are two cases, which we also see in Fig. 4: Either $\text{type}(P)=0$, then the local index of S is 2 and that of S' is 3, or $\text{type}(P)=1$, in which case the local index of S is 0 and that of S' is 1. Thus, in both cases the TM-index of S must be smaller than that of S'. We suppose now that S and S' have different parents, which implies $L \geq 2$, and denote these different level $L-1$ subsimplices by P and P'. The only possible combination is that $\text{type}(P)=1$ and $\text{type}(P')=0$, and that P and P' are neighbors along P's top face. Therefore, by an induction argument, $m(P) < m(P')$, and since the TM-index preserves the local order under refinement, each child of P has a smaller TM-index than each child of P'. In particular we find $m(S) < m(S')$. □

Let us now show a 2D analogue to Proposition 7.

Lemma 11 *Consider a triangle T that is uniformly refined to level L. If T has type 0, then a contiguous segment of the SFC ending in the last level L subsimplex has just one face-connected component. If T has type 1, then this holds for segments starting in the first level L subsimplex.*

Proof We present the proof for $\text{type}(T)=0$, since we can then use the symmetry of the 2D curve (Remark 6) to obtain the result for the case $\text{type}(T)=1$. We proceed by induction over L.

For $L=0$ there is only one possible segment and it is connected. For $L=1$ we obtain the result by investigating all 10 cases. For $L>1$, let $j \in \{0,1,2,3\}$ be the

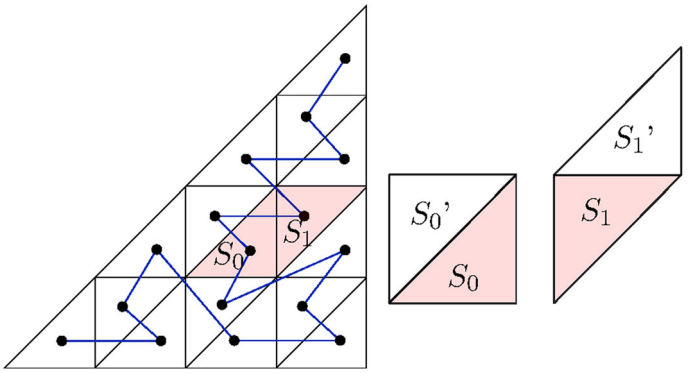


Fig. 10 Illustration of Lemma 10. In 2D, choose any subsimplex S_*. If its neighbor along the top face S'_* is inside the root triangle, then $m(S_*) < m(S'_*)$. This condition is always fulfilled by any type 1 triangle and by type 0 triangles that are descendants of the middle level 1 subtriangle

local index of the level 1 subtree T' of T in which the first level L subsimplex of the segment lies. If $j \in \{0, 1, 3\}$, then the type of T' is 0 and the statement follows by induction with the same argument as in the proof of Proposition 7. Thus, let $j = 2$, i.e., the segment starts in the type 1 subtree of T. The part of the segment that is not inside T' is the full last subtree of T (local index 3) and thus it is face-connected in itself. With Lemma 10 we conclude that each subsimplex in the subsegment in T' is face-connected to a simplex with greater TM-index. Iterating this process, we conclude that each of these subsimplices is face-connected to a subsimplex of the full last subtree of T. Thus, the whole segment is face-connected. □

For all other segments beginning with the first or ending in the last level L subsimplex, and notably for all of those segments in 3D, we obtain an upper bound of $L + 1$ face-connected components, which we show in the next two lemmas.

Lemma 12 *Let a segment of the space-filling curve for a uniform level L refined d-simplex consist of several full level 1 subsimplices plus one single level L simplex either at the end or at the beginning, then this segment has at most two face-connected components.*

Proof Similarly to the last paragraph in the proof of Proposition 9, and in analogy to Fig. 8, we can show this claim by enumerating all possible cases (no induction required). □

Lemma 13 *If a d-simplex is uniformly refined to level L, then any segment of the space-filling curve ending in the last subsimplex or starting in the first has at most $L + 1$ face-connected components.*

Proof Consider the case that the segment starts in the first simplex. For $L = 0$ there is only one possible segment consisting of the unique level 0 subsimplex and it is thus connected. Let now $L > 0$. Since the segment begins at the very first level L subsimplex, we can separate it into two parts. The first part at the beginning consists

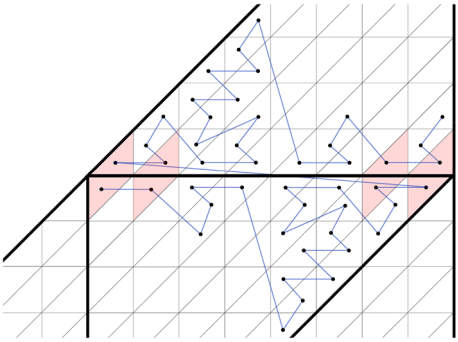


Fig. 11 An illustration of the 2D case in the proof of Proposition 14 for $L = 4$. The bottom segment has the maximal number of L face-connected components. Since its first and second triangles (on the left, shaded in pink) are connected with the top segment, the possible number of connected components is reduced by two. If the second segment has L face-connected components then its last two triangles (on the right) are connected with the bottom segment. Thus, the number of face-connected components is less than or equal to $2L - 2$

of 0 to $2^d - 1$ full level 1 subtrees, and the second part is one possibly incomplete level 1 subtree.

By the induction assumption, the second part has at most L face-connected components. From Lemma 12 we obtain that the first part together with the first level L subsimplex of the second part has at most two face-connected components. Since this first level L subsimplex is contained in one of the components of the second part, we obtain

$$L + 2 - 1 = L + 1 \tag{8}$$

components in total.

If the segments ends in the last simplex, the order of parts is reversed. The first part of the segment is the part in the level 1 subtree where the segment starts, and the second part consists of the remaining full level 1 subtrees. We obtain the bound on the number of face-connected components using the same inductive reasoning as above.

We have so far argued the connectivity of specific kinds of SFC segments. This suffices to proceed to arbitrary segments of the tetrahedral Morton SFC.

Proposition 14 *Any contiguous segment of the space-filling curve of a uniform level $L \geq 2$ refinement of a type 0 simplex has at most $2(L-1)$ face-connected components in 2D and $2L + 1$ face-connected components in 3D. For $L = 1$, there are at most two face-connected components, and one for $L = 0$ (this applies to both 2D and 3D).*

Proof Again, the cases $L = 0$ and $L = 1$ follow by inspecting all cases. Thus, let $L \geq 2$. We first show that for $d \leq 3$ the number of face-connected components is bounded by $2L + 1$:

If a given segment is contained in a level 1 subtree, we are done by induction. Otherwise we can divide the segment into three (possibly empty) pieces: First, the segment in one incomplete level 1 subtree ending at its last level L subsimplex, then one contiguous segment of full level 1 subtrees and finally a segment in one (possibly

incomplete) level 1 subtree that starts at its first level L subsimplex. Lemma 13 implies that the first and the last piece have at most L face-connected components each. By Lemma 12, the second piece has one or two face-connected components, and if the number is two, then it is face-connected to the first or to the third piece. Thus, it adds only one face-connected component to the total number, and we obtain at most

$$L + 1 + L = 2L + 1 \tag{9}$$

face-connected components.

Let us now specialize to 2D. We conclude from Lemma 11 that the first subsegment only adds more than one face-connected component if it is contained in the only level 1 subtree of type 1 (local index 2). Similarly, the third subsegment only adds more than one face-connected component if it is contained in a level 1 subtree of type 0. In particular, if both subsegments add more than one connected component, the third subsegment is contained in the last level 1 subtree (local index 3). Thus, the second subsegment is empty in this case.

If both of these subsegments have less than L face-connected components, there is nothing left to show since the overall number of components is then less than or equal to $2(L-1)$. So suppose that one of the subsegments has L face-connected components and the other one has at least $L-1$. We depict this situation in Fig. 11. We observe that the first and second level L simplex in this first segment are face-connected to the first and second level L simplex in the second segment. If, however, the second subsegment has L connected components then its last two level L simplices are face-connected to the last two level L simplices of the first subsegment.

We thus can subtract two connected components from the total count, which leads to at most

$$L + L - 2 = 2(L - 1) \tag{10}$$

face-connected components in total. □

We briefly discuss whether we can sharpen these bounds. In 2D, this is not possible by counterexample; see Fig. 9. In 3D, we construct a segment with $2L$ face-connected components using the SFC-indices 22–25 of a uniform level 2 refinement of a type 0 tetrahedron. We believe that the case that the first and the last piece described in the proof of Proposition 14 have L face-connected components each and that additionally the middle piece adds one component does not occur.

Conjecture 15 *In 3D, the number of face-connected components is bounded by* $2L$. *This estimate is sharp.*

4 Proofs for Arbitrary Dimension

The construction of the cubical Morton curve generalizes readily to arbitrary space dimension d. We should suppose that the main result (that any contiguous segment consists of at most two face-connected subdomains) generalizes as well. Indeed, we propose two different ways to prove this in the following Sects. 4.1 and 4.2. The

first is closer to the geometric approach we have been using in Sect. 3, while the second is more formal and paves the way for quantitative studies of the frequency of disconnections in Sect. 5.

We close this section with the extension of the proofs for both cubical and tetrahedral curves from uniform to adaptive meshes (see Sect. 4.3), which is the remaining step to establish Theorems 1 and 2, and discuss implications for a forest of octrees (Sect. 4.4).

4.1 Induction Proofs for *d*-Cubes

We use induction over both the dimension and the level of subdivision to prove the main statement for all dimensions $d > 0$. These proofs imply the statements of the previous Sect. 3.1 as special cases. For convenience we denote any d-tant as a quadrant. We make use of the following definition of subtree ranges.

Definition 16 Let the space dimension be $d > 0$. For any $0 \leq d' \leq d$ and $0 \leq k < 2^{d-d'}$, we define the following interval containing $2^{d'}$ integers,

$$I_k^{d'} = 2^{d'}[k, k+1). \tag{11}$$

We use this interval to denote a specific contiguous range of subtree indices.

We define the following auxiliary statements, first considering a one-sided segment and then a general two-sided one.

Proposition 17 *If a segment of a Morton curve is fully contained in the level 1 subtrees enumerated by a given $I_k^{d'}$ and contains the first or last subquadrant in this range of subtrees, then it corresponds to one face-connected subvolume.*

Proof By symmetry of the Morton curve, we can restrict the discussion to the case of the first subquadrant. Let us begin by proving the statement for subdivision level $L = 1$. By (11) the lowest subtree index in the segment is $k2^{d'}$. This number has d' zero bits from the right. All other indices in $I_k^{d'}$ have one or more ones in the lower d' bits while being bitwise identical in the higher bits. For any of these indices we can flip the low bits to zero one by one, effectively transitioning through face neighbors and monotonously decreasing the index until we reach $k2^{d'}$. This whole sequence of face-connected subtrees is contained in $I_k^{d'}$. In conclusion, all trees in $I_k^{d'}$ are face-connected to $k2^{d'}$ and thus to each other.

Now let $L > 1$ and assume the above statement for $L - 1$. To prove it for L we make an induction over d'. If $d' = 0$, we have a single subtree and can readily invoke the induction assumption for $L - 1$. Else there are two possible cases: Either the segment is fully contained in one of $I_{2k}^{d'-1}$ or $I_{2k+1}^{d'-1}$ and we apply the induction over d'. Otherwise $I_{2k}^{d'-1}$ contains full subtrees only and the segment reaches into $I_{2k+1}^{d'-1}$. Each non-empty subtree j in the latter interval must contain its first subquadrant, which has a face connection to the full tree $j - 2^{d'-1} \in I_{2k}^{d'-1}$. Since by the proof for $L = 1$ all subtrees in $I_{2k}^{d'-1}$ are face-connected, we are done. □

Proposition 18 *If a segment of a Morton curve is contained in the level 1 subtrees $I_k^{d'}$, it produces no more than two distinct face-connected subvolumes.*

Proof Again let us prove the statement first for $L = 1$. If $d' = 0$ we have just one level 1 subquadrant that clearly satisfies our claim. For positive d' we distinguish the following cases. If the segment is fully contained in either $I_{2k}^{d'-1}$ or $I_{2k+1}^{d'-1}$, we apply the induction on d'. Else we know that the last subquadrant of $I_{2k}^{d'-1}$ and the first of $I_{2k+1}^{d'-1}$ are in the segment. By Proposition 17 we have at most two disconnected pieces and the statement holds.

If $L > 1$ the case $d' = 0$ reduces to the same statement for $L - 1$ and we are done by applying the induction over L. Else, the proof proceeds unchanged as above with the desired result. □

We have implicitly proved the main result for any uniform level L subdivision, since a level 0 subtree trivially satisfies our claim, and otherwise the root cube is the union of the level 1 subtrees I_0^d.

4.2 A Non-inductive Proof for *d*-Cubes

In this section we elaborate on the formalism of the Morton index (see Sect. 2.1) to obtain the result without induction. The tool we use is the map Ω from the index $Q = (q^1 \dots q^L)_2$ to a subset of $\mathbb{R}^d$ stated in (3). For $1 \le r \le d$, we define the coordinate along axis r,

$$Q_r = (q_r^1 \dots q_r^L)_2. \tag{12}$$

The map $\Omega(Q)$ may be written as

$$\Omega(Q) = [2^{-L}Q_1, 2^{-L}(Q_1 + 1)] \times \cdots \times [2^{-L}Q_d, 2^{-L}(Q_d + 1)]. \tag{13}$$

Lemma 19 *If Q and $\tilde{Q}$ are such that $\tilde{Q}_k \le Q_k$ for all $1 \le k \le d$, then $\tilde{Q} \le Q$.*

Proof The order of the bits in Q_k ($\tilde{Q}_k$) is the same as their order in Q ($\tilde{Q}$). □

Theorem 20 *For any index Q^{end}, the interior of $Y_0 = \cup_{Q=0}^{Q^{end}} \Omega(Q)$ is star-shaped (and thus Y_0 is face-connected and contractible).*

Proof If $Q \in \{0, \dots, Q^{\text{end}}\}$, then so are all $\tilde{Q}$ such that $\tilde{Q}_k \le Q_k$ for all $1 \le k \le d$. The domains of these quadrants define a box between the origin and the corner of $\Omega(Q)$ farthest from the origin,

$$B(Q) := [0, 2^{-L}(Q_1 + 1)] \times \cdots \times [0, 2^{-L}(Q_d + 1)]. \tag{14}$$

Indeed, Y_0 is the union of these boxes, $Y_0 = \cup_{Q=0}^{Q^{\text{end}}} B(Q)$, and the union of their interiors is the interior of Y_0. Each of these boxes contains the midpoint of $\Omega(0)$ in its interior and is star-shaped with respect to it. Therefore the interior of Y_0 is star-shaped with respect to that point as well. □

Corollary 21 *If $Q|Q^{start} = Q$ (| means bitwise-or) for all $Q \in \{Q^{start}, \ldots, Q^{end}\}$, then the interior of $Y = \cup_{Q=Q^{start}}^{Q^{end}} \Omega(Q)$ is star-shaped.*

Proof If $Q|Q^{\text{start}} = Q$, then the 1-bits of $Q - Q^{\text{start}}$ are a subset of the 1-bits of Q. One can then verify that $(Q - Q^{\text{start}})_k = Q_k - Q_k^{\text{start}}$ for all $1 \le k \le d$. Therefore $\cup_{Q=Q^{\text{start}}}^{Q^{\text{end}}} \Omega(Q - Q^{\text{start}})$ is Y translated by the vector $(-2^{-L} Q_1^{\text{start}}, \ldots, -2^{-L} Q_d^{\text{start}})$. This is the same as $\cup_{Q=0}^{Q^{\text{end}}-Q^{\text{start}}} \Omega(Q)$, which is star-shaped by Theorem 20. □

Corollary 22 *If Q & $Q^{end} = Q$ (& denotes the bitwise and operator) for all $Q \in \{Q^{start}, \ldots, Q^{end}\}$, then the interior of Y is star-shaped.*

Proof Mirroring every quadrant about the midpoint of the unit cube does not change the shape of Y. The mirror of $\Omega(Q)$ is $\Omega(R(Q))$, where $R(Q)$ denotes the bitwise negation (4). Therefore $R(Q)|R(Q^{\text{end}}) = R(Q \;\&\; Q^{\text{end}}) = R(Q)$ for all $R(Q) \in \{R(Q^{\text{end}}), \ldots, R(Q^{\text{start}})\}$. □

Theorem 23 *The interior of $Y = \cup_{Q=Q^{start}}^{Q^{end}} \Omega(Q)$ is star-shaped, or Y is the union of two sets whose interiors are star-shaped.*

Proof Let $\tilde{q}$ be the most significant bits common to all of $\{Q^{\text{start}}, \ldots, Q^{\text{end}}\}$. We can split the segment into $\{Q^{\text{start}}, \ldots, (\tilde{q}011\ldots1)_2\}$ and $\{(\tilde{q}100\ldots0)_2, \ldots, Q^{\text{end}}\}$. The interior of the domain of the first segment is star-shaped by Corollary 22; the interior of the domain of the second segment is star-shaped by Corollary 21. □

4.3 From Uniform to Adaptive Meshes

We have completed the necessary proofs for a uniform space division, in the case of cubical refinement for any space dimension d, and previously for triangular and tetrahedral refinement (see Sect. 3.2). As we state in this section, an adaptive space division does not require any more effort (see also [3, page 176]).

Proof (Proof of Theorems 1 and 2) Any adaptive tree of quadrants with level $\le L$ can be refined into level L quadrants exclusively. This operation does not change the connectivity between boundaries of the designated subdomain. In particular, the number of face-connected subdomains remains unchanged and the proof reduces to applying Propositions 9, 14 (only $d \le 3$) or 18 (any d) above.

4.4 From one Tree to a Forest

A forest of octrees as used in [4,8,20] refers to a collection of tree roots, each understood as a geometric primitive, that form a tessellation of a possibly non-trivial domain. Inside each root, understood as a coarse element, a space-filling curve may be defined. While the tessellation of roots (the coarse mesh) remains fixed throughout a simulation, each tree may be refined and coarsened adaptively to build the fine (forest) mesh.

We define a space-filling curve on the whole forest by concatenating the curves of individual trees. Thus, a contiguous segment of the Morton curve may traverse more

than one tree. In this case, the segment necessarily contains the last subquadrant of any predecessor tree as well as the first subquadrant of any successor tree in the segment. In the cubical case, we know by Proposition 17 that no jumps can occur at all (when not counting the transition between two successive trees as a jump). For the simplicial case, we may use Lemmas 11 and 13 to obtain the bounds $L + 1$ (2D) and $2L + 1$ (3D).

We note that it is not known a-priori how the coarse elements are numbered and which of them are face-connected and how they are rotated against their neighbors, thus we cannot make any further statements about the connectedness between any two trees.

5 Enumeration of Face-Connected Segments

We would like to examine not only how many pieces an SFC segment can have, but also how frequently segments of different numbers of pieces occur. To this end, we propose a theoretical lower bound for the cubical case and supply numerical studies for both cubical and tetrahedral SFCs.

5.1 Lower Bound on Fraction of Face-Connected Segments

Theorem 24 *The fraction of face-connected segments of length l of the level L d-dimensional, cubical Morton curve is*

$$\phi_{d,L,l} \geq \frac{1}{2^d - 1}. \tag{15}$$

Proof Let $s = \{Q^{\text{start}}, \ldots, Q^{\text{end}} = Q^{\text{start}} + l - 1\}$ be the first disconnected segment of length l. As in the proof of Theorem 23, s divides into two connected segments, $\{Q^{\text{start}}, \ldots, (\tilde{q}011\ldots1)_2\}$ and $\{(\tilde{q}100\ldots0)_2, \ldots, Q^{\text{end}}\}$, where $\tilde{q}$ are significant bits that are common to all numbers in the segment. By the same reasoning as used in Corollary 21, the shapes of the two segments are not affected by $\tilde{q}$: changing $\tilde{q}$ translates the whole domain. Therefore, as this is the first disconnected segment, $\tilde{q}$ must be $(0\ldots0)$, i.e., the two connected pieces of the segment are $s^- = \{Q^{\text{start}}, \ldots, 2^k - 1\}$ and $s^+ = \{2^k, \ldots, Q^{\text{end}}\}$ for some k. Since Q^{end} cannot have a more significant 1-bit than 2^k, $Q^{\text{end}} \leq 2^{k+1} - 1$. We note that bitwise negation of the first $k + 1$ bits, $\tilde{R}(Q) := 2^{k+1} - 1 - Q$, induces a map $\Omega(Q) \mapsto \Omega(\tilde{R}(Q))$ that is the reflection about the midpoint of the box formed by $\cup_{Q=0}^{2^{k+1}-1} \Omega(Q)$.

We first want to find a lower bound for Q^{start}. Let $\Omega(Q^*)$ be a face-adjacent neighbor of $\Omega(2^k)$: because s^- and s^+ are disconnected, $Q^* \notin s^-$, so if $Q^* < 2^k$, then by the definition of s^-, $Q^* < Q^{\text{start}}$.

Let $j = k \mod d$ and let $j' = d - j - 1$. The neighbor $\Omega(Q^*)$ of $\Omega(2^k)$ that plays the crucial role is the neighboring quadrant that is closer to the origin in the $(j + 1)$th direction. To identify Q^*, we have to subtract one from the $(j + 1)$th coordinate of 2^k: i.e., $(2^k)_{j+1} = 2^{\lfloor k/d \rfloor}$ in the notation of (12), so $Q^*_{j+1} = 2^{\lfloor k/d \rfloor} - 1$, while the other

coordinates are the same, viz. zero. Therefore

$$\begin{aligned} Q^* &= (0\ldots0\overbrace{\underbrace{0\ldots0}_{j'\text{times}}1\underbrace{0\ldots0}_{j\text{times}}}^{\lfloor k/d\rfloor\text{times}})_2 & &= \sum_{i=1}^{\lfloor k/d\rfloor} 2^{(i-1)d+j} \\ &= \sum_{i=1}^{\lfloor k/d\rfloor} 2^{(\lfloor k/d\rfloor-i)d+j} & &= 2^k \sum_{i=1}^{\lfloor k/d\rfloor} 2^{-id} \\ &= 2^k\left(\sum_{i=1}^{\infty}2^{-id} - \sum_{i=\lfloor k/d\rfloor+1}^{\infty}2^{-id}\right) & &= 2^k\frac{1}{2^d-1} - 2^k\sum_{i=\lfloor k/d\rfloor+1}^{\infty}2^{-id} \\ &= 2^k\frac{1}{2^d-1} - 2^j\sum_{i=1}^{\infty}2^{-id} & &= \frac{2^k-2^j}{2^d-1}. \end{aligned} \tag{16}$$

By the definition of Q^{start}, there is a connected segment of length l that starts at each $Q \in \{0, \ldots, Q^*\}$. For each of these, there is another connected segment, obtained by the reflection map $\tilde{R}$, that ends with $Q \in \{\tilde{R}(Q^*), \ldots, 2^{k+1}-1\}$. We want to show that these two sets of connected segments are distinct, i.e., that there is no segment of length l that starts with $Q \le Q^*$ and ends with $Q + l - 1 \ge \tilde{R}(Q^*)$. We thus have to show that the shortest segment with endpoints in each set, $\{Q^*, \ldots, \tilde{R}(Q^*)\}$ is longer than l, i.e., $l < \tilde{R}(Q^*) - Q^* + 1 = 2^{k+1} - 2Q^*$.

We will prove this bound by finding an upper bound for Q^{end}. We note that $\tilde{R}(2^k) = 2^k - 1 \in s^-$, so $\Omega(\tilde{R}(Q^*))$ is a face neighbor of $\Omega(2^k-1)$, so by the same reasoning as above, $\tilde{R}(Q^*)$ must be greater than Q^{end}, and thus

$$l = Q^{\text{end}} - Q^{\text{start}} + 1 \le (\tilde{R}(Q^*) - 1) - (Q^* + 1) + 1 = 2^{k+1} - 2Q^* - 2. \tag{17}$$

We have shown that there are at least $2(Q^*+1)$ connected segments in the first 2^{k+1} segments, each of which begins and ends in the range $\{0, \ldots, 2^{k+1}-1\}$. None of the numbers in these segments has more than $k+1$ significant bits, so by the same reasoning as in Corollary 21, adding a multiple of 2^{k+1} to each number in one of these segments is a translation of its domain and preserves its connectedness. Therefore there are at least $2(Q^*+1)$ connected segments for *every* 2^{k+1} segments, and thus

$$\begin{aligned} \phi_{d,L,l} &\ge \tfrac{2(Q^*+1)}{2^{k+1}} = \tfrac{Q^*+1}{2^k} \\ &= \tfrac{2^k-2^j}{2^k(2^d-1)} + \tfrac{1}{2^k} \\ &= \tfrac{2^k-2^j+2^d-1}{2^k(2^d-1)} \\ &\ge \tfrac{2^k}{2^k(2^d-1)} \qquad (j < d \text{ by definition}) \\ &= \tfrac{1}{2^d-1}. \end{aligned} \tag{18}$$

□

5.2 Computational Studies: Cubical Morton Curve

Having shown that a segment of a Morton curve is composed of one or two face-connected subdomains, a natural question to ask is how many of each type there are. More formally, we ask: for a given dimension d, recursive level L, and segment length l, what fraction $\phi_{d,L,l}$ of the $2^{dL} - (l-1)$ possible segments are in one face-connected piece?

This question can be answered recursively. Every segment s^h on level $(L+1)$ can be associated with a segment s^H on level L, such that the end quadrants $\{q^h_{\text{start}}, q^h_{\text{end}}\}$ of s^h are contained in the end quadrants $\{q^H_{\text{start}}, q^H_{\text{end}}\}$ of s^H, as illustrated in Fig. 12. This means that each s^H with length $l \geq 1$ is associated with 2^{2d} segments (corresponding to each pairing of a child of q^H_{start} with a child of q^H_{end}): we say that s^H "refines" to the set $\{s^h\}$ of its associations.

Clearly each s^h associated with it cannot be longer than $2^d l$; because $s^h \backslash \{q^h_{\text{start}}, q^h_{\text{end}}\}$ covers $s^h \backslash \{q^h_{\text{start}}, q^h_{\text{end}}\}$, the length of s^h is at least $2^d(l-2)+2$.

We divide connected segments into two categories: weakly connected, when the first and last quadrants in the segment are (face-)adjacent, and strongly connected, when they are not. Disconnected segments only refine to disconnected segments. Strongly connected segments only refine to strongly connected segments. Weakly connected segments refine to all three types: how many of each depends on the direction in which the first and last quadrants are adjacent.

We give pseudocode for this recursive calculation in the function Enumerate (Algorithm 1). This algorithm is implemented in the Python script `morton.py`. [1]

Enumerate calls on some lookup tables: RefineOne(d,l) (Fig. 13) counts how many disconnected, strongly and weakly connected segments of length l are refined from one d-dimensional quadrant (the weakly connected segments are broken down by the direction in which the first and last quadrant are adjacent); RefineWeak(d,j,r) (Fig. 14) counts how many disconnected, strongly and weakly connected segments with r quadrants in the end-families are refined from one weakly connected segment in direction j (a weakly connected segment only produces weakly connected segments in the same direction).

In Fig. 15, we use Enumerate to calculate the fraction of connected segments $\phi_{d,L,l}$ for $d = 2$ and $d = 3$ for large values of L. We observe that $\phi_{d,L,l}$ tends to vary between $1/2$ and $1/(2^d - 1)$. (We had proved in Sect. 5.1 that the latter is indeed a lower bound.)

5.3 Computational Studies: Simplicial Morton Curve

In the preceding section, we investigate the fraction of connected hypercube SFC segments of a given particular length l among all segments of length l. For simplices, we enumerate all possible SFC segments for a given uniform refinement level and compute the number of their face-connected components. We achieve this by performing a depth-first search on the connectivity graph of the submesh generated by the segment. [2]

[1] https://github.com/cburstedde/p4est/tree/develop/doc/morton/morton.py.

[2] https://github.com/holke/sfc_conncomp.

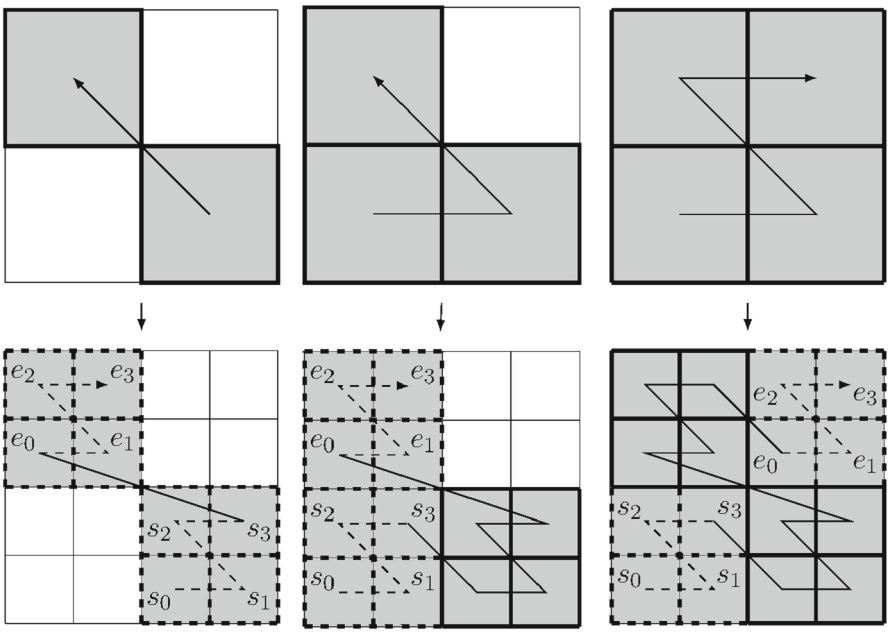
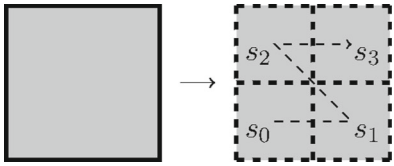


Fig. 12 The refinement of disconnected (left), weakly connected (middle), and strongly connected segments (right). Each coarse segment (top) refines to one of $2^{2d} = 16$ possible refined segments (bottom), each starting with $s \in \{s_0, \ldots, s_3\}$ and ending with $e \in \{e_0, \ldots, e_3\}$. Since either of these start and end segments is a possibility, the lines are dashed. Disconnected segments refine to disconnected segments; strongly connected segments refine to strongly connected segments; weakly connected segments refine to disconnected (e.g., $\{s_3, \ldots, e_0\}$), weakly connected (e.g., $\{s_2, \ldots, e_0\}$), and strongly connected segments (e.g., $\{s_0, \ldots, e_3\}$)



RefineOne(2,1)	$n_s = 4$ $(\{s_0\}, \{s_1\}, \{s_2\}, \{s_3\})$
RefineOne(2,2)	$n_d = 1$ $(\{s_1, s_2\})$, $n_{w,1} = 2$ $(\{s_0, s_1\}, \{s_2, s_3\})$
RefineOne(2,3)	$n_{w,2} = 2$ $(\{s_0, s_1, s_2\}, \{s_1, s_2, s_3\})$
RefineOne(2,4)	$n_s = 1$ $(\{s_0, s_1, s_2, s_3\})$

Fig. 13 We list the RefineOne(d,l) tables used in Enumerate (Algorithm 1) for $d = 2$ as an example. Unlisted values are zero

RefineWeak(2,2,2)	$n_d = 1$ $(\{s_3, \ldots, e_0\})$
RefineWeak(2,2,3)	$n_{w,2} = 2$ $(\{s_2, \ldots, e_0\}, \{s_3, \ldots, e_1\})$
RefineWeak(2,2,4)	$n_s = 3$ $(\{s_1, \ldots, e_0\}, \{s_2, \ldots, e_1\}, \{s_3, \ldots, e_2\})$
RefineWeak(2,2,5 $\le r \le$ 8)	$n_s = 9 - r$ $(\{s_0, \ldots, e_{r-5}\}, \ldots, \{s_{8-r}, \ldots, e_3\})$

Fig. 14 We list the RefineWeak(d,j,r) tables used in Enumerate (Algorithm 1) for $d = 2$ and $j = 2$ as an example. The start points and end points refer to Fig. 12 (middle). Unlisted values are zero

Algorithm 1: Enumerate (d, L, l)

Data: dimension $d \geq 1$, level $L \geq 0$, segment length $1 \leq l \leq 2^{dL}$.
Result: $(n_d, n_s, n_{w,1}, \ldots, n_{w,d})$, the number of segments of length l that are disconnected, strongly connected, and weakly connected in each direction.

if $l = 1$ **or** $L = 0$ **then** [define segments of length 1 to be strongly connected]
  return $(0, 2^{dL}, 0, \ldots, 0)$ [one segment for each quadrant]
end
$(n_d, n_s, n_{w,1}, \ldots, n_{w,d}) \leftarrow (0, \ldots, 0)$
$c \leftarrow \lceil l/2^d \rceil$ [compute the shortest length that can refine to length l]
if $l \mod 2^d < 2$ **then** $C \leftarrow 1 + \lfloor l/2^d \rfloor$ **else** $C \leftarrow 2 + \lfloor l/2^d \rfloor$ […longest …]
for $k \in \{c, \ldots, C\}$ **do**
  if $k = 1$ **then**
    $N \leftarrow 2^{d(L-1)}$ [# of coarse quadrants]
    $(n_d, n_s, n_{w,1}, \ldots, n_{w,d}) \leftarrow (n_d, n_s, n_{w,1}, \ldots, n_{w,d}) + N *$ RefineOne (d,l)
  else
    $r \leftarrow l - (k-2)2^d$ [# of children in two coarse end quadrants]
    $m \leftarrow \min\{(r-1), 2^{d+1} - (r-1)\}$ [# of ways to split between two families]
    $(N_d, N_s, N_{w,1}, \ldots, N_{w,d}) \leftarrow$ Enumerate $(d, L-1, k)$
    $n_d \leftarrow n_d + N_d * m$ [disconnected → disconnected]
    $n_s \leftarrow n_s + N_s * m$ [strongly connected → strongly connected]
    for $1 \leq j \leq d$ **do**
      $(n_d, n_s, n_{w,j}) \leftarrow (n_d, n_s, n_{w,j}) + N_{w,j} *$ RefineWeak (d, j, r)
    end
  end
end
return $(n_d, n_s, n_{w,1}, \ldots, n_{w,d})$

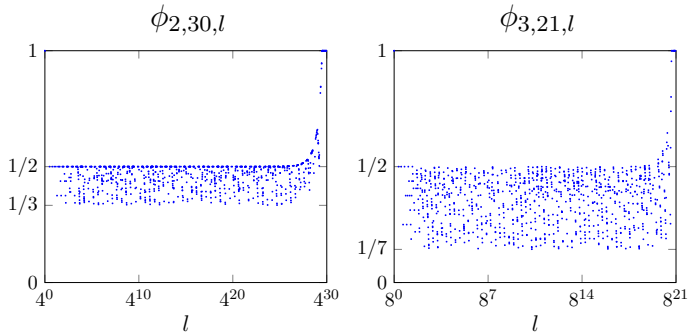


Fig. 15 We plot the fraction of face-connected segments of length l, $\phi_{d,L,l}$, for $d = 2$ and $L = 30$, (left) and $d = 3$ and $L = 21$ (right), for one thousand log-uniformly randomly sampled lengths

By binning the lengths occurring by powers of two, we obtain a diagram that resembles that for hypercubes; see Fig. 16. A comparison of the fractions of connected vs. non-connected segments between hypercubes and simplices is given in Table 2.

Figure 17 provides an alternative perspective by binning by the number of components. This is motivated by the observation that all possible lengths of SFC segments can occur in an application. On the one hand, we could have a forest consisting of a single tree. If the number of participating processes is of the same magnitude than the

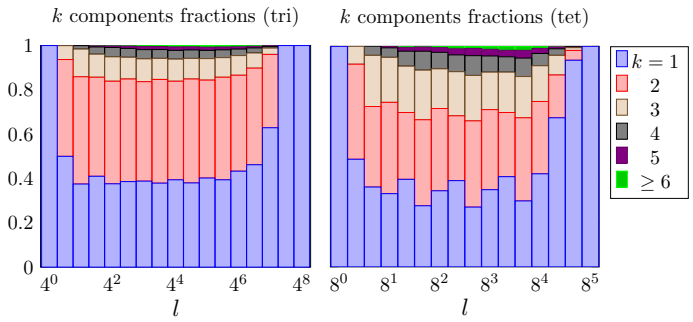


Fig. 16 We take all simplicial Morton segments of length $2^j \leq l < 2^{j+1}$ and compute the fraction that have k face-connected components. These data are like the data for the hypercube Morton curve in Fig. 15, but averaged in bins over l, and with $k > 2$ appearing. Left: uniform level 8 refined triangle. Right: level 5 refined tetrahedron

Table 2 The relative counts of connected and non-connected segments across all possible SFC segments of a uniform level 5 and level 8 (2D only) refinement

	Level 5				Level 8	
	Quads	Cubes	Triangles	Tets	Quads	Triangles
Connected	71.6%	60.0%	63.9%	61.0%	71.4%	63.7%
Non-connected	28.4%	40.0%	36.1%	39.0%	28.6%	36.3%

number of elements in that tree, then very short segments can occur, possibly even segments consisting only of a single element. On the other hand, consider a setting where we have many trees, possibly as many or more trees than processes. In this case, the lengths of SFC segments within a single tree can be large (on the order of the number of elements in one tree).

6 Conclusion

We prove in this document that the classical Morton or z-curve does not lead to a fragmentation of the root cube into more than two disconnected subdomains. Its loss of continuity in comparison with the Hilbert curve is thus controlled. This is in line with experimental results that establish the suitability of the Morton curve for numerical applications.

We show that the bound for the recently proposed tetrahedral Morton cube is roughly $2L$ and thus growing with the level of refinement. Yet, we can demonstrate numerically that the fraction of connected to non-connected segments is close to the cubical case. Even though it is still open whether our bounds for simplicial tessellations can be improved by considering other families of space-filling curves, we expect the simplicial and hypercube approaches to behave similarly.

Our result would appear relevant to make informed choices about the type of space-filling curve to use, for example in writing a new element-based parallel code for the

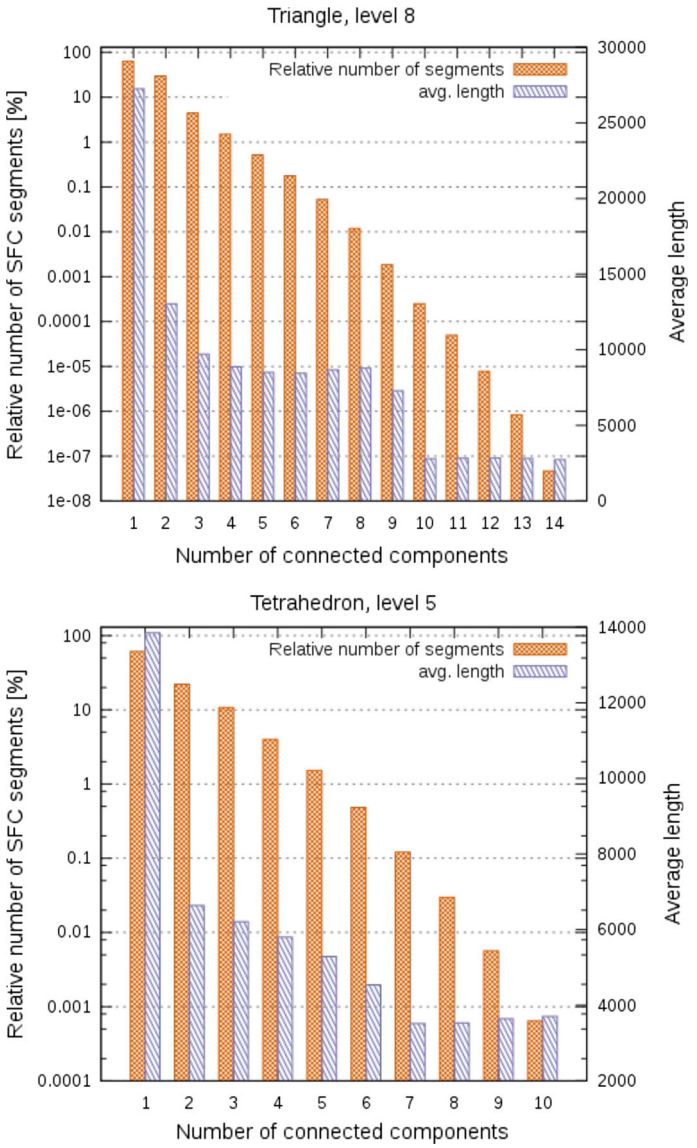


Fig. 17 The relative count of SFC segments by number of connected components and the average length (right y-axis) of these segments. Left: the distribution for a uniform level 8 refined triangle. We observe that almost 98% of all SFC segments have three connected components or less. 63.7% are connected, 29.7% have two connected components and 4.4% have three connected components. Right: the distribution for a uniform level 5 refined tetrahedron. Here, more than 93% of the segments have three connected components or less with 61.0% having exactly one connected component, 22.1% with two connected components and 10.7% with three connected components. The highest number of segments occurring are $14 = 2(8-1)$ in 2D and $10 = 2 \times 5$ in 3D. This is in agreement with Proposition 14 (2D) and Conjecture 15 (3D)

numerical solution of partial differential equations, or any other code that benefits from a recursive subdivision of space. It should be noted that the number of components is not necessarily the only metric of quality, since the relative sizes of the connected components and the number of individual facets matter to some applications, too. So far however, our theory and experiments support the existing numerical evidence that a fragmentation of the parallel partition is not observed.

Acknowledgements Burstedde would like to thank Andreas Dedner for the invitation to the ICMS workshop on Galerkin methods with applications in weather and climate forecasting, which provided motivation to get going proving this conjecture. The authors would like to thank Michael Bader and Herman Haverkort for suggesting additional relevant literature. Burstedde and Holke acknowledge travel support by the Hausdorff Center for Mathematics (HCM) at Bonn University funded by the German Research Foundation (DFG). Isaac gratefully acknowledges the support of the Intel Parallel Computing Center at the University of Chicago. Holke gratefully acknowledges the scholarship support by the Bonn International Graduate School for Mathematics (BIGS) as part of HCM. We would also like to thank two anonymous referees for their thoughtful and constructive comments.

References

1. Ahimian, A., Lashuk, I., Veerapaneni, S., Aparna, C., Malhotra, D., Moon, I., Sampath, R., Shringarpure, A., Vetter, J., Vuduc, R., Zorin, D., Biros, G.: Petascale direct numerical simulation of blood flow on 200k cores and heterogeneous architectures. In: SC10: Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis. ACM/IEEE (2010)
2. Akçelik, V., Bielak, J., Biros, G., Epanomeritakis, I., Fernandez, A., Ghattas, O., Kim, E.J., Lopez, J., O'Hallaron, D.R., Tu, T., Urbanic, J.: High resolution forward and inverse earthquake modeling on terascale computers. In: SC03: Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis. ACM/IEEE (2003)
3. Bader, M.: Space-Filling Curves: An Introduction with Applications in Scientific Computing. Texts in Computational Science and Engineering. Springer (2012)
4. Bangerth, W., Hartmann, R., Kanschat, G.: deal.II – a general-purpose object-oriented finite element library. ACM Transactions on Mathematical Software **33**(4), 24 (2007). https://doi.org/10.1145/1268776.1268779
5. Bey, J.: Der BPX-Vorkonditionierer in drei Dimensionen: Gitterverfeinerung, Parallelisierung und Simulation. Universität Heidelberg (1992). Preprint
6. Burstedde, C., Ghattas, O., Gurnis, M., Isaac, T., Stadler, G., Warburton, T., Wilcox, L.C.: Extreme-scale AMR. In: SC10: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. ACM/IEEE (2010)
7. Burstedde, C., Holke, J.: A tetrahedral space-filling curve for nonconforming adaptive meshes. SIAM Journal on Scientific Computing **38**(5), C471–C503 (2016). http://dx.doi.org/10.1137/15M1040049
8. Burstedde, C., Wilcox, L.C., Ghattas, O.: `p4est`: Scalable algorithms for parallel adaptive mesh refinement on forests of octrees. SIAM Journal on Scientific Computing **33**(3), 1103–1133 (2011). 10.1137/100791634
9. de Berg, M., Haverkort, H., Thite, S., Toma, L.: Star-quadtrees and guard-quadtrees: I/O-efficient indexes for fat triangulations and low-density planar subdivisions. Computational Geometry **43**, 493–513 (2010)
10. Finkel, R.A., Bentley, J.L.: Quad trees A data structure for retrieval on composite keys. Acta Informatica **4**, 1–9 (1974)
11. Griebel, M., Zumbusch, G.W.: Parallel multigrid in an adaptive PDE solver based on hashing and space-filling curves. Parallel Computing **25**, 827–843 (1999)
12. Haverkort, H.: Sixteen space-filling curves and traversals for d-dimensional cubes and simplices. arXiv preprint arXiv:1711.04473 (2017). http://arxiv.org/abs/1711.04473
13. Haverkort, H., van Walderveen, F.: Locality and bounding-box quality of two-dimensional space-filling curves. Computational Geometry **43**(2), 131–174 (2010)

14. Hilbert, D.: Über die stetige Abbildung einer Linie auf ein Flächenstück. Mathematische Annalen **38**, 459–460 (1891)
15. Lebesgue, H.L.: Leçons sur l'intégration et la recherche des fonctions primitives. Gauthier-Villars (1904)
16. Meagher, D.: Geometric modeling using octree encoding. Computer Graphics and Image Processing **19**(2), 129–147 (1982). 10.1016/0146-664X(82)90104-6
17. Morton, G.M.: A computer oriented geodetic data base; and a new technique in file sequencing. Tech. rep., IBM Ltd. (1966)
18. Peano, G.: Sur une courbe, qui remplit toute une aire plane. Math. Ann. **36**(1), 157–160 (1890)
19. Sagan, H.: Space-Filling Curves. Springer (1994)
20. Stewart, J.R., Edwards, H.C.: A framework approach for developing parallel adaptive multiphysics applications. Finite Elements in Analysis and Design **40**(12), 1599–1617 (2004). 10.1016/j.finel.2003.10.006
21. Tropf, H., Herzog, H.: Multidimensional range search in dynamically balanced trees. Angewandte Informatik **2**, 71–77 (1981)
22. Weinzierl, T., Mehl, M.: Peano—a traversal and storage scheme for octree-like adaptive Cartesian multiscale grids. SIAM Journal on Scientific Computing **33**(5), 2732–2760 (2011). http://link.aip.org/link/?SCE/33/2732