Convexity preserving deformations of digital sets: Characterization of removable and insertable pixels

Lama Tarsissi [a,b,*], Yukiko Kenmochi [c], Pascal Romon [d], David Coeurjolly [e], Jean-Pierre Borel [f]

[a] LIGM, Univ Gustave Eiffel, CNRS, ESIEE Paris, F-77454 Marne-la-Vallée, France
[b] Dept.Sciences and Engineering-Sorbonne University Abu Dhabi, United Arab Emirates
[c] Normandie Univ, UNICAEN, ENSICAEN, CNRS, GREYC, F-14000 Caen, France
[d] LAMA, Univ Gustave Eiffel, UPEC, CNRS, F-77454 Marne-la-Vallée, France
[e] Univ Lyon, CNRS, INSA Lyon, UCBL, LIRIS, UMR5205, Villeurbanne, France
[f] Mathis-XLim, Univ. Limoges - CNRS, F-87060 Limoges, France

ARTICLE INFO

ABSTRACT

In this paper, we are interested in digital convexity. This notion is applied in several domains like image processing and discrete tomography. We choose to study the inflation and deflation of digital convex sets while maintaining the convexity property. Knowing that any digital convex set can be read and identified by its boundary word, we use combinatorics on words perspective instead of a purely geometric approach. In this context, we characterize the points that can be added or removed over the digital convex sets without losing their convexity. Some algorithms are given at the end of each section with examples of each process.

© 2023 Elsevier B.V. All rights reserved.

1. Introduction

In this paper, we provide characterizations of convexity-preserving removable and insertable points of digital convex sets. In $\mathbb{Z}^2$, even a simple transformation (such as rotation) of a digital convex set can cause the loss of the digital convexity property. We aim to provide the discrete equivalent of infinitesimal transform, from which we can derive convexity-preserving set deformations. In order to study such atomic deformations that preserve digital convexity, we need to investigate deflation and inflation processes [30]. The theory of combinatorics on words provides useful tools and techniques for our investigation [22]. Relying on this theory, we are able to characterize the removable points x of a digital convex set C, such that $C \setminus \{x\}$ is still digitally convex. Similarly, we provide the characterization of the insertable points $x \in \overline{C}$ (the complement of C in $\mathbb{Z}^2$) such that $C \cup \{x\}$ is still digitally convex. These two actions can only be applied at some specific points on the digital convex set if we want to preserve digital convexity. As a matter of fact, deflation is easy; finding the correct point to remove from a digital set is simple from a geometrical point of view. On the other hand, inflation is more involved. Propositions and theorems for inflation and deflation are naturally associated with algorithms in the mathematical sense. In [31], we provide an implementable algorithm, containing all the necessary details.

The plan of the paper is the following. Section 2 provides an overview of digital convexity, from combinatorics on words perspective, and gives the basic notations needed to understand our results. In Section 3, we characterize removable

* Corresponding author at: Dept.Sciences and Engineering-Sorbonne University Abu Dhabi, United Arab Emirates.
  E-mail addresses: lama.tarsissi@esiee.fr (L. Tarsissi), yukiko.kenmochi@unicaen.fr (Y. Kenmochi), pascal.romon@univ-eiffel.fr (P. Romon), david.coeurjolly@cnrs.fr (D. Coeurjolly), jean-pierre.borel@unilim.fr (J.-P. Borel).
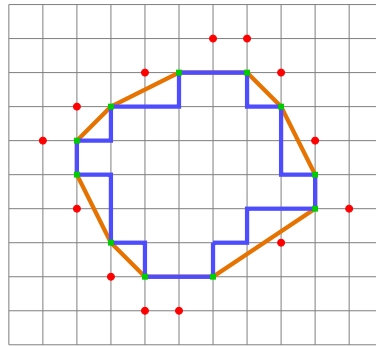
Lowres version

Fig. 1. The green points are the removable points, which are the vertices of the convex hull of a digital convex set C. The red points are the insertable points, which some are the exterior Bézout point of each segment of the convex hull. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
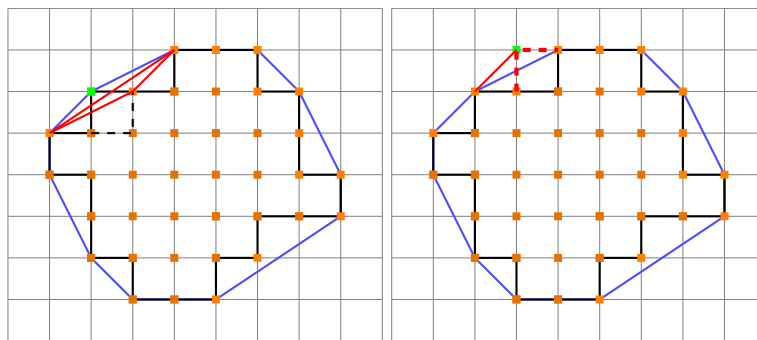

Fig. 2. The first picture shows an example of deflation. The green point x is the point to be removed from a digital convex set (orange points) C. The second picture shows an example of inflation. The green point x is the new point to be added. The blue and red segments in both pictures are respectively $Conv(C)$ and their modifications after each process. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

points and prove that such a point can be any simple point that is a corner of the convex hull of C. As said before, finding a characterization of insertable points is not an easy task. We provide necessary and sufficient conditions to determine candidate points in Section 4. The necessary condition is based on a result from [13]: the authors proved that adding the closest outer point of a segment maintains its digital convexity. Based on this result, we provide in this paper the characterization of all insertable points for the whole digital convex set. We provide two main results for the sufficient condition. The first one is the general case and leads to propagation after inflation. The second result imposes a strong constraint on the sufficient condition. Fig. 1 shows an example of a digitally convex set with some removable and insertable points. For both procedures and after each iteration, we must consider an update for the segments of the convex hull. In this paper, we discuss all the possible cases for this update. One of them is presented in Fig. 2. The last section is left for the conclusions and perspectives.

2. Digital convexity and combinatorics on words

In this section, we first give the definition of digital convex sets based on convex hull [19], which are also called H-convex sets [16]. We then show that their boundaries can be expressed by words represented with the Freeman chain code [18], called a boundary word. After recalling basic notions of combinatorics on words, we present the characterization of digital convexity along boundary words using those notions [8].

2.1. Definitions of digital convex sets

In $\mathbb{R}^2$, a subset R is said to be convex if for any pair of points $x, y \in R$, every point on the straight line segment joining x and y is also within R. This notion, however, cannot be straightforwardly applied to subsets in $\mathbb{Z}^2$. In order to tackle this issue, various notions of convexity of a set S of $\mathbb{Z}^2$ have been proposed. They are categorized into four approaches; the first approach is based on no existence of triplet of collinear points (p, q, r) such that $p, r \in S$ and $q \in \bar{S}$ [25]. The second one is based on the existence of a convex set $R \subset \mathbb{R}^2$ such that the digitization of R is S [29]. Kim proposed
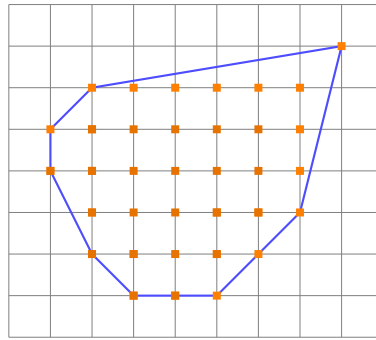
Fig. 3. An example of a digital set C that is digitally convex but not 4-connected. The orange points represent the elements $x \in C$ while the blue line represents $\mathcal{C}onv(C)$.

another definition, which can be characterized by verifying if the digitization of the convex hull of S is equal to S, and also showed its equivalence with the two former [19]; the definition based on the convex hull is later called H-convex by Eckhardt [16]. The fourth approach is based on the notion of digital line segments [20], such that S is digitally convex if the digital line segment joining any pair of points of S belongs to S. Note that this last notion guarantees the connectivity of S contrary to the other three. Under the connectivity assumption, it is then shown that the above different concepts coincide [16].

The definition of connectivity on $\mathbb{Z}^2$ is based on the notion of neighborhood. Given a point $p \in \mathbb{Z}^2$, the neighborhood of p is defined by $\mathcal{N}(p) = \{q \in \mathbb{Z}^2 : \|p - q\|_1 \leq 1\}$, also called 4-neighborhood. We say that a point q is 4-adjacent to p if $q \in \mathcal{N}(p) \setminus \{p\}$. From the reflexive–transitive closure of this adjacency relation on a finite subset $X \subset \mathbb{Z}^2$, we derive the 4-connectivity relation on X, which is an equivalence relation. If there is exactly one equivalence class for this relation, then we say that X is 4-connected. In this article, we consider finite and 4-connected sets of $\mathbb{Z}^2$. As this 4-connectivity assumption yields the coincidence of the above different definitions of digital convexity [16],[1] we use the one based on convex hull [16,19], which is defined for a finite subset $Y \subset \mathbb{R}^2$ as:

$$\mathcal{C}onv(Y) := \{x \in \mathbb{R}^2 \mid x = \sum_{y \in Y} \lambda_y y \ \wedge \ \sum_{y \in Y} \lambda_y = 1 \ \wedge \ \forall y \in Y, \ \lambda_y \geq 0\}.$$

Definition 1 (*Digital Convexity [19]*)**.** A finite subset S of $\mathbb{Z}^2$ is *digitally convex* if $\mathcal{C}onv(S) \cap \mathbb{Z}^2 = S$.

From this definition, we have the following remark, which is contrary to the implication of connectivity in the concept of convexity in $\mathbb{R}^2$ (see Fig. 3 for an example).

Remark 1. Digital convexity does not imply connectivity in $\mathbb{Z}^2$.

The above definition of digital convexity is also used in [12], whose aim is verifying if a given finite 4-connected set S of $\mathbb{Z}^2$ is digitally convex. Their approach focuses rather on the boundary of S, on which maximal line segments and their arithmetic properties are analyzed [12]. On the other hand, properties of the boundary of S based on combinatorics on words are also studied [8]. The problems treated in this article are different from the ones from [8,12], but are related. Our problem is stated as follows; given a finite, 4-connected and digitally convex set C of $\mathbb{Z}^2$ and a point p of C (resp. the complement $\overline{C}$), we would like to verify if $C \setminus \{p\}$ (resp. $C \cup \{p\}$) is still digitally convex (and 4-connected). In order to answer such questions, we use the boundary properties based on combinatorics on words that are presented in [8].

2.2. Boundary words of digitally convex sets

Let $C \subset \mathbb{Z}^2$ be a finite, 4-connected digitally convex set. The border points of C can be tracked by a classical border following algorithm (for example, see [1] for "left-hand-on-wall" border following, i.e. placing the left hand on a wall (border) and then following the wall by maintaining contact between the hand and the wall), which generates a 4-connected sequence of the border points of C. Note that the sequence can include dead-ends and thus sometimes turnaround sub-sequences. Such a sequence is also called the boundary path of C, denoted by $Bd(C)$, and represented by a word obtained by Freeman chain code [18], denoted by $W(Bd(C))$ or simply $W(C)$ and called the boundary word of C [17]. Boundary words are thus defined over an alphabet of four letters $0, 1, \overline{0}, \overline{1}$, which are associated to the right, up, left and down steps, respectively. See Fig. 4 for an example.

---

[1] In [16], 8-connectivity, to which the 4-connectivity leads, is assumed to show the coincidence of the different concepts of digital convexity.
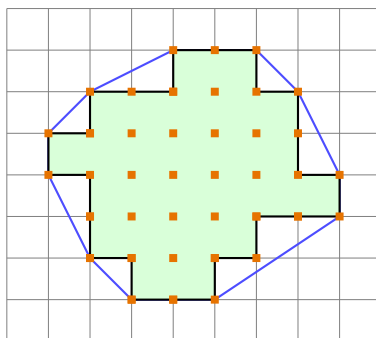
Fig. 4. A digital convex set C is represented by the orange points and its boundary path $Bd(C)$ is represented by the ordered orange points linked by black line segments. The boundary word $W(C)$ is then given by 10100100101101001010010110 while the blue polygonal line represents the convex hull $conv(C)$.
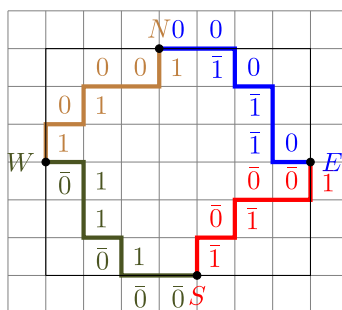

Fig. 5. The four parts $\mathcal{WN}$, $\mathcal{NE}$, $\mathcal{ES}$ and $\mathcal{SW}$ of the boundary of a digital convex set C are represented in four different colors. The word that codes the $\mathcal{WN}$-path is $w = 101001$.

It is then observed that the boundary word $W(C)$ of any digital convex set C can be factorized into four sub-words, such that each sub-word is a binary word, i.e., it contains only two letters. For such a factorization, we first enclose the boundary path $Bd(C)$ by its bounding box, and cut $Bd(C)$ into four parts at the four intersections with the bounding box; for example, on the left side of the bounding box, we consider the lowest intersection point W as a cutting position. Similarly, we can find the other three cutting positions, denoted by N, E, and S, as seen in Fig. 5. Starting from W in the clockwise direction and ending at N, we obtain the $\mathcal{WN}$-path as a part of $Bd(C)$, which is associated with the $\mathcal{WN}$-word of $W(C)$. Similarly, we obtain $\mathcal{NE}-$, $\mathcal{ES}-$, and $\mathcal{SW}$-paths and their associated words. Fig. 5 illustrates the factorization of boundary words and shows that each word is a binary word. This factorization allows us to treat the four parts of the boundary of any digital convex set independently, and to introduce the notion of digital convexity adapted to each part [8].

Definition 2 ([8])**.** A word w is said $\mathcal{WN}$-*convex* if it codes the $\mathcal{WN}$-word of the boundary word of some finite, 4-connected and digitally convex set of $\mathbb{Z}^2$.

Similarly, we can define $\mathcal{NE}$-, $\mathcal{ES}$-, $\mathcal{SW}-$convex words.

Our aim is to deform digital convex sets with preserving their digital convexity and our approach is based on combinatorics on words. In order to present the important characterization of digital convexity by combinatorics on words [8], we first recall the necessary notions of combinatorics on words in the following.

2.3. Basic notations of words

We first present some terminologies of words that can be found in [22]. An *alphabet* A is a nonempty finite set of symbols called letters; in this article, we have four letters $0, 1, \bar{0}, \bar{1}$ as mentioned above. A word w is a sequence of concatenated letters from A. The empty word ϵ is a sequence of zero symbols. A^* denotes the set of all finite words over A. We let $|w|$ represent the length of a word w, while $|w|_a$ represents the number of occurrences of a in w. For all $a \in A$, we thus have $|w| = \sum_{a \in A} |w|_a$. The n-times concatenation of w is denoted by w^n. The sub-word s of w from position k to position l is denoted by: $s = w[k : l]$.

A word is said *primitive* if it is not the power of a nonempty word. We say that w and w' are *conjugate*, denoted by $w \equiv w'$, if there exist two factors u, v in both w and w', such that $w = uv$ and $w' = vu$. The *reversal* of a word

Table 1
The alphabet, the lexicographic order and the slope's calculation used in each of the 4 parts of the boundary word.

Words	Alphabet	Order	Slopes
$\mathcal{WN}$	$\{0, 1\}$	$0 < 1$	$\frac{\|w\|_1}{\|w\|_0}$
$\mathcal{NE}$	$\{0, \bar{1}\}$	$\bar{1} < 0$	$\frac{\|w\|_0}{\|w\|_{\bar{1}}}$
$\mathcal{ES}$	$\{\bar{0}, \bar{1}\}$	$\bar{0} < \bar{1}$	$\frac{\|w\|_{\bar{1}}}{\|w\|_{\bar{0}}}$
$\mathcal{SW}$	$\{\bar{0}, 1\}$	$1 < \bar{0}$	$\frac{\|w\|_{\bar{0}}}{\|w\|_1}$

$w = a_1 a_2 \ldots a_n$ is $\tilde{w} = a_n \ldots a_2 a_1$ where each a_i is a letter. If $\tilde{w} = w$, the word is called a *palindrome*. In this paper, we use the total *lexicographic order*, denoted by $<$.

We recall that the boundary word over $\{0, 1, \bar{0}, \bar{1}\}$ can be divided into 4 parts and each of them belongs to a different binary alphabet. For each of these parts, we give the lexicographic order between the two letters in such a way we preserve the decreasing order of the Lyndon factorization (see below for the definition). In this case, the slope of each factor will be calculated depending on which part it belongs to. Table 1 shows this information with respect to each part of the boundary word.

In the following, we introduce the two families of words needed in this article, Lyndon and Christoffel words.

2.4. Lyndon words

We introduce the lexicographic order over the words in order to talk about the Lyndon family that was introduced by R. C. Lyndon [23].

Definition 3 ([22])**.** Let u and v be two words. We say $u > v$ if and only if:

$$\exists\, r \in A^* \; u = vr \quad or \quad \exists\, m, l, e \in A^* \; u = m1e \text{ and } v = m0l$$

Definition 4 ([23])**.** A word w is *Lyndon* if it is the smallest among its conjugates using its lexicographic order.

We then have the following unique factorization of any word, which is called *Lyndon factorization*, introduced by Lyndon, Chen and Fox in 1958.

Theorem 1 ([28])**.** *Every non-empty word w admits a unique factorization as a lexicographical decreasing sequence of Lyndon words, $w = \ell_1^{n_1} \ell_2^{n_2} \cdots \ell_k^{n_k}$, where every $n_i \in \mathbb{N}$ and every ℓ_i is a primitive Lyndon word such that $\ell_1 > \ell_2 > \cdots > \ell_k$.*

A linear-time algorithm to compute the Lyndon factorization was proposed by Duval [15], while there also exists a $O(\log n)$-time parallel algorithm proposed by Apostolico and Crochemore [2], where n is the length of the word.

Definition 5. Let $C \subset \mathbb{Z}^2$ be a finite 4-connected, digitally convex set. The points on $Bd(\mathcal{C})$ that separate different Lyndon factors of $W(\mathcal{C})$ are called *Lyndon points* of $\mathcal{C}$.

Geometrically, Lyndon points correspond to the vertices of the convex hull of C. They are used in the next section and you can see them in Fig. 1 as the green points in C.

2.5. Christoffel words

Christoffel words are regarded as the discretization of line segments with rational slopes.

Geometrically, a Christoffel word is determined by encoding with Freeman chain code the discretization of a line segment of rational slope [4]. More precisely, for any two non-negative co-prime integers a and b, the discretization of a line segment of rational slope $\frac{a}{b}$ is the closest digital path below the line segment such that no integer point exists between the path and the line segment. This digital path is called the Christoffel path. If a and b are positive co-prime numbers, a Christoffel word w of slope $\frac{a}{b}$, denoted by $\mathscr{C}\left(\frac{a}{b}\right)$, is a sequence of $a + b$ letters chosen from the binary alphabet $\{0, 1\}$. The choice of letters is not random, but it is obtained by assigning the letter 0 (resp. 1), to each increasing (resp. decreasing) step in the sequence of all the multiples of a modulo $(a + b)$ as given below.

Definition 6 ([9,28])**.** Given a pair of non-negative co-prime integers a and b, the *Christoffel word* w of slope $\frac{a}{b}$ is the sequence of a + b letters of $\{0, 1\}^*$, such that the ith letter of w is given by:

$$\forall i \in \{1, \ldots, n\} \; w[i] := \begin{cases} 0 & \text{if } r_{i-1} < r_i, \\ 1 & \text{otherwise} \end{cases}$$

Table 2
The Christoffel word w of slope $\frac{5}{8}$ with the remainder sequence $(r_i)_{0 \leq i \leq 13}$.

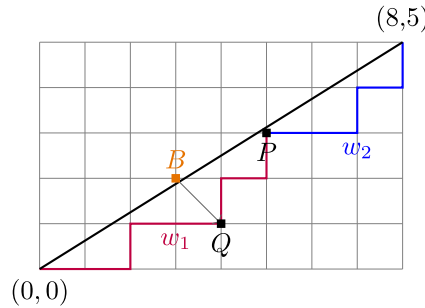| i	0	1	2	3	4	5	6	7	8	9	10	11	12	13
r_i	0	5	10	2	7	12	4	9	1	6	11	3	8	0
$w[i]$		0	0	1	0	0	1	0	1	0	0	1	0	1

Fig. 6. The standard factorization of the Christoffel word w of slope $\frac{5}{8}$. P corresponds to the closest point $\mathfrak{cl}(w)$ while Q corresponds to the furthest point $\mathfrak{fu}(w)$. P is also called the interior Bézout point, while B is the exterior Bézout point. We have: $\mathscr{C}\left(\frac{5}{8}\right) = \mathscr{C}\left(\frac{3}{5}\right)\mathscr{C}\left(\frac{2}{5}\right)$.

where the remainder r_i is defined by

$$r_i = ia \bmod (a + b)$$

for $0 \leq i \leq a + b$.

The remainder sequence r_i and the ith letter of $\mathscr{C}\left(\frac{5}{8}\right)$ are shown in Table 2.

It is also noticed that each Christoffel word starts with a horizontal step, *i.e.* 0, and ends with a vertical step, *i.e.* 1, while the central part is a palindrome.

Property 1 (*[3]*). *Let $\mathscr{C}\left(\frac{a}{b}\right)$ be the Christoffel word of slope $\frac{a}{b}$ with a and b two positive co-prime numbers, we can write $\mathscr{C}\left(\frac{a}{b}\right) = 0w'1$ where w' is a palindrome.*

As the slope of a Christoffel word is exactly the number of occurrences of the letter 1 over the number of occurrences of the letter 0, the slope of each Christoffel word is defined as follows.

Definition 7 (*[6]*). Given a non-empty primitive Christoffel word $w \in A^*$, the morphism $\rho(w)$ from A^* to $\mathbb{Q}^+ \cup \{\infty\}$ is defined by

$$\rho(w) := \begin{cases} \infty & \text{if } w = 1 \\ \frac{|w|_1}{|w|_0} & \text{otherwise.} \end{cases}$$

Property 2 (*[6]*). *For any two non-empty Christoffel words u and v, we have*

$$u > v \Leftrightarrow \rho(u) > \rho(v).$$

For this family of words, Borel and Laubie introduced the *standard factorization* [6,7], which allows writing any Christoffel word as the concatenation of two other Christoffel words in a unique way as follows.

Theorem 2 (*[6,7]*). *Any Christoffel word w of length greater than or equal to 1 can be written in a unique way such that $w = uv$ where u and v are both primitive Christoffel words. The couple (u, v) is called the standard factorization of w.*

Geometrically, this factorization can be seen as the decomposition of w into u and v at the *closest* point of the path corresponding to w with respect to the line segment between the origin to $(|w|_0, |w|_1)$. It is equivalent to say that this factorization corresponds exactly to the position where $r_i = 1$. This closest point of w is denoted by $\mathfrak{cl}(w)$. Fig. 6 illustrates the standard factorization of the Christoffel word in Table 2, that is (00100101, 00101).

A direct application of Theorem 2 and Property 1 induces the *palindromic factorization*, which allows to write any Christoffel word as a concatenation of two palindromes [6,28].

Property 3 (*[6,10,28]*). *Any primitive Christoffel word w of length strictly greater than 1 can be written in a unique way as $w = p_1p_2$ where p_1 and p_2 are palindromes.*

This property comes from the fact that the central part of each factor of the couple (u, v), which forms the standard factorization of w, is a palindrome (*i.e.* u can be written as $0u_1 1$ with u_1 a palindrome). Namely, we have:

$$w = uv = 0u_1 10v_1 1 = 0p1 = 0v_1 01u_1 1 = p_1 p_2$$

where p, p_1 and p_2 are palindromes (Property 1 and Theorem 2).

Corollary 1 ([6]). *Let w be a Christoffel word of length greater than 1. If its palindromic factorization is given by $w = p_1 p_2$, then $\rho(p_1) < \rho(w) < \rho(p_2)$.*

Corollary 2 ([13]). *Let w be a Christoffel word of length greater than 1. Then the palindromic factorization decomposes w into p_1 and p_2 exactly at the furthest point of the Christoffel path of w with respect to the line segment from the origin to the grid point $(|w|_0, |w|_1)$, which is unique respect to w.*

This furthest point is denoted $\mathfrak{fu}(w)$. Note that the position of the grid point $\mathfrak{fu}(w)$ over w is $r_i = |w| - 1$ as seen in Fig. 6. The two types of points, the closest point $\mathfrak{cl}(w)$ and the furthest point $\mathfrak{fu}(w)$, on the Christoffel path are highlighted due to particular interest for the rest of the paper; $\mathfrak{fu}(w)$ is used in List 4 for inflation of digital convex sets. We call a *closest outer point* of a word, a point which is the diagonally opposite to $\mathfrak{fu}(w)$. In order to link Lyndon and Christoffel words, we need the notion of k-balanced words [24]; a word $w \in \{0, 1\}^*$ is k-balanced if and only if for every pair of sub-words s, t of w, we have:

$$|s| = |t| \implies \left| |s|_1 - |t|_1 \right| \le k.$$

It is shown that Christoffel words are 1-balanced [6,7], and the following theorem connects the two families of words.

Theorem 3 ([24]). *A word w is a Christoffel word if and only if it is a 1-balanced Lyndon word.*

2.6. Digital convexity interpreted by combinatorics on words

The authors in [8] gave a characterization for the boundary word $W(C)$ using the notions of combinatorics on words, in particular Lyndon and Christoffel words.

Theorem 4 ([8]). *A word of $\{0, 1\}^*$ is WN-convex if and only if its Lyndon factorization is unique, and its factors are all primitive Christoffel words.*

Thanks to Theorem 3, if each factor of Lyndon factorization is in addition 1-balanced, we can say that the factors are also Christoffel words. From Theorem 4, we can also conclude that the slopes of the Christoffel factors are in decreasing order.

Corollary 3. *If $w = w_1^{n_1} w_2^{n_2} \ldots w_k^{n_k}$ be the Lyndon factorization of the boundary word of a certain digital convex set C then $\rho(w_1) > \rho(w_2) > \cdots > \rho(w_k)$.*

The following example shows a $\mathcal{WN}$-path and the factorization of its boundary word.

Example 1. Let us consider a $\mathcal{WN}$-convex word w whose Lyndon factorization is:

$$w = (1)(011101111)(0111)(01)(001)(00010001001)(0001).$$

The slopes ρ of the factors are decreasing: $\frac{1}{0}, \frac{7}{2}, \frac{3}{1}, \frac{1}{1}, \frac{1}{2}, \frac{3}{8}, \frac{1}{3}$.

In [8], we can find a linear-time algorithm over the word length that checks the $\mathcal{WN}$-convexity of a path encoded by a binary alphabet. Then, Theorem 4 allows us to detect the exact position of the vertices of the convex hull of $Bd(C)$ over the boundary word w.

Property 4 ([8]). *Given a digital convex 4-connected set C, the vertices of the convex hull of C or its boundary $Bd(C)$ corresponds to the Lyndon points of $Bd(C)$.*

Fig. 7 shows the result of Theorem 4 and Property 4 for the digitally convex set with boundary word $W(C) = 1010010010110100101010010110$.

In the next sections, we show how to deflate and inflate a digital convex 4-connected set using the above tools of combinatorics on words. In such a context, we characterize removable points for the deflation process in Section 3 and similarly identify insertable points for inflation in Section 4.
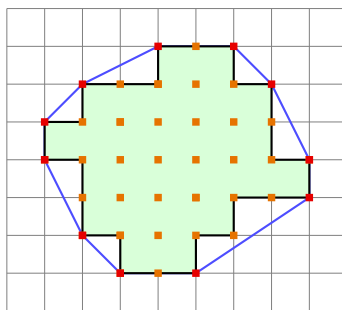
Fig. 7. Let us consider the digital convex set C of the orange and red points, whose boundary $Bd(C)$ is drawn as the integer points on the black rectilinear polygonal line. The Lyndon points of $Bd(C)$ are colored in red. Each Lyndon point corresponds to the end of each distinct factor of the Lyndon factorization of the boundary word $W(C) = (1)(01)(001)(0)^2(\bar{1}0)(\bar{1}\bar{1}0)(\bar{1})(\bar{0}\bar{0}\bar{1}0\bar{1})(\bar{0})^2(1\bar{0})(1\bar{1}0)$. We have $\rho(1) > \rho(01) > \rho(001) > \rho(0)$ in $\mathcal{WN}$-path. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3. Deflation of digital convex sets

In this section, we first define *removable points* that allow deflating a digital convex 4-connected set C while preserving its convexity. We then characterize such points in order to make a list of all removable points using concepts of digital topology and combinatorics on words. They are defined along the boundary path $Bd(C)$ via the boundary word $W(C)$ (see Section 2.2). Applying iteratively such point-wise deflation operation may be required in various applications of computer imagery. This implies that we need each step to choose a removable point among the list of removable points and then to update the list, which is made by the new Lyndon factorization of $W(C)$. We show that this update can be made locally. In practice, we need to choose one removable point in the list by using some priority. In this article, however, we do not discuss how to define such priority and simply assume that we have such a priority a priori [30]. We focus on the characterization of removable points and the update of the list of removable points, which are critical issues in the deflation algorithm.

3.1. Removable points

We first give a definition of removable points, which are used for deflating a 4-connected digital convex set C while preserving its convexity.

Definition 8. Given a digitally convex, 4-connected set C, a point $x \in C$ is removable if $C \setminus \{x\}$ is still 4-connected and digitally convex.

First of all, let us consider the preservation of connectivity. In digital topology, given a subset $X \subset \mathbb{Z}^2$, a point $x \in X$ is said to be *simple* if deleting it from X preserves the topological characteristics of X, i.e. the number of connected components of both X and its complement in $\mathbb{Z}^2$ [11,21]. As we consider X to be 4-connected and digitally convex here, preserving the topological characteristics of X is equivalent to preserving the connectivity of X; X can have no hole due to the digital convexity. Therefore, removable points must be simple if we would like to preserve the connectivity of C. Note that digital convexity does not imply connectivity (see Remark 1).

The definition of simple points relies on the notion of connected components, which is a global characteristic such that the whole object must be taken into account. However, it is well known that simple points can be characterized locally [11,26], for example, using the connectivity number defined in the 8-neighborhood [5].

Given a digital convex 4-connected set C, we recall that the boundary path $Bd(C)$ is decomposed into four parts $Bd_{\mathcal{WN}}(C)$, $Bd_{\mathcal{NE}}(C)$, $Bd_{\mathcal{ES}}(C)$ and $Bd_{\mathcal{SW}}(C)$ whose boundary words are respectively $W_{\mathcal{WN}}(C)$, $W_{\mathcal{NE}}(C)$, $W_{\mathcal{ES}}(C)$ and $W_{\mathcal{SW}}(C)$. Let us consider removing a point on $W_{\mathcal{WN}}(C)$, except its extremities, with preserving the convexity. Then we observe that if boundary points are removable, then the points and their neighboring points back and forth in the path $Bd_{\mathcal{WN}}(C)$ may form the sub-word 10 in $W_{\mathcal{WN}}(C)$. Removing a point from C means switching a factor of the form 10 into 01 in $W_{\mathcal{WN}}(C)$ (resp. $\bar{1}0$, $\bar{0}\bar{1}$, $1\bar{0}$ into $0\bar{1}$, $\bar{1}0$, $0\bar{1}$ in $W_{\mathcal{NE}}(C)$, $W_{\mathcal{ES}}(C)$, $W_{\mathcal{SW}}(C)$) as seen in Fig. 8. For the extremity points, we usually use the same concept. A particular study will be considered further in this paragraph. This *switch* operator can be defined over the alphabet $\{0, 1, \bar{0}, \bar{1}\}$ so that two consecutive letters are exchanged at a position k as follows.

Note that any point all over $Bd(C)$, whose neighboring points back and forth in $Bd(C)$ form a sub-word of the same letter can be removed. But, in this case, we lose the convexity as we can see in Fig. 9.

Definition 9 ([30])**.** Given a word $w = a_1 a_2 \ldots a_n$, the *switch operator* at position k, $k < n$, on w is defined by:

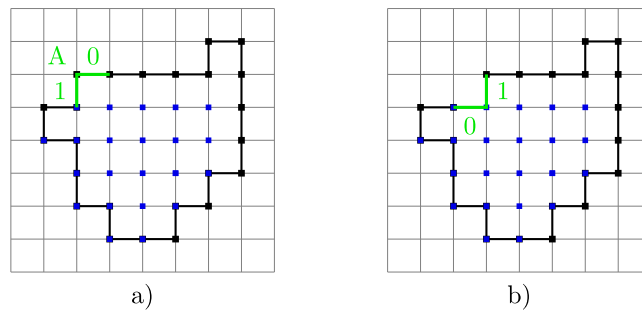$$switch_k(w) := a_1 a_2 \ldots a_{k+1} a_k \ldots a_n.$$

Fig. 8. (a) A is a point in the $\mathcal{WN}$-path, where its neighboring points back and forth in $Bd_{\mathcal{WN}}(C)$ form the sub-word 10. (b) Removing the point A corresponds to switching the sub-word 10 into 01, i.e. applying the switch operator at the position of the point A.
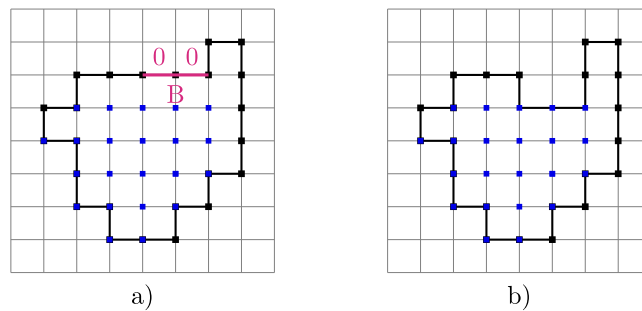

Fig. 9. (a) B is a point in the $\mathcal{WN}$-path, where its neighboring points back and forth in $Bd_{\mathcal{WN}}(C)$ form the sub-word 00. (b) The digital convexity is lost when removing the point B.
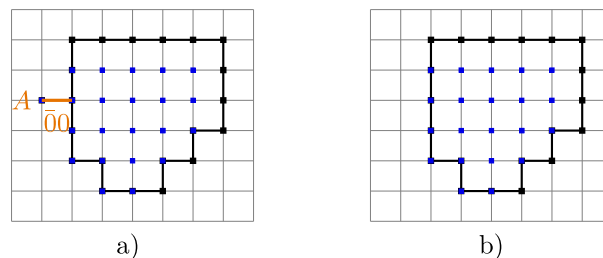

Fig. 10. (a) The digital set C whose boundary word is: $W(C) = 0110000\bar{1}\bar{1}\bar{1}0\bar{1}0\bar{1}0\bar{0}\bar{1}0110\bar{0}$. The point $A \in C$ creates the sub-word $\bar{0}0$ in $W(C)$. (b) Removing A means replacing this factor by ϵ.

We should mention that sometimes the same grid points appear twice in the boundary path $Bd(C)$ so that the boundary word $W(C)$ contains two consecutive letters of opposite directions. They are always positioned at the junction of two of four decomposed paths, $Bd_{\mathcal{WN}}(C)$, $Bd_{\mathcal{NE}}(C)$, $Bd_{\mathcal{ES}}(C)$ and $Bd_{\mathcal{SW}}(C)$, such that the two consecutive letters are in different paths. We can also remove such a grid point, so that we replace the sub-word consisting of two letters of opposite directions by ϵ, as seen in Fig. 10.

Over $W(C)$, we can find several points that form with their neighboring points either sub-word 10, 0$\bar{1}$, $\bar{1}0$, $\bar{0}1$, $\bar{0}0$, 1$\bar{1}$, 0$\bar{0}$ or $\bar{1}1$. According to Definition 5, they correspond to Lyndon points; we recall that Lyndon points are geometrically the vertices of $Conv(C)$, each of which is at the end of each factor of the Lyndon factorization of $W(C)$ according to Property 4. However, not all of Lyndon points are removable. Indeed, the switch operation on such factors may lead to losing the connectivity. To avoid this problem, Theorem 5 gives the characterization of removable points for any C. In order to prove this theorem, we recall that the Christoffel words, which are the one-balanced Lyndon words of the Lyndon factorization of $W(C)$ (Theorem 4), have the following form: $0u1$, $\bar{1}k0$, $\bar{0}\ell\bar{1}$ and $1m\bar{0}$, where u, k, ℓ and m are palindromes in $W_{\mathcal{WN}}(C)$, $W_{\mathcal{NE}}(C)$, $W_{\mathcal{ES}}(C)$ and $W_{\mathcal{SW}}(C)$ (Property 1).

Theorem 5. *Given a digital convex 4-connected set C of $\mathbb{Z}^2$, a point $x \in C$ is removable if and only if x is a simple point with respect to C and a Lyndon point of the boundary $Bd(C)$.*

Proof. Let us consider the boundary word $W(C)$ of C, which is decomposed by the Lyndon factorization such that $W(C) = \ell_1^{n_1} \ell_2^{n_2} \ldots \ell_s^{n_s}$. Since C is digitally convex, each ℓ_i, $1 \leq i \leq s$, is a Christoffel word (Theorems 3 and 4). We give the proof only for the binary sub-word $W_{\mathcal{WN}}(C)$ of $W(C)$ as follows; the similar proofs are found to the three other sub-words $W_{\mathcal{NE}}(C)$, $W_{\mathcal{ES}}(C)$ and $W_{\mathcal{SW}}(C)$. As mentioned above, removing a point from the boundary means applying the switch operator at a Lyndon point, so that the corresponding sub-word 10 is replaced by 01, or $\bar{0}0$, $1\bar{1}$ by ϵ. The simplicity of a point x guarantees the 4-connectivity of $C \setminus \{x\}$. If a point is not a Lyndon point, i.e. the boundary sub-word belongs to one of the Lyndon factors ℓ_i of $W_{\mathcal{WN}}(C)$, it cannot be removed as we lose the $\mathcal{WN}$-convexity. Let us consider any boundary sub-word 10 such that 1 appears at the end of one of the Lyndon factors in $W_{\mathcal{WN}}(C)$, and 0 appears at the beginning of the consecutive Lyndon factor. If these two Lyndon factors are identical, then switching the pair 10 makes us lose the $\mathcal{WN}$-convexity. Now, let us focus on sub-words 10 that are obtained by two consecutive distinct Lyndon factors $\ell_i \ell_{i+1}$ of $W_{\mathcal{WN}}(C)$. By Property 1, we can write $\ell_i = 0u1$ and $\ell_{i+1} = 0v1$ where u and v are both palindromes, so that $\ell_i \ell_{i+1} = 0u10v1$. By applying the switch operator on $\ell_i \ell_{i+1}$ at position $|\ell_i|$, we obtain: $switch_k(\ell_i \ell_{i+1}) = 0u01v1$, which is also a Christoffel word if $u01v$ is a palindrome, as seen in Property 3; otherwise, we get several other Christoffel words. This effect will be studied and detailed in the next subsection where we give the update's effect on the sub-word after removing a point (see also the proof of Theorem 1 in [30]). Finally, when the corresponding sub-word of the point to remove is $\bar{0}0$ (resp. $1\bar{1}$), this means that we are at the intersection between the $\mathcal{SW}$ and $\mathcal{WN}$ paths (resp. $\mathcal{WN}$ and $\mathcal{NE}$ paths). Replacing this sub-word with ϵ means that the position of the point W will change and the $Bd_{\mathcal{WN}}(C)$ will be longer and starts from the first letter after the last $\bar{0}$ in $Bd_{\mathcal{SW}}(C)$. $\square$

Property 4 provides the geometrical interpretation of Theorem 5: any vertex of the convex hull of C corresponds to a Lyndon point on $Bd(C)$. We note that detecting all the removable points for C is performed in linear complexity with respect to $|W(C)|$ as it is done via Lyndon factorization of $W(C)$.

3.2. Updating the removable points

Given a digital convex 4-connected set C of $\mathbb{Z}^2$, the first iteration of our deflation process of C is started by making the list of all the possible removable points with respect to C. Once a candidate has been chosen among them, the boundary word $W(C)$ is modified and the Lyndon factorization is made again. This means that new candidates can arise or disappear from the list of removable points.

Example 2. Let $w_1 = C(\frac{2}{5})C(\frac{1}{3})$, $w_2 = C(\frac{3}{7})C(\frac{1}{3})$, $w_3 = C(\frac{3}{5})C(\frac{1}{3})$ and $w_4 = C(\frac{14}{19})C(\frac{4}{17})$ be three different factors of the boundary word $W(C)$ of a digital convex 4-connected set C. As proved before, the removable points with respect to C are simple and Lyndon points. From the definition, the Lyndon points of each w_i are positioned at the joint between consecutive distinct Christoffel words. If we apply the switch operator on each w_i at the Lyndon point, we obtain: $switch_7(w_1) = C(\frac{3}{8})$; $switch_{10}(w_2) = (C(\frac{2}{5}))^2$; $switch_8(w_3) = C(\frac{2}{3})(C(\frac{1}{2}))^2$ and $switch_{33}(w_4) = C(\frac{11}{15})C(\frac{2}{3})C(\frac{1}{2})(C(\frac{1}{4}))^4$. These examples show that Lyndon points can disappear or newly appear due to the switch operation (see Fig. 11).

The above deflation step is iterated in general, and we remark that several removable point candidates exist at each iteration step. Various heuristics can be considered to choose one of the removable points [30]. It is also important to know what happens after each iteration because of the update of the Lyndon factorization. The Property 5 shows all the possible effects that arise after removing a point. We can also see that the update of removable points are made locally, which means that the Lyndon factorization before the factor u and after the consecutive factor v is unchanged if the Lyndon point is removed at the joint of u and v.

Property 5 ([30]). *Let the binary word $w = uv$ such that u and v are two Christoffel words with $\rho(u) > \rho(v)$. By applying the switch operator to w at position $|u|$, we obtain $\mathrm{switch}_{|u|}(w) = m_1^{\ell_1} \ldots m_k^{\ell_k}$, $k \geq 1$ such that $\rho(u) > \rho(m_1) > \cdots > \rho(m_k)$.*

3.3. Algorithm

We give a small recap about all the previous results considered in combinatorics on words point of view. In order to deflate a digital convex 4-connected set C, we must start the procedure by considering the boundary word w of C. By Property 4, it is sufficient to apply the Lyndon factorization on w in order to get the positions of all the Lyndon points. At each step of the deflation procedure, several points are considered good candidates. In order to know which point is removable, we need to be sure that this point is also a simple point thanks to Theorem 5. We then choose a certain heuristic in order to stock these candidates in a priority queue. Once a point is selected from the queue, we apply the switch operator at its position, given in Definition 9. Some local updates must be considered over the Lyndon factorization. This will certainly affect the priority queue of removable points, which requires an update too since the Lyndon points may vary. This procedure can be repeated for a fixed number of times. Knowing that we are removing a single point at each step, the number of iterations is limited by the cardinality of the digital set. To remove k points from C while preserving digital convexity, the procedure must be iterated k times. Algorithm 1 represents such an iterative pointwise deflation over a 4-connected digital convex set C.

Fig. 12 shows the deflation procedure applied on a digital convex 4-connected set using the heuristic of area-change, which minimize or maximize the area difference between the convex hull of C and $C \setminus \{p\}$ (see [30] for more details).
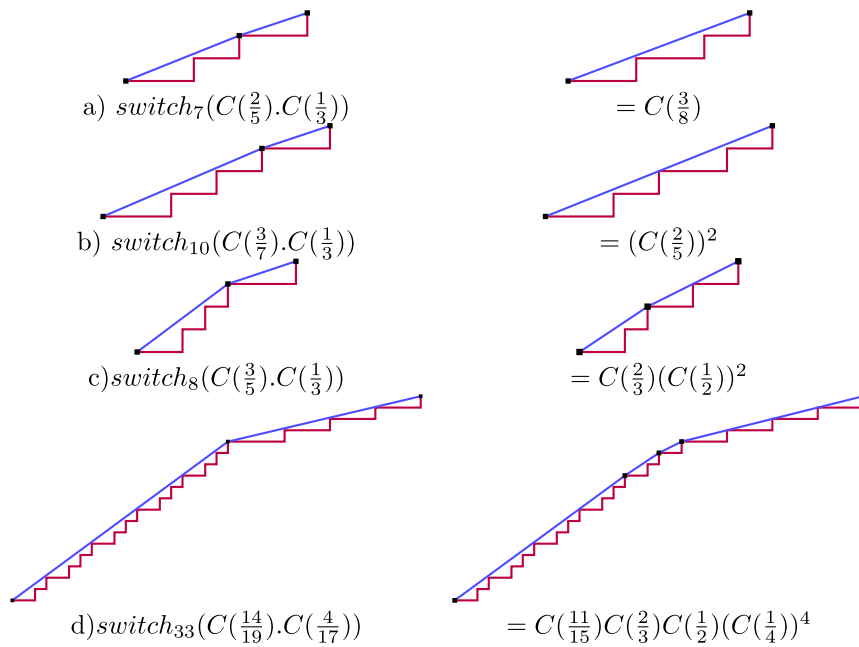
Fig. 11. Before and after the pixel-wise deflation of (a) w_1, (b) w_2, (c) w_3 and (d) w_4 and their Lyndon factorization (left and right); (a) and (b) no removable point remains after the update; (c) a removable point arises at a different position; (d) several new removable points appear.
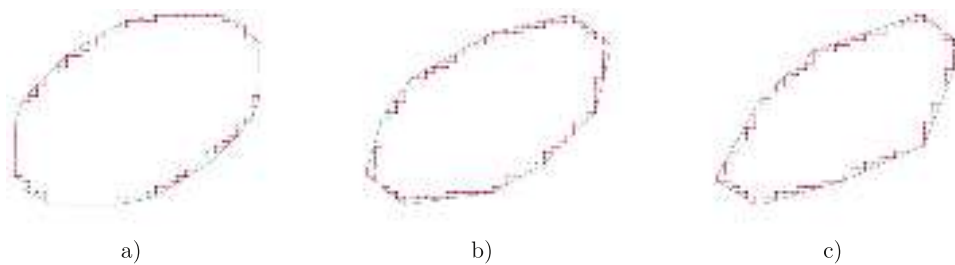

Fig. 12. The deflation process of a digital convex 4-connected set (a) represented after 150 and 250 iterations respectively in (b) and (c) using the heuristic of area-change.

Algorithm 1 Point-wise deflation

Input: a digital convex 4-connected set C, a number of removing points k
Output: a sequence of removed points R

1: Compute the Lyndon Factorization F of the boundary word of C
2: Insert all the Lyndon points in a priority queue L
3: **while** $k > 0$ **do**
4:     Pull the highest-priority simple point p from L and add p to R
5:     Let u and v be the two distinct Christoffel words of F around p
6:     Compute $w = switch_{|u|}(uv)$
7:     Compute the Lyndon Factorization of w and update F and L
8:     Decrement k
9: **end while**
10: **return** R

---

4. Inflation of digital convex sets

In this section, we study the pixel-wise inflation of any digital convex 4-connected set C with the preservation of digital convexity, which is more complex than the deflation process as seen in the previous section. We first give a necessary condition and then a sufficient condition for characterizing insertable points for such inflation. We will show that such

insertable points cannot be at any arbitrary place around C. Here the key process, if such inflation is applied iteratively, is also the update of Lyndon factorization over the boundary word $W(C)$. The question raised is whether the induced factorization can be made locally and efficiently.

Here, we focus on the closest outer points, each of which is uniquely defined for every Lyndon factor of the boundary word $W(C)$, with length greater than 1 (see Corollary 2). Lyndon factors of length 1 have no closest outer point, and this case will be treated differently later on in this section. In order to inflate C, we need to determine the *insertable points*, which are necessarily the closest outer points indeed. In order to verify the insertability of a closest outer point, we have to take into consideration the local and global effects on $W(C)$ after adding this point, which requires the Lyndon factorization update of $W(C)$. In fact, we might lose the digital convexity after adding a furthest point if we do not choose the right one. We characterize insertable points with some iterative local digital convexity verification. In order to avoid such an expensive iterative verification, we also propose some strong constraints to locally characterize a subset of insertable points.

4.1. Definition of insertable points

We define insertable points, which are used for pixel-wise inflation of a digital convex 4-connected set C without losing its digital convexity.

Definition 10. Given a digital convex 4-connected set C, a point $x \in \overline{C}$ is *insertable* if $C \cup \{x\}$ is still 4-connected and digitally convex.

In this section and unlike Section 3, there is no need to verify if an insertable point refers to a simple point or not. We only need to verify the digital convexity as the simple 4-connectedness of $C \cup \{x\}$ is kept. We recall that the Lyndon factorization of the boundary word $W(C)$ is of the following form: $W(C) = \ell_1^{n_1} \ldots \ell_s^{n_s}$ where all ℓ_i are made of primitive Christoffel words of length greater or equal to 1, with the decreasing slope order, and every n_i is a positive integer. If $n_i > 1$, this signifies that a Christoffel word ℓ_i is repeated n_i times in $W(C)$. To find insertable points, we need to consider the following two cases for each Lyndon factor ℓ_i of $W(C)$: $|\ell_i| > 1$ and $|\ell_i| = 1$.

4.2. The link between insertable points and furthest points

In this part, we show the link between insertable points and furthest points associated with primitive Christoffel words. Let us call the diagonal opposite point in $\overline{C}$ of each furthest point in C *the closest outer point*.[2] We show that for each primitive Christoffel word of length strictly greater than 1, we can find a unique position in the boundary word $W(C)$ where a point is added at its closest outer point. We first present the definition of the *split operator* applied on primitive Christoffel words, proposed in [13]. This operator helps us to see the local modification over the boundary word $W(C)$, which can influence the Lyndon factorization of $W(C)$ globally sometimes.

4.2.1. Necessary condition for insertable points

Adding a point to a digital convex 4-connected set C at neither its vertical nor horizontal parts of the bounding box of C correlates, in a viewpoint of combinatorics on words, to applying the switch operator, given in Definition 9, to a sub-word of form 01 (resp. $\bar{1}0$, $0\bar{1}$, $1\bar{0}$) over the boundary word $W_{\mathcal{WN}}(C)$ (resp. $W_{\mathcal{NE}}(C)$, $W_{\mathcal{ES}}(C)$, $W_{\mathcal{SW}}(C)$); on the vertical and horizontal sides, the letters $0\bar{1}0$ (resp. $10\bar{1}$, $0\bar{1}\bar{0}$, $\bar{1}0\bar{1}$) replace 1 (resp. 0, $\bar{1}$, $\bar{0}$) instead. We aim at inflating C without losing its digital convexity. This means that after adding a point, the updated Lyndon factorization of $W(C)$ must remain made of factors of Christoffel words whose slopes are in decreasing order, as seen in Corollary 3. In general, this condition is not satisfied if we add a point randomly. We will show that any insertable point must be a closest outer point, which is the diagonally opposite point of a furthest point. Note that the reverse is not always true.

Definition 11 ([13]). The *split operator* applied on a primitive Christoffel word w decomposes w into two Christoffel words w^+ and w^-, defined as

$$split(w) := (w^+, w^-)$$

such that:

- when $|w| \geq 2$,

    $$(w^+, w^-) := (w'[1 : k], w'[k + 1 : |w|])$$

    where $w' = switch_k(w)$; with k the position of the furthest point of w,

---

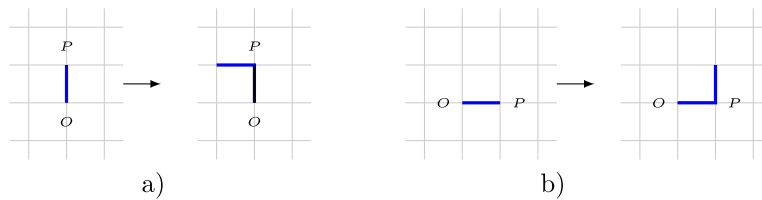2 This point corresponds exactly to the exterior Bézout point of a line segment in $\mathbb{Z}^2$.

Fig. 13. Split operator applied to Christoffel paths of slopes (a) ∞ so that $split(1) = (1\bar{0}, 0)$, and (b) 0 such that $split(0) = (01, \bar{1})$.
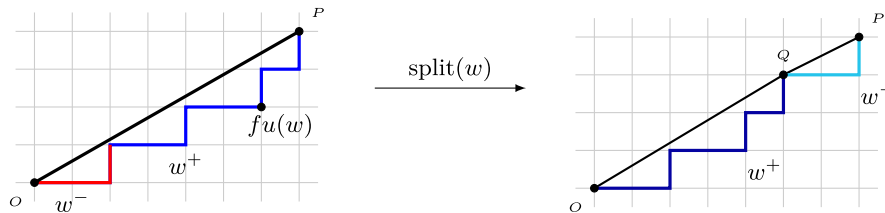

Fig. 14. Christoffel word (left) whose standard factorization is $w = w^- w^+$, and the result of the split operator (right), $split(w) = w^+ w^- = switch_8(w)$.

- when $|w| = 1$,

$$
(w^+, w^-) := \begin{cases}
(1\bar{0}, 0) & \text{if } w = 1, \\
(01, \bar{1}) & \text{if } w = 0, \\
(\bar{1}0, \bar{0}) & \text{if } w = \bar{1}, \\
(\bar{0}\bar{1}, 1) & \text{if } w = \bar{0}.
\end{cases}
$$

In Fig. 13, we show examples of the case that the length of a Christoffel word is equal to 1. Note that, if the Christoffel word w is not primitive, i.e. $w = \ell_i^{n_i}$ with $n_i > 1$, we can apply the split operator on any of these ℓ_i.

Corollary 2 shows that the split operator applied at the furthest point position of a Christoffel word of length greater than 1 gives two other Christoffel words. These two words are of decreasing order with respect to their slopes. Lemma 1 also shows that the concatenation of these two new factors gives the same result for the switch operator, defined previously when applied at the same position. This can be seen in Example 3 and illustrated in Fig. 14.

Lemma 1 (*[13]*). *Given a Christoffel word w with $|w| > 1$ such that $w = uv$ and u and v are the factors of the standard factorization of w. Then we have*

$$split(w) = (v, u)$$

where $\rho(v) > \rho(u)$.

From Theorem 2, we know that u and v, the two factors of the standard factorization of w, are both primitive Christoffel words. Lemma 1 then shows that the result of the split operator exactly consists of the two primitive Christoffel words v and u, which are in the reverse order of the standard factorization.

Example 3 (*Example of The Split Operator*). Let $w = 00100100101$ be the Christoffel word of slope $\frac{4}{7}$ with its standard factorization $w = w^- w^+$ where $w^- = \mathscr{C}\left(\frac{1}{2}\right)$, $w^+ = \mathscr{C}\left(\frac{3}{5}\right)$ and $\rho(w^-) < \rho(w^+)$. By applying the split operator on w, we obtain

$$split(w) = (w^+, w^-) = (00100101, 001).$$

Property 3 ensures the uniqueness of the furthest point for each Christoffel word. Based on that, it is shown in [13,14] that any Christoffel word of length greater than 1 can only be split at this position.

Proposition 1 (*[14]*). *Let w be a Christoffel word of length n and k the position of $\mathfrak{fu}(w)$.*

1. *The words $u = w[1, k - 1]1$ and $v = 0w[k + 2, n]$, are two Christoffel words, where the notation $w[i, j]$ indicates the subword of w from position i to j, with $1 \leq i \leq j \leq n$.*
2. *For each non-negative integer k' different from k, the words $u' = w[1, k' - 1]1$ and $v' = 0w[k' + 2, n]$ are not both Christoffel words.*

Proof. Let w be a Christoffel word of length n and k is the position of $\mathfrak{fu}(w)$. From Property 1, we know that $w = 0p1$ where p is a palindrome. From Corollary 2, we can write w in a unique way as $w = p_1 p_2$, where p_1 and p_2 are two
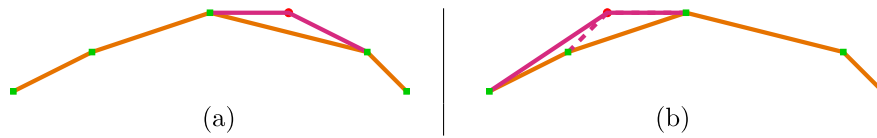
Fig. 15. Figure (*a*) shows that the inflation at this position maintains the convexity. Figure (*b*) shows that there is an additional step of propagation to verify the convexity on the left side of the segment.

palindromes with $|p_1| = k$. This gives $w = 0s01t1$ where the length of s is $k - 2$. Therefore, $u = w[1, k - 1]1 = 0s1$ and $v = 0w[k + 2, n] = 0t1$ are two Christoffel words. The unicity of the position of $\mathfrak{f}u(w)$ ensures the unicity of u and v. □

A useful consequence follows.

Corollary 4. *Let w, u and v be as defined in* Proposition 1. *It holds $\rho(u) > \rho(v)$.*

In other words, if we need to add a point from $\overline{C}$ to C in order to inflate a certain digital line segment of the boundary of C, we must choose a closest outer point associated to the digital line segment.

So far, we have shown how to split one Christoffel word with or without multiplicity; Lemma 1 ensures that the order of the slopes for the new Christoffel words, obtained after splitting, is still decreasing. The next question is the following: will this order be preserved also around the boundary word? In other words, does this operation affect the order of the slopes around a chosen Christoffel word to be split? These questions are answered in the following part. They help us to give the characterization of insertable points.

4.2.2. Example of closest outer points that are not insertable

Let us consider a digital convex 4-connected set C and its boundary word whose Lyndon factorization is given by $W(C) = \ell_1^{n_1} \dots \ell_s^{n_s}$. Each factor $\ell_i^{n_i}$ represents one of the polygonal line segments of $\mathcal{C}onv(C)$ and each ℓ_i is a primitive Christoffel word. As we have seen, in order to add a point around C, we choose one of the factors of $W(C)$ together with its closest outer point. This must be done by respecting the conditions given in Definition 11 and Lemma 1. Two examples are illustrated in Fig. 15.

During the inflation process, we can also get the case where we can completely lose the convexity property. In this case, we know that this closest outer point, cannot be chosen as an insertable point. This is illustrated in Example 4.

Example 4 (*Example of a closest outer point that is not insertable*). Let $w_1 = \mathscr{C}\left(\frac{30}{41}\right)$ and $w_2 = \mathscr{C}\left(\frac{5}{7}\right)$ be two consecutive Christoffel words on the boundary word of a certain digital convex 4-connected set. From Theorem 4, we have $\rho(w_1) > \rho(w_2)$. If we apply the split operator on w_2, *i.e.* add the closest outer point of w_2, we lose the digital convexity while it is not the case if we apply it on w_1:

$$split(w_1)w_2 = \left(\mathscr{C}\left(\frac{11}{15}\right), \mathscr{C}\left(\frac{19}{26}\right)\right)\mathscr{C}\left(\frac{5}{7}\right)$$

$$w_1 split(w_2) = \mathscr{C}\left(\frac{30}{41}\right)\left(\mathscr{C}\left(\frac{3}{4}\right), \mathscr{C}\left(\frac{2}{3}\right)\right).$$

This example indicates that a closest outer point is not always insertable. In the following, we study sufficient conditions for insertable points in detail.

4.2.3. Characterization of insertable points and Lyndon factorization update

As seen before, not all closest outer points correspond to insertable points. In order to verify such insertability, we need to update, after adding a closest outer point, the Lyndon factorization over the boundary word $W(C)$. As mentioned before, we might need to do some propagation for certain cases. This propagation can be made on the right and/or left side of the Christoffel word where we split, and can reach the beginning or the end of the sub-word $W_{\mathcal{WN}}(C)$, $W_{\mathcal{NE}}(C)$, $W_{\mathcal{ES}}(C)$ or $W_{\mathcal{SW}}(C)$, in the worst case. We give the characterization of an insertable point in Theorem 6. The proof will be given at the end of this section since we need to show some notions before. Definition 12 introduces two types of insertability verification for grid points, one on the left side and the other one on the right side. A point is insertable on the left (resp. right) if we are able to concatenate a finite number of the previous (resp. next) consecutive Christoffel words. This concatenation is called the *propagation* and is possible under some conditions as shown in Definition 12.

Definition 12. Given a digital convex 4-connected set C, let us consider the boundary word $W(C)$ and its Lyndon factorization $W(C) = \ell_1^{n_1} \dots \ell_m^{n_m}$. Let x be the closest upper point in $\overline{C}$ of the jth Lyndon factor ℓ_i for $i \in [1, m]$ and $j \in [1, n_i]$ in $W(C)$ such that $split(\ell_i) = (\ell_i^+, \ell_i^-)$, we say that:

- x is insertable on the left if there exists some non-negative integer k ($k \in \mathbb{Z}^+$), such that $\ell_{i-k-1} \geq L_k$, where L_h for $h \leq k$ is recursively defined by

$$L_h := \begin{cases} \ell_i^{j-1}\ell_i^+ & \text{for } h = 0 \\ \ell_{i-h}^{n_{i-h}} L_{h-1} & \text{for } h \geq 1 \text{ if } \exists\, m_{h-1} \in \mathbb{Z}^+,\, \ell_{i-h} = \ell_{i-h-1} L_{h-1}^{m_{h-1}} \end{cases}$$

- x is insertable on the right if there exists some non-negative integer k such that $\ell_{i+k+1} \leq R_k$, where R_h for $h \leq k$ is recursively defined by

$$R_h := \begin{cases} \ell_i^- \ell_i^{n_i-j} & \text{for } h = 0 \\ R_{h-1} \ell_{i+h}^{n_{i+h}} & \text{for } h \geq 1 \text{ if } \exists\, m_{h-1} \in \mathbb{Z}^+,\, \ell_{i+h} = R_{h-1}^{m_{h-1}} \ell_{i+h-1} \end{cases}$$

Example 5. Let us consider the three consecutive Christoffel words: $w_1 = C(\frac{3}{2})$, $w_2 = C(\frac{4}{3})$ and $w_3 = C(\frac{14}{11})$, where $w_1 > w_2 > w_3$. By splitting w_1, we get: $split(w_1) = (C(\frac{2}{1}), C(\frac{1}{1}))$. The point added is insertable on the right since we can do a finite propagation:

- $R_0 = C(\frac{1}{1})$, but $w_2 > R_0$. So we construct $R_1 = R_0 w_2 = C(\frac{5}{4})$, since $w_2 = R_0 C(\frac{4}{3})$.
- We now have, $w_3 > R_1$, but $w_3 = R_1^2 w_2$, then we can construct $R_2 = R_1 w_3 = C(\frac{19}{15})$.

After this finite propagation, we end up with $C(\frac{2}{1})R_2$. For more details check Fig. 16.

Theorem 6 is one of the main results of this paper. It provides the characterization of the insertability of such a closest upper point $x \in \overline{C}$.

Theorem 6. *Given a digital convex 4-connected set C, let x be the closest upper point in $\overline{C}$ with respect to one of the Lyndon factors of the boundary word $W(C)$. Then, x is insertable if and only if x is insertable on both the left and right sides.*

Theorem 6 characterizes insertable points. Using this characterization, we can obtain the positions of all the insertable points to inflate a digital convex set by preserving its digital convexity. Once one of these eligible candidates is chosen and added, we update $W(C)$, which also updates the list of insertable points for the next step of inflation. Theorem 6 and Definition 12 show all the possible cases we can face after adding an insertable point. In order to prove this theorem, we first need to define the following morphism that maps the set of Christoffel words to the set of binary words.

Definition 13 ([7]). Given an ordered pair of Christoffel words over $\{0,1\}^*$ $B = (\mathscr{C}(\frac{a}{b}), \mathscr{C}(\frac{c}{d}))$, we define the *Christoffel morphism* Θ_B from the set of Christoffel words to A^* such that $\Theta_B(0) = \mathscr{C}(\frac{a}{b})$ and $\Theta_B(1) = \mathscr{C}(\frac{c}{d})$.

Being a morphism, $\Theta_B(uv) = \Theta_B(u)\Theta_B(v)$, and it is ordered as we can see in the following property.

Property 6 ([7]). *If $\mathscr{C}(\frac{a}{b}) < \mathscr{C}(\frac{c}{d})$, then the Christoffel morphism Θ_B with $B = (\mathscr{C}(\frac{a}{b}), \mathscr{C}(\frac{c}{d}))$ is an increasing morphism. In other words, for any two Christoffel words w_1 and w_2 such that $w_1 < w_2$, we have: $\Theta_B(w_1) < \Theta_B(w_2)$.*

Example 6. Let $B = (\mathscr{C}(\frac{3}{5}), \mathscr{C}(\frac{2}{3}))$, $w_1 = \mathscr{C}(\frac{3}{4})$ and $w_2 = \mathscr{C}(\frac{3}{2})$. From the lexicographic order, we have that $\mathscr{C}(\frac{3}{5}) < \mathscr{C}(\frac{2}{3})$ and $w_1 < w_2$. By applying Definition 13 and Property 6, we get the following words:

$$\Theta_B(w_1) = \mathscr{C}\left(\frac{3}{5}\right)^2 \mathscr{C}\left(\frac{2}{3}\right) \mathscr{C}\left(\frac{3}{5}\right) \mathscr{C}\left(\frac{2}{3}\right) \mathscr{C}\left(\frac{3}{5}\right) \mathscr{C}\left(\frac{2}{3}\right)$$

$$\Theta_B(w_2) = \mathscr{C}\left(\frac{3}{5}\right) \mathscr{C}\left(\frac{2}{3}\right) \mathscr{C}\left(\frac{3}{5}\right) \mathscr{C}\left(\frac{2}{3}\right)^2$$

so that $\Theta_B(w_1) < \Theta_B(w_2)$.

Lemma 2 studies the effect of the split operator of a certain Christoffel word with multiplicity higher than 1.

Lemma 2. *Let us consider a $\mathcal{WN}$-convex word w and its Lyndon factorization $w = \ell_1^{n_1} \ldots \ell_s^{n_s}$. For any $j \in [1, s]$ such that $|\ell_j| > 1$, if we split the ith element of $\ell_j^{n_j}$ where $1 \leq i \leq n_j$, we obtain $split(\ell_j) = (\ell_j^+, \ell_j^-)$ and the factors $\ell_j^{i-1} \ell_j^+$ and $\ell_j^- \ell_j^{n_j-i}$ are two decreasing Christoffel words, i.e. $\ell_j^{i-1} \ell_j^+ > \ell_j^- \ell_j^{n_j-i}$.*

Proof. First, we prove that the two factors $\ell_j^{i-1} \ell_j^+$ and $\ell_j^- \ell_j^{n_j-i}$ are Christoffel words. This is obtained by applying the Christoffel morphism Θ_A over $A = (\ell_j^-, \ell_j^+)$ on the words $(01)^{i-1}1$ and $0(01)^{n_j-i}$ respectively. Second, we must prove that $\ell_j^{i-1} \ell_j^+ > \ell_j^- \ell_j^{n_j-i}$. This inequality comes from the fact that Θ_A is an increasing morphism, as seen in Property 6; since $\ell_j^- < \ell_j^+$ and $0(01)^{n_j-i} < (01)^{i-1}1$, we have $\Theta_A((01)^{i-1}1) < \Theta_A(0(01)^{n_j-i})$. $\square$
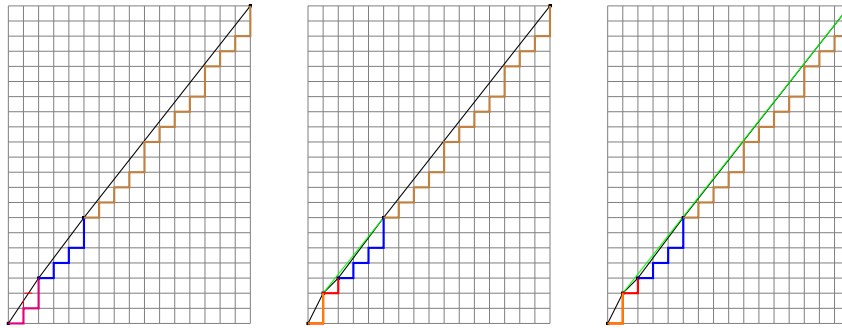
Fig. 16. Procedure of insertability verification on the right with propagation: the closest upper point $(1, 2)$ of ℓ_i (ℓ_i in pink) is inserted (left); as $\ell_{i+1} \leq R_0$ (ℓ_{i+1} in blue and $R_0 = \ell^-$ in red) is not satisfied but we have $\ell_{i+1} = R_0\ell_i$, we obtain $R_1 = R_0\ell_{i+1}$ (R_1 in green (=red + blue)) (center); as $\ell_{i+2} \leq R_1$ (ℓ_{i+2} in brown) is not satisfied but we have $\ell_{i+2} = R_1^2\ell_{i+1}$, we obtain $R_2 = R_1\ell_{i+2}$ (R_2 in green (=red + blue + brown)) (right). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

When the length of a Christoffel word is equal to one, the split operator gives two Christoffel words with different binary alphabets such as $split(1) = (1\bar{0}, 0)$. These two new Christoffel words cannot be compared with each other as the two Christoffel words do not belong to the same part of the four parts of the boundary path (ex. $\mathcal{WN}$-path). When they belong to different parts of the boundary word of a digital convex 4-connected set C, the right insertability of the associated point is verified in the initial binary boundary sub-word while the left insertability is verified in the previous one. In the case of $split(1) = (1\bar{0}, 0)$, the right insertability verification is made in $W_{\mathcal{WN}}(C)$ while the left one is made in $W_{\mathcal{SW}}(C)$.

To simplify the proof of Theorem 6, we treat only the right insertability as the left one can be proved similarly. We consider all the possible situations that can arise after applying the split operation.

Proof of Theorem 6. Let us consider the boundary word $W(C)$ and its Lyndon factorization $W(C) = \ell_1^{n_1} \ldots \ell_m^{n_m}$. Let x be the closest upper point in $\bar{C}$ of the jth Lyndon factor ℓ_i for $i \in [1, m]$ and $j \in [1, n_i]$ in $W(C)$ such that $split(\ell_i) = (\ell_i^+, \ell_i^-)$. From Lemma 2, we know that $\ell_i^- \ell_i^{n_i-1}$ is a Christoffel word and that $\ell_i^+ > \ell_i^- \ell_i^{n_i-1}$. Let us consider that R_{i+1} is in $W_{\mathcal{WN}}(C)$. To check the $\mathcal{WN}$-convexity, we must compare $R_0 = \ell_i^- \ell_i^{n_i-1}$ to ℓ_{i+1}. We get the following cases:

1. If $\ell_{i+1} < R_0$, so no further verification is needed and the digital convexity is maintained.
2. If $\ell_{i+1} = R_0$, the convexity is also preserved in this case since the line segments discretized by R_0 and $\ell_{i+1}^{n_{i+1}}$ will be aligned, and the multiplicity of ℓ_{i+1} is increased by 1 in the factorization of w.
3. If $\ell_{i+1} = R_0^{m_i} \ell_i$, then $R_1 = R_0 \ell_{i+1}^{n_{i+1}}$ is set and R_1 is now compared with to $\ell_{i+2}^{n_{i+2}}$ using the same reasoning. This propagation is kept until the right insertability verification is done with one of the cases 1, 2 and 4 or until the end of $W_{\mathcal{WN}}(C)$. Hence the sequence R_k is constructed.
4. If $\ell_{i+1} > R_0$ and $\ell_{i+1} \neq R_0^{m_i} \ell_i$, then we lose the decreasing order, namely the $\mathcal{WN}$-convexity. Hence, the point x is not insertable.

A similar proof can be given for the left insertability. Note that if the length of ℓ_i is equal to one, the left insertability should be verified in the previous binary boundary sub-word. □

Note that the propagation test is limited by the extremity of the $\mathcal{WN}$ side. In Fig. 16, we show an example of the propagation following Theorem 6.

With Theorem 6, we have the characterization of the insertable pixels over $W(C)$. We give in Examples 7–9, several numerical cases showing the inflation process.

Example 7. Let us consider a $\mathcal{WN}$-convex word $w = \left(\mathscr{C}\left(\frac{3}{4}\right)\right)^4 \mathscr{C}\left(\frac{3}{5}\right)\mathscr{C}\left(\frac{1}{2}\right)$ and the split of the second factor. As $split(\mathscr{C}\left(\frac{3}{5}\right)) = \left(\mathscr{C}\left(\frac{2}{3}\right), \mathscr{C}\left(\frac{1}{2}\right)\right)$, we obtain the new $\mathcal{WN}$-convex word w';

$$w' = \left(\mathscr{C}\left(\frac{3}{4}\right)\right)^4 \mathscr{C}\left(\frac{2}{3}\right)\left(\mathscr{C}\left(\frac{1}{2}\right)\right)^2.$$

Example 8. Let us consider a $\mathcal{WN}$-convex word $w = (\mathscr{C}(\frac{3}{4}))^4(\mathscr{C}(\frac{3}{5}))^3\mathscr{C}(\frac{1}{2})$ and the split of the second word of the second factor ℓ_2. As $split(\mathscr{C}(\frac{3}{5})) = (\mathscr{C}(\frac{2}{3}), \mathscr{C}(\frac{1}{2}))$ and $\ell_3 = \ell_2^-$, we obtain:

$$w' = \left(\mathscr{C}\left(\frac{3}{4}\right)\right)^4 \mathscr{C}\left(\frac{3}{5}\right)\mathscr{C}\left(\frac{2}{3}\right)\mathscr{C}\left(\frac{1}{2}\right)\mathscr{C}\left(\frac{3}{5}\right)\mathscr{C}\left(\frac{1}{2}\right)$$

whose Lyndon factorization is $(\mathscr{C}(\frac{3}{4}))^4\mathscr{C}(\frac{5}{8})\mathscr{C}(\frac{4}{7})\mathscr{C}(\frac{1}{2})$

Example 9. Let us consider a $\mathcal{WN}$-convex word $w = (\mathscr{C}(\frac{3}{4}))^4 (\mathscr{C}(\frac{3}{5}))^5 \mathscr{C}(\frac{24}{41})$, so that $\ell_3 = (\ell_2^- \ell_2^2)^3 \ell_2$, and consider splitting the third word of the second factor $\mathscr{C}(\frac{3}{5})$. We get:

$$w' = \left(\mathscr{C}\left(\frac{3}{4}\right)\right)^4 \mathscr{C}\left(\frac{2}{3}\right)\mathscr{C}\left(\frac{13}{22}\right)\mathscr{C}\left(\frac{24}{41}\right)$$

whose Lyndon factorization is $(\mathscr{C}(\frac{3}{4}))^4\mathscr{C}(\frac{8}{13})\mathscr{C}(\frac{31}{53})$.

4.3. Strong sufficient condition to insertable points

Given a digital convex 4-connected set C, till now, for each closest upper point of C, we must check the conditions given in Theorem 6 in order to know if it is an insertable point or not. However, checking if a point is insertable, using Theorem 6 may require processing the whole boundary word $W(C)$ due to the propagation process for the left and right insertability verification (see Definition 12). In this section, we consider some extra local constraints on a chosen Lyndon factor, which ensures that the associated closest upper point is always insertable without the propagation process. For this aim, the authors in [7] restricted the study within the case where all the factors of the Lyndon factorization of the boundary word are primitive. They choose the closest upper point of the segment that is correlated to a primitive Christoffel word of maximal length with respect to the previous and next primitive Christoffel words. With this constraint, removing the corresponding closest upper point of the locally longest Christoffel word always preserves the digital convexity. This result was not proved and neither generalized in the case where the factors of the boundary word are not primitive. We provide here the general result with a full study for all the possible cases and updates.

In fact, we would like to show that splitting a locally longest primitive Christoffel word of $W(C)$ guarantees inflation at any step, while preserving the digital convexity. In other terms, the closest upper points of all the primitive Christoffel words of local maximal length correspond to insertable points. Before reaching this theorem, we prove first that if we have three consecutive decreasing Christoffel words such that the first one is longer than the neighbors, then its split preserves this decreasing order in the local part.

From Definition 13 and Property 6, we recall that the Christoffel morphism Θ_B induces an increasing bijection between the set of Christoffel words and itself. Based on this we can get the following Corollary.

Corollary 5. *Any Christoffel word C such that $\mathscr{C}(\frac{a}{b}) < C < \mathscr{C}(\frac{c}{d})$ satisfies $|C| \geq |\mathscr{C}(\frac{a}{b})| + |\mathscr{C}(\frac{c}{d})| > \max(|\mathscr{C}(\frac{a}{b})|, |\mathscr{C}(\frac{c}{d})|)$.*

Proof. Let $B = (\mathscr{C}(\frac{a}{b}), \mathscr{C}(\frac{c}{d}))$, and $\Theta = \Theta_B$ be the Christoffel morphism defined by $\Theta(0) = \mathscr{C}(\frac{a}{b})$ and $\Theta(1) = \mathscr{C}(\frac{c}{d})$. If $\mathscr{C}(\frac{a}{b}) < C < \mathscr{C}(\frac{c}{d})$ then $C = \Theta(U)$, where U is a Christoffel word other than 0 or 1. Hence, U contains one letter 0 and one letter 1. Therefore the image contains two disjoint factors $\mathscr{C}(\frac{a}{b})$ and $\mathscr{C}(\frac{c}{d})$. $\square$

Corollary 5 allows us to give an algorithm that inflates a digital convex set C while preserving its digital convexity. We start by defining the notion of maximal primitive length of a Christoffel word with respect to the previous and next one.

Definition 14. Let $W(C) = \ell_1^{n_1} \ldots \ell_s^{n_s}$ be the Lyndon factorization of a digital convex set C, and let ℓ_j, $1 \leq j \leq s$, be one of the Christoffel words, seen in a cyclic way at $j = 1$ and $j = s$. We say that ℓ_j has a local maximal primitive length if $|\ell_j| > \max(|\ell_{j-1}|, |\ell_{j+1}|)$.

By inflating the Christoffel word ℓ_j, with a local maximal primitive on the closest upper point, we preserve the digital convexity, thanks to Theorem 7.

Theorem 7. *Given a digital convex 4-connected set C, if $x \in \overline{C}$ is the closest upper point of a local maximal primitive length Christoffel word of the Lyndon factorization of the boundary word $W(C)$ then x is an insertable point.*

Proof. The proof is deduced from Corollary 5. $\square$

This means that adding this constraint assures that this particular closest upper point is an insertable one.

After inflating C with respect to this strong constraint, we only can face one out of four cases, when we update $W(C)$. They are mentioned in Lemma 3.

Lemma 3. *Let ℓ_j be one of the Christoffel words of local maximal primitive length in the boundary word, $W(C) = \ell_1^{n_1} \ldots \ell_s^{n_s}$, of a 4-connected digital convex set C. By applying the split operator on the ith ℓ_j for any $1 \le i \le n_j$ such that $split(\ell_j) = (\ell_j^+, \ell_j^-)$, the Lyndon factorization can be updated by the following local replacement:*

1. *if $(i > 1$ or $\ell_{j-1} > \ell_j^+)$ and $(i < n_j$ or $\ell_{j+1} < \ell_j^-)$:*

$$\ell_1^{n_1} \ell_2^{n_2} \ldots \ell_{j-1}^{n_{j-1}} (\ell_j^{i-1} \ell_j^+)(\ell_j^- \ell_j^{n_j-i}) \ell_{j+1}^{n_{j+1}} \ldots \ell_k^{n_k},$$

2. *if $(i = 1$ and $\ell_{j-1} = \ell_j^+)$ and $(i < n_j$ or $\ell_{j+1} < \ell_j^-)$:*

$$\ell_1^{n_1} \ell_2^{n_2} \ldots \ell_{j-1}^{n_{j-1}+1} (\ell_j^- \ell_j^{n_j-i}) \ell_{j+1}^{n_{j+1}} \ldots \ell_k^{n_k},$$

3. *if $(i > 1$ or $\ell_{j-1} > \ell_j^+)$ and $(i = n_j$ and $\ell_{j+1} = \ell_j^-)$:*

$$\ell_1^{n_1} \ell_2^{n_2} \ldots \ell_{j-1}^{n_{j-1}} (\ell_j^{i-1} \ell_j^+) \ell_{j+1}^{n_{j+1}+1} \ldots \ell_k^{n_k},$$

4. *if $(i = 1$ and $\ell_{j-1} = \ell_j^+)$ and $(i = n_j$ and $\ell_{j+1} = \ell_j^-)$:*

$$\ell_1^{n_1} \ell_2^{n_2} \ldots \ell_{j-1}^{n_{j-1}+1} \ell_{j+1}^{n_{j+1}+1} \ldots \ell_k^{n_k}.$$

Proof. The proof of this lemma relies on the following two points:

1. Showing that $u = \ell_j^{i-1} \ell_j^+$ and $v = \ell_j^- \ell_j^{n_j-i}$ are Christoffel words. This follows from Lemma 2 by taking the base $B = (\ell_j^-, \ell_j^+)$. We obtain $u = \Theta_B((01)^{i-1}1)$ and $v = \Theta_B(0(01)^{n_j-i})$.
2. Proving the following inequalities: $\ell_{j-1} > \ell_j^{i-1} \ell_j^+ > \ell_j^- \ell_j^{n_j-i} > \ell_{j+1}$.

   - The inequality in the middle comes from the fact that Θ_B defined earlier is increasing and $0(01)^{n_j-i} < (01)^{i-1}1$.
   - If the last inequality is not correct, we have: $\ell_j^- \le \ell_j^- \ell_j^{n_j-i} \le \ell_{j+1} < \ell_j$. Then ℓ_j is a Christoffel word in the B and the condition does not allow the equality $\ell_{j+1} = \ell_j$; in this case it has to be longer than ℓ_j, contradicting the condition that ℓ_j is longer than ℓ_{j+1}.
   - The first inequality is treated in a symmetric way as the previous one. $\square$

From Lemma 3, we can remark that in all the four cases, the digital convexity is preserved. The inflated segment of the $conv(C)$ is replaced by two others. For the second case, ℓ_j^+ is the same as $\ell_j^{n_{j-1}}$ and a concatenation from the left side arises. Similarly and by symmetry, for the third case, ℓ_j^- is the same as $\ell_{j+1}^{n_{j+1}}$ and a concatenation from the right side arises. For the last case, ℓ_j^+ (resp. ℓ_j^-) is the same as $\ell_j^{n_{j-1}}$ (resp. $\ell_{j+1}^{n_{j+1}}$), and concatenations from both sides arise. In other words, we lose one of the segments of $conv(C)$. For all the cases, we can note that the propagation does not exceed the neighboring Christoffel words $\ell_{j+1}^{n_{j+1}}$ and $\ell_{j-1}^{n_{j-1}}$.

Corollary 6. *Splitting the Christoffel word of the Lyndon factorization which has the local maximal primitive length limits the propagation and bounds it only by the previous and next factor.*

4.4. Algorithm

We give now a general algorithm to inflate C based on the previous results: for the general case or the case with strong constraint. In Algorithm 2, we determine the list of all the possible insertable points obtained either by applying Theorems 6 or 7. For each iteration, we choose the point with the highest priority queue and we apply the necessary updates on the Lyndon factorization. We recall that for the general case, this update can lead to a propagation, while for the case with strong constraint, this update is local. Two different and more detailed algorithms, for each of the cases separately, will be given in future work. By using this approach for inflation, if we keep choosing the side with primitive local maximal length at each iteration, we remark that the horizontal and vertical segments of $conv(C)$ will not be chosen. In fact, the discretization of these segments is either factors of the form 0^p, 1^q, $\overline{0}^r$ or $\overline{1}^s$ for certain integer numbers p, q, r, or s. This means that their primitive length is always equal to 1. Hence, at any step of the inflation they cannot be designated. We see in this case, that the inflation will happen at the beginning on the $\mathcal{WN}$, $\mathcal{NE}$, $\mathcal{ES}$ and $\mathcal{SW}$ sides without considering the horizontal and vertical sides which leads to an octagonal shape as seen in Fig. 17. Once we are at this step, all the remaining factors are of length 2 and 1. The inflation will continue until we reach the form: $1^k 0^l \overline{1}^k \overline{0}^l$ for certain integers l, k, which is our rectangle-bounded box.
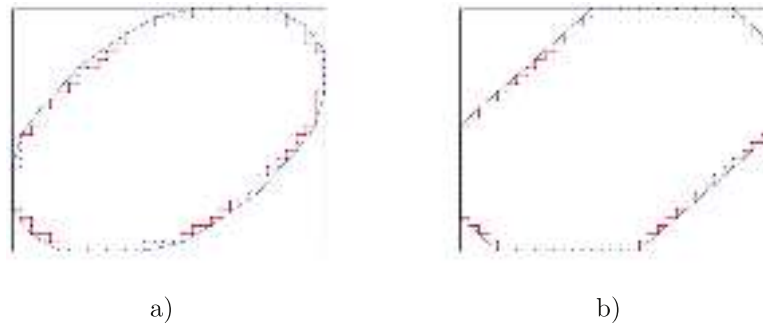
a)   b)

Fig. 17. Figure (b) shows the inflation of the digital convex set represented in (a) by applying the algorithm with the stronger constraint based on local maximal primitive length of a Lyndon factor.

Algorithm 2 Point-wise inflation

Input: a digital convex 4-connected set C, a number of adding points k
Output: a sequence of inserted points I

1: Compute the Lyndon factorization F of the boundary word w of C
2: Insert all insertable points in a priority queue P
3: **while** $k \geq 0$ **do**
4:      Pull the highest-priority insertable point p from P and add p to I
5:      Let ℓ_j be the associated primitive Christoffel word of p in F;
6:      Compute $split(\ell_j)$ and update F and P following Lemmas 2 or 3
7:      Decrement k
8: **end while**
9: **return** I

---

5. Conclusion

We have proposed a combinatorics-on-words study of the points that can be chosen to inflate and deflate a 4-connected digital convex set C while preserving its digital convexity property. The approach relies on Christoffel words and Lyndon factorizations of the boundary word of C that is represented on a four letter alphabet. Some update procedures are to be done on these factors, in order to maintain the Lyndon factorizations while adding/removing specific points. For both operations, we have characterized for C the set of points that can be inserted or removed, while maintaining the convexity. For the deflation process, the updates and modifications are local. In contrast, for the inflation process, the updates, using the general procedure, can be global. In worst case, these updates do not go past the side where the inflation is applied. Adding the strong condition on the choice of the insertable points corresponding to the local maximal primitive Christoffel word, the updates during this procedure become local.

In this work, we have focused on the characterization and theoretical properties of geometrical set operations on boundary word factorizations. The algorithmic details and optimization can be found in [31]. Another question that can arise is to determine if there exists an optimal heuristic for deflating a digital convex set. The choice of the heuristic is crucial, since for each choice we get a different convergent shape. Another stimulating perspective would be to apply these algorithms on non-convex shapes by studying the locally convex boundary using combinatorics on words [12,27].

Data availability

No data was used for the research described in the article.

Acknowledgments

References

[1] J.C. Alexander, A.I. Thaler, The boundary count of digital pictures, J. ACM 18 (1) (1971) 105–112.
[2] A. Apostolico, M. Crochemore, Fast parallel lyndon factorization with applications, Math. Syst. Theory 28 (1995) 89–108.
[3] J. Berstel, A. Lauve, C. Reutenauer, F. Saliola, Combinatorics on Words: Christoffel Words and Repetitions in Words, volume 27, American Mathematical Society, 2009.

[4] J. Berstel, A. De Luca, Sturmian words, lyndon words and trees, Theoret. Comput. Sci. 178 (1) (1997) 171–203.
[5] G. Bertrand. Simple points, Topological numbers and geodesic neighborhoods in cubic grids, Pattern Recognit. Lett. 15 (1994) 1003–1011.
[6] J.-P. Borel, F. Laubie, Quelques mots sur la droite projective réelle, J. Théor. Nombres Bordx. 5 (1) (1993) 23–51.
[7] J-P. Borel, L. Tarsissi, L. Vuillon, Inflation of digitally convex polyominoes, in: JournÉE Montoises D'Informatique ThÉOrique, Bordeaux, France, 2018.
[8] S. Brlek, J.-O. Lachaud, X. Provençal, C. Reutenauer, Lyndon Christoffel = digitally convex, Pattern Recognit. 42 (10) (2009) 2239–2246.
[9] E. Christoffel, Observatio arithmetica, Ann. Mat. Pura Appl. (1867-1897) 6 (1) (1875) 148–152.
[10] W.-F. Chuan, α-words and factors of characteristic sequences, Discrete Math. 177 (1–3) (1997) 33–50.
[11] M. Couprie, G. Bertrand, New characterizations of simple points in 2D, 3D and 4D discrete spaces, IEEE Trans. Pattern Anal. Mach. Intell. 31 (2009) 637–648.
[12] I. Debled-Rennesson, J.L. Rémy, J. Rouyer-Degli, Detection of the discrete convexity of polyominoes, Discrete Appl. Math. 125 (1) (2003) 115–133.
[13] P. Dulio, A. Frosini, S. Rinaldi, L. Tarsissi, L. Vuillon, First steps in the algorithmic reconstruction of digital convex sets, in: International Conference on Combinatorics on Words, Springer, 2017, pp. 164–176.
[14] P. Dulio, A. Frosini, S. Rinaldi, L. Tarsissi, L. Vuillon, Further steps on the reconstruction of convex polyominoes from orthogonal projections, J. Comb. Optim. 44 (4) (2022) 2423–2442.
[15] J.P. Duval, Factorizing words over an ordered alphabet, J. Algorithms 4 (1983) 363–381.
[16] U. Eckhardt, Digital lines and digital convexity, Lecture Notes in Comput. Sci. 2243 (2001) 209–228.
[17] P. Flajolet, R. Sedgewick, Analytic Combinatorics, Cambridge University Press, 2009.
[18] H. Freeman, On the encoding of arbitrary geometric configurations, IRE Trans. Electron. Comput. 2 (1961) 260–268.
[19] C.E. Kim, On the cellular convexity of complexes, IEEE Trans. Pattern Anal. Mach. Intell. 3 (6) (1981) 617–625.
[20] C.-E. Kim, A. Rosenfeld, Digital straight lines and convexity of digital regions, IEEE Trans. Pattern Anal. Mach. Intell. 2 (1982) 149–153.
[21] T.Y. Kong, A. Rosenfeld, Digital topology: Introduction and survey, Comput. Vis. Graph. Image Process. 48 (3) (1989) 357–393.
[22] M. Lothaire, Applied combinatorics on words, in: Encyclopedia of Mathematics and its Applications, vol. 90, Cambridge University Press, Cambridge, 2002.
[23] R. Lyndon, Identities in finite algebras, Proc. Amer. Math. Soc. 5 (1) (1954) 8–9.
[24] G. Melançon, Lyndon words and singular factors of sturmian words, Theor. Comput. Sci. 218 (1) (1999) 41–59.
[25] M. Minsky, S. Papert, Perceptrons, MIT Press, 1969.
[26] A. Rosenfeld, Connectivity in digital pictures, J. ACM 17 (1) (1970) 146–160.
[27] T. Roussillon, I. Sivignon, Faithful polygonal representation of the convex and concave parts of a digital curve, Pattern Recognit. 44 (10–11) (2011) 2693–2700.
[28] M.I. Shamos, Problems in computational geometry, Comput. Geom. (1975).
[29] J. Sklansky, Recognition of convex blobs, Pattern Recognit. 2 (1970) 3–10.
[30] L. Tarsissi, D. Coeurjolly, Y. Kenmochi, P. Romon, Convexity preserving contraction of digital sets, in: Pattern Recognition, Auckland, New Zealand, 2019.
[31] L. Tarsissi, Y. Kenmochi, H. Djerroumi, D. Coeurjolly, P. Romon, J.-P. Borel, Algorithms for pixelwise shape deformations preserving digital convexity, Discret. Geom. Math. Morphol. (2022).