SN Computer Science (2023) 4:8
https://doi.org/10.1007/s42979-022-01320-9

ORIGINAL RESEARCH

Clustering Analyses of Two-Dimensional Space-Filling Curves: Hilbert and *z*-Order Curves

H. K. Dai[1] · **H. C. Su**[2]

Received: 15 April 2022 / Accepted: 27 June 2022 / Published online: 15 October 2022

Abstract
A discrete space-filling curve provides a linear traversal or indexing of a multi-dimensional grid space. This paper presents two analytical studies on clustering analyses of the 2-dimensional Hilbert and z-order curve families. The underlying measure is the mean number of cluster over all identically shaped subgrids. We derive the exact formulas for the clustering statistics for the 2-dimensional Hilbert and z-order curve families. The two exact analytical results yield: (1) a comparison of their relative performances with respect to this measure: when the grid-order is sufficiently larger than the subgrid-order (typical scenario for most applications), Hilbert curve family performs significantly better than z-order curve family, and (2) good agreements with asymptotic analyses of continuous space-filling curves and of (non-continuous) z-order curve in the literature.

Keywords Index structure · Space-filling curve · Hilbert curve · z-Order curve · Clustering

Preliminaries

Discrete space-filling curves have a wide range of applications in databases, parallel computation, algorithms, in which linearization techniques of multi-dimensional arrays or computational grids are needed. Sample applications include heuristics for combinatorial algorithms and data structures: traveling salesperson algorithm [22] and nearest-neighbor finding [6], multi-dimensional space-filling indexing methods [2, 5, 11, 17], image compression [18], dynamic unstructured mesh partitioning [15], and linearization and traversal of sensor networks [4, 24]. Some recent diverse applications of space-filling curves extend to statistical sampling [13] and bioinformatics [16]. For a comprehensive historical development of classical space-filling curves, see [3, 23].

For positive integer n, denote $[n] = \{1, 2, \ldots, n\}$. An m-dimensional (discrete) space-filling curve of length n^m is a bijective mapping $C : [n^m] \to [n]^m$, thus providing a linear indexing/traversal or total ordering of the grid points in $[n]^m$. An m-dimensional grid is said to be of order k if it has side-length $n = 2^k$; a space-filling curve has order k if its codomain is a grid of order k. An m-dimensional space-filling curve C is continuous if the Euclidean distance between $C(i)$ and $C(i+1)$ is 1 for all $i \in [n^m - 1]$. The generation of a sequence of multi-dimensional space-filling curves of successive orders usually follows a recursive framework (on the dimensionality and order), which results in a few classical families, such as Gray-coded curves, Hilbert curves, Peano curves, and z-order curves (see, for examples, [1, 20]).

Denote by H_k^m and Z_k^m an m-dimensional Hilbert and z-order, respectively, space-filling curve of order k. Figure 1 illustrates the recursive constructions of H_k^m and Z_k^m for $m = 2$ and $k = 1, 2$, and $m = 3$ and $k = 1$.

We measure the applicability of a family of space-filling curves based on their common structural characteristics: clustering, inter-clustering, and locality, which are informally described as follows. Clustering performance measures the distribution of continuous runs of grid points

This article is part of the topical collection "Future Data and Security Engineering 2021" guest edited by Tran Khanh Dang.

✉ H. K. Dai
dai@cs.okstate.edu

H. C. Su
suh@astate.edu

[1] Computer Science Department, Oklahoma State University, Stillwater, OK 74078, USA

[2] Department of Computer Science, Arkansas State University, Jonesboro, AR 72401, USA

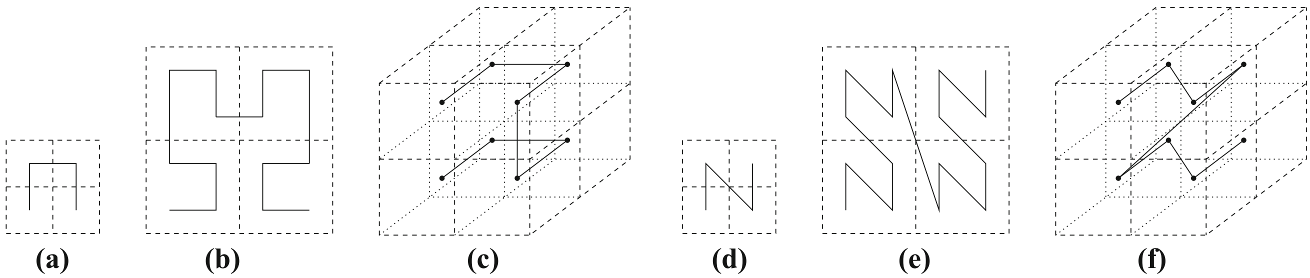


Fig. 1 Recursive self-similar generations of Hilbert and z-order curves of higher order (respectively, H_k^m and Z_k^m) by interconnecting symmetric subcurves, via reflection and/or rotation, of lower order (respectively, H_{k-1}^m and Z_{k-1}^m) along an order-1 subcurve (respectively, H_1^m and Z_1^m): **a** H_1^2; **b** H_2^2; **c** H_1^3; **d** Z_1^2; **e** Z_2^2; **f** Z_1^3

(clusters) over identically shaped subspaces of $[n]^m$, which can be characterized by the mean number of clusters and the average inter-cluster distance (in $[n^m]$) within a subspace. Locality preservation reflects proximity between the grid points of $[n]^m$, that is, close-by points in $[n]^m$ are mapped to close-by indices/numbers in $[n^m]$, or vice versa.

For an m-dimensional space-filling curve $C : [n^m] \to [n]^m$ and a subgrid G of $[n]^m$, a cluster of G induced by C is a maximal contiguous subinterval I of $[n^m]$ such that $C(I) \subseteq G$. We can partition and order $C^{-1}(G)$ into disjoint union of clusters. An inter-cluster gap of G is a contiguous subinterval of $[n^m]$ delimited by two consecutive clusters of G, and the corresponding inter-cluster distance is the length of the inter-cluster gap. Thus, the space-filling curve C induces the following statistics: (1) average clustering: the mean number of clusters of $C^{-1}(G)$ over all identically shaped subgrids G of $[n]^m$, (2) average inter-clustering: the (universe) mean inter-cluster distance over all inter-cluster gaps from all identically shaped subgrids G of $[n]^m$, and (3) average total inter-clustering: the mean total inter-cluster distance (in a subgrid) over all identically shaped subgrids G of $[n]^m$.

Empirical and analytical studies of clustering and inter-clustering performances of various low-dimensional space-filling curves have been reported in the literature (see [20] for details). Generally, the Hilbert and z-order curve families exhibit good performance in these respects.

Related Work in Clustering, Inter-clustering, and Locality

Jagadish [14] derives exact formulas for the mean numbers of clusters over all rectangular 2×2 and 3×3 subgrids of an H_k^2-structural grid space. Moon et al. [20] prove that in a sufficiently large m-dimensional H_k^m-structural grid space, the asymptotic mean number of clusters over all rectilinear polyhedral queries with common surface area $S_{m,k}$ approaches $\frac{1}{2}\frac{S_{m,k}}{m}$ as k approaches ∞. They also extend the work in [14] to obtain the exact formula for the mean number of clusters over all rectangular $2^q \times 2^q$ subgrids of an H_k^2-structural grid space.

Xu and Tirthapura [25] generalize the above asymptotic mean number of clusters over all rectilinear polyhedral queries with common surface area from m-dimensional Hilbert curves to arbitrary continuous space-filling curves. Note that rectangular queries with common volume yield the optimal asymptotic mean number of clusters for a continuous space-filling curve.

Dai and Su [7] obtain the exact formulas for the following three statistics for H_k^2 and Z_k^2: (1) the summation of all inter-cluster distances over all $2^q \times 2^q$ query subgrids, (2) the universe mean inter-cluster distance over all inter-cluster gaps from all $2^q \times 2^q$ subgrids, and (3) the mean total inter-cluster distance over all $2^q \times 2^q$ subgrids. Based on the analytical results, the asymptotic comparisons indicate that z-order curve family performs better than Hilbert curve family with respect to the statistics.

A few locality measures have been proposed and analyzed for space-filling curves in the literature. Denote by d and d_p the Euclidean metric and p-normed metric (rectilinear ($p = 1$) and maximum metric ($p = \infty$)), respectively.

Let $\mathcal{C}$ denote a family of m-dimensional curves of successive orders. For quantifying the proximity preservation of close-by points in the m-dimensional space $[n]^m$, Pérez et al. [21] employ an average locality measure:

$$L_{\mathrm{PKK}}(C) = \sum_{i,j\in[n^m]|i<j} \frac{|i-j|}{d(C(i),C(j))} \quad \text{for } C \in \mathcal{C},$$

and provide a hierarchical construction for a 2-dimensional $\mathcal{C}$ with good but suboptimal locality with respect to this measure.

Mitchison and Durbin [19] use a more restrictive locality measure parameterized by q:

$$L_{\mathrm{MD,q}}(C) = \sum_{i,j\in[n^m]|i<j \text{ and } d(C(i),C(j))=1} |i-j|^q \quad \text{for } C \in \mathcal{C}$$

to study optimal 2-dimensional mappings for $q \in [0,1]$. For the case $q = 1$, the optimal mapping with respect to $L_{\mathrm{MD,1}}$ is very different from that in [21].

Dai and Su [8] consider a locality measure similar to $L_{\mathrm{MD},1}$ conditional on a 1-normed distance of δ between points in $[n]^m$:

$$L_\delta(C) = \sum_{i,j\in[n^m] \mid i<j \text{ and } d_1(C(i),C(j))=\delta} |i-j| \quad \text{for } C \in \mathcal{C}.$$

They derive exact formulas for L_δ for the Hilbert curve family $\{H_k^m \mid k=1,2,\ldots\}$ and z-order curve family $\{Z_k^m \mid k=1,2,\ldots\}$ for $m=2$ and arbitrary δ that is an integral power of 2, and $m=3$ and $\delta=1$. With respect to the locality measure L_δ and for sufficiently large k and $\delta \ll 2^k$, the z-order curve family performs better than the Hilbert curve family for $m=2$ and over the δ-spectrum of integral powers of 2. When $\delta=2^k$, the domination reverses. The superiority of the z-order curve family persists but declines for $m=3$ with unit 1-normed distance for L_δ.

For measuring the proximity preservation of close-by points in the indexing space $[n^m]$, Gotsman and Lindenbaum [12] consider the following measures for $C \in \mathcal{C}$:

$$L_{\mathrm{GL,min}}(C) = \min_{i,j\in[n^m] \mid i<j} \frac{d(C(i),C(j))^m}{|i-j|}, \quad \text{and}$$

$$L_{\mathrm{GL,max}}(C) = \max_{i,j\in[n^m] \mid i<j} \frac{d(C(i),C(j))^m}{|i-j|}.$$

They show that for arbitrary m-dimensional curve C,

$$L_{\mathrm{GL,min}}(C) = O(n^{1-m}), \quad \text{and}$$

$$L_{\mathrm{GL,max}}(C) > (2^m-1)(1-\frac{1}{n})^m.$$

For the m-dimensional Hilbert curve family $\{H_k^m \mid k=1,2,\ldots\}$, they prove that:

$$L_{\mathrm{GL,max}}(H_k^m) \le 2^m(m+3)^{\frac{m}{2}}.$$

Alber and Niedermeier [1] generalize $L_{\mathrm{GL,max}}$ to $L_{\mathrm{AN,p}}$ by employing the p-normed metric d_p in place of the Euclidean metric d. They improve and extend the above tight bounds for the 2-dimensional Hilbert curve family to:

$$L_{\mathrm{AN},1}(H_k^2) \le 9\frac{3}{5},$$

$$6(1-O(2^{-k})) \le L_{\mathrm{AN},2}(H_k^2) \le 6\frac{1}{2}, \quad \text{and}$$

$$6(1-O(2^{-k})) \le L_{\mathrm{AN},\infty}(H_k^2) \le 6\frac{2}{5}.$$

Dai and Su [9, 10] provide analytical studies on the locality measure $L_{\mathrm{AN,p}}$ for the 2-dimensional Hilbert curve family, and obtain exact formulas for $L_{\mathrm{AN,p}}(H_k^2)$ for $p=1$ and all reals $p \ge 2$. In addition, they identify all the representative grid-point pairs (which realize $L_{\mathrm{AN,p}}(H_k^2)$) for $p=1$ and all reals $p \ge 2$. A practical implication of their results on $L_{\mathrm{AN,p}}(H_k^2)$ is that the exact formulas provide good bounds on measuring the loss in data locality in the index space, while spatial correlation exists in the 2-dimensional grid space.

The studies of clustering and inter-clustering performances for space-filling curves are motivated by the applicability of multi-dimensional space-filling indexing methods, in which an m-dimensional data space is mapped onto a 1-dimensional data space (external storage structure) by adopting a 1-dimensional indexing method based on an m-dimensional space-filling curve. The space-filling index structure can support efficient query processing (such as range queries) provided that we minimize the average number of external fetch/seek operations, which is related to the clustering and inter-clustering statistics. An example study/application of the optimization of range queries over space-filling index structures is detailed in [2], which aims at minimizing the number of seek operations.

This paper presents two analytical studies on the clustering analyses of two representative 2-dimensional space-filling curve families: Hilbert and z-order. For each studied space-filling curve family, we derive the exact formulas for its clustering statistics that includes the mean number of clusters for the underlying curve family over all identically shaped rectangular subgrids of $[n]^2$. Note that our analytical and combinatorial approach in deriving the clustering statistics, uniform for both curve-families, are three-fold: (1) extending the work for the 2-dimensional Hilbert curve family in [20] while completing the relative-performance study of the 2-dimensional Hilbert and z-order curve families with respect to the clustering statistics, (2) complementing our earlier analytical and approximation studies [7] of inter-clustering performances of the 2-dimensional Hilbert and z-order curve families, and (3) more importantly, this common analytical approach provides us a natural geometric transition in computing the clustering statistics for their 3- and higher-dimensional curve families.

The computation of the mean-clustering statistics proceeds to establishing many systems of recurrences and their closed-form solutions over two geometric regions of 2-dimensional Hilbert and z-order curves (boundary edge- and vertex-regions). For each studied space-filling curve family, we provide the overall and successive stepwise ideas in developing the recurrence systems for the boundary edge- and vertex-regions and the final exact formula for the desired clustering statistics. Closed-form solutions for most systems of recurrences in our studies are solved analytically and/or are computed via the mathematical and analytical software Maple. Complete detailed formulations of all recurrence systems, derivations/proofs of their solutions, and verifying computer programs for Hilbert and z-order curve families are available from the authors.

Our Analytical and Combinatorial Approach

For an H_k^2- and Z_k^2-structural grid spaces, we obtain the exact formulas for the mean number of clusters over all rectangular $2^q \times 2^q$ subspaces by computing the edge cuts in and between its subgrids that are decomposed recursively. The idea behind this derivation is to count the total number of edges that are cut by the sides of all possible identically shaped $2^q \times 2^q$ subspaces—by noting that the entry and exit grid points of a cluster connect to grid points outside of this subspace (two cuts by side(s) of this subspace) and every cluster has two cuts by the subspace. A cut on an edge by a side of a subspace is called an "edge cut". We give an overview of the derivation for both curve families as follows.

1. Compute the number of edges that are cut by the sides of all possible $2^q \times 2^q$ subspaces, which are exactly inside of one of the four quadrants, and the number of edges cut by the sides of subspaces across different quadrants, respectively. The number of cuts on edges is twice the number of clusters over all identically shaped subspaces.
2. Categorize the edge cuts caused by subspaces across quadrants into: edge cuts within $2^q \times 2^q$ corner boundaries of the quadrants and within side boundaries (of 2^q rows/columns) of the quadrants. (Note that the edges that are cut by subspaces across quadrants only in the boundary regions (sides and corners) of the quadrants.) By decomposing these corner and side boundaries, we derive recurrences for each of them:
   (a) For edge cuts within one of the four corner boundaries: edge cuts within upper (left or right) corner and lower (left or right) corner boundaries are inter-recurrence related.
   (b) For edge cuts within one of the four side boundaries:
      (i) Edge cuts in left boundary (same as right boundary) consists of substructures of left boundary, bottom boundary and two lower-corner boundaries,
      (ii) Edge cuts in bottom boundary consists of substructures of two left boundaries and two upper-corner boundaries, and
      (iii) Edge cuts in the top boundary consists of substructures of two top boundaries and two upper-corner boundaries. These inter-recurrent relations are based upon the construction of canonical H_k^2 and Z_k^2 (see Fig. 1a, b).
3. After obtaining the numbers of edge cuts in boundaries that are derived from step 2, we solve the recurrence for the total number of edge cuts, then divide it by 2 to get the total number of clusters. To obtain the mean number of clusters, the total number of clusters needs to be divided by the total number of subspaces of size $2^q \times 2^q$, which is $(2^k - 2^q + 1)^2$.

Clustering Statistics of 2-Dimensional Hilbert Curves

With respect to the canonical orientation of H_k^2 shown in Fig. 1a, we cover the 2-dimensional k-order grid with 2^k rows $(R_{k,1}, R_{k,2}, \ldots, R_{k,2^k})$, indexed from the bottom, and 2^k columns $(C_{k,1}, C_{k,2}, \ldots, C_{k,2^k})$, indexed from the left. We denote:

1. The x-y coordinate system addressed in our study is the usual Cartesian x-y coordinate system: horizontal x-axis and vertical y-axis.
2. The $(+\frac{\pi}{2})$-rotation and $(-\frac{\pi}{2})$-rotation correspond to the 90°-clockwise rotation from vertical y-axis to horizontal x-axis and the 90°-counterclockwise from horizontal x-axis to vertical y-axis, respectively.
3. For a grid point $v \in [2^k]^2$, its x- and y-coordinate by $X(v)$ and $Y(v)$, respectively (that is, v is the intersection grid point of the row $R_{k,Y(v)}$ and the column $C_{k,X(v)}$),
4. For a rectangular query subgrid with its lower-left corner at grid point (x, y) and upper-right corner at grid point (x', y') $(1 \leq x \leq x' \leq 2^k$ and $1 \leq y \leq y' \leq 2^k)$ covering $\cup_{\alpha=x}^{x'} C_{k,\alpha} \cap \cup_{\beta=y}^{y'} R_{k,\beta}$ by $G(x, y, x', y')$ $(= \{v \in [2^k]^2 \mid x \leq X(v) \leq x'$ and $y \leq Y(v) \leq y'\})$. The size of the query subgrid $G(x, y, x', y')$ is $(x' - x + 1) \times (y' - y + 1)$.

***Remark* 1** For most self-similar m-dimensional order-k space-filling curve C_k^m indexing the grid $[2^k]^m$, we can view C_k^m as a C_{k-q}^m-curve interconnecting $2^{2(k-q)}$ C_q^m-subcurves for all $q \in [k]$.

The remark above motivates our analytical study of clustering performances to be based upon query subgrids of size $2^q \times 2^q$.

For a 2-dimensional order-k Hilbert curve H_k^2, let $S_{k,2^q}(H_k^2)$ denote the summation of all numbers of clusters over all $2^q \times 2^q$ query subgrids of an H_k^2-structural grid space $[2^k]^2$.

***Remark* 2** Within a query subgrid G, the number of clusters is half of the number of edges of underlying space-filling curve that are cut by the sides of G.

Denote by $\bar{n}(C, G)$ the number of edges within C that are cut by sides of subgrid G (without counting edge(s) connecting between C and other subcurves). Remark 2 translates the computation of the summation of all numbers of clusters over all identically shaped subgrids G to the computations of $\frac{1}{2}(\sum_{\text{all G}} \bar{n}(C, G) + 2)$ (the contribution of 2 is the number of cuts for connecting edges of C to other curves). For $2^q \times 2^q$ subgrids G, we denote by $E_q(C)$ all numbers of clusters over all identically shaped $2^q \times 2^q$ subgrids G, which is $\sum_{\text{allG}} \bar{n}(C, G) + 2$.

The recursive decomposition of H_k^2 (see Fig. 1b) gives that:

$$E_q(H_k^2) = 4E_q(H_{k-1}^2) + \varepsilon_{k,q}(H_k^2),$$

where $\varepsilon_{k,q}(H_k^2)$ denotes the summation of all edge cuts over all identically shaped $2^q \times 2^q$ query subgrids, each of which overlaps with more than one quadrant (that is, two or four). These query subgrids are contained in the boundary regions of neighboring quadrants as shown in Fig. 2.

***Remark* 3** For a 2-dimensional Hilbert curve H_k^2, the connecting edge between $Q_1(H_k^2)$ and $Q_2(H_k^2)$ is on the first column (left-most column), that between $Q_2(H_k^2)$ and $Q_3(H_k^2)$ is on the $(2^{k-1} + 1)$-st row (the lowest row of these two quadrants), and that between $Q_3(H_k^2)$ and $Q_4(H_k^2)$ is on the 2^k-th column (right-most column).

We denote the connecting edge between two quadrants $Q_i(H_k^2)$ and $Q_j(H_k^2)$ by a pair $(Q_i(H_k^2), Q_j(H_k^2))$. The previous remark tells the locations of the connecting edges. In addition to the cuts on connecting edges, the computation of $\varepsilon_{k,q}(H_k^2)$ is divided into two parts according to the overlaps of subspaces:

For a $2^q \times 2^q$ query subgrid G, G overlaps with:

1. exactly $Q_i(H_k^2)$ and $Q_{i \bmod 4+1}(H_k^2)$,
2. $Q_i(H_k^2)$ for all $i \in \{1, 2, 3, 4\}$.

We develop combinatorial lemmas in the following two subsections to support the computations. We denote by $G(x, y, x', y')$ the query subgrid, in which the lower-left grid point is (x, y) and the upper-right grid point (x', y').

$\sum \bar{n}(H_k^2, G)$ over Subgrids G Overlapping with Two Quadrants

Consider an arbitrary $2^q \times 2^q$ query subgrid G that exactly overlaps two quadrants $Q_i(H_k^2)$ and $Q_{i \bmod 4+1}(H_k^2)$, where $i \in \{1, 2, 3, 4\}$. The side-length is from 1 to $2^q - 1$ for the side across two quadrants. Since the quadrants are isomorphic to a canonical H_{k-1}^2 via symmetry (reflection and rotation), we consider the following system of summations

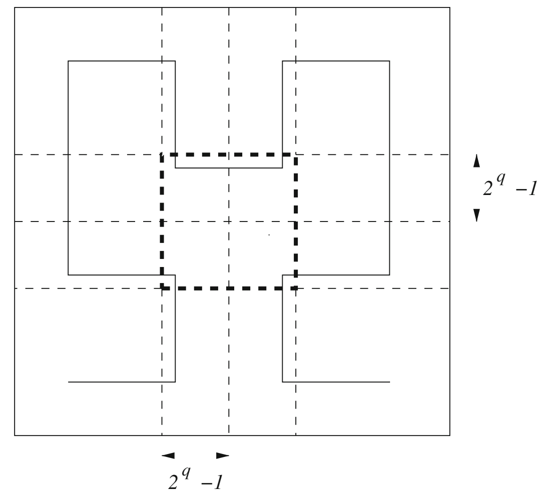


Fig. 2 The boundary regions of neighboring quadrants are organized into nine regions

$\Omega_{k,2^q} = (\Omega^L_{k,2^q}, \Omega^R_{k,2^q}, \Omega^B_{k,2^q}, \Omega^T_{k,2^q})$ in a general context of a canonical H_k^2:

$$\Omega^L_{k,2^q} = \sum_{x=1}^{2^q-1} \sum_{y=1}^{2^k-2^q+1} \bar{n}(H_k^2, G(1, y, x, y + 2^q - 1))$$

(for left boundary; see Fig. 3a),

$$\Omega^R_{k,2^q} = \sum_{x=2^k-2^q+2}^{2^k} \sum_{y=1}^{2^k-2^q+1} \bar{n}(H_k^2, G(x, y, 2^k, y + 2^q - 1))$$

(for right boundary),

$$\Omega^B_{k,2^q} = \sum_{x=1}^{2^k-2^q+1} \sum_{y=1}^{2^q-1} \bar{n}(H_k^2, G(x, 1, x + 2^q - 1, y))$$

(for bottom boundary),

$$\Omega^T_{k,2^q} = \sum_{x=1}^{2^k-2^q+1} \sum_{y=2^k-2^q+2}^{2^k} \bar{n}(H_k^2, G(x, y, x + 2^q - 1, 2^k))$$

(for top boundary),

$$\mathcal{N}^S_{k,2^q} = \sum_{x=1}^{2^q-1} \sum_{y=1}^{2^k-2^q+1} 1$$

(for the number of rectangular subgrids in a boundary for $\Omega_{k,2^q}$).

We establish a system of recurrences (in k) for $\Omega_{k,2^q}$ (see Lemma 4 below). The system of recurrence involves another system of summations as prerequisites, as demonstrated in the following example. Consider a recursive decomposition of $\Omega^L_{k,2^q}$, illustrated in Fig. 3b, into four parts: (1) $\Omega^B_{k-1,2^q}$, (2) $\Omega^{c_4}_{k-1,2^q}$, $\Omega^{c_1}_{k-1,2^q}$, (3) $\Omega^L_{k-1,2^q}$, and (4) the number of cuts on connecting edges. The part $\Omega^{c_1}_{k-1,2^q}$ ($\Omega^{c_4}_{k-1,2^q}$) computes $\sum \bar{n}(H_k^2, G)$ over all "incomplete" rectangular subgrids G (with one side-length at most $2^q - 1$) overlapping both $Q_1(H_k^2)$ and $Q_2(H_k^2)$. Each of the three parts $\Omega^B_{k-1,2^q}$, $\Omega^{c_1}_{k-1,2^q}$ ($\Omega^{c_4}_{k-1,2^q}$), and $\Omega^L_{k-1,2^q}$ is defined

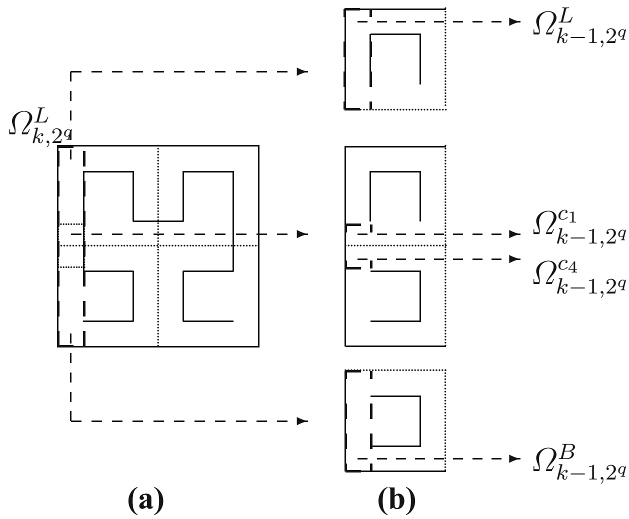


Fig. 3 **a** $\Omega^L_{k,2^q}$ for a canonical H^2_k; **b** its recursive decomposition

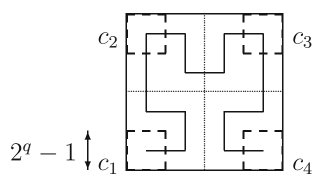


Fig. 4 The four $(2^q-1)\times(2^q-1)$ corners of a canonical H^2_k

with respect to a canonical H^2_{k-1}. Note that $\Omega^L_{k,2^q}=\Omega^R_{k,2^q}$ because of the left-right symmetry property of H^2_k.

The recursive decompositions of all four parts in $\Omega^L_{k,2^q}$, $\Omega^R_{k,2^q}$, $\Omega^B_{k,2^q}$, and $\Omega^T_{k,2^q}$ lead us to consider a prerequisite system of summations $\Omega^c_{k,2^q}=(\Omega^{c_1}_{k,2^q},\Omega^{c_2}_{k,2^q},\Omega^{c_3}_{k,2^q},\Omega^{c_4}_{k,2^q})$ in a more general context of a canonical H^2_k (see Fig. 4):

$$\Omega^{c_1}_{k,2^q}=\sum_{x=1}^{2^q-1}\sum_{y=1}^{2^q-1}\bar{n}(H^2_k,G(1,1,x,y))$$
(for lower-left corner),

$$\Omega^{c_2}_{k,2^q}=\sum_{x=1}^{2^q-1}\sum_{y=2^k-2^q+2}^{2^k}\bar{n}(H^2_k,G(1,y,x,2^k))$$
(for upper-left corner),

$$\Omega^{c_3}_{k,2^q}=\sum_{x=2^k-2^q+2}^{2^k}\sum_{y=2^k-2^q+2}^{2^k}\bar{n}(H^2_k,G(x,y,2^k,2^k))$$
(for upper-right corner),

$$\Omega^{c_4}_{k,2^q}=\sum_{x=2^k-2^q+2}^{2^k}\sum_{y=1}^{2^q-1}\bar{n}(H^2_k,G(x,1,2^k,y))$$
(for lower-right corner),

$$\mathcal{N}^c_{k,2^q}=\sum_{x=1}^{2^q-1}\sum_{y=1}^{2^q-1}1$$
(for the number of incomplete rectangular subgrids in a corner).

These four summations involve rectangular subgrids contained in $(2^q-1)\times(2^q-1)$ corners. Note that $\Omega^{c_1}_{k,2^q}=\Omega^{c_4}_{k,2^q}$ and $\Omega^{c_2}_{k,2^q}=\Omega^{c_3}_{k,2^q}$ because of the left-right symmetry property of H^2_k. As suggested by Remark 1, we zoom in on the $2^q\times 2^q$ H^2_q-structural corners, and consider the following system of summations $\overline{\Omega}^c_{q,2^q}=(\overline{\Omega}^{c_1}_{q,2^q},\overline{\Omega}^{c_2}_{q,2^q})$:

$$\overline{\Omega}^{c_1}_{q,2^q}=\sum_{x=1}^{2^q}\sum_{y=1}^{2^q}\bar{n}(H^2_q,G(1,1,x,y))$$
(for lower(-left) corner),

$$\left(=\sum_{x=1}^{2^q}\sum_{y=1}^{2^q}\bar{n}(H^2_q,G(x,1,2^q,y))\right)$$
(for lower(-right) corner),

$$\overline{\Omega}^{c_2}_{q,2^q}=\sum_{x=1}^{2^q}\sum_{y=1}^{2^q}\bar{n}(H^2_q,G(1,y,x,2^q))$$
(for upper(-left) corner),

$$\left(=\sum_{x=1}^{2^q}\sum_{y=1}^{2^q}\bar{n}(H^2_q,G(x,y,2^q,2^q))\right)$$
(for upper(-right) corner),

$$\overline{\mathcal{N}}^c_{q,2^q}=\sum_{x=1}^{2^q}\sum_{y=1}^{2^q}1$$
(for the number of rectangular subgrids in a $2^q\times 2^q$ corner).

Thus far, we learn that the system of recurrences for $\Omega_{k,2^q}$ can be defined and solved via the prerequisite system $\Omega^c_{k,2^q}$, which is related to the system $\overline{\Omega}^c_{q,2^q}$ (see Lemma 3 below). The system $\overline{\Omega}^c_{q,2^q}$, which involves subgrids (with both side-lengths at most 2^q) of a canonical H^2_q, represents the basis of the recursive decompositions (in k to q) of $\Omega_{k,2^q}$ and $\Omega^c_{k,2^q}$. Similar to the reduction of $\Omega_{k,2^q}$ to $\Omega^c_{k,2^q}$, we develop a system of recurrences (in q) for $\overline{\Omega}^c_{q,2^q}$ via a prerequisite system, as demonstrated in the following example.

Consider a recursive decomposition of $\overline{\Omega}^{c_1}_{q,2^q}=\sum_{x=1}^{2^q}\sum_{y=1}^{2^q}\bar{n}(H^2_q,G(1,1,x,y))$ into four parts, based upon the overlapping scenario of the rectangular subgrid G (1, 1, x, y) with the four quadrants of a canonical H^2_q (see Fig. 5):

Case 1: $G(1,1,x,y)$ is contained in $Q_1(H^2_q)$ (see Fig. 5a). This part is reduced to $\overline{\Omega}^{c_1}_{q-1,2^{q-1}}$ after $(-\frac{\pi}{2})$-rotating and then left-right reflecting $Q_1(H^2_q)$ into a canonical H^2_{q-1} and all numbers of cuts on connecting edge $(Q_1(H^2_q),Q_2(H^2_q))$ that are caused by the top (horizontal) side of G.

Case 2: $G(1,1,x,y)$ overlaps with exactly $Q_1(H^2_q)$ and $Q_2(H^2_q)$ (see Fig. 5b). This part is reduced to $\overline{\Omega}^{c_1}_{q-1,2^{q-1}}$ and all numbers of cuts on horizontal edge within $Q_1(H^2_q)$ and on connecting edge $(Q_2(H^2_q),Q_3(H^2_q))$ that are caused by the right (vertical) side of G.

Case 3: $G(1,1,x,y)$ overlaps with exactly $Q_1(H^2_q)$ and $Q_4(H^2_q)$ (see Fig. 5d). This part is reduced to $\overline{\Omega}^{c_2}_{q-1,2^{q-1}}$ after

$(+\frac{\pi}{2})$-rotating and then left-right reflecting $Q_4(H_q^2)$ into a canonical H_{q-1}^2 and all numbers of cuts on vertical edges within $Q_1(H_q^2)$ and on connecting edges $(Q_1(H_q^2), Q_2(H_q^2))$ and $(Q_3(H_q^2), Q_4(H_q^2))$ that are caused by the top (horizontal) side of G.

Case 4: $G(1, 1, x, y)$ overlaps with exactly all the quadrants (see Fig. 5c). This part is reduced to $\overline{\Omega}_{q-1,2^{q-1}}^{c_1}$ and all numbers of cuts on vertical edges within $Q_2(H_q^2)$ and horizontal edges within $Q_4(H_q^2)$ that are caused by the top (horizontal) and right (vertical), respectively, sides of G.

The recursive decompositions of $\overline{\Omega}_{q,2^q}^{c_1}$, and $\overline{\Omega}_{q,2^q}^{c_2}$ lead us to consider a prerequisite system of summations $\overline{\Pi}_q = (\overline{\Pi}_q^v, \overline{\Pi}_q^h)$ in a general context of a canonical H_q^2:

$$\overline{\Pi}_q^v = \sum_{y=1}^{2^q} \bar{n}(H_q^2, G(1, 1, 2^q, y))$$

(number of vertical edges (edges cut by top (horizontal) sides of G that covers lower part of H_q^2)),

$$\overline{\Pi}_q^h = \sum_{x=1}^{2^q} \bar{n}(H_q^2, G(1, 1, x, 2^q))$$

(number of horizontal edges (edges cut by right (vertical) sides of G that covers left part of H_q^2)).

We develop and solve a system of recurrences for $\overline{\Pi}_q$ and reverse the sequence of reductions to obtain the closed-form solutions for $\Omega_{k,2^q}$, which are summarized in the following four lemmas.

Lemma 1 *For a canonical* H_q^2,

$$\overline{\Pi}_q^v = \begin{cases} 2\overline{\Pi}_{q-1}^v + 2\overline{\Pi}_{q-1}^h + 2 & \text{if } q > 1, \\ 2 & \text{if } q = 1; \end{cases}$$

$$\overline{\Pi}_q^h = \begin{cases} 2\overline{\Pi}_{q-1}^v + 2\overline{\Pi}_{q-1}^h + 1 & \text{if } q > 1, \\ 1 & \text{if } q = 1. \end{cases}$$

Proof For $\overline{\Pi}_q^v$, the number of edge cuts by top (horizontal) sides of G can be computed from cuts within the four quadrants and cuts on the two connecting edges $(Q_1(H_q^2), Q_2(H_q^2))$ and $(Q_3(H_q^2), Q_4(H_q^2))$:

$$\begin{aligned}
\overline{\Pi}_q^v &= \sum_{y=1}^{2^q} \bar{n}(H_q^2, G(1, 1, 2^q, y)) \\
&= \sum_{y=1}^{2^{q-1}} \bar{n}(H_q^2, G(1, 1, 2^q, y)) \\
&\qquad \text{(cuts on vertical edge within } Q_1(H_q^2), Q_4(H_q^2) \\
&\qquad \text{plus cuts on two connecting edges when } x = 2^{q-1}) \\
&\quad + \sum_{y=2^{q-1}+1}^{2^q} \bar{n}(H_q^2, G(1, 1, 2^q, y)) \\
&\qquad \text{(cuts on vertical edges within } Q_2(H_q^2), Q_3(H_q^2)) \\
&= (\sum_{y=1}^{2^{q-1}} \bar{n}(Q_1(H_q^2), G(1, 1, 2^{q-1}, y)) \\
&\qquad \text{(cuts on vertical edges within } Q_1(H_q^2)) \\
&\quad + \sum_{y=1}^{2^{q-1}} \bar{n}(Q_4(H_q^2), G(2^{q-1}+1, 1, 2^q, y)) \\
&\qquad \text{(cuts on vertical edges within } Q_4(H_q^2)) \\
&\quad + 2) \\
&\qquad \text{(cuts on connecting edges } (Q_1(H_q^2), Q_2(H_q^2)), \\
&\qquad (Q_3(H_q^2), Q_4(H_q^2)) \text{ when } x = 2^{q-1}) \\
&\quad + (\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Q_2(H_q^2), G(1, 2^{q-1}+1, 2^{q-1}, y)) \\
&\qquad \text{(cuts on vertical edges within } Q_2(H_q^2))
\end{aligned}$$

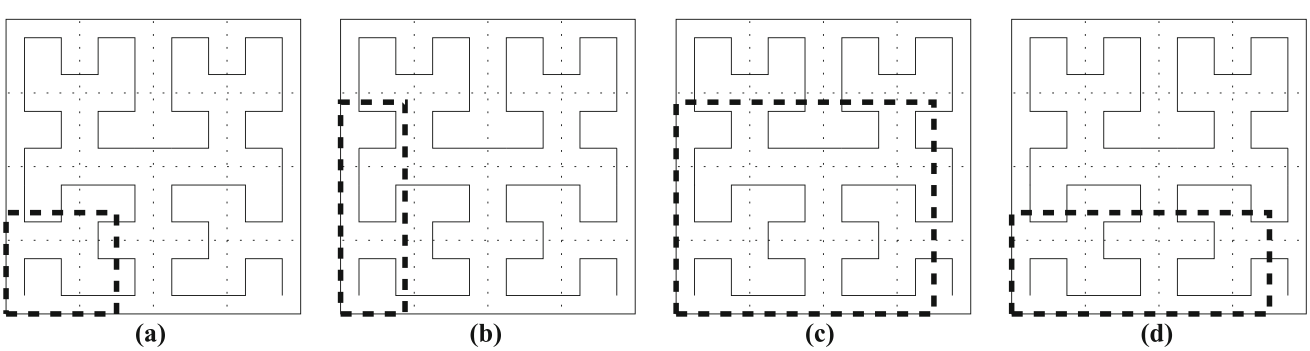


Fig. 5 Four overlapping scenarios when decomposing $\overline{\Omega}_{q,2^q}^{c_1}$ in a canonical H_q^2: **a** contained in $Q_1(H_q^2)$; **b** and **d** overlapping with exactly two quadrants; **c** overlapping with all quadrants

$$+\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Q_3(H_q^2), G(2^{q-1}+1, 2^{q-1}+1, 2^q, y)))$$
(cuts on vertical edges within $Q_3(H_q^2)$)
$$=(\sum_{x=1}^{2^{q-1}} \bar{n}(H_{q-1}^2, G(1,1,x,2^{q-1}))$$
(after $(-\frac{\pi}{2})$-rotating and left-right reflecting $Q_1(H_q^2)$ into a canonical H_{q-1}^2)
$$+\sum_{x=1}^{2^{q-1}} \bar{n}(H_{q-1}^2, G(1,1,x,2^{q-1}))$$
(after $(+\frac{\pi}{2})$-rotating and left-right reflecting $Q_4(H_q^2)$ into a canonical H_{q-1}^2)
$$+2)$$
$$+(\sum_{y=1}^{2^{q-1}} \bar{n}(H_{q-1}^2, G(1,1,2^{q-1},y))$$
($Q_2(H_q^2)$: a canonical H_{q-1}^2)
$$+\sum_{y=1}^{2^{q-1}} \bar{n}(H_{q-1}^2, G(1,1,2^{q-1},y)))$$
($Q_3(H_q^2)$: a canonical H_{q-1}^2)
$$=(\overline{\Pi}_{q-1}^h+\overline{\Pi}_{q-1}^h+2)+(\overline{\Pi}_{q-1}^v+\overline{\Pi}_{q-1}^v)$$
$$=2\overline{\Pi}_{q-1}^v+2\overline{\Pi}_{q-1}^h+2.$$

The proof of $\overline{\Pi}_q^h$ is similar to that of $\overline{\Pi}_q^v$. □

The closed-form solutions for $\overline{\Pi}_q$ are employed to establish a system of recurrences for $\overline{\Omega}_{q,2^q}^c$.

Lemma 2 *For a canonical H_q^2,*

$$\overline{\Omega}_{q,2^q}^{c_1}=\begin{cases} 3\overline{\Omega}_{q-1,2^{q-1}}^{c_1}+\overline{\Omega}_{q-1,2^{q-1}}^{c_2}+3\cdot 2^{q-1}\cdot\overline{\Pi}_{q-1}^v+2^{q-1}\cdot\overline{\Pi}_{q-1}^h \\ \qquad +3\cdot 2^{q-1}+1 & \text{if } q>1, \\ 4 & \text{if } q=1; \end{cases}$$
$$\overline{\Omega}_{q,2^q}^{c_2}=\begin{cases} \overline{\Omega}_{q-1,2^{q-1}}^{c_1}+3\overline{\Omega}_{q-1,2^{q-1}}^{c_2}+2^{q-1}\cdot\overline{\Pi}_{q-1}^v+3\cdot 2^{q-1}\cdot\overline{\Pi}_{q-1}^h \\ \qquad +3\cdot 2^{q-1}+2 & \text{if } q>1, \\ 5 & \text{if } q=1. \end{cases}$$

Proof As in Fig. 5 and the case discussion for $\Omega_{q,2^q}^{c_1}$, we can split $\Omega_{q,2^q}^{c_1}$ into four parts. Thus,

$$\overline{\Omega}_{q,2^q}^{c_1}=\sum_{x=1}^{2^q}\sum_{y=1}^{2^q} \bar{n}(H_q^2, G(1,1,x,y))$$
$$=\sum_{x=1}^{2^{q-1}}\sum_{y=1}^{2^{q-1}} \bar{n}(H_q^2, G(1,1,x,y))$$
(Fig. 5a)
$$+\sum_{x=1}^{2^{q-1}}\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(H_q^2, G(1,1,x,y))$$
(Fig. 5b)
$$+\sum_{x=2^{q-1}+1}^{2^q}\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(H_q^2, G(1,1,x,y))$$
(Fig. 5c)
$$+\sum_{x=2^{q-1}+1}^{2^q}\sum_{y=1}^{2^{q-1}} \bar{n}(H_q^2, G(1,1,x,y))$$
(Fig. 5d)
$$=(\sum_{x=1}^{2^{q-1}}\sum_{y=1}^{2^{q-1}} \bar{n}(Q_1(H_q^2), G(1,1,x,y))$$
(cuts within $Q_1(H_q^2)$)
$$+\sum_{x=1}^{2^{q-1}}\sum_{y=2^{q-1}}^{2^{q-1}} 1)$$
(cuts on connecting edge $(Q_1(H_q^2), Q_2(H_q^2))$)
$$+(\sum_{x=1}^{2^{q-1}}\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Q_2(H_q^2), G(1,2^{q-1}+1,x,y))$$
(cuts within $Q_2(H_q^2)$)
$$+\sum_{x=1}^{2^{q-1}}\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Q_1(H_q^2), G(1,1,x,2^{q-1}))$$
(cuts within $Q_1(H_q^2)$ by vertical side of G)
$$+\sum_{x=2^{q-1}}^{2^{q-1}}\sum_{y=2^{q-1}+1}^{2^q} 1)$$
(cuts on connecting edge $(Q_2(H_q^2), Q_3(H_q^2))$)

$$+ (\sum_{x=2^{q-1}+1}^{2^q} \sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Q_3(H_q^2), G(2^{q-1}+1, 2^{q-1}+1, x, y))$$

(cuts within $Q_3(H_q^2)$)

$$+ \sum_{x=2^{q-1}+1}^{2^q} \sum_{y=1}^{2^{q-1}} \bar{n}(Q_2(H_q^2), G(1, 2^{q-1}+1, 2^{q-1}, y))$$

(cuts within $Q_2(H_q^2)$ by horizontal side of G)

$$+ \sum_{x=2^{q-1}+1}^{2^q} \sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Q_4(H_q^2), G(2^{q-1}+1, 1, x, 2^{q-1})))$$

(cuts within $Q_4(H_q^2)$ by vertical side of G)

$$+ (\sum_{x=2^{q-1}+1}^{2^q} \sum_{y=1}^{2^{q-1}} \bar{n}(Q_4(H_q^2), G(2^{q-1}+1, 1, x, y))$$

(cuts within $Q_4(H_q^2)$)

$$+ \sum_{x=2^{q-1}+1}^{2^q} \sum_{y=1}^{2^{q-1}} \bar{n}(Q_1(H_q^2), G(1, 1, 2^{q-1}, y))$$

(cuts within $Q_1(H_q^2)$ by horizontal side of G)

$$+ \sum_{x=2^{q-1}+1}^{2^q} \sum_{y=2^{q-1}}^{2^{q-1}} 1$$

(cuts on connecting edge $(Q_1(H_q^2), Q_2(H_q^2))$)

$$+ \sum_{x=2^q}^{2^q} \sum_{y=2^{q-1}}^{2^{q-1}} 1)$$

(cuts on connecting edge $(Q_3(H_q^2), Q_4(H_q^2))$)

$$= (\sum_{x=1}^{2^{q-1}} \sum_{y=1}^{2^{q-1}} \bar{n}(H_{q-1}^2, G(1, 1, 2^{q-1}, y))$$

(after $(-\frac{\pi}{2})$-rotating (and left-right reflecting) $Q_1(H_q^2)$ into a canonical H_{q-1}^2)

$$+ (2^{q-1}))$$

$$+ (\sum_{x=1}^{2^{q-1}} \sum_{y=1}^{2^{q-1}} \bar{n}(H_{q-1}^2, G(1, 1, x, y))$$

($Q_2(H_q^2)$ a canonical H_{q-1}^2)

$$+ \sum_{x=1}^{2^{q-1}} \sum_{y=1}^{2^{q-1}} \bar{n}(H_{q-1}^2, G(1, 1, 2^{q-1}, y))$$

(after $(-\frac{\pi}{2})$-rotating (and left-right reflecting) $Q_1(H_q^2)$ into a canonical H_{q-1}^2)

$$+ (2^{q-1}))$$

$$+ (\sum_{x=1}^{2^{q-1}} \sum_{y=1}^{2^{q-1}} \bar{n}(H_{q-1}^2, G(1, 1, x, y))$$

($Q_3(H_q^2)$ a canonical H_{q-1}^2)

$$+ \sum_{x=1}^{2^{q-1}} \sum_{y=1}^{2^{q-1}} \bar{n}(H_{q-1}^2, G(1, 1, 2^{q-1}, y))$$

($Q_2(H_q^2)$: a canonical H_{q-1}^2)

$$+ \sum_{x=1}^{2^{q-1}} \sum_{y=1}^{2^{q-1}} \bar{n}(H_{q-1}^2, G(1, 1, 2^{q-1}, y)))$$

(after $(+\frac{\pi}{2})$-rotating (and left-right reflecting) $Q_4(H_q^2)$ into a canonical H_{q-1}^2)

$$+ (\sum_{x=1}^{2^{q-1}} \sum_{y=1}^{2^{q-1}} \bar{n}(H_{q-1}^2, G(1, 1, x, y))$$

(after $(+\frac{\pi}{2})$-rotating (and left-right reflecting) $Q_4(H_q^2)$ into a canonical H_{q-1}^2)

$$+ \sum_{x=1}^{2^{q-1}} \sum_{y=1}^{2^{q-1}} \bar{n}(H_{q-1}^2, G(1, 1, x, 2^{q-1}))$$

(after $(-\frac{\pi}{2})$-rotating (and left-right reflecting) $Q_1(H_q^2)$ into a canonical H_{q-1}^2)

$$\begin{aligned}
&+ (2^{q-1}) + 1) \\
&= (\overline{\Omega}_{q-1,2^{q-1}}^{c_1} + 2^{q-1}) \\
&\quad + (\overline{\Omega}_{q-1,2^{q-1}}^{c_1} + \overline{\Pi}_{q-1}^{v} \cdot 2^{q-1} + 2^{q-1}) \\
&\quad + (\overline{\Omega}_{q-1,2^{q-1}}^{c_1} + \overline{\Pi}_{q-1}^{v} \cdot 2^{q-1} + \overline{\Pi}_{q-1}^{v} \cdot 2^{q-1}) \\
&\quad + (\overline{\Omega}_{q-1,2^{q-1}}^{c_2} + \overline{\Pi}_{q-1}^{h} \cdot 2^{q-1} + 2^{q-1} + 1) \\
&= 3\overline{\Omega}_{q-1,2^{q-1}}^{c_1} + \overline{\Omega}_{q-1,2^{q-1}}^{c_2} \\
&\quad + 3 \cdot 2^{q-1} \cdot \overline{\Pi}_{q-1}^{v} + 2^{q-1} \cdot \overline{\Pi}_{q-1}^{h} + 3 \cdot 2^{q-1} + 1.
\end{aligned}$$

The proof of $\overline{\Omega}_{q,2^q}^{c_2}$ is similar to that of $\overline{\Omega}_{q,2^q}^{c_1}$. □

The closed-form solutions for $\overline{\Omega}_{q,2^q}^{c}$ and $\overline{\Pi}_q$ are employed to obtain exact formulas for $\Omega_{k,2^q}^{c}$.

Lemma 3 *For a canonical H_k^2 structured as an H_{k-q}^2-curve interconnecting $2^{2(k-q)}$ H_q^2-subcurves,*

$$\Omega_{k,2^q}^{c_1} = \overline{\Omega}_{q,2^q}^{c_1} - \overline{\Pi}_q^{h} - \overline{\Pi}_q^{v},$$
$$\Omega_{k,2^q}^{c_2} = \overline{\Omega}_{q,2^q}^{c_2} - \overline{\Pi}_q^{h} - \overline{\Pi}_q^{v}.$$

Proof By the definition, we have:

$$\begin{aligned}
\Omega_{k,2^q}^{c_1} &= \sum_{x=1}^{2^q-1} \sum_{y=1}^{2^q-1} \bar{n}(H_k^2, G(1, 1, x, y)) \\
&= \sum_{x=1}^{2^q} \sum_{y=1}^{2^q-1} \bar{n}(H_k^2, G(1, 1, x, y)) \\
&\quad - \sum_{x=2^q}^{2^q} \sum_{y=1}^{2^q-1} \bar{n}(H_k^2, G(1, 1, x, y))
\end{aligned}$$

$$=(\sum_{x=1}^{2^q}\sum_{y=1}^{2^q}\bar{n}(H_k^2,G(1,1,x,y))$$
$$-\sum_{x=1}^{2^q}\sum_{y=2^q}^{2^q}\bar{n}(H_k^2,G(1,1,x,y)))$$
$$-(\sum_{x=2^q}^{2^q}\sum_{y=1}^{2^q}\bar{n}(H_k^2,G(1,1,x,y))$$
$$-\sum_{x=2^q}^{2^q}\sum_{y=2^q}^{2^q}\bar{n}(H_k^2,G(1,1,x,y)))$$
$$=\overline{\Omega}_{q,2^q}^{c_1}-\overline{\Pi}_q^h-\overline{\Pi}_q^v.$$

The proof of $\Omega_{k,2^q}^{c_2}$ is similar to that of $\Omega_{k,2^q}^{c_1}$. □

The exact formulas for $\Omega_{k,2^q}^{c}$ are employed to establish a system of recurrences for $\Omega_{k,2^q}$.

Lemma 4 *For a canonical H_k^2 structured as an H_{k-q}^2-curve interconnecting $2^{2(k-q)}$ H_q^2-subcurves,*

$$\Omega_{k,2^q}^L=\begin{cases}\Omega_{k-1,2^q}^B+\Omega_{k-1,2^q}^L+2\Omega_{k-1,2^q}^{c_1}+2(2^q-1) & \text{if } k>q,\\ \overline{\Pi}_q^h & \text{if } k=q;\end{cases}$$
$$\Omega_{k,2^q}^B=\begin{cases}2\Omega_{k-1,2^q}^L+2\Omega_{k-1,2^q}^{c_2} & \text{if } k>q,\\ \overline{\Pi}_q^v & \text{if } k=q;\end{cases}$$
$$\Omega_{k,2^q}^T=\begin{cases}2\Omega_{k-1,2^q}^T+2\Omega_{k-1,2^q}^{c_2} & \text{if } k>q,\\ \overline{\Pi}_q^v & \text{if } k=q.\end{cases}$$

Proof Similar to the proof of Lemma 3, from the definition, we have (see Fig. 3):

$$\Omega_{k,2^q}^L=\sum_{x=1}^{2^q-1}\sum_{y=1}^{2^k-2^q+1}\bar{n}(H_k^2,G(1,y,x,y+2^q-1))$$
$$=\sum_{x=1}^{2^q-1}\sum_{y=1}^{2^{k-1}-2^q+1}\bar{n}(H_k^2,G(1,y,x,y+2^q-1))$$
$$(G \text{ in } Q_1(H_k^2))$$
$$+\sum_{x=1}^{2^q-1}\sum_{y=2^{k-1}-2^q+2}^{2^{k-1}}\bar{n}(H_k^2,G(1,y,x,y+2^q-1))$$
$$(G \text{ across } Q_1(H_k^2),Q_2(H_k^2))$$
$$+\sum_{x=1}^{2^q-1}\sum_{y=2^{k-1}+1}^{2^k-2^q+1}\bar{n}(H_k^2,G(1,y,x,y+2^q-1))$$
$$(G \text{ in } Q_2(H_k^2))$$
$$=(\sum_{x=1}^{2^q-1}\sum_{y=1}^{2^{k-1}-2^q+1}\bar{n}(Q_1(H_k^2),G(1,y,x,y+2^q-1))$$
$$(\text{cuts within } Q_1(H_k^2))$$
$$+\sum_{x=1}^{2^q-1}\sum_{y=2^{k-1}-2^q+1}^{2^{k-1}-2^q+1}1)$$
$$(\text{cuts on connecting edge } (Q_1(H_q^2),Q_2(H_q^2)))$$
$$+(\sum_{x=1}^{2^q-1}\sum_{y=2^{k-1}-2^q+2}^{2^{k-1}}\bar{n}(Q_1(H_k^2),G(1,y,x,2^{k-1}))$$
$$(\text{zooming in } Q_1(H_k^2))$$
$$+\sum_{x=1}^{2^q-1}\sum_{y=2^{k-1}-2^q+2}^{2^{k-1}}\bar{n}(Q_2(H_k^2),G(1,2^{k-1}+1,x,y+2^q-1)))$$
$$(\text{zooming in } Q_2(H_k^2))$$
$$+(\sum_{x=1}^{2^q-1}\sum_{y=2^{k-1}+1}^{2^k-2^q+1}\bar{n}(Q_2(H_k^2),G(1,y,x,y+2^q-1))$$
$$(\text{cuts within } Q_2(H_k^2))$$
$$+\sum_{x=1}^{2^q-1}\sum_{y=2^{k-1}+1}^{2^{k-1}+1}1)$$
$$(\text{cuts on connecting edge } (Q_1(H_q^2),Q_2(H_q^2)))$$
$$=(\sum_{x=1}^{2^{k-1}-2^q+1}\sum_{y=1}^{2^q-1}\bar{n}(H_{k-1}^2,G(x,1,x+2^q-1,y))$$
$$(\text{after } (-\frac{\pi}{2})\text{-rotating and left-right reflecting } Q_1(H_k^2)$$
$$\text{into a canonical } H_{k-1}^2)$$
$$+(2^q-1))$$
$$+(\sum_{x=2^{k-1}-2^q+2}^{2^{k-1}}\sum_{y=1}^{2^q-1}\bar{n}(H_{k-1}^2,G(x,1,2^{k-1},y))$$
$$(\text{after } (-\frac{\pi}{2})\text{-rotating and left-right reflecting } Q_1(H_k^2)$$
$$\text{into a canonical } H_{k-1}^2)$$
$$+\sum_{x=1}^{2^q-1}\sum_{y=1}^{2^q-1}\bar{n}(H_{k-1}^2,G(1,1,x,y)))$$
$$(Q_2(H_k^2)\text{: a canonical } H_{k-1}^2)$$
$$+(\sum_{x=1}^{2^q-1}\sum_{y=1}^{2^{k-1}-2^q+1}\bar{n}(H_{k-1}^2,G(1,y,x,y+2^q-1))$$
$$(Q_2(H_k^2)\text{: a canonical } H_{k-1}^2)$$
$$+(2^q-1))$$
$$=(\Omega_{k-1,2^q}^B+(2^q-1))$$
$$+(\Omega_{k-1,2^q}^{c_4}+\Omega_{k-1,2^q}^{c_1})$$
$$(\Omega_{k-1,2^q}^{c_1}=\Omega_{k-1,2^q}^{c_4})$$
$$+(\Omega_{k-1,2^q}^L+(2^q-1))$$
$$=\Omega_{k-1,2^q}^L+\Omega_{k-1,2^q}^B+2\Omega_{k-1,2^q}^{c_1}+2(2^q-1).$$

For $\Omega_{k-1,2^q}^B$ and $\Omega_{k-1,2^q}^T$, the proofs are similar to that of $\Omega_{k,2^q}^L$. □

We obtain the closed-form solutions for $\Omega_{k,2^q}$ by using the mathematical software Maple.

Query Subgrids Overlapping with All Quadrants

For a $2^q \times 2^q$ query subgrid G that overlaps four quadrants around the center of H_k^2, when zooming in on the incomplete rectangular subgrid $G \cap G_1$ (with both side-lengths at most $2^q - 1$), where G_1 denotes the subspace of $Q_1(H_k^2)$, we reduce $\sum_{\mathrm{allG}\cap\mathrm{G}_1} \bar{n}(H_k^2, G \cap G_1)$ to $\Omega^{c_3}_{k-1,2^q} (= \Omega^{c_2}_{k-1,2^q})$ after $(-\frac{\pi}{2})$-rotating and left-right reflecting $Q_1(H_k^2)$ into a canonical H^2_{k-1}. Similar consideration leads to reductions of $\sum_{\mathrm{allG}\cap\mathrm{G}'} \bar{n}(H_k^2, G \cap G')$ to $\Omega^{c_4}_{k-1,2^q} (= \Omega^{c_1}_{k-1,2^q})$, $\Omega^{c_1}_{k-1,2^q}$ and $\Omega^{c_2}_{k-1,2^q}$ when $G \cap G'$ denotes the subspace for G overlapping $Q_2(H_k^2)$, $Q_3(H_k^2)$, or $Q_4(H_k^2)$, respectively.

Thus, the summation of numbers of edge cuts for all $2^q \times 2^q$ query subgrids G that overlap all four quadrants is:

$$2\Omega^{c_2}_{k-1,2^q} + 2\Omega^{c_1}_{k-1,2^q}.$$

The Big Picture: Computing $E_q(H_k^2)$

The results in the previous three subsections yield $\varepsilon_{k,q}(H_k^2)$. Hence, we have the following lemma for $E_q(H_k^2)$.

Lemma 5 *For a canonical H_k^2, the recurrence for total number of cuts on edges by all $2^q \times 2^q$ subgrids G:*

$$E_q(H_k^2) = \begin{cases} 4E_q(H^2_{k-1}) + (\Omega^L_{k-1,2^q} + \Omega^B_{k-1,2^q} + (2^q - 1)) \\ \quad +(\Omega^L_{k-1,2^q} + \Omega^L_{k-1,2^q}) \\ \quad +(\Omega^L_{k-1,2^q} + \Omega^B_{k-1,2^q} + (2^q - 1)) \\ \quad +(\Omega^T_{k-1,2^q} + \Omega^T_{k-1,2^q}) \\ \quad +(2\Omega^{c_1}_{k-1,2^q} + 2\Omega^{c_2}_{k-1,2^q}) & \text{if } k > q, \\ \\ 2 & \text{if } k = q. \end{cases}$$

Proof Similar to the proofs of Lemmas 1 to 4:

Case 1: G overlaps with exactly $Q_1(H_k^2)$ and $Q_2(H_k^2)$. This part is reduced to $\Omega^R_{k-1,2^q} (= \Omega^L_{k-1,2^q})$ (cuts on $Q_1(H_k^2)$), $\Omega^B_{k-1,2^q}$ (cuts on $Q_2(H_k^2)$), and $2^q - 1$ cuts on the connecting edge $(Q_2(H_k^2), Q_3(H_k^2))$.

Case 2: G overlaps with exactly $Q_2(H_k^2)$ and $Q_3(H_k^2)$. This part is reduced to $\Omega^R_{k-1,2^q} (= \Omega^L_{k-1,2^q})$ (cuts on $Q_2(H_k^2)$), and $\Omega^L_{k-1,2^q}$ (cuts on $Q_3(H_k^2)$).

Case 3: G overlaps with exactly $Q_3(H_k^2)$ and $Q_4(H_k^2)$. This part is reduced to $\Omega^B_{k-1,2^q}$ (cuts on $Q_3(H_k^2)$), $\Omega^L_{k-1,2^q}$ (cuts on $Q_4(H_k^2)$), and $2^q - 1$ cuts on the connecting edge $(Q_2(H_k^2), Q_3(H_k^2))$.

Case 4: G overlaps with exactly $Q_1(H_k^2)$ and $Q_4(H_k^2)$. This part is reduced to $\Omega^T_{k-1,2^q}$ (cuts on $Q_1(H_k^2)$), and $\Omega^T_{k-1,2^q}$ (cuts on $Q_4(H_k^2)$).

Case 5: G overlaps with exactly all four quadrants. This part is reduced to $\Omega^{c_3}_{k-1,2^q} (= \Omega^{c_2}_{k-1,2^q})$ (cuts on $Q_1(H_k^2)$), $\Omega^{c_4}_{k-1,2^q} (= \Omega^{c_1}_{k-1,2^q})$ (cuts on $Q_2(H_k^2)$), $\Omega^{c_1}_{k-1,2^q}$ (cuts on $Q_3(H_k^2)$), and $\Omega^{c_2}_{k-1,2^q}$ (cuts on $Q_4(H_k^2)$).

Combining all the five cases, we complete the recurrence.

For the case of $k = q$, there are two cuts that are the edge cut between entry grid point and other curve, and the edge cut between exit grid point and other curve. □

Therefore, the exact formula for the summation of all numbers of edge cuts over all identically shaped $2^q \times 2^q$ subgrids, $E_q(H_k^2)$, is:

$$E_q(H_k^2) = 2^{2k+q+1} - 2^{k+2q+2} + 2^{k+q+1} + 2^{k-q+1} + 2^{3q+1} - 2^{2q+1}.$$

The summation of all numbers of clusters over all identically shaped $2^q \times 2^q$ query subgrids of an H_k^2-structural grid space $[2^k]^2$ is:

$$\mathcal{S}_{k,2^q}(H_k^2) = \frac{E_q(H_k^2)}{2}.$$

The mean number of cluster within a subspace of size $2^q \times 2^q$ for H_k^2 is:

$$\frac{\mathcal{S}_{k,2^q}(H_k^2)}{(2^k - 2^q + 1)^2} = \frac{E_q(H_k^2)}{2(2^k - 2^q + 1)^2}.$$

Thus, the exact formula for the mean number of cluster within a subspace $2^q \times 2^q$ for H_k^2 is consequently derived.

Theorem 1 *The mean number of cluster over all identical subspaces $2^q \times 2^q$ for H_k^2 is:*

$$\frac{2^{2k+q+1} - 2^{k+2q+2} + 2^{k+q+1} + 2^{k-q+1} + 2^{3q+1} - 2^{2q+1}}{2(2^k - 2^q + 1)^2}.$$

Clustering Statistics of 2-Dimensional *z*-Order Curves

With respect to the canonical orientation of Z_k^2 shown in Fig. 1e, we apply the same approach and notations as in previous section to derive the exact formula for $\mathcal{S}_{k,2^q}(Z_k^2)$, which is the summation of all numbers of clusters over all identically shaped $2^q \times 2^q$ query subgrids of an Z_k^2-structural grid space $[2^k]^2$.

Like the approach in previous section, we compute the summation of all numbers of edge cuts over all identically shaped $2^q \times 2^q$ subgrids G by the recursive decomposition of Z_k^2:

$$E_q(Z_k^2) = 4E_q(Z_{k-1}^2) + \varepsilon_{k,q}(Z_k^2),$$

where $\varepsilon_{k,q}(Z_k^2)$ denotes the summation of all edge cuts over all $2^q \times 2^q$ query subgrids, each of which overlaps with more than one quadrant (that is, two or four). These query subgrids are contained in the boundary regions of neighboring quadrants.

We set up the systems for $E_q(Z_k^2)$ similar to that for $E_q(H_k^2)$ in previous section.

$\sum \bar{n}(Z_k^2, G)$ over Subgrids G Overlapping with Two Quadrants

Consider an arbitrary $2^q \times 2^q$ query subgrid G that exactly overlaps two quadrant $Q_i(Z_k^2)$ and $Q_j(Z_k^2)$, where $(i,j) \in \{(1,2),(1,3),(2,4),(3,4)\}$. The side-length is from 1 to $2^q - 1$ for the side across two quadrants. Since all the quadrants are isomorphic to a canonical Z_{k-1}^2, we consider the following system of summations $\Omega_{k,2^q} = (\Omega_{k,2^q}^L, \Omega_{k,2^q}^R, \Omega_{k,2^q}^B, \Omega_{k,2^q}^T)$ in a general context of a canonical Z_k^2:

$$\Omega_{k,2^q}^L = \sum_{x=1}^{2^q-1} \sum_{y=1}^{2^k-2^q+1} \bar{n}(Z_k^2, G(1,y,x,y+2^q-1))$$

(for left boundary; see Fig. 6a),

$$\Omega_{k,2^q}^R = \sum_{x=2^k-2^q+2}^{2^k} \sum_{y=1}^{2^k-2^q+1} \bar{n}(Z_k^2, G(x,y,2^k,y+2^q-1))$$

(for right boundary),

$$\Omega_{k,2^q}^B = \sum_{x=1}^{2^k-2^q+1} \sum_{y=1}^{2^q-1} \bar{n}(Z_k^2, G(x,1,x+2^q-1,y))$$

(for bottom boundary),

$$\Omega_{k,2^q}^T = \sum_{x=1}^{2^k-2^q+1} \sum_{y=2^k-2^q+2}^{2^k} \bar{n}(Z_k^2, G(x,y,x+2^q-1,2^k))$$

(for top boundary),

$$\mathcal{N}_{k,2^q}^S = \sum_{x=1}^{2^q-1} \sum_{y=1}^{2^k-2^q+1} 1$$

(for the number of rectangular subgrids in a boundary for $\Omega_{k,2^q}$).

We establish a system of recurrences (in k) for $\Omega_{k,2^q}$ (see Lemma 9 below). Note, $\Omega_{k,2^q}^L = \Omega_{k,2^q}^R$ and $\Omega_{k,2^q}^B = \Omega_{k,2^q}^T$ because of the property of symmetry for Z_k^2.

The recursive decompositions of all four parts in $\Omega_{k,2^q}^L$, $\Omega_{k,2^q}^R$, $\Omega_{k,2^q}^B$, and $\Omega_{k,2^q}^T$ require a prerequisite system of summations $\Omega_{k,2^q}^c = (\Omega_{k,2^q}^{c_1}, \Omega_{k,2^q}^{c_2}, \Omega_{k,2^q}^{c_3}, \Omega_{k,2^q}^{c_4})$ in a more general context of a canonical Z_k^2 (see Fig. 7):

$$\Omega_{k,2^q}^{c_1} = \sum_{x=1}^{2^q-1} \sum_{y=1}^{2^q-1} \bar{n}(Z_k^2, G(1,1,x,y))$$

(for lower-left corner),

$$\Omega_{k,2^q}^{c_2} = \sum_{x=1}^{2^q-1} \sum_{y=2^k-2^q+2}^{2^k} \bar{n}(Z_k^2, G(1,y,x,2^k))$$

(for upper-left corner),

$$\Omega_{k,2^q}^{c_3} = \sum_{x=2^k-2^q+2}^{2^k} \sum_{y=1}^{2^q-1} \bar{n}(Z_k^2, G(x,1,2^k,y))$$

(for lower-right corner),

$$\Omega_{k,2^q}^{c_4} = \sum_{x=2^k-2^q+2}^{2^k} \sum_{y=2^k-2^q+2}^{2^k} \bar{n}(Z_k^2, G(x,y,2^k,2^k))$$

(for upper-right corner),

$$\mathcal{N}_{k,2^q}^c = \sum_{x=1}^{2^q-1} \sum_{y=1}^{2^q-1} 1$$

(for the number of incomplete rectangular subgrids in a corner).

Note that $\Omega_{k,2^q}^{c_1} = \Omega_{k,2^q}^{c_4}$ and $\Omega_{k,2^q}^{c_2} = \Omega_{k,2^q}^{c_3}$ because of the symmetry property of Z_k^2. To compute $\Omega_{k,2^q}^c$, we set up the following system of summations $\overline{\Omega}_{q,2^q}^c = (\overline{\Omega}_{q,2^q}^{c_1}, \overline{\Omega}_{q,2^q}^{c_2})$:

$$\overline{\Omega}_{q,2^q}^{c_1} = \sum_{x=1}^{2^q} \sum_{y=1}^{2^q} \bar{n}(Z_q^2, G(1,1,x,y))$$

(for lower-left corner),

$$\left(= \sum_{x=1}^{2^q} \sum_{y=1}^{2^q} \bar{n}(Z_q^2, G(x,y,2^q,2^q))\right)$$

(for upper-right corner),

$$\overline{\Omega}_{q,2^q}^{c_2} = \sum_{x=1}^{2^q} \sum_{y=1}^{2^q} \bar{n}(Z_q^2, G(1,y,x,2^q))$$

(for upper-left corner),

$$\left(= \sum_{x=1}^{2^q} \sum_{y=1}^{2^q} \bar{n}(Z_q^2, G(x,1,2^q,y))\right)$$

(for lower-right corner),

$$\overline{\mathcal{N}}_{q,2^q}^c = \sum_{x=1}^{2^q} \sum_{y=1}^{2^q} 1$$

(for the number of rectangular subgrids in a $2^q \times 2^q$ corner).

Similar to the reduction of $\Omega_{k,2^q}$ to $\Omega^c_{k,2^q}$, we develop a system of recurrences (in q) for $\overline{\Omega}^c_{q,2^q}$ via a prerequisite system $\overline{\Pi}_q = (\overline{\Pi}^v_q, \overline{\Pi}^h_q)$.

This prerequisite system of summations in a general context of a canonical Z^2_q:

$$\overline{\Pi}^v_q = \sum_{y=1}^{2^q} \bar{n}(Z^2_k, G(1,1,2^q,y))$$
(number of edge cuts by horizontal lines),
$$\overline{\Pi}^h_q = \sum_{x=1}^{2^q} \bar{n}(Z^2_k, G(1,1,x,2^q))$$
(number of edge cuts by vertical lines).

We develop and solve a system of recurrences for $\overline{\Pi}_q$ and reverse the sequence of reductions to obtain the closed-form solutions for $\Omega_{k,2^q}$, which are summarized in the following four lemmas.

Lemma 6 *For a canonical* Z^2_q,

$$\overline{\Pi}^v_q = \begin{cases} 4\overline{\Pi}^v_{q-1} + 2^q + 1 & \text{if } q > 1, \\ 3 & \text{if } q = 1; \end{cases}$$
$$\overline{\Pi}^h_q = \begin{cases} 4\overline{\Pi}^h_{q-1} + 2(2^{q-1} - 1) + 1 & \text{if } q > 1, \\ 1 & \text{if } q = 1. \end{cases}$$

Proof The number of edges that are cut by horizontal lines can be computed from the edge cuts within four quadrants plus cuts on three connecting edges $(Q_1(Z^2_q), Q_2(Z^2_q))$, $(Q_3(Z^2_q), Q_4(Z^2_q))$, and $(Q_2(Z^2_q), Q_3(Z^2_q))$:

$$\begin{aligned}
\overline{\Pi}^v_q &= \sum_{y=1}^{2^q} \bar{n}(Z^2_q, G(1,1,2^q,y)) \\
&= \sum_{y=1}^{2^{q-1}} \bar{n}(Z^2_q, G(1,1,2^q,y)) \\
&\quad \text{(cuts within } Q_1(Z^2_q), Q_3(Z^2_q) \text{ plus the cuts on connecting edges)} \\
&\quad + \sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Z^2_q, G(1,1,2^q,y)) \\
&\quad \text{(cuts within } Q_2(Z^2_q), Q_4(Z^2_q) \text{ plus the cuts on connecting edge)} \\
&= (\sum_{y=1}^{2^{q-1}} \bar{n}(Q_1(Z^2_q), G(1,1,2^{q-1},y)) \\
&\quad \text{(cuts within } Q_1(Z^2_q)) \\
&\quad + \sum_{y=1}^{2^{q-1}} \bar{n}(Q_3(Z^2_q), G(2^{q-1}+1,1,2^q,y)) \\
&\quad \text{(cuts within } Q_3(Z^2_q)) \\
&\quad + \sum_{y=2^{q-1}}^{2^{q-1}} 1 \\
&\quad \text{(cuts on connecting edge } (Q_1(Z^2_q), Q_2(Z^2_q))) \\
&\quad + \sum_{y=1}^{2^{q-1}} 1 \\
&\quad \text{(cuts on connecting edge } (Q_2(Z^2_q), Q_3(Z^2_q))) \\
&\quad + \sum_{y=2^{q-1}}^{2^{q-1}} 1) \\
&\quad \text{(cuts on connecting edge } (Q_3(Z^2_q), Q_4(Z^2_q))) \\
&\quad + (\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Q_2(Z^2_q), G(1, 2^{q-1}+1, 2^{q-1}, y)) \\
&\quad \text{(cuts within } Q_2(Z^2_q)) \\
&\quad + \sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Q_4(Z^2_q), G(2^{q-1}+1, 2^{q-1}+1, 2^q, y)) \\
&\quad \text{(cuts within } Q_4(Z^2_q)) \\
&\quad + \sum_{y=2^{q-1}+1}^{2^q-1} 1) \\
&\quad \text{(cuts on connecting edges } (Q_2(Z^2_q), Q_3(Z^2_q))) \\
&= (\sum_{y=1}^{2^{q-1}} \bar{n}(Z^2_{q-1}, G(1,1,2^{q-1},y))) \\
&\quad (Q_1(Z^2_q)\text{: a canonical } Z^2_{q-1}) \\
&\quad + (\sum_{y=1}^{2^{q-1}} \bar{n}(Z^2_{q-1}, G(1,1,2^{q-1},y)) \\
&\quad (Q_3(Z^2_q)\text{: a canonical } Z^2_{q-1}) \\
&\quad + 1 + 2^{q-1} + 1) \\
&\quad + (\sum_{y=1}^{2^{q-1}} \bar{n}(Z^2_{q-1}, G(1,1,2^{q-1},y))) \\
&\quad (Q_2(Z^2_q)\text{: a canonical } Z^2_{q-1}) \\
&\quad + (\sum_{y=1}^{2^{q-1}} \bar{n}(Z^2_{q-1}, G(1,1,2^{q-1},y)) \\
&\quad (Q_4(Z^2_q)\text{: a canonical } Z^2_{q-1}) \\
&\quad + 2^{q-1} - 1) \\
&= \overline{\Pi}^v_{q-1} + (\overline{\Pi}^v_{q-1} + 2^{q-1} + 2) + \overline{\Pi}^v_{q-1} + (\overline{\Pi}^v_{q-1} + 2^{q-1} - 1) \\
&= 4\overline{\Pi}^v_{q-1} + 2^q + 1.
\end{aligned}$$

The proof of $\overline{\Pi}^h_q$ is similar to that of $\overline{\Pi}^v_q$. □

The closed-form solutions for $\overline{\Pi}_q$ are employed to establish a system of recurrences for $\overline{\Omega}^c_{q,2^q}$.

Lemma 7 *For a canonical* Z^2_q,

$$\overline{\Omega}^{c_1}_{q,2^q} = \begin{cases} 4\overline{\Omega}^{c_1}_{q-1,2^{q-1}} + 2^q \cdot \overline{\Pi}^{v}_{q-1} + 2^q \cdot \overline{\Pi}^{h}_{q-1} + 2^{2q} - 2^q + 3 & \text{if } q > 1, \\ 5 & \text{if } q = 1; \end{cases}$$

$$\overline{\Omega}^{c_2}_{q,2^q} = \begin{cases} 4\overline{\Omega}^{c_2}_{q-1,2^{q-1}} + 2^q \cdot \overline{\Pi}^{v}_{q-1} + 2^q \cdot \overline{\Pi}^{h}_{q-1} + 2^{2q} + 2^q & \text{if } q > 1, \\ 6 & \text{if } q = 1. \end{cases}$$

Proof As in Fig. 8 and the case discussion for $\Omega^{c_1}_{q,2^q}$, we can split $\Omega^{c_1}_{q,2^q}$ into four parts:

$$\begin{aligned}
\overline{\Omega}^{c_1}_{q,2^q} &= \sum_{x=1}^{2^q}\sum_{y=1}^{2^q} \bar{n}(Z_q^2, G(1,1,x,y)) \\
&= \sum_{x=1}^{2^{q-1}}\sum_{y=1}^{2^{q-1}} \bar{n}(Z_q^2, G(1,1,x,y)) \\
&\quad \text{(Fig. 8a)} \\
&\quad + \sum_{x=1}^{2^{q-1}}\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Z_q^2, G(1,1,x,y)) \\
&\quad \text{(Fig. 8b)} \\
&\quad + \sum_{x=2^{q-1}+1}^{2^q}\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Z_q^2, G(1,1,x,y)) \\
&\quad \text{(Fig. 8c)} \\
&\quad + \sum_{x=2^{q-1}+1}^{2^q}\sum_{y=1}^{2^{q-1}} \bar{n}(Z_q^2, G(1,1,x,y)) \\
&\quad \text{(Fig. 8d)} \\
&= (\sum_{x=1}^{2^{q-1}}\sum_{y=1}^{2^{q-1}} \bar{n}(Q_1(Z_q^2), G(1,1,x,y)) \\
&\quad \text{(cuts within } Q_1(Z_q^2)) \\
&\quad + \sum_{x=2^{q-1}}^{2^{q-1}}\sum_{y=2^{q-1}}^{2^{q-1}} 1) \\
&\quad \text{(cut on connecting edge } (Q_1(Z_q^2), Q_2(Z_q^2))) \\
&\quad + (\sum_{x=1}^{2^{q-1}}\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Q_2(Z_q^2), G(1,1,x,y)) \\
&\quad \text{(cuts within } Q_2(Z_q^2)) \\
&\quad + \sum_{x=1}^{2^{q-1}}\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Q_1(Z_q^2), G(1,1,x,2^{q-1})) \\
&\quad \text{(cuts within } Q_1(Z_q^2) \text{ by vertical side of } G) \\
&\quad + \sum_{x=1}^{2^{q-1}-1}\sum_{y=2^{q-1}+1}^{2^q} 1 \\
&\quad \text{(cuts on connecting edge } (Q_1(Z_q^2), Q_2(Z_q^2))) \\
&\quad + \sum_{x=2^{q-1}}^{2^{q-1}}\sum_{y=2^q}^{2^q} 1) \\
&\quad \text{(cut on connecting edge } (Q_2(Z_q^2), Q_3(Z_q^2))) \\
&\quad + (\sum_{x=2^{q-1}+1}^{2^q}\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Q_4(Z_q^2), G(1,1,x,y)) \\
&\quad \text{(cuts within } Q_4(Z_q^2)) \\
&\quad + \sum_{x=2^{q-1}+1}^{2^q}\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Q_2(Z_q^2), G(1,1,2^{q-1},y)) \\
&\quad \text{(cuts within } Q_2(Z_q^2) \text{ by horizontal side of } G) \\
&\quad + \sum_{x=2^{q-1}+1}^{2^q}\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Q_3(Z_q^2), G(1,1,x,2^{q-1})) \\
&\quad \text{(cuts within } Q_3(Z_q^2) \text{ by vertical side of } G) \\
&\quad + \sum_{x=2^{q-1}+1}^{2^q}\sum_{y=2^{q-1}+1}^{2^q-1} 1 \\
&\quad \text{(cuts on connecting edge } (Q_2(Z_q^2), Q_3(Z_q^2))) \\
&\quad + \sum_{x=2^{q-1}+1}^{2^q-1}\sum_{y=2^{q-1}+1}^{2^q} 1) \\
&\quad \text{(cuts on connecting edge } (Q_3(Z_q^2), Q_4(Z_q^2))) \\
&\quad + (\sum_{x=2^{q-1}+1}^{2^q}\sum_{y=1}^{2^{q-1}} \bar{n}(Q_3(Z_q^2), G(1,1,x,y)) \\
&\quad \text{(cuts within } Q_3(Z_q^2)) \\
&\quad + \sum_{x=1}^{2^{q-1}}\sum_{y=2^{q-1}+1}^{2^q} \bar{n}(Q_1(Z_q^2), G(1,1,2^{q-1},y)) \\
&\quad \text{(cuts within } Q_1(Z_q^2) \text{ by horizontal side of } G) \\
&\quad + \sum_{x=1}^{2^{q-1}}\sum_{y=2^{q-1}}^{2^{q-1}} 1 \\
&\quad \text{(cuts on connecting edge } (Q_1(Z_q^2), Q_2(Z_q^2))) \\
&\quad + \sum_{x=1}^{2^{q-1}}\sum_{y=2^{q-1}+1}^{2^q} 1 \\
&\quad \text{(cuts on connecting edge } (Q_2(Z_q^2), Q_3(Z_q^2))) \\
&\quad + \sum_{x=2^q}^{2^q}\sum_{y=2^{q-1}}^{2^{q-1}} 1) \\
&\quad \text{(cut on connecting edge } (Q_3(Z_q^2), Q_4(Z_q^2)))
\end{aligned}$$

$$
\begin{aligned}
=&(\sum_{x=1}^{2^{q-1}}\sum_{y=1}^{2^{q-1}}\bar{n}(Z^2_{q-1},G(1,1,x,y))\\
&(Q_1(Z^2_q)\text{: a canonical } Z^2_{q-1})\\
&+1)\\
&+(\sum_{x=1}^{2^{q-1}}\sum_{y=1}^{2^{q-1}}\bar{n}(Z^2_{q-1},G(1,1,x,y))\\
&\quad(Q_2(Z^2_q)\text{: a canonical } Z^2_{q-1})\\
&+(2^{q-1})\sum_{x=1}^{2^{q-1}}\bar{n}(Z^2_q,G(1,1,x,2^{q-1}))\\
&\quad(Q_1(Z^2_q)\text{: a canonical } Z^2_{q-1})\\
&+(2^{q-1}(2^{q-1}-1))\\
&+1)\\
&+(\sum_{x=1}^{2^{q-1}}\sum_{y=1}^{2^{q-1}}\bar{n}(Z^2_{q-1},G(1,1,x,y))\\
&\quad(Q_4(Z^2_q)\text{: a canonical } Z^2_{q-1})\\
&+(2^{q-1})\sum_{y=1}^{2^{q-1}}\bar{n}(Z^2_{q-1},G(1,1,2^{q-1},y))\\
&\quad(Q_2(Z^2_q)\text{: a canonical } Z^2_{q-1})\\
&+(2^{q-1})\sum_{x=2^{q-1}+1}^{2^q}\bar{n}(Z^2_{q-1},G(1,1,x,2^{q-1}))\\
&\quad(Q_3(Z^2_q)\text{: a canonical } Z^2_{q-1})\\
&+(2^{q-1}(2^{q-1}-1))\\
&+(2^{q-1}(2^{q-1}-1)))\\
&+(\sum_{x=1}^{2^{q-1}}\sum_{y=1}^{2^{q-1}}\bar{n}(Z^2_{q-1},G(1,1,x,y))\\
&\quad(Q_3(Z^2_q)\text{: a canonical } Z^2_{q-1})\\
&+\sum_{x=1}^{2^{q-1}}\sum_{y=1}^{2^{q-1}}\bar{n}(Z^2_{q-1},G(1,1,2^{q-1},y))\\
&\quad(Q_1(Z^2_q)\text{: a canonical } Z^2_{q-1})\\
&+2^{q-1}+(2^{q-1})^2+1)\\
=&(\overline{\Omega}^{c_1}_{q-1,2^{q-1}}+1)\\
&+(\overline{\Omega}^{c_1}_{q-1,2^{q-1}}+2^{q-1}\cdot\overline{\Pi}^h_{q-1}+2^{q-1}(2^{q-1}-1)+1)\\
&+(\overline{\Omega}^{c_1}_{q-1,2^{q-1}}+2^{q-1}\cdot\overline{\Pi}^v_{q-1}+2^{q-1}\cdot\overline{\Pi}^h_{q-1}\\
&+2^{q-1}(2^{q-1}-1)+2^{q-1}(2^{q-1}-1))\\
&+(\overline{\Omega}^{c_1}_{q-1,2^{q-1}}+2^{q-1}\cdot\overline{\Pi}^v_{q-1}+2^{q-1}+(2^{q-1})^2+1)\\
=&4\overline{\Omega}^{c_1}_{q-1,2^{q-1}}+2^q\cdot\overline{\Pi}^v_{q-1}+2^q\cdot\overline{\Pi}^h_{q-1}+2^{2q}-2^q+3.
\end{aligned}
$$

The proof for $\overline{\Omega}^{c_2}_{q,2^q}$ is similar to that of $\overline{\Omega}^{c_1}_{q,2^q}$. □

The closed-form solutions for $\overline{\Omega}^{c}_{q,2^q}$ and $\overline{\Pi}_q$ are employed to obtain exact formulas for $\Omega^c_{k,2^q}$.

Lemma 8 *For a canonical Z^2_k structured as an Z^2_{k-q}-curve interconnecting $2^{2(k-q)}$ Z^2_q-subcurves,*

$$
\begin{aligned}
\Omega^{c_1}_{k,2^q}&=\overline{\Omega}^{c_1}_{q,2^q}-\overline{\Pi}^h_q-\overline{\Pi}^v_q,\\
\Omega^{c_2}_{k,2^q}&=\overline{\Omega}^{c_2}_{q,2^q}-\overline{\Pi}^h_q-\overline{\Pi}^v_q.
\end{aligned}
$$

Proof By the definition, we have:

$$
\begin{aligned}
\Omega^{c_1}_{k,2^q}=&\sum_{x=1}^{2^q-1}\sum_{y=1}^{2^q-1}\bar{n}(Z^2_k,G(1,1,x,y))\\
=&\sum_{x=1}^{2^q}\sum_{y=1}^{2^q-1}\bar{n}(Z^2_k,G(1,1,x,y))-\sum_{x=2^q}^{2^q}\sum_{y=1}^{2^q-1}\bar{n}(Z^2_k,G(1,1,x,y))\\
=&(\sum_{x=1}^{2^q}\sum_{y=1}^{2^q}\bar{n}(Z^2_k,G(1,1,x,y))-\sum_{x=1}^{2^q}\sum_{y=2^q}^{2^q}\bar{n}(Z^2_k,G(1,1,x,y)))\\
&-(\sum_{x=2^q}^{2^q}\sum_{y=1}^{2^q}\bar{n}(Z^2_k,G(1,1,x,y))-\sum_{x=2^q}^{2^q}\sum_{y=2^q}^{2^q}\bar{n}(Z^2_k,G(1,1,x,y)))\\
=&\overline{\Omega}^{c_1}_{q,2^q}-\overline{\Pi}^h_q-\overline{\Pi}^v_q.
\end{aligned}
$$

The proof of $\Omega^{c_2}_{k,2^q}$ is similar to that of $\Omega^{c_1}_{k,2^q}$. □

The exact formulas for $\Omega^c_{k,2^q}$ are employed to establish a system of recurrences for $\Omega_{k,2^q}$.

Lemma 9 *For a canonical Z^2_k structured as an Z^2_{k-q}-curve interconnecting $2^{2(k-q)}$ Z^2_q-subcurves,*

$$
\Omega^L_{k,2^q}=\begin{cases}2\Omega^L_{k-1,2^q}+\Omega^{c_1}_{k-1,2^q}+\Omega^{c_2}_{k-1,2^q}+(2^q-1)^2+(2^q-1) & \text{if } k>q,\\ \overline{\Pi}^h_q & \text{if } k=q;\end{cases}
$$

$$
\Omega^B_{k,2^q}=\begin{cases}2\Omega^B_{k-1,2^q}+\Omega^{c_1}_{k-1,2^q}+\Omega^{c_2}_{k-1,2^q}+(2^q-1)^2+(2^q-1) & \text{if } k>q,\\ \overline{\Pi}^v_q & \text{if } k=q.\end{cases}
$$

Proof Similar to the proof of Lemma 8 and noting the parallel of Lemmas 3 and 4 for Hilbert curves with Lemmas 8 and 9 for z-order curves, from the definition, we have (see Fig. 6):

$$
\begin{aligned}
\Omega^L_{k,2^q}&=\sum_{x=1}^{2^q-1}\sum_{y=1}^{2^k-2^q+1}\bar{n}(Z^2_k,G(1,y,x,y+2^q-1))\\
&=\sum_{x=1}^{2^q-1}\sum_{y=1}^{2^{k-1}-2^q+1}\bar{n}(Z^2_k,G(1,y,x,y+2^q-1))\\
&\quad(\text{in } Q_1(Z^2_k))
\end{aligned}
$$

$$
\begin{aligned}
&+\sum_{x=1}^{2^q-1}\sum_{y=2^{k-1}-2^q+2}^{2^{k-1}} \bar{n}(Z_k^2, G(1,y,x,y+2^q-1))\\
&\quad(\text{across } Q_1(Z_k^2), Q_2(Z_k^2))\\
&+\sum_{x=1}^{2^q-1}\sum_{y=2^{k-1}+1}^{2^k-2^q+1} \bar{n}(Z_k^2, G(1,y,x,y+2^q-1))\\
&\quad(\text{in } Q_2(Z_k^2))\\
=&\Big(\sum_{x=1}^{2^q-1}\sum_{y=1}^{2^{k-1}-2^q+1} \bar{n}(Q_1(Z_k^2), G(1,y,x,y+2^q-1))\Big)\\
&(\text{cuts within } Q_1(Z_k^2))\\
&+\Big(\sum_{x=1}^{2^q-1}\sum_{y=2^{k-1}-2^q+2}^{2^{k-1}} \bar{n}(Q_1(Z_k^2), G(1,y,x,2^{k-1}))\\
&\quad(\text{zooming in } Q_1(Z_k^2))\\
&+\sum_{x=1}^{2^q-1}\sum_{y=2^{k-1}-2^q+2}^{2^{k-1}} \bar{n}(Q_2(Z_k^2), G(1,2^{k-1}+1,x,y+2^q-1))\\
&\quad(\text{zooming in } Q_2(Z_k^2))\\
&+\sum_{x=1}^{2^q-1}\sum_{y=2^{k-1}-2^q+2}^{2^{k-1}} 1\Big)\\
&\quad(\text{cuts on connecting edge } (Q_1(Z_k^2), Q_2(Z_k^2)))\\
&+\Big(\sum_{x=1}^{2^q-1}\sum_{y=2^{k-1}+1}^{2^k-2^q+1} \bar{n}(Q_2(Z_k^2), G(1,y,x,y+2^q-1))\\
&\quad(\text{cuts within } Q_2(Z_k^2))\\
&+\sum_{x=1}^{2^{q-1}-1}\sum_{y=2^{k-1}+1}^{2^{k-1}+1} 1\Big)\\
&\quad(\text{cuts on connecting edge } (Q_1(Z_k^2), Q_2(Z_k^2)))\\
=&\Big(\sum_{x=1}^{2^q-1}\sum_{y=1}^{2^{k-1}-2^q+1} \bar{n}(Z_{k-1}^2, G(1,y,x,y+2^q-1))\Big)\\
&\quad(Q_1(Z_k^2)\text{: a canonical } Z_{k-1}^2)\\
&+\Big(\sum_{x=1}^{2^q-1}\sum_{y=2^{k-1}-2^q+2}^{2^{k-1}} \bar{n}(Z_{k-1}^2, G(1,y,x,2^{k-1}))\\
&\quad(Q_1(Z_k^2)\text{: a canonical } Z_{k-1}^2)\\
&+\sum_{x=1}^{2^q-1}\sum_{y=1}^{2^q-1} \bar{n}(Z_{k-1}^2, G(1,1,x,y))\\
&\quad(Q_2(Z_k^2)\text{: a canonical } Z_{k-1}^2)\\
&+(2^q-1)^2\Big)\\
&+\Big(\sum_{x=1}^{2^q-1}\sum_{y=1}^{2^{k-1}-2^q+1} \bar{n}(Z_{k-1}^2, G(1,y,x,y+2^q-1))\\
&\quad(Q_2(Z_k^2)\text{: a canonical } Z_{k-1}^2)\\
&+(2^q-1)\Big)\\
=&\Omega_{k-1,2^q}^L\\
&+(\Omega_{k-1,2^q}^{c_1}+\Omega_{k-1,2^q}^{c_2}+(2^q-1)^2)\\
&+(\Omega_{k-1,2^q}^L+(2^q-1))\\
=&2\Omega_{k-1,2^q}^L+\Omega_{k-1,2^q}^{c_1}+\Omega_{k-1,2^q}^{c_2}+(2^q-1)^2+(2^q-1).
\end{aligned}
$$

For $\Omega_{k-1,2^q}^B$, the proof is similar to that of $\Omega_{k,2^q}^L$. □

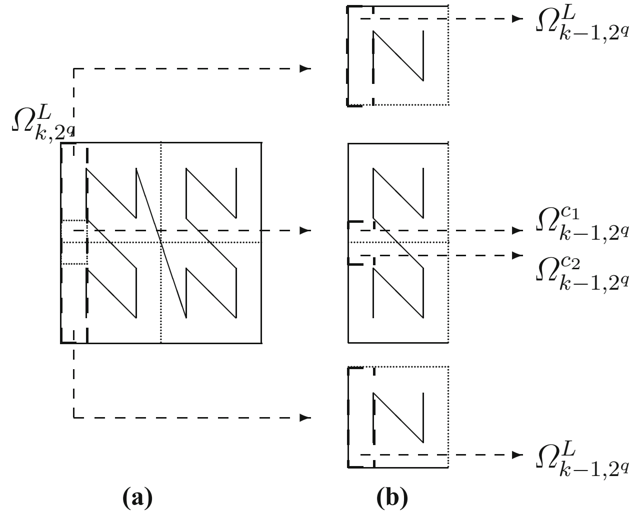
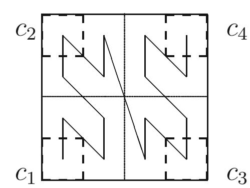


Fig. 6 **a** $\Omega_{k,2^q}^L$ for a canonical Z_k^2; **b** its recursive decomposition

Fig. 7 The four $(2^q-1)\times(2^q-1)$ corners of a canonical Z_k^2



We obtain the closed-form solutions for $\Omega_{k,2^q}$ by using the mathematical software Maple.

Query Subgrids Overlapping with All Quadrants

For a $2^q\times 2^q$ query subgrid G that overlaps four quadrants around the center of Z_k^2, when zooming in on the incomplete rectangular subgrid $G\cap G_1$, where G_1 denotes the subspace of $Q_1(Z_k^2)$ (with both side-lengths at most 2^q-1), we reduce $\sum_{\text{all } G\cap G_1}\bar{n}(Z_k^2, G\cap G_1)$ to $\Omega_{k-1,2^q}^{c_4}(=\Omega_{k-1,2^q}^{c_1})$. Similar consideration leads to a reduction of $\sum_{\text{all} G\cap G'}\bar{n}(Z_k^2, G\cap G')$ to $\Omega_{k-1,2^q}^{c_3}(=\Omega_{k-1,2^q}^{c_2})$, $\Omega_{k-1,2^q}^{c_2}$ and $\Omega_{k-1,2^q}^{c_1}$ when $G\cap G'$ denotes the subspace for G overlapping $Q_2(Z_k^2)$, $Q_3(Z_k^2)$, or $Q_3(Z_k^2)$, respectively.

Thus, the summation of numbers of edge cuts for all identically shaped $2^q\times 2^q$ query subgrids G that overlap all four quadrants is:

$$2\Omega_{k-1,2^q}^{c_2}+2\Omega_{k-1,2^q}^{c_1}.$$

The Big Picture: Computing $E_q(Z_k^2)$

The results in the previous three subsections yield $\varepsilon_{k,q}(Z_k^2)$. Hence, we have the following lemma for $E_q(Z_k^2)$.

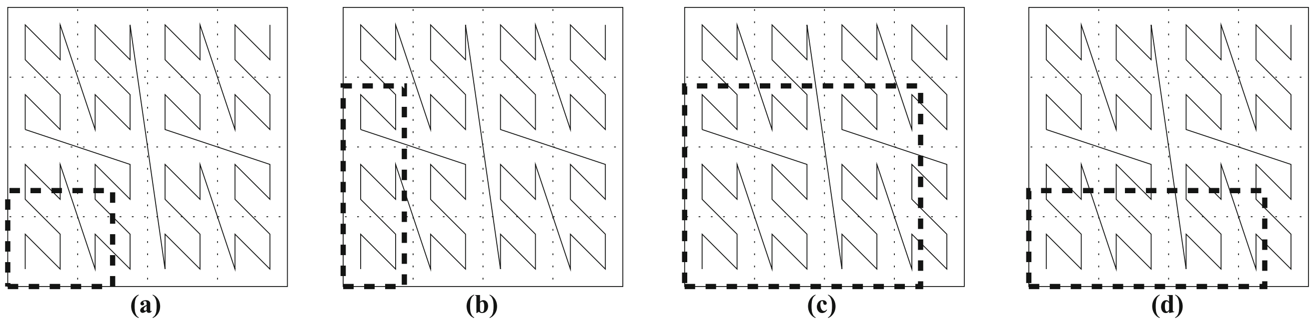


Fig. 8 Four overlapping scenarios when decomposing $\overline{\Omega}^{c_1}_{q,2^q}$ in a canonical Z^2_q: **a** contained in $Q_1(Z^2_q)$; **b** and **d** overlapping with exactly two quadrants; **c** overlapping with all quadrants

Lemma 10 *For a canonical Z^2_k, the recurrence for total number of cuts on edges by all identically shaped $2^q \times 2^q$ subgrids G:*

$$E_q(Z^2_k) = \begin{cases} 4E_q(Z^2_{k-1}) + 2(\Omega^B_{k-1,2^q} + \Omega^B_{k-1,2^q} + 2(2^q-1)) \\ \quad +2(\Omega^L_{k-1,2^q} + \Omega^L_{k-1,2^q} + 2(2^q-1)) \\ \quad +(2\Omega^{c_1}_{k-1,2^q} + 2\Omega^{c_2}_{k-1,2^q} + 2(2^q-1)^2) & \text{if } k > q+1, \\ 4E_q(Z^2_{k-1}) + 2(\Omega^B_{k-1,2^q} + \Omega^B_{k-1,2^q}) \\ \quad +2(\Omega^L_{k-1,2^q} + \Omega^L_{k-1,2^q} + 2(2^q-1)) \\ \quad +(2\Omega^{c_1}_{k-1,2^q} + 2\Omega^{c_2}_{k-1,2^q} + 2(2^q-1)^2) & \text{if } k = q+1, \\ 2 & \text{if } k = q. \end{cases}$$

Proof For $k > q+1$, the proof is similar to the proof of Lemma 5:

Case 1: G overlaps with exactly $Q_1(Z^2_k)$ and $Q_2(Z^2_k)$. This part is reduced to $\Omega^T_{k-1,2^q} (= \Omega^B_{k-1,2^q})$ (cuts on $Q_1(Z^2_k)$), $\Omega^B_{k-1,2^q}$ (cuts on $Q_2(Z^2_k)$), $2(2^q-1)$ cuts on the connecting edge $(Q_1(Z^2_k), Q_2(Z^2_k))$ when subgrids G align left-most and right-most side of these two quadrants.

Case 2: G overlaps with exactly $Q_2(Z^2_k)$ and $Q_4(Z^2_k)$. This part is reduced to $\Omega^R_{k-1,2^q} (= \Omega^L_{k-1,2^q})$ (cuts on $Q_2(Z^2_k)$), $\Omega^L_{k-1,2^q}$ (cuts on $Q_4(Z^2_k)$), (2^q-1) cuts on the connecting edge $(Q_2(Z^2_k), Q_3(Z^2_k))$ when subgrids G align top-most side of these two quadrants, and (2^q-1) cuts on the connecting edge $(Q_3(Z^2_k), Q_4(Z^2_k))$ when subgrids G align bottom-most side of these two quadrants.

Case 3: G overlaps with exactly $Q_3(Z^2_k)$ and $Q_4(Z^2_k)$. This part is reduced to $\Omega^T_{k-1,2^q} (= \Omega^B_{k-1,2^q})$ (cuts on $Q_3(Z^2_k)$), $\Omega^B_{k-1,2^q}$ (cuts on $Q_4(Z^2_k)$), and $2(2^q-1)$ cuts on the connecting edge $(Q_3(Z^2_k), Q_4(Z^2_k))$ when subgrids G align left-most and right-most side of these two quadrants.

Case 4: G overlaps with exactly $Q_1(Z^2_k)$ and $Q_3(Z^2_k)$. This part is reduced to $\Omega^R_{k-1,2^q} (= \Omega^L_{k-1,2^q})$ (cuts on $Q_1(Z^2_k)$), $\Omega^L_{k-1,2^q}$ (cuts on $Q_3(Z^2_k)$), (2^q-1) cuts on the connecting edge $(Q_2(Z^2_k), Q_3(Z^2_k))$ when subgrids G align bottom-most side of these two quadrants, and (2^q-1) cuts on the connecting edge $(Q_1(Z^2_k), Q_2(Z^2_k))$ when subgrids G align top-most side of these two quadrants.

Case 5: G overlaps with exactly all four quadrants. This part is reduced to $\Omega^{c_4}_{k-1,2^q} (= \Omega^{c_1}_{k-1,2^q})$ (cuts on $Q_1(Z^2_k)$), $\Omega^{c_3}_{k-1,2^q} (= \Omega^{c_2}_{k-1,2^q})$ (cuts on $Q_2(Z^2_k)$), $\Omega^{c_2}_{k-1,2^q}$ (cuts on $Q_3(Z^2_k)$), $\Omega^{c_1}_{k-1,2^q}$ (cuts on $Q_4(Z^2_k)$), and $2(2^q-1)^2$ cuts on the connecting edges $(Q_1(Z^2_k), Q_2(Z^2_k))$ and $(Q_3(Z^2_k), Q_4(Z^2_k))$.

Combining all the five cases, we complete the recurrence for $k > q+1$.

For $k = q+1$, there are no cuts on connecting edges in Cases 1 and 3 because the connecting edges are inside of all the subspaces that are across exactly two quadrants (first and second, or third and fourth quadrants).

For the case of $k = q$, there are two cuts that are the edge cut between entry grid point and other curve, and the edge cut between exit grid point and other curve. □

The exact formula for the summation of all numbers of edge cuts over all identically shaped $2^q \times 2^q$ subgrids, $E_q(Z^2_k)$, is:

$$\begin{aligned} E_q(Z^2_k) &= 2^{2k+q+2} - 2^{2k+2} + 3 \cdot 2^{2k-q} - 2^{2k-2q} \\ &\quad - 2^{k+2q+3} + 3 \cdot 2^{k+q+2} \\ &\quad - 2^{k+3} + 2^{k-q+2} + 2^{3q+2} - 2^{2q+3} + 2^{q+2}. \end{aligned}$$

The summation of all numbers of clusters over all identically shaped $2^q \times 2^q$ query subgrids of an Z^2_k-structural grid space $[2^k]^2$ is:

$$\mathcal{S}_{k,2^q}(Z_k^2) = \frac{E_q(Z_k^2)}{2}.$$

The mean number of cluster within a subspace $2^q \times 2^q$ is:

$$\frac{\mathcal{S}_{k,2^q}(Z_k^2)}{(2^k - 2^q + 1)^2} = \frac{E_q(Z_k^2)}{2(2^k - 2^q + 1)^2}.$$

Thus, the exact formula for the mean number of cluster within a subspace $2^q \times 2^q$ for Z_k^2 is corollarily derived.

Theorem 2 *The mean number of cluster over all identical subspaces $2^q \times 2^q$ for Z_k^2 is*:

$$\begin{aligned}&(2^{2k+q+2} - 2^{2k+2} + 3 \cdot 2^{2k-q} + 2^{2k-2q}\\ &\quad - 2^{k+2q+3} + 3 \cdot 2^{k+q+2} + 2^{k+3}\\ &\quad + 2^{k-q+2} + 2^{3q+2} + 2^{2q+3} + 2^{q+2})/(2(2^k - 2^q + 1)^2).\end{aligned}$$

Comparisons and Validation

For a space-filling curve C_k indexing the grid space $[2^k]^2$, denote by $\mathcal{S}_{k,q}(C_k)$ the mean number of clusters over all $2^q \times 2^q$ subgrids of the C_k-structural grid space.

The exact formulas for $E_q(H_k^2)$ and $E_q(Z_k^2)$ give the exact formulas for $\mathcal{S}_{k,2^q}(H_k^2)$ and $\mathcal{S}_{k,2^q}(Z_k^2)$. We simplify the exact results asymptotically as follows. For sufficiently large k and q with $k \gg q$ (typical scenario for range queries),

$$\frac{\mathcal{S}_{k,2^q}(Z_k^2)}{\mathcal{S}_{k,2^q}(H_k^2)} = \frac{E_q(Z_k^2)}{E_q(H_k^2)} \approx 2$$

With respect to the $\mathcal{S}_{k,q}$-statistics, the Hilbert curve family clearly performs better than the z-order curve family over the considered ranges for k and q.

Conclusion

The analytical study of the clustering performances of 2-dimensional order-k Hilbert and z-order curve families are based upon the clustering statistics $\mathcal{S}_{k,2^q}$—mean number of clusters over all $2^q \times 2^q$ identically shaped subgrids, respectively. By taking advantage of self-similar properties of Hilbert and z-order curve, we derive their exact formulas for $\mathcal{S}_{k,2^q}$.

Our two exact analytical results yield the followings: (1) a comparison of relative performances of Hilbert and z-order curve families with respect to the clustering measure: when the grid-order is sufficiently larger than the subgrid-order (typical scenario for most applications), that is, for sufficiently large k and q with $k \gg q$, Hilbert curve family performs significantly better than z-order curve family with respect to $\mathcal{S}_{k,2^q}$, and (2) good agreements with asymptotic analyses of continuous space-filling curves [20, 26] and of (non-continuous) z-order curve [26] in the literature.

Funding This study was not supported by any funding.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Alber J, Niedermeier R. On multi-dimensional curves with Hilbert property. Theory Comput Syst. 2000;33(4):295–312.
2. Asano T, Ranjan D, Roos T, Welzl E, Widmayer P. Space-filling curves and their use in the design of geometric data structures. Theor Comput Sci. 1997;181(1):3–15.
3. Bader M. Space-Filling Curves—An Introduction with Applications in Scientific Computing, vol. 9. Texts in Computational Science and Engineering. Berlin: Springer; 2013.
4. Ban X, Goswami M, Zeng W, Gu X, Gao J. Topology dependent space filling curves for sensor networks and applications. In: Proceedings of the IEEE INFOCOM 2013, Turin, Italy, April 14–19, 2013. New York: IEEE; 2013. p. 2166–74.
5. Böhm C, Berchtold S, Keim DA. Searching in high-dimensional spaces—index structures for improving the performance of multimedia databases. ACM Comput Surv. 2001;33(3):322–73.
6. Chen H-L, Chang Y-I. Neighbor-finding based on space-filling curves. Inf Syst. 2005;30(3):205–26.
7. Dai HK, Su HC. Approximation and analytical studies of inter-clustering performances of space-filling curves. In: Proceedings of the international conference on discrete random walks (Discrete mathematics and theoretical computer science, vol. AC (2003)). Strasbourg, France: Discrete Mathematics and Theoretical Computer; 2003. p. 53–68.
8. Dai HK, Su HC. On the locality properties of space-filling curves. In: Ibaraki T, Katoh N, Ono H, editors. Lecture notes in computer science (2906): algorithms and computation: 14th international symposium. ISAAC 2003 proceedings. Berlin: Springer; 2003. p. 385–94.
9. Dai HK, Su HC. Norm-based locality measures of two-dimensional Hilbert curves. In: Dondi R, Fertin G, Mauri G, editors. Algorithmic aspects in information and management—11th international conference. AAIM 2016, Bergamo, Italy, July 18–20, 2016, Proceedings, volume 9778 of Lecture Notes in Computer Science. Berlin: Springer; 2016. p. 14–25.
10. Dai HK, Su HC. On norm-based locality measures of 2-dimensional discrete Hilbert curves. In: Dang TK, Küng J, Takizawa M, Chung TM, editors. Lecture notes in computer science (12466): future data and security engineering. 7th International conference, FDSE 2020, Quy Nhon, Vietnam, November 25–27, 2020, proceedings. Berlin: Springer; 2020. p. 169–84.
11. Gaede V, Günther O. Multidimensional access methods. ACM Comput Surv. 1998;30(2):170–231.
12. Gotsman C, Lindenbaum M. On the metric properties of discrete space-filling curves. IEEE Trans Image Process. 1996;5(5):794–7.
13. He Z, Owen AB. Extensible grids: uniform sampling on a space filling curve. J R Stat Soc Ser B. 2016;78(4):917–31.

14. Jagadish HV. Analysis of the Hilbert curve for representing two-dimensional space. Inf Process Lett. 1997;62(1):17–22.
15. Kaddoura M, Ou C-W, Ranka S. Partitioning unstructured computational graphs for non-uniform and adaptive environments. IEEE Parallel Distrib Technol. 1995;3(3):63–9.
16. Kinney N, Hickman M, Anandakrishnan R, Garner HR. Crossing complexity of space-filling curves reveals entanglement of S-phase DNA. Public Libr Sci One. 2020;15(8):1–20.
17. Lawder JK. The application of space-filling curves to the storage and retrieval of multi-dimensional data. PhD thesis, Birkbeck College, University of London, December 1999.
18. Lempel A, Ziv J. Compression of two-dimensional images. In: Apostolico A, Galil Z, editors. Combinatorial algorithms on words, vol. F12. ASI series. Berlin: Springer; 1984. p. 141–56.
19. Mitchison G, Durbin R. Optimal numberings of an $N \times N$ array. SIAM J Algebraic Discrete Methods. 1986;7(4):571–82.
20. Moon B, Jagadish HV, Faloutsos C, Saltz JH. Analysis of the clustering properties of the Hilbert space-filling curve. IEEE Trans Knowl Data Eng. 2001;13(1):124–41.
21. Pérez A, Kamata S, Kawaguchi E. Peano scanning of arbitrary size images. In: Proceedings of the international conference on pattern recognition. New York: IEEE Computer Society; 1992. p. 565–8.
22. Platzman LK, Bartholdi JJ III. Spacefilling curves and the planar travelling salesman problem. J ACM. 1989;36(4):719–37.
23. Sagan H. Space-filling curves. New York: Springer; 1994.
24. Wang C, Jiang H, Dong Y. Connectivity-based space filling curve construction algorithms in high genus 3D surface WSNs. ACM Trans Sens Netw. 2016;12(3):22:1-22:29.
25. Xu P, Tirthapura S. On the optimality of clustering through a space filling curve. In: PODS '12 proceedings of the 31st symposium on principles of database systems. New York: Association for Computing Machinery; 2012. p. 215–24.
26. Xu P, Tirthapura S. Optimality of clustering properties of space-filling curves. ACM Trans Database Syst. 2014;39(2):1–27.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.