# Routing by matching on convex pieces of grid graphs

H. Alpert[a,*], R. Barnes[b], S. Bell[c], A. Mauro[d], N. Nevo[e], N. Tucker[f], H. Yang[g]

[a] *Auburn University, 221 Parker Hall, Auburn, AL 36849*
[b] *Harvey Mudd College, 320 East Foothill Boulevard, Claremont, CA 91711*
[c] *Willamette University, 900 State Street, Salem, OR 97301*
[d] *Stanford University, Building 380, Stanford, CA 94305*
[e] *Colorado College, 14 E. Cache La Poudre St., Colorado Springs, CO 80903*
[f] *Juniata College, 1700 Moore Street, Huntingdon, PA 16652*
[g] *MIT, 77 Massachusetts Avenue, Cambridge, MA 02139*

**Abstract**

The routing number is a graph invariant introduced by Alon, Chung, and Graham in 1994, and it has been studied for trees and other classes of graphs such as hypercubes. It gives the minimum number of routing steps needed to sort a set of distinct tokens, placed one on each vertex, where each routing step swaps a set of disjoint pairs of adjacent tokens. Our main theorem generalizes the known estimate that a rectangular grid graph $R$ with width $w(R)$ and height $h(R)$ satisfies $\mathrm{rt}(R) \in O(w(R)+h(R))$. We show that for the subgraph $P$ of the infinite square lattice enclosed by any convex polygon, we have $\mathrm{rt}(P) \in O(w(P)+h(P))$.

*Keywords:* routing number, makespan, motion planning, token graph, parallel sorting

*2010 MSC:* 68U05, 05C85, 68M10

## 1. Introduction

Routing number is an invariant of graphs, defined by Alon, Chung, and Graham [1]. Given a connected graph $G$ on $n$ vertices, we imagine tokens labeled 1 through $n$ sitting on the vertices of $G$ in some order, with exactly

*Corresponding author
*Email addresses:* hcalpert@auburn.edu (H. Alpert), rjbarnes@hmc.edu (R. Barnes), scbell@willamette.edu (S. Bell), amauro@stanford.edu (A. Mauro), n_nevo@coloradocollege.edu (N. Nevo), tuckent18@juniata.edu (N. Tucker), hannay@mit.edu (H. Yang)

one token per vertex. That is, the tokens specify a bijection between the set of vertices and the set of token labels $\{1, 2, \ldots, n\}$. We refer to each of the $n!$ bijections as a ***configuration*** of labeled tokens on $G$. We imagine transforming one token configuration into another using a sequence of ***routing steps***, defined as follows. In each routing step, we may select any set of disjoint edges in $G$, and for each edge, swap the tokens on the two vertices of that edge. The result is another token configuration.

For any two token configurations, there is a sequence of routing steps we can apply to the first configuration to transform it into the second configuration. To see this, take a spanning tree of $G$, and move the tokens into position one at a time, starting with the leaves and moving inward. Thus, we may define the ***distance*** between any two configurations to be the least possible number of routing steps that can be used to transform one configuration into the other. The ***routing number*** of $G$, denoted $\mathrm{rt}(G)$, is the maximum, over all pairs of token configurations, of the distance between the two configurations.

In our main theorem, the graphs we consider are induced subgraphs of the infinite grid graph, which has vertex set $\mathbb{Z} \times \mathbb{Z}$ and an edge between each pair of vertices with Euclidean distance 1. Given a convex polygon $P \subseteq \mathbb{R}^2$, we define the ***convex grid piece cut out by*** $P$ to be the graph $G_P$ with vertices at all lattice points in and on $P$, and edges between pairs of lattice points of distance 1. In the remainder of the paper, we use the letter $P$ for both the polygon and the graph, using the notation $\mathrm{rt}(P)$ to mean $\mathrm{rt}(G_P)$. Although there are some convex polygons $P$ for which the graph $G_P$ is disconnected, the routing number is defined only when $G_P$ is connected. Note also that when $P$ is translated or rotated, the graph changes, and so the routing number may change slightly.

Our main theorem bounds $\mathrm{rt}(P)$ in terms of the width and height of $P$. The width $w(P)$ and height $h(P)$ are the maximum absolute differences in $x$-coordinates and in $y$-coordinates, respectively, of any pair of points in $P$.

**Theorem 1.** *Let $P$ be a connected convex grid piece. Then the routing number of $P$ satisfies the bound $\mathrm{rt}(P) \leq C(w(P) + h(P))$ for some universal constant*

$C$.

In other words, we have $\mathrm{rt}(P) \in O(w(P) + h(P))$. The reverse inequality is immediate: the diameter of $P$ is within a constant factor of $w(P) + h(P)$, and the routing number of any graph is at least its diameter, because a token may need to travel between two farthest vertices. Thus we may estimate $\mathrm{rt}(P)$ as $\Theta(w(P) + h(P))$.

One motivation for studying the routing number of convex grid pieces is as a discrete model of configuration spaces of disks. Given a region $R$ in the plane, such as a convex polygon, the configuration space $\mathrm{Conf}_{n,r}(R)$ as defined in [2] is the space of all ways to arrange $n$ disjoint, labeled disks of radius $r$ inside $R$; we refer to each of these arrangements as a configuration. If the configuration space is connected, we can define the distance between two configurations to be the amount of time it takes to move between them if the disks can move simultaneously, each with speed at most 1. Roughly, the maximum distance between two configurations corresponds to the routing number of the grid piece cut out by $R$; one major difference is that the routing number does not account for what proportion of $R$ is covered by disks, simplifying the problem.

Whereas the routing number of graphs has clear significance in terms of routing information through computer networks, configuration spaces of disks have their own concrete applications. The 3-dimensional version of disk configuration spaces is the hard spheres gas model, in which the disks (or spheres) represent individual molecules moving around in a container; see [3, 4] for exposition on the hard spheres model. If the molecules are densely packed, they can only rattle in place, as in a solid; if there is a lot of space, they can move almost independently, as in a gas, and at intermediate densities the configuration space is somehow like that of a liquid. Another interpretation of configuration spaces of disks imagines each disk as a robotic car, moving around in an enclosed room such as a warehouse floor. The geometry and topology of the configuration space constrains what instructions may be used to coordinate the motion of the robots, as in Farber's "topological complexity" [5]. All of these applications in-

volve a requirement to avoid collisions, which corresponds in the routing number setup to the requirement that no two tokens can occupy the same vertex.

Researchers interested in the robotic car interpretation have made various discrete models of configuration spaces of disks; see, for instance, [6, 7, 8]. Typically a discrete result is proved for a rectangular grid, and then the discrete result implies a continuous result about configurations of disks in a rectangular region. Although restricting attention to a rectangular regions may seem like a minor assumption, the proof structure of the discrete results tends to rely on the rectangular shape. The reason is that rectangles are self-similar: a rectangular grid is a union of smaller rectangular blocks, with the blocks arranged again in a rectangular grid pattern. For robotic cars moving in a round disk, for example, these self-similarity properties do not apply.

Thus, the purpose of our theorem is to prove a discrete result for regions that are not necessarily rectangular. The proof is for convex regions because the claim is not true for arbitrary nonconvex regions; for grid pieces cut out by nonconvex polygons, the bound on routing number is about as bad as for arbitrary trees, which are the hardest to route of all graphs. Although considering routing number of convex grid pieces is just one possible discrete model for configuration spaces of disks, we hope that the proof method suggests the steps needed to prove such a result for other discrete models as well. Some models allow a large loop of tokens to cycle by one step simultaneously, rather than only allowing adjacent pairs to swap. This is similar to the continuous case, where a long line of disks can slide simultaneously toward an open space. Because our main theorem is an upper bound on routing number, it also implies an upper bound on routing in the more permissive model where cycling requires fewer steps.

To prove the theorem, we first construct an algorithm for routing tokens on a special class of convex grid pieces, which we call ramp-like polygons. This class generalizes both rectangles and right triangles, and the recursive algorithm is fairly technical. Then, we prove that bounds on routing number for some graphs imply bounds on routing number for other graphs: if we can route ramp-like polygons, then we can route polygons cut into two (and then four) ramp-like

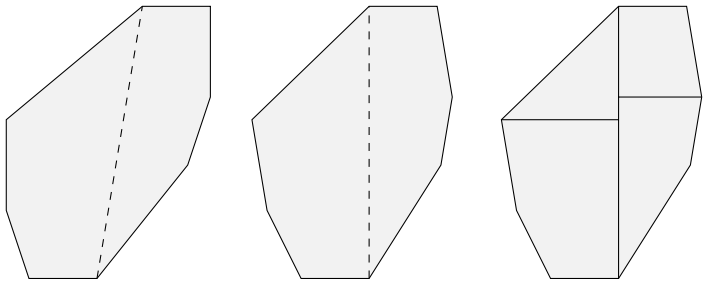


Figure 1: Given an arbitrary convex polygon, as shown on the left, our overall proof strategy is to shear it by at most 45° to get a "burger bun" polygon, where the top and bottom are vertically aligned as shown in the middle. This burger bun polygon can be cut vertically into two pieces, and then horizontally into four "ramp-like" pieces, as shown on the right. We show how to route within each of the four pieces, then within each of the two pieces, then within the burger bun polygon, and then within the original polygon.

pieces, and then if we shear these polygons by at most 45 degrees, we can still route the result. Using these reductions we show the bound for all convex grid pieces. Figure 1 depicts the conceptual flow of the strategy.

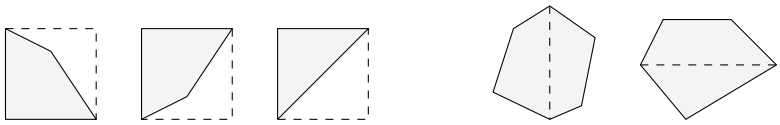
Section 2 contains definitions and lemmas needed for the rest of the paper. In Section 3, we prove the routing number bound for the class of ramp-like polygons (defined in Section 2). In Section 4, we extend the bound to a more general class which we call burger bun polygons (defined in Section 2), each of which can be cut into four ramp-like pieces. In Section 5 we extend the bound to arbitrary convex polygons, using the fact that they can be obtained from burger bun polygons using a shear transformation of at most 45 degrees. Section 6 contains some questions for further study.

*Acknowledgments.* This research was performed at the MathILy-EST 2020 REU, supported by the National Science Foundation under Award No. DMS 1851842. H. Alpert was also supported by National Science Foundation Award No. DMS 1802914. We thank the referees for many helpful suggestions that improved the paper substantially.

## 2. Preliminaries

In this section, first we give definitions needed for the rest of the paper. Then in Theorems 2, 3, and 4 we state the known results on routing number that we need. Finally in Lemmas 5 and 6 we prove two lemmas that we use in multiple later sections, showing that adding a small number of vertices to a graph does not increase the routing number by too much.

We define ramp-like polygons and burger bun polygons to be special classes of convex polygons in the plane. A ***ramp-like polygon*** is a convex polygon that shares two edges with its bounding box. That is, there is a rectangle containing our polygon with vertices $(x_1, y_1), (x_1, y_2), (x_2, y_1), (x_2, y_2)$, such that (at least) three of these vertices are vertices of our convex polygon. A ***burger bun polygon*** either has top and bottom points on the same vertical line, or has leftmost and rightmost points on the same horizontal line. That is, either there are two points $(x, y_1)$ and $(x, y_2)$ such that all the $y$-coordinates in the polygon are in the interval $[y_1, y_2]$, or there are two points $(x_1, y)$ and $(x_2, y)$ such that all the $x$-coordinates in the polygon are in the interval $[x_1, x_2]$. Some examples of each of these polygons can be seen in Figure 2.



(a) (b)

Figure 2: (a) Ramp-like polygons share at least two sides with their bounding boxes (drawn with dashed lines) and (b) burger-bun polygons either have top and bottom points on the same vertical line, or have leftmost and rightmost points on the same horizontal line.

We have defined the routing number $\mathrm{rt}(G)$ of a graph $G$ to be the minimum number of routing steps needed to get from any permutation of labeled tokens on the vertices of $G$ to any other permutation. Sometimes, instead of having a different label for each token, it helps to consider just two distinct types of tokens, for instance, black tokens and white tokens. Equivalently, we can

consider all of the tokens to be identical, but have some vertices with no tokens on them, so that instead of black tokens and white tokens, we have vertices with tokens and vertices without tokens. We define the ***unlabeled routing number*** of $G$, denoted $\mathrm{urt}(G)$, to be the minimum number of routing steps needed to get from any arrangement of black and white tokens, one token per vertex of $G$, to any other arrangement with the same number of black and white tokens. That is, we take the maximum, over all $k$ and all pairs of arrangements with $k$ black tokens and $|V(G)| - k$ white tokens, of the minimum number of routing steps needed to get from one arrangement to the other. As before, each routing step consists of selecting a set of disjoint edges of $G$ and swapping the two tokens on the ends of each edge. Sometimes we refer to routing number as ***labeled routing number*** to distinguish it from unlabeled routing number.

For reference we state the theorems estimating the routing numbers of paths, trees, and rectangular grids. The versions that follow are sufficient for our use in this paper, and the proofs can be found in [1].

**Theorem 2** (Path bound)**.** *For a path $P$ with $n$ vertices, we have* $\mathrm{rt}(P) = n$.

**Theorem 3** (Tree bound)**.** *For a tree with $n$ vertices, and thus for any connected graph $G$ with $n$ vertices, we have* $\mathrm{rt}(G) \leq 3n$.

**Theorem 4** (Rectangle bound)**.** *For a $p$ by $q$ rectangular grid graph $R_{p,q}$, we have*

$$\mathrm{rt}(R_{p,q}) \leq \frac{3}{2}(p+q).$$

The two lemmas in the remainder of this section are stated in terms of ***lattice graphs***, which we define to be graphs $G$ such that the vertex set is a set of points $(x, y) \in \mathbb{Z} \times \mathbb{Z}$, and there is an edge between two vertices whenever the Euclidean distance between them is exactly 1. Our notation sometimes conflates polygons, graphs, and their vertex sets. When we use set operations on graphs, we typically mean that the operation should be done on the vertex sets, and then we should consider the induced subgraph of $\mathbb{Z} \times \mathbb{Z}$ determined by the resulting set of vertices. We use the notation $w(G)$ and $h(G)$ for the width

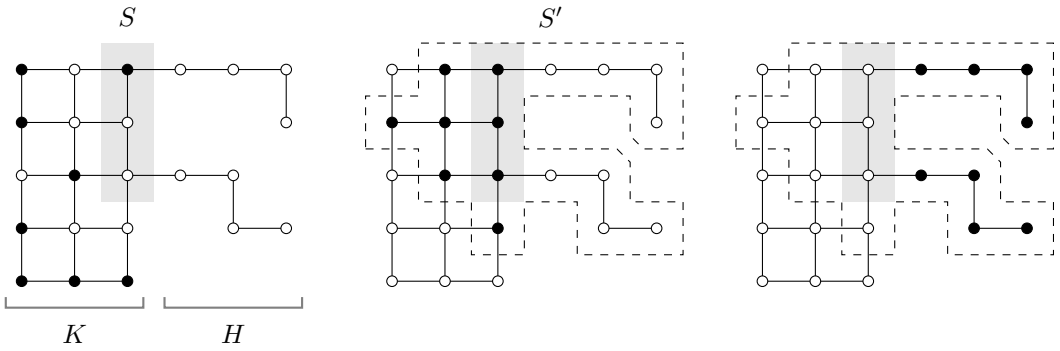



Figure 3: Left: The skin set $S$ (shown with gray background) is part of the core set $K$, and its union with the hair set $H$ is connected. In the starting token configuration, the tokens belonging in $H$ are shown in black. Middle: In the first routing phase, we move the tokens within $K$ so that those belonging in $H$ occupy $S$ and the vertices nearest to $S$. The union of these vertices with $H$ is $S'$ (shown outlined by dashes). Right: Next, we route within $S'$ so that the tokens belonging in $H$ move to their home positions.

and height of a lattice graph $G$, defined similarly to the width and height of a polygon.

In the rest of the paper, we will often trim off, or add on, small or skinny parts of our graphs that do not significantly change the routing numbers. The following lemma makes that operation precise. In the statement of the lemma, the words "core", "hair", and "skin" do not have any definition beyond what appears in the lemma statement, but are meant to aid in remembering the roles of the various subgraphs.

**Lemma 5.** *Let $G$ be a connected lattice graph, with its vertices partitioned into sets $K$ ("core") and $H$ ("hair"). Suppose that*

1. *There are at most $c_1 \cdot (w(G) + h(G))$ vertices in $H$;*
2. *There is a set of vertices $S \subseteq K$ ("skin") containing at most $c_2 \cdot (w(G) + h(G))$ vertices, such that the induced subgraph $S \cup H$ is connected; and*
3. *The routing number of $K$ is at most $c_3 \cdot (w(G) + h(G))$.*

*Then, we have* $\mathrm{rt}(G) \leq (6c_1 + 3c_2 + 2c_3)(w(G) + h(G))$.

*Proof.* Figure 3 shows an example of a core $K$, skin $S \subseteq K$, and hair $H$, and summarizes the routing steps.

First we route within $K$ so that the tokens in $K$ that belong in $H$ have the following property: no token in $K$ that belongs in $H$ has a greater distance to $S$ than a token in $K$ that does not belong in $H$. That is, for the tokens in $K$ belonging in $H$, we move them to fill $S$ first, then to fill the vertices at distance 1 from $S$, and so on. This first phase takes at most $c_3(w(G) + h(G))$ routing steps, by the hypothesis on the routing number of $K$.

Next, we consider the induced subgraph $S'$ of $G$ containing $H$, $S$, and any other vertices with tokens belonging in $H$. Because of the previous step and the fact that $H \cup S$ is connected, we know that $S'$ is connected. It has at most $2|H| + |S| \leq (2c_1 + c_2)(w(G) + h(G))$ vertices, by the hypotheses on the sizes of $H$ and $S$. Thus, if we use the routing method in Theorem 3 to route along the spanning tree of $S'$, we may use at most $3(2c_1 + c_2)(w(G) + h(G))$ routing steps on $S'$ to move all the tokens belonging in $H$ to their home vertices.

Finally, we route within $K$ to move all the tokens belonging in $K$ to their home vertices. The total number of routing steps is at most $(6c_1 + 3c_2 + 2c_3)(w(G) + h(G))$. □

Sometimes the "skin" set $S$ is very easy to describe, but in the final proof we need to be able to find the skin set of an arbitrary convex grid piece (as defined in the introduction). The following theorem describes how to do so.

**Lemma 6.** *Let $P \subseteq \mathbb{R}^2$ be a convex polygon, and suppose that the corresponding convex grid piece $K$ is a connected graph. Then there is a connected subgraph $S$ of $K$ containing at most $2(w(K) + h(K))$ vertices, with the following property: if $G$ is a connected lattice graph containing $K$, then $S \cup (G \setminus K)$ is connected.*

*Proof.* The process of constructing $S$ is shown in Figure 4; there, $S$ is the subgraph shown with thick edges inside the polygon. We think of $S$ as the circuit enclosing $K$; informally, if we cut the plane along all the edges of $K$, we get many unit square pieces and one unbounded piece, and $S$ is the boundary of the unbounded piece. To be more precise, we start with the loop $P$, which

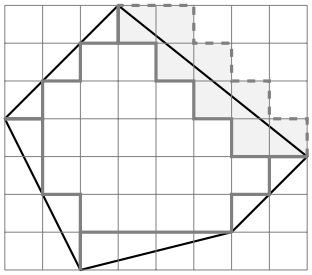


Figure 4: The circuit $S$ is obtained by replacing each edge of $P$ with a path in $K$ which is the boundary of the grid square polygon formed by taking the union of grid squares that $E$ passes through.

we may assume is the boundary of the convex hull of $K$. Then for each edge $E$ of $P$, we modify the loop in the following way. The two ends of $E$ are lattice points, and we consider the union of grid squares that $E$ passes through. This union is some polygon with 180° rotational symmetry (probably non-convex), and $E$ cuts it into two halves.

We claim that one of the rotationally symmetric halves of the boundary of this grid-square polygon is completely contained in the graph $K$. To see this, suppose to the contrary that another edge $F$ of $P$ also passes through a grid square that $E$ passes through. Then we can draw a line segment from one point on $E$ to one point on $F$ in the interior of this grid square, and this line segment cuts $P$ into two pieces in a way that separates the two vertices of $E$ but does not intersect the graph $K$. This contradicts the assumption that $K$ is connected.

Thus, one half of the boundary of the grid-square polygon determined by $E$ is a path in $K$, and in our loop $P$, we may replace $E$ by this path in $K$. By doing these replacements on all edges of $P$, we obtain a circuit in $K$, and we let $S$ be the set of all vertices and edges in this circuit. $S$ is connected, by following the circuit.

For each row of vertical edges in $K$, our subgraph $S$ contains only the leftmost and rightmost, and for each column of horizontal edges in $K$, our subgraph $S$ contains only the topmost and bottommost. Thus the circuit traverses $2h(K)$ vertical edges and $2w(K)$ horizontal edges, and the same number of vertices, so

there are at most $2(w(K) + h(K))$ vertices in $S$.

Suppose that $G$ is a connected lattice graph containing $K$, and let $H = G \backslash K$. Consider any edge $e$ from $H$ to $K$. We claim that its vertex $v$ in $K$ is in $S$. In the case where $v$ is on the boundary of the polygon $P$, we know that $v$ is in $S$ because our replacement process to transform from the boundary loop of $P$ to the circuit $S$ does not touch the vertices on the loop. Otherwise, the edge $e$ crosses from inside $P$ to outside $P$, so it crosses an edge $E$ of $P$. Then $E$ crosses the two grid squares containing $e$, and so the vertex $v$ on their boundary is part of $S$. Then, to show that $S \cup H$ is connected, consider any path in $G$ between two vertices of $S \cup H$. It alternates between sequences of vertices in $H$ and sequences of vertices in $K$, and we have just shown that each $K$ sequence begins and ends with vertices in $S$. Thus, we may replace each $K$ sequence by an $S$ sequence to get a walk in $S \cup H$ connecting the same two vertices. □

## 3. Ramp-like polygons

The purpose of this section is to prove Theorem 11, which bounds the routing number of ramp-like polygons. Subsection 3.1 contains all of the proof except for two big lemmas, which we save for their own subsections: the monotonic configuration theorem (Theorem 9) is proved in Subsection 3.2, and the column preparation lemma (Lemma 10) is proved in Subsection 3.3.

### *3.1. Ramp-like routing overview*

When routing within ramp-like polygons, the way we use the ramp-like geometry is by defining a slightly more general property of the induced graph, and then using that graph property in the routing. The following definition is illustrated by Figure 5.

**Definition 1.** A ***ramp-like graph*** $R$ is a finite induced subgraph of the infinite lattice graph $\mathbb{Z} \times \mathbb{Z}$ with the following properties:

1. Rows are contiguous, and start at $x$-coordinate 0: if $(x, y) \in V(R)$, then $(i, y) \in V(R)$ for $0 \leq i \leq x$. When we give row numbers, we number

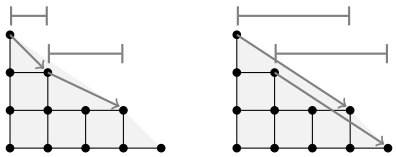


Figure 5: A ramp-like graph is shown above. The convex hull of its vertices, in gray, is a ramp-like polygon. The discretely convex border property says that the differences between row lengths roughly increase from bottom to top, failing to increase by at most 1.

the rows by their $y$-coordinates, so that the row numbers increase from bottom to top rather than from top to bottom.

2. Columns are contiguous, and start at $y$-coordinate 0: if $(x, y) \in V(R)$, then $(x, i) \in V(R)$ for $0 \leq i \leq y$. When we give column numbers, we number the columns by their $x$-coordinates.
3. $R$ has discretely convex border: if $n_i$ denotes the greatest $x$-coordinate among all vertices with $y$-coordinate $i$, then for all $i > j \geq c > 0$, we have $n_{i-c} - n_i \geq n_{j-c} - n_j - 1$.

The intuition behind the third property is to think of $R$ as being cut out from the first quadrant by the sideways graph $(f(y), y)$ of a function $f$. If $f$ cuts out a convex shape, then for any constant $c$, the function $f(y-c) - f(y)$ is increasing in $y$. However, the resulting lattice graph only satisfies the inequality with the term of $-1$ included, as in Figure 5.

**Lemma 7.** *The convex grid piece $R$ cut out by any ramp-like polygon $P$ is isomorphic to a ramp-like graph.*

*Proof.* We may assume that $P$ is the convex hull of its enclosed lattice points. By translating $P$ and rotating by some multiple of a right angle, we may assume that the two sides of $P$ that coincide with sides of its bounding box are along the positive $x$-axis and the positive $y$-axis. Then the rows and columns are contiguous and start at 0, so it remains to check the third property, about having discretely convex border.

We consider the rightmost vertices in the rows of $R$ with the $y$-coordinates $i, i-c, j$, and $j-c$, with $i > j \geq c > 0$. They are $(n_i, i), (n_{i-c}, i-c), (n_j, j)$,

and $(n_{j-c}, j-c)$, respectively.

We note that by convexity of the ramp-like polygon $P$, the convex hull of $(n_i, i)$, $(n_{j-c}, j-c)$, $(n_i, 0)$, $(n_{j-c}, 0)$ is contained in $P$. We know that $(n_{i-c}+1, i-c)$ and $(n_j+1, j)$ are not in $P$, so they also must not be in the convex hull of $(n_i, i)$, $(n_{j-c}, j-c)$, $(n_i, 0)$, $(n_{j-c}, 0)$. Thus, the lattice points $(n_{i-c}+1, i-c)$ and $(n_j+1, j)$ must be to the right of the line segment between $(n_i, i)$ and $(n_{j-c}, j-c)$. We let $s$ be the slope of this segment, and compare this slope to the slopes of the segments from $(n_i, i)$ to $(n_{i-c}+1, i-c)$ and from $(n_j+1, j)$ to $(n_{j-c}, j-c)$, as in Figure 6.

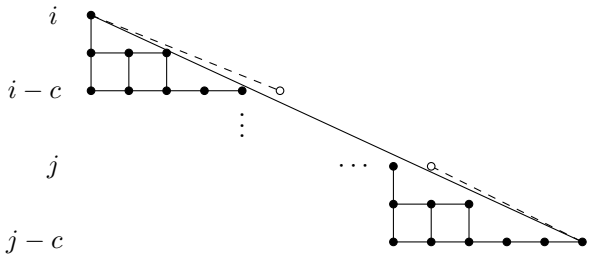



Figure 6: The slope of the solid line, $s$ must be steeper (more negative) than the slope of the upper dashed line, $\frac{-c}{(n_{i-c}+1)-n_i}$, and shallower (less negative) than the slope of the lower dashed line, $\frac{-c}{n_{j-c}-(n_j+1)}$.

We note that if $s$ is undefined, then each row between $i$ and $j-c$ must contain $n_i = n_{j-c}$ vertices, so the inequality must be true. Otherwise, the slope of the segment from $(n_i, i)$ to $(n_{i-c}+1, i-c)$ is $\frac{-c}{(n_{i-c}+1)-n_i}$, and the slope of the segment from $(n_j+1), j)$ to $(n_{j-c}, j-c)$ is $\frac{-c}{n_{j-c}-(n_j+1)}$, and the three negative slopes are ordered as

$$\frac{-c}{n_{j-c}-(n_j+1)} < s < \frac{-c}{(n_{i-c}+1)-n_i}.$$

Taking absolute values and comparing the denominators, we have

$$n_{j-c}-(n_j+1) < (n_{i-c}+1)-n_i,$$

and since both sides are integers, this is equivalent to our desired inequality

$$n_{i-c}-n_i \geq n_{j-c}-n_j-1.$$

□

The main goal of this section is to prove the following bound on the unlabeled routing number of ramp-like graphs. This bound implies the corresponding bound for labeled routing number relatively easily.

**Theorem 8** (Unlabeled ramp-like bound)**.** *There is a constant $C$ such that for any ramp-like graph $R$, the unlabeled routing number of $R$ satisfies the bound*

$$\mathrm{urt}(R) \leq C \cdot (w(R) + h(R)),$$

*where $w(R)$ and $h(R)$ denote the width and height of $R$ as a lattice graph.*

Given any $k$, we define the ***row-major order*** configuration of $k$ black tokens and $|V(R)| - k$ white tokens on the graph $R$ as follows: we order the vertices in $R$ by row from top to bottom, and within each row from left to right, and we take the configuration in which all of the black tokens appear before all of the white tokens. To prove the unlabeled ramp-like bound, we start with an arbitrary configuration of black and white tokens on $R$, and describe how to route from this configuration to row-major order.

The process for arbitrary ramp-like graphs is considerably more complicated than it is for rectangular grids. On a rectangular grid, given a configuration of black and white tokens, we can move them within their rows to get the right number of each color into each column, and then move them within their columns to achieve row-major order. An arbitrary ramp-like graph may be much narrower at the top than at the bottom, so our starting configuration could be a few wide rows of black tokens along the bottom, which we want to move to form several narrow rows at the top. In this case we would need to alternate between horizontal and vertical motion several times to move between the configurations. Whatever the starting configuration is, our first phase is to route to what we call a monotonic configuration.

**Definition 2.** We say a configuration of tokens on a ramp-like graph $R$ is ***left-aligned*** if for every black token on some $(x, y) \in V(R)$, there is also a

black token on $(x-1, y) \in V(R)$, or $(x-1, y) \notin V(R)$. Similarly, we say a configuration of tokens is ***up-aligned*** if for every black token on some $(x, y) \in V(R)$, either there is also a black token on $(x, y+1) \in V(R)$, or $(x, y+1) \notin V(R)$. Additionally, we say a given black token is left-aligned or up-aligned if it satisfies the corresponding condition stated above.

We say a configuration is ***monotonic*** if it is both left-aligned and up-aligned. Examples of configurations with each of these properties can be seen in Figure 7.

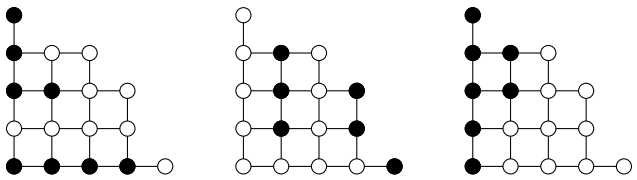


Figure 7: (a) A left-aligned configuration of tokens, (b) an up-aligned configuration of tokens, and (c) a monotonic configuration of tokens.

We note that routing all black tokens such that they are as far up or left as possible within their column or row results in an up-aligned or left-aligned configuration, respectively. We abbreviate the process of routing all black tokens as far up as possible (within their column) with the phrase "pushing up," and routing all black tokens as far left as possible (within their row) with "pushing left."

The following theorem describes a process sufficient for moving the tokens to a monotonic configuration.

**Theorem 9** (Monotonic configuration)**.** *For any configuration $X_0$ of black and white tokens on a ramp-like graph, after pushing the black tokens up, then left, then up, then left, the new configuration of tokens is monotonic.*

Proving the monotonic configuration theorem (Theorem 9) is the most technical aspect of routing within ramp-like polygons, and the proof appears in Subsection 3.2. After routing to a monotonic configuration, we are in a better position to move the tokens to the correct columns, after which all that remains is to push up all the black tokens to get to row major order. Our proof that we

can move the tokens to the correct columns sufficiently quickly is by induction on the number of rows in our graph. To do this, we want to partition the graph into a top slice (denoted by $R_{m_1-1}$ in the next lemma) and a bottom slice and route the two slices separately; that is, we want a way to move the tokens to the correct columns without crossing between the two slices. The next lemma states that this division into top and bottom is possible. For the lemma statement, we denote the number of vertices in the $y = i$ row by $1 + n_i$, so that as before the $x$-coordinates of those vertices range from 0 to $n_i$.

**Lemma 10** (Column preparation)**.** *Let $R$ be a ramp-like graph with $m$ rows (that is, from $y = 0$ through $y = m-1$), and let $R_i$ denote the induced subgraph of $R$ consisting of the top $i$ rows (that is, from $y = m - i$ through $y = m - 1$). Let $X$ be a monotonic configuration on $R$ with $t$ black tokens, and let $m_1$ be such that*

$$\sum_{i=1}^{m_1-1}(1 + n_{m-i}) < t \leq \sum_{i=1}^{m_1}(1 + n_{m-i}),$$

*meaning that $t$ black tokens can fit on the vertices of $R_{m_1}$, but are not able to fit on the vertices of $R_{m_1-1}$. Then, there exists a configuration $Y$ on $R$ such that*

1. *$R_{m_1-1}$ contains the same number of tokens of each color in $Y$ as in $X$, and*
2. *In $Y$, each column of $R$ contains the same number of tokens of each color as there are in row major order.*

The proof of the column preparation lemma (Lemma 10) is also fairly technical, and it appears in Subsection 3.3. Outside of the proof of the lemma, we do not need to remember the definition of $m_1$; rather, what is important is the conclusion of the lemma, which makes $R_{m_1-1}$ the top slice and the remainder of $R$ the bottom slice. It would be more intuitive to choose the top slice to be $R_{m_1}$, but then in the case where $R_{m_1}$ is the whole graph $R$, we would not be able to apply the inductive hypothesis to it. Thus, we choose the top slice to be $R_{m_1-1}$. Assuming the monotonic configuration theorem (Theorem 9) and the column preparation lemma (Lemma 10), we can finish proving the upper bound on unlabeled routing number of ramp-like graphs.

*Proof of unlabeled ramp-like bound (Theorem 8).* Let $m$ be the number of rows in our ramp-like graph $R$, and let $\mu$ be the number of columns. If $m$ or $\mu$ is 1, then $R$ is a path, which we already know how to route by the path bound (Theorem 2). Thus, we may assume that $w(R)$ and $h(R)$ are both at least 1. In this case, we have $\mu = w(R) + 1 \leq 2w(R)$ and $m = h(R) + 1 \leq 2h(R)$, so it suffices to find a constant $C$ such that we can move an arbitrary configuration of black and white tokens to row-major order in at most $C(m+\mu)$ routing steps.

Let $C_0$ be the constant from the rectangle bound (Theorem 4) such that a rectangular grid with $p$ rows and $q$ columns has unlabeled routing number at most $C_0(p+q)$.

First we show by induction on $m$ that the process in the column preparation lemma (Lemma 10) of moving from configuration $X$ to configuration $Y$ can be accomplished in at most $C_0(m+\mu)$ moves. In the base case $m = 1$, our ramp-like graph is a path of length $\mu$, which is already rectangular. For $m > 1$, we partition the ramp-like graph $R$ into three parts: the upper ramp-like shape $R_{m_1-1}$, the rectangular piece containing the row below $R_{m_1-1}$ and all vertices directly below this row, and the ramp-like graph consisting of columns to the right of the rectangular piece (Figure 8).

To move from configuration $X$ to configuration $Y$, we apply the inductive hypothesis to $R_{m_1-1}$ while simultaneously routing within the rectangular piece; the graph to the right of the rectangular piece has no tokens in it, so it does not need any routing. Because $R_{m_1-1}$ has strictly fewer rows than the original graph $R$, the inductive hypothesis applies. The rectangular piece has at most $m$ rows and at most $\mu$ columns, so it requires at most $C_0(m+\mu)$ routing steps to move from $X$ to $Y$, and by the inductive hypothesis, $R_{m_1-1}$ also requires at most $C_0(m+\mu)$ routing steps. Performing the steps simultaneously completes the induction.

Given an arbitrary token configuration $X_0$ on the ramp-like graph $R$, we start by pushing all the tokens up, then left, then up, then left, which by the monotonic configuration theorem (Theorem 9) gives a monotonic configuration. We apply the column preparation lemma (Lemma 10) to the resulting monotonic

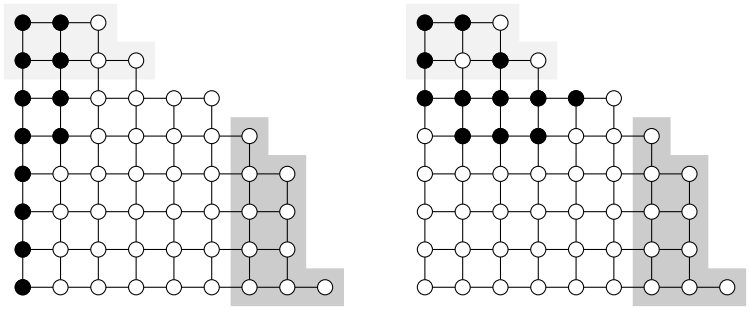



Figure 8: $R$ is partitioned into $R_{m_1-1}$ (light gray), a rectangle, and a subgraph to the right of the rectangle (dark gray). $R_{m_1-1}$ is routed using the inductive hypothesis, and in parallel, the rectangular piece is routed to produce a configuration, which, when pushed up, is in row major order. The subgraph to the right of the rectangle contains no black tokens either in the monotonic configuration $X$ or in row-major order.

configuration, and then we push all the tokens up to get to row-major order. By the path bound (Theorem 2) it takes at most $m$ routing steps to push up and at most $\mu$ routing steps to push left, so the total number of routing steps to get to row-major order is at most $3m + 2\mu + C_0(m + \mu) \leq (3 + C_0)(m + \mu)$.

Given any two configurations of the same set of white and black tokens on $R$, we can route from one to the other by routing the first into row-major order, and then routing from row-major order to the second. Thus, if we choose $C = 4(3+C_0)$, we can route between the two configurations in $2(3+C_0)(m+\mu) \leq C(w(R) + h(R))$ routing steps. □

We can use the bound on unlabeled routing number of ramp-like graphs to give a bound on labeled routing number, by dividing the graph into four quadrants and applying recursion.

**Theorem 11** (Labeled ramp-like bound)**.** *There is a constant $C$ such that for any ramp-like polygon $P$, the routing number of $P$ satisfies the bound*

$$\mathrm{rt}(P) \leq C \cdot (w(P) + h(P)).$$

*Proof.* The idea of the proof is to divide $P$ into quadrants, each with width at

most $\frac{1}{2}w(P)$ and height at most $\frac{1}{2}h(P)$. Using the unlabeled ramp-like bound (Theorem 8), we can move each token into the correct quadrant. Then, each quadrant is ramp-like, so we can apply recursion to route within all four quadrants simultaneously, as shown in Figure 9.

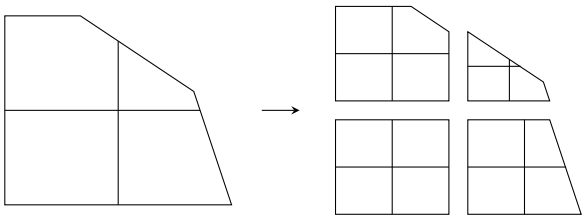


Figure 9: $P$ is divided into four quadrants, and all tokens are routed into the correct quadrant using unlabeled routing twice. Then, this process is repeated recursively for each quadrant, simultaneously.

We may assume that the vertical and horizontal sides that $P$ shares with its bounding box have rational $x$-coordinate and $y$-coordinate, respectively, and that the width and height of $P$ are both irrational. These assumptions guarantee that when we cut $P$ in half, the cut does not go through any lattice points. If $P$ does not have these properties already, we can make $P$ very slightly bigger so that it does, which increases the right-hand side of the desired inequality very slightly; taking the limit of a shrinking sequence of approximations gives the desired inequality.

Let $C_0$ be the constant for unlabeled routing from the unlabeled ramp-like bound (Theorem 8). We divide $P$ into halves with a vertical line bisecting the width, and then into quadrants with a horizontal line bisecting the height. One of the quadrants may be empty. It takes at most $C_0(w(P) + h(P))$ routing steps to move the tokens so that those that belong in the left half go to the left half, and those that belong to the right half go to the right half. Then each half is ramp-like and has width $\frac{1}{2}w(P)$ and height $h(P)$, so it takes at most $C_0\left(\frac{1}{2}w(P) + h(P)\right)$ additional routing steps to move the tokens into the quadrants where they belong, working with both halves simultaneously.

We select $C = 4C_0$, and prove the theorem by induction on $\lceil w(P)+h(P)\rceil$. If

$\lceil w(P)+h(P)\rceil = 1$, then $P$ has only one vertex, so the routing number is 0, which is certainly at most $C \cdot (w(P)+h(P))$. Otherwise, we have $\lceil \frac{1}{2}w(P)+\frac{1}{2}h(P)\rceil < \lceil w(P)+h(P)\rceil$, so we may apply the inductive hypothesis to find that the routing number of each quadrant of $P$ is at most $C \cdot \left(\frac{1}{2}w(P)+\frac{1}{2}h(P)\right)$. Then the total number of steps to route an arbitrary configuration of tokens on $P$ to a home configuration is at most

$$\begin{aligned} C_0(w(P)+h(P)) + C_0\left(\frac{1}{2}w(P)+h(P)\right) + C\left(\frac{1}{2}w(P)+\frac{1}{2}h(P)\right) \\ \leq \left(2C_0+\frac{1}{2}C\right)\cdot(w(P)+h(P)) \\ = C\cdot(w(P)+h(P)). \end{aligned}$$

□

*3.2. Moving to a monotonic configuration*

In this subsection we prove the monotonic configuration theorem (Theorem 9). To understand the strategy, we observe that if our ramp-like graph were a rectangular grid, then pushing the black tokens up and then left would already give a monotonic configuration. This is because after pushing up from an arbitrary configuration, every black token below the top row has another black token directly above it, so the number of tokens in each row is non-increasing as we consider the rows from top to bottom. However, for an arbitrary ramp-like graph, after pushing up from an arbitrary configuration, a row might have more black tokens than the row above it, because the lower row might have black tokens in columns to the right of all columns in the upper row.

The proof of the theorem is based on Lemmas 12 and 13, which together show that after pushing up, left, and up on a ramp-like graph, the result is similar to what we would get from simply pushing up on a rectangular grid. Namely, if we consider the rightmost column that contains black tokens, then we show that all of the rows that are too short to extend to that column are completely full of black tokens. The configuration below these rows looks like

an up-aligned configuration on a rectangular grid, so pushing left one more time results in a monotonic configuration.

Lemma 12 shows how the shape of our ramp-like graph affects the possible numbers of black tokens per row after our first step of pushing up, and thus after our second step of pushing left as well. Then Lemma 13 describes the result of our third step of pushing up. For both lemmas, we use the following notation. Given a ramp-like graph $R$, we let $\#(R, i)$ denote the number of vertices in the row of $R$ with $y$-coordinate $i$. Given a configuration $X$ on $R$, we let $\#(X, i)$ denote the number of black tokens in $X$ with $y$-coordinate $i$.

**Lemma 12.** *Let $X_1$ be an up-aligned configuration on a ramp-like graph $R$. Then for all rows $b$ and $b + c > b$ of $R$, we have*

$$\#(X_1, b) - \#(X_1, b + c) \leq \#(R, b) - \#(R, b + c).$$

*Proof.* Let $d$ be the number of black tokens in row $b$ of $X_1$ that are to the right of column $n_{b+c}$, the rightmost column of row $b + c$. Using the fact that $X_1$ is up-aligned, we have

$$\#(X_1, b) - \#(X_1, b + c) \leq d,$$

because every token in row $b$ has a token directly above it in row $b+c$, except for those in the $d$ columns to the right of column $n_{b+c}$. Because the total number of columns in row $b$ to the right of $n_{b+c}$ is $\#(R, b) - \#(R, b + c)$, we also have

$$d \leq \#(R, b) - \#(R, b + c).$$

Together, these inequalities give the desired inequality

$$\#(X_1, b) - \#(X_1, b + c) \leq \#(R, b) - \#(R, b + c).$$

□

Geometrically, the lemma says that after we push up to form $X_1$ and then left to form a configuration $X_2$, the right boundary of the cluster of black tokens is steeper than the right boundary of the graph $R$. The next lemma starts with this configuration $X_2$ that results from pushing up and left, and describes what happens after pushing up again.

**Lemma 13.** *Let $X_2$ be the left-aligned configuration on a ramp-like graph $R$ which was produced by starting from some up-aligned configuration and pushing left. Let $(x_2, b)$ be the coordinates of $\tau$, a black token in $X_2$, and let $r > b$ be a row number such that $n_r < x_2$, if such a row exists. Then when we push up to reach another configuration $X_3$, all rows with y-coordinate at least $r$ will have only black tokens.*

*Proof.* We imagine translating all of the tokens upward in the lattice so that row $b$ moves up to row $r$ and some tokens may occupy lattice points that are not in the graph $R$. If all rows $r$ and above are covered by black tokens in this arrangement, they also have only black tokens in $X_3$. Thus, it suffices to show that for all $c \geq 0$, the number of black tokens in row $b + c$ of $X_2$ is at least the number of vertices in row $r + c$.

To show this, we make the following note: from Lemma 12, for all rows $b$ and $b + c > b$ of $R$ in configuration $X_1$, we have

$$\#(X_1, b) - \#(X_1, b + c) \leq \#(R, b) - \#(R, b + c).$$

Because when we push left to get to $X_2$, every token is in the same row in $X_2$ as in $X_1$, so the same inequality is true of $X_2$.

We then combine this inequality with an inequality resulting from the discretely convex border property of $R$: because $r > b$, we have

$$n_r - n_{r+c} \geq n_b - n_{b+c} - 1,$$

or equivalently,

$$\#(R, r) - \#(R, r + c) + 1 \geq \#(R, b) - \#(R, b + c).$$

Putting these inequalities together with the assumption $n_r < x_2$ (or equivalently $\#(R, r) < \#(X_2, b)$), we have

$$\begin{aligned}
\#(X_2, b + c) &= \#(X_2, b) - [\#(X_2, b) - \#(X_2, b + c)] \\
&> \#(R, r) - [\#(R, b) - \#(R, b + c)] \\
&\geq \#(R, r) - [\#(R, r) - \#(R, r + c) + 1] \\
&= \#(R, r + c) - 1,
\end{aligned}$$

and so because all of the quantities are integers, we obtain our desired inequality $\#(X_2, b+c) \geq \#(R, r+c)$. □

Having described the configuration that results from pushing up, left, and up, we are ready to prove that pushing this configuration left results in a monotonic configuration.

*Proof of monotonic configuration theorem (Theorem 9).* We label the sequence of configurations as follows: let

- the starting configuration be $X_0$,
- the configuration after pushing up be $X_1$,
- the configuration after pushing left be $X_2$,
- the configuration after pushing up a second time be $X_3$, and
- the configuration after pushing left a second time be $X_4$, which is also the final configuration.

Note that $X_4$ is left-aligned, so in order to show $X_4$ is monotonic, we just need to show it is up-aligned. We will do this by showing that all black tokens in $X_4$ are up-aligned.

We consider an arbitrary black token $\tau$ in $X_0$. Let the vertex which $\tau$ is on in $X_2$ be $(x_2, b)$. Then by Lemma 13, we have that when we push up to get to $X_3$, all rows with $y$-coordinate at least $r$ will contain only black tokens.

Let $s$ be the row such that $\tau$ is in row $s-1$ in $X_3$. Because $\tau$ is in the same column $x_2$ in $X_2$ and $X_3$, and this column is to the right of $n_r$, we have $s-1 < r$, or in other words $s \leq r$. If $s = r$, then $\tau$ is up-aligned in $X_4$ because the row above $\tau$ is row $r$, and we have shown that rows $r$ and above are all black in $X_3$.

If $s < r$, then because of how $r$ is defined we have $x_2 \leq n_s$. Thus, in $X_3$ (which is up-aligned), for $\tau$ and every black token to the left of it in row $s-1$, there is a corresponding black token immediately above, in row $s$. When we

push left to get $X_4$, there are at least as many black tokens in row $s$ as there are black tokens in row $s-1$ to the left of and including $\tau$, so $\tau$ is up-aligned.

Thus, for any black token $\tau$ in $X_0$, in $X_4$, there is a black token (or no vertex) above it and a black token (or no vertex) to the left of it. □

### *3.3. Distributing tokens among columns*

Proving the column preparation lemma (Lemma 10) is the last piece needed to complete the proof of the unlabeled ramp-like bound (Theorem 8) and thus the labeled ramp-like bound (Theorem 11). The goal is to get the right number of tokens of each color into each column, without moving tokens between the top slice $R_{m_1-1}$ and the bottom slice $R \setminus R_{m_1-1}$. For the top slice, we choose to move the tokens to row-major order; this determines how many black tokens we want in each column of the bottom slice. The only thing that could potentially go wrong is if we have somehow assigned more black tokens to a column of the bottom slice than its number of vertices. We show this does not happen, roughly because the black tokens in the bottom slice are more evenly spaced, among at least as many columns, in our target configuration $Y$ as in our starting configuration $X$.

*Proof of column preparation lemma (Lemma 10).* Let $t_1$ be the number of black tokens in $R_{m_1-1}$ in configuration $X$, and let $m_0 \leq m_1 - 1$ be the number such that

$$\sum_{i=1}^{m_0-1} (1 + n_{m-i}) < t_1 \leq \sum_{i=1}^{m_0} (1 + n_{m-i}),$$

meaning that $t_1$ black tokens can fit on the vertices of $R_{m_0}$, but are not able to fit on the vertices of $R_{m_0-1}$.

On any ramp-like graph, we denote the configuration of $t$ black tokens (and the remainder white tokens) in row-major order by $RM(t)$. We set $Y$ to be equal to $RM(t_1)$ on $R_{m_1-1}$. Let $Z$ be the configuration of $t - t_1$ black tokens, one at every vertex where $RM(t)$ has a black token but $RM(t_1)$ does not, and let $z_j$ be the number of black tokens in $Z$ that are in the $x = j$ column of $R$.

We note that $z_j = 0$ for $j > n_{m_1}$. To prove the lemma, we need to show that $z_j$ black tokens can fit into column $j$ of $R \setminus R_{m_1-1}$; that is, $z_j \leq m - m_1 + 1$.

First we address the case where $X$ contains a black token in a column strictly to the right of the subgraph $R_{m_0}$. We claim that in this case, $z_j \leq 1 \leq m-m_1+1$ for all $j$, so we are done. Because $X$ is monotonic, if $X$ contains a black token to the right of $R_{m_0}$, then $R_{m_0}$ must be entirely full of black tokens, and the row below it must also contain black tokens. Thus, $R_{m_0+1}$ contains more than $t_1$ black tokens. Because $R_{m_1-1}$ contains only $t_1$ black tokens, this then implies that $R_{m_1-1}$ must be smaller than $R_{m_0+1}$. Because we always have $m_0 \leq m_1 - 1$, we conclude that in this case we have $m_1 - 1 = m_0$, with $R_{m_1-1}$ entirely full of black tokens. The definition of $m_1$ implies that the black tokens not in $R_{m_1-1}$ all fit into the row just below $R_{m_1-1}$, so $z_j \leq 1$.

Thus, we may assume that we are in the case where all black tokens in $X$ are in columns 0 through $n_{m_0}$. In this case, the idea of the proof is that if we were to distribute the tokens in the bottom slice $R \setminus R_{m_1-1}$ as evenly as possible among the columns 0 through $n_{m_0}$, the column with the most black tokens would have at least as many as in $Z$, because $Z$ may use the columns to the right of $n_{m_0}$ as well.

More precisely, the portion of $X$ in $R \setminus R_{m_1-1}$ has $t - t_1$ black tokens, all of which are in columns 0 through $n_{m_0}$, so we have

$$t - t_1 \leq (1 + n_{m_0})(m - m_1 + 1). \tag{1}$$

Also, because $Z$ has $t - t_1$ black tokens in total, we have

$$t - t_1 = \sum_{0 \leq j \leq m_1} z_j.$$

Let $\delta$ be the greatest $x$-coordinate of the black tokens in the $y = m - m_1$ row of $RM(t)$ (the $m_1$st row from the top, and the last not-all-white row), and let $\delta_1$ be the greatest $x$-coordinate of the black tokens in the $y = m - m_0$ row of $RM(t_1)$ (the $m_0$th row from the top, and the last not-all-white row). We now estimate $z_j$ in the three possible cases for how $\delta_1$ and $\delta$ compare, depicted in Figure 10.

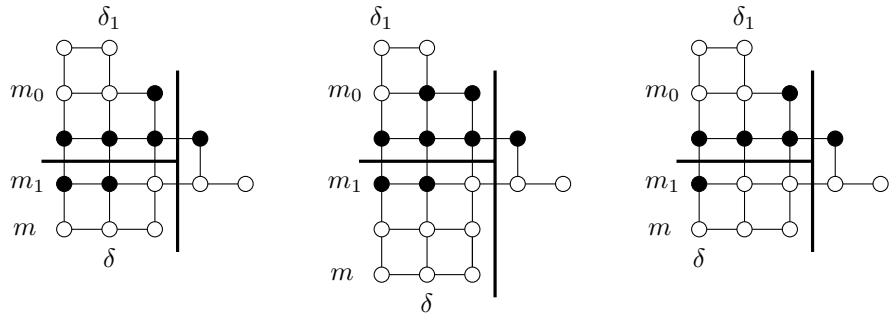



Figure 10: In every case, the black tokens of $Z = RM(t) \setminus RM(t_1)$ have enough vertical space to slide down below subgraph $R_{m_1-1}$, because the total area of columns 0 through $n_{m_0}$ in $R \setminus R_{m_1-1}$ is large enough for the black tokens in columns 0 through $n_{m_0}$ of $Z$.

If $\delta_1 = \delta$, or if $n_{m_0} = \delta_1 < \delta$, then every column $j$ with $j \leq n_{m_0}$ has $z_j = m_1 - m_0$. For $j > n_{m_0}$, because row $m - m_0$ has no vertices in column $j$, we have $z_j \leq m_1 - m_0$. The total number $t - t_1$ of black tokens in $Z$ is at least the number in columns 0 through $n_{m_0}$, so we have $t - t_1 \geq (1 + n_{m_0})(m_1 - m_0)$. Combining this inequality with inequality (1), we have

$$(1 + n_{m_0})(m_1 - m_0) \leq t - t_1 \leq (1 + n_{m_0})(m - m_1 + 1),$$

so $m_1 - m_0 \leq m - m_1 + 1$, and so $z_j \leq m - m_1 + 1$ for all $j$.

If $\delta_1 < \delta$ and $\delta_1 < n_{m_0}$, then for $j \leq \delta_1$ and for $\delta < j \leq n_{m_0}$ we have $z_j = m_1 - m_0$, and for $\delta_1 < j \leq \min(\delta, n_{m_0})$ we have $z_j = 1 + m_1 - m_0$. For $j > n_{m_0}$, we have $z_j \leq 1 + m_1 - m_0$ as well. Thus we have

$$(1 + n_{m_0})(m_1 - m_0) < t - t_1 \leq (1 + n_{m_0})(m - m_1 + 1),$$

which implies that $m_1 - m_0 < m - m_1 + 1$, and so $z_j \leq m - m_1 + 1$ for all $j$.

If, finally, $\delta_1 > \delta$, then for $j \leq \delta$ and for $\delta_1 < j \leq n_{m_0}$ we have $z_j = m_1 - m_0$, and for $\delta < j \leq \delta_1$ we have $z_j = -1 + m_1 - m_0$. For $j > n_{m_0}$, we have $z_j \leq m_1 - m_0$ as well. Thus we have

$$(1 + n_{m_0})(-1 + m_1 - m_0) < t - t_1 \leq (1 + n_{m_0})(m - m_1 + 1),$$

which implies that $-1 + m_1 - m_0 < m - m_1 + 1$, and so $z_j \leq m - m_1 + 1$ for all $j$.

Thus, in every case, $R \setminus R_{m_1-1}$ has enough space to fit the same number of black tokens in each column as $Z$. We set $Y$ to be any configuration that in $R_{m_1-1}$ has the same number of black tokens in each column as $RM(t_1)$, and that in $R \setminus R_{m_1-1}$ has $z_j$ black tokens in column $j$ for each $j$. □

## 4. Burger bun polygons

In this section we prove Theorem 18, the bound on routing number of burger bun polygons. Our strategy is to divide the burger bun polygon in half, then to divide each half into two ramp-like pieces. We know that we can route a single ramp-like piece from the ramp-like bound (Theorem 11), so our first step is to use this to prove that we can route a pair of ramp-like pieces and thus a half of a burger bun polygon. Then, using a similar argument, we show that this implies that we can route a whole burger bun polygon.

We have defined a ramp-like polygon to be a convex polygon such that two of its edges coincide with edges of its bounding box; we refer to each of these edges as a ***spine*** of the ramp-like polygon. If two otherwise disjoint ramp-like polygons have a common spine, then their union is also a convex polygon. Similarly, every burger bun polygon is divided in two by a spine. Specifically, if a burger bun has two points of maximum and minimum $y$-coordinate with equal $x$-coordinate, then we refer to the segment between those two points as its (vertical) spine, and if it has two points of maximum and minimum $x$-coordinate with equal $y$-coordinate, then we refer to the segment between those two points as its (horizontal) spine.

Our strategy for this section is as follows. In our situation, we have two polygons with a common spine, and we may assume that we know how to route within each of the two polygons. Given an arbitrary configuration of labeled tokens on the union of the polygons, we want to route those tokens to their home positions. It suffices to get each token into the half where it belongs, because

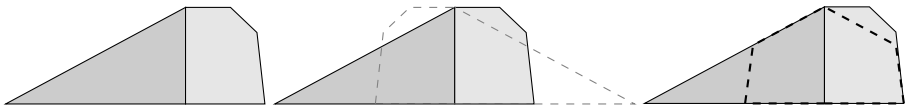


Figure 11: Left: The two ramp-like polygons shaded in different shades of gray share a spine, forming half of a burger bun polygon. Middle: To find the shared subpolygon of each half, we intersect it with the reflection (gray dashed) of the other half across the spine. Right: The shared subpolygons are mirror images (dashed line).

then we can route within the halves separately to get each token to its home vertex. Thus, we have an unlabeled routing problem, thinking of the tokens belonging in the first half as black, and the tokens belonging in the second half as white.

In the special case where the two polygons, and their corresponding graphs, are mirror images across the spine, we can route as follows: first we route within the second half so that the configuration is a color-reversed mirror image of the first half. This is possible because when the two halves have the same number of vertices, the number of black tokens in the second half is equal to the number of white tokens in the first half. Once the two halves are color-reversed mirror images, each row (if the spine is vertical) has the same number of white tokens as black tokens, so we may route all rows simultaneously to get each token into the half where it belongs.

In the general case, where the two polygons are two ramp-like polygons or two halves of a burger bun polygon, the polygons may not be mirror images. Instead, we find a subpolygon inside each side, called the shared subpolygon, such that the two shared subpolygons are mirror images and still contain a significant fraction of the vertices. Informally, to find the shared subpolygons, we imagine reflecting either polygon over the spine and intersecting the other polygon with the reflection, as in Figure 11. This construction only makes sense when the spine has an integer or half-integer coordinate, so for the general case of two ramp-like polygons sharing an arbitrary spine, the following theorem finds those shared subpolygons and shows that they contain enough vertices.

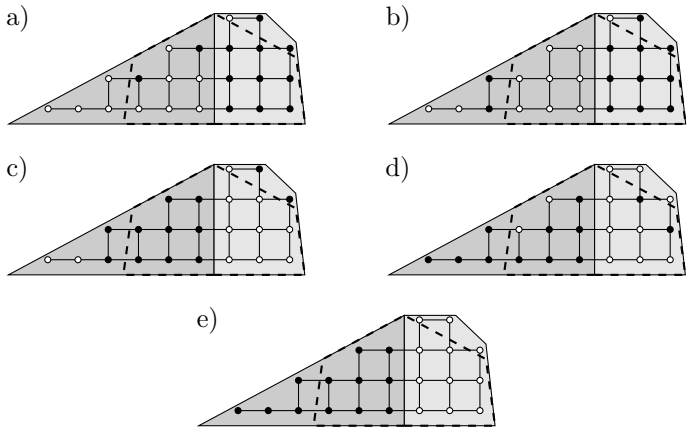



Figure 12: The tokens belonging in the left ramp-like polygon (shaded in darker gray) are colored black, and in the right ramp-like polygon (shaded in lighter gray) are colored white. The union of the shared polygons of each side is marked with a dashed line. Part a) shows the starting configuration. In b), as many improper tokens as possible, one for each vertex, are routed into the union of the shared polygons. Note that the tokens in this union are color-reversed mirror images of each other along the spine. Along each horizontal path of vertices, the improper tokens are routed to the proper half to produce c). Then, once again, as many improper tokens as possible are routed into the union, such that the tokens in it are color-reversed mirror images of each other, shown in d). In e), again, along each horizontal path of vertices, the improper tokens are routed to the proper half, completing the routing process.

We consider both the case where the ramp-like polygons together form half of a burger bun (that is, their non-shared spines are collinear) and the case where they do not, because this latter case turns out to be useful in the next part of the proof, where the two polygons are halves of a burger bun. Once we have found the shared subpolygons, then we can solve the unlabeled routing problem by repeatedly applying the mirror-image technique to these shared subpolygons; Figure 12 illustrates the routing process, which is presented in more detail in the routing between ramp-like theorem (Theorem 17).

**Theorem 14** (Intersection magnitude)**.** *Let $P_1$ and $P_2$ be two ramp-like polygons with common vertical spine $E$, such that the widths of $P_1$ and $P_2$ and the*

*length of $E$ are all at least 41. Then there exist subgraphs $G_1$ and $G_2$ of $P_1$ and $P_2$, respectively, both disjoint from $E$, such that $G_2$ is a reflection of $G_1$ over some vertical line, and the equal number of vertices in $G_1$ or $G_2$ is at least $\frac{1}{20}\min\{|P_1|, |P_2|\}$, where $|P_i|$ denotes the number of lattice points in the interior and boundary of $P_i$.*

The proof of this theorem is based on two lemmas: the spine alignment lemma (Lemma 15), and the triangle trimming lemma (Lemma 16). Roughly, the idea is that to find the subgraphs $G_1$ and $G_2$, we should reflect $P_1$ over the shared spine and intersect it with $P_2$ to find $G_2$, or reflect $P_2$ over the shared spine and intersect it with $P_1$ to find $G_1$. The spine alignment lemma (Lemma 15) accounts for the fact that the shared spine might not be at an integer or half-integer coordinate, so reflecting across it might not take lattice points to lattice points. Then the triangle trimming lemma (Lemma 16) starts from a quick estimate of the area of the polygon intersection, and produces an estimate of the number of lattice points inside that polygon.

**Lemma 15** (Spine alignment)**.** *Let $P$ be a burger bun polygon with vertical spine $E$ dividing $P$ into left side $P_1$ and right side $P_2$. Suppose that the $x$-coordinate of $E$ is not an integer. Then there is another burger bun $P'$ with vertical spine $E'$, dividing $P'$ into left side $P_1'$ and right side $P_2'$, with the following properties:*

- *The $x$-coordinate of $E'$ is an integer.*
- *$P_1$ and $P_1'$ contain the same lattice points.*
- *The set of lattice points inside $P_2'$ is obtained by translating the set of lattice points inside $P_2$ one unit to the right.*

*Proof.* Figure 13 shows the relationship between the polygons. Let $P_1$ and $P_2$ be arbitrary convex polygons sharing a vertical edge $E$ with a non-integer $x$-coordinate $x_0$. Construct a new vertical line segment $E'$ with the same length as $E$ and integer $x$-coordinate $\lceil x_0 \rceil$. Translate all polygon vertices of $P_2$ one unit to the right, and take the convex hull of these vertices with the endpoints of

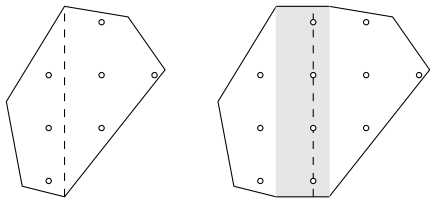


Figure 13: If we cut a burger bun polygon along its spine and insert a rectangle of width 1 instead, the result is a burger bun polygon with one additional column and a spine along that column.

$E'$ to form congruent polygon $P_2'$. Similarly, take the convex hull of all polygon vertices of $P_1$ with the endpoints of $E'$ to form congruent polygon $P_1'$. The set of lattice points inside $P_1'$ disjoint from $E'$ is equal to the set of lattice points inside $P_1$, and the set of lattice points inside $P_2'$ disjoint from $E'$ is equal to the set of lattice points inside $P_2$ translated by 1 in the positive $x$-direction. □

To prove the intersection magnitude theorem (Theorem 14), we find a triangle in each ramp-like piece that covers at least half the area, then intersect these triangles to get a smaller triangle that covers at least $\frac{1}{4}$ of the area of the smaller ramp-like piece. Once we have this triangle in common, we need to show that it has sufficiently many lattice points. The following lemma estimates the number of lattice points in such a triangle.

**Lemma 16** (Triangle trimming)**.** *Let $P$ be a triangle with at least one side parallel to an axis. Then the number of lattice points strictly inside $P$ is at least*

$$\text{Area}(P) - 2 \cdot \text{Perimeter}(P) + 1,$$

*if this quantity is at least* 1*.*

*Proof.* The strategy is to use Pick's theorem, which relates the area of a lattice triangle to the number of enclosed lattice points and the number of boundary lattice points. Our triangle $P$ does not necessarily have vertices at lattice points, so our goal is to find a large enough lattice triangle inside $P$. First we construct a parallel line 2 units inward from each side of $P$. We call the similar triangle defined by these parallel lines the "trimmed triangle", denoted $P_t$. We show

below that $P_t$ has area at least $\text{Area}(P) - 2 \cdot \text{Perimeter}(P)$. (In the case that there is no triangle left after the trimming process, we show that $\text{Area}(P) - 2 \cdot \text{Perimeter}(P) < 0$ and so the lemma is vacuously true.) If we can find an "intermediate triangle" $P_i$ that is a lattice triangle and is strictly between $P_t$ and $P$, then Pick's theorem states

$$\text{Area}(P_i) = \#(\text{interior lattice points}) + \frac{1}{2} \cdot \#(\text{boundary lattice points}) - 1,$$

so we have

$$\#(\text{total lattice points of } P_i) \geq \text{Area}(P_i) + 1,$$

giving our goal inequality

$$\#(\text{interior lattice points of } P) \geq \text{Area}(P_t) + 1 \geq \text{Area}(P) - 2 \cdot \text{Perimeter}(P) + 1.$$

We begin by estimating the area of $P_t$. The region inside $P$ and outside $P_t$ consists of three trapezoids, each with height 2 and one base a side of $P$. Because the two angles bordering that side add up to less than 180°, the other base of each trapezoid—that is, the corresponding side of $P_t$—must be shorter. Thus, the total area of the trapezoids is less than $2 \cdot \text{Perimeter}(P)$, giving the estimate

$$\text{Area}(P_t) \geq \text{Area}(P) - 2 \cdot \text{Perimeter}(P).$$

Suppose there is no triangle left after the trimming. Then the inradius $r$ of $P$ is at most 2, and connecting the vertices of $P$ to the incenter divides $P$ into three triangles, each with height $r$ and base equal to one side of $P$. Thus we have

$$\text{Area}(P) = \frac{r}{2} \cdot \text{Perimeter}(P) < 2 \cdot \text{Perimeter}(P),$$

and so the quantity $\text{Area}(P) - 2 \cdot \text{Perimeter}(P)$ is negative.

650 At each corner of the triangle, there is a parallelogram enclosed by the two sides of the triangle and the lines parallel to each side at distance 2, as in Figure 14. It suffices to find a lattice point inside each of these corner parallelograms, because these three points determine a triangle for which none of the sides crosses either a side of the original triangle $P$ or a side of the trimmed

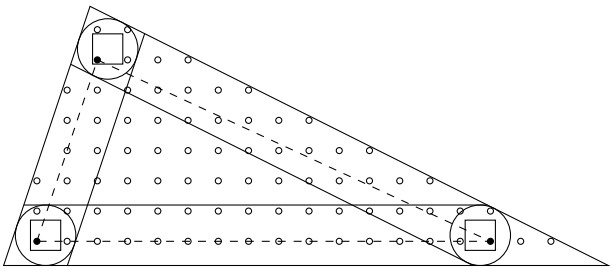


Figure 14: The parallelogram trimmed off near each vertex of the original triangle contains a lattice point, because it contains a circle of radius 1, which contains a square of side length 1. The lattice points from the three parallelograms form the intermediate triangle $P_i$ (dashed).

triangle $P_t$; thus, we can choose that triangle to be our intermediate triangle $P_i$.

To find the lattice point, first we observe that there is an inscribed circle of radius 1 inside each corner parallelogram; this is because the parallelogram is the intersection of two infinite strips of width 2, and the center lines of the two strips intersect at the center of the circle. We also know that every square with sides parallel to the axes and of length 1 must contain a lattice point, because tiling the plane with such squares gives lattice points at the same relative locations in each square. Any circle of radius 1 contains such a square of side length 1—in fact, it contains a square of side length $\sqrt{2}$, because the diagonal has the same length 2 as the diameter of the circle.

Thus every corner parallelogram does contain a lattice point in its interior, so we can select one such lattice point from each corner parallelogram to define the intermediate triangle $P_i$. Because the area of $P_i$ is greater than that of the trimmed triangle $P_t$, Pick's theorem implies that $P_i$ must have enough lattice points in its interior and boundary. □

Using these lemmas, we can finish proving that our pair of ramp-like polygons contains a pair of subgraphs, one on each side of the spine, that are mirror images.

*Proof of intersection magnitude theorem (Theorem 14).* If the common vertical

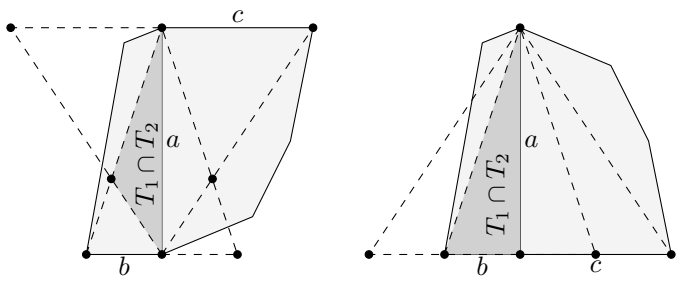



Figure 15: The shared subpolygon $T_1 \cap T_2$ is formed by taking the largest right triangles $T_1$ and $T_2$ inside the two ramp-like polygons, and intersecting $T_1$ with the reflection of $T_2$. In the case where the non-shared spines of the two ramp-like polygons are collinear, this intersection $T_1 \cap T_2$ is equal to the smaller right triangle, $T_1$.

edge $E$ does not have an integer $x$-coordinate, we apply the spine alignment lemma (Lemma 15) to replace $P_1$ and $P_2$ by polygons that cut out the same subgraphs in their interiors. Thus, we may assume that $E$ has an integer $x$-coordinate.

Let $a$ be the length of $E$, and let $b$ and $c$ be the widths of graphs $P_1$ and $P_2$, respectively. Without loss of generality we assume $b \leq c$. We observe that $P_1$ and $P_2$ are contained in their bounding boxes, which have $(a+1)(b+1)$ lattice points and $(a+1)(c+1)$ lattice points, respectively. Thus it suffices to construct subgraphs $G_1$ and $G_2$ with at least $\frac{1}{20}(a+1)(b+1)$ vertices each.

Let $T_1$ and $T_2$ be right triangles constructed from the endpoints of $E$ and the vertices of $P_1$ and $P_2$ (respectively) with the greatest horizontal distance from $E$. We construct $G_1$ by reflecting $T_2$ over $E$ and taking all the vertices in the interior of $T_1$ that are also in the interior of the reflected $T_2$; similarly, we construct $G_2$ by reflecting $T_1$ over $E$ and taking all the vertices in the interior of $T_2$ that are also in the interior of the reflected $T_1$. Abusing notation, we let $T_1 \cap T_2$ denote the triangle formed by intersecting $T_1$ with the reflection of $T_2$, as in Figure 15. Once we estimate its area and perimeter, we can use the triangle trimming lemma (Lemma 16) to get a lower bound on the number of vertices of $G_1$, and hence of $G_2$ as well.

First we claim

$$\text{Area}(T_1 \cap T_2) \geq \frac{ab}{4}.$$

To prove this area bound, we observe that given lengths $a$, $b$, and $c$, the case where $\text{Area}(T_1 \cap T_2)$ is the least is the case where the third vertex of $T_1$—that is, the vertex not on the common spine $E$—shares a $y$-coordinate with the top vertex of $E$, and the third vertex of $T_2$ shares a $y$-coordinate with the bottom vertex of $E$, or vice versa. In this case, if $b = c$ then the width of the intersection triangle is exactly $\frac{b}{2}$ so we have $\text{Area}(T_1 \cap T_2) = \frac{ab}{4}$. If $c > b$ then the intersection triangle is larger. Thus, in every case we have the desired area bound.

We also claim

$$\text{Perimeter}(T_1 \cap T_2) \leq 2(a + b).$$

This is because the perimeter of $T_1 \cap T_2$ is less than the perimeter of its bounding box, which has height $a$ and width at most $b$.

We put together the area and perimeter bounds with the triangle trimming lemma (Lemma 16) to estimate the number of lattice points enclosed by $T_1 \cap T_2$. It is algebraically true that for all $a, b \geq 41$, we have

$$\frac{ab}{4} - 4(a + b) + 1 \geq \frac{1}{20}(a + 1)(b + 1).$$

(To check this, we can use $ab = \frac{b}{2}a + \frac{a}{2}b > 20(a+b)$.) Thus, using our hypothesis that $a, b, c \geq 41$, we see that the number of vertices in the interior of $T_1 \cap T_2$ satisfies the inequalities

$$|G_1| \geq \frac{ab}{4} - 4(a + b) + 1 \geq \frac{1}{20}(a + 1)(b + 1),$$

as desired. □

Having proved this estimate on the size of the shared subpolygons, we can finish proving a bound on the routing number of the union of two ramp-like polygons along a shared spine.

**Theorem 17** (Routing between ramp-like)**.** *Consider two ramp-like pieces $P_1$ and $P_2$ with common vertical spine $E$. There exists a constant $C > 0$ such that* $\text{rt}(P_1 \cup P_2) \leq C \cdot (w(P_1 \cup P_2) + h(P_1 \cup P_2))$.

*Proof.* First consider the case where $w(P_1)$, $w(P_2)$, and the length of $E$ are all at least 41, so the intersection magnitude theorem (Theorem 14) applies. Fix a home configuration of tokens on the vertices of $P_1 \cup P_2$, and consider an arbitrary starting configuration. The ramp-like bound (Theorem 11) implies that we can efficiently route the tokens within $P_1$ and the tokens within $P_2$. Thus, what we need to show is that we can efficiently route the tokens into their home halves—that is, those that belong in $P_1$ should go to $P_1$ and those that belong in $P_2$ should go to $P_2$.

We label each token either black or white indicating whether it belongs in $P_1$ or $P_2$, respectively, in the home configuration. If the common spine $E$ has an integer $x$-coordinate, then some lattice points are shared between $P_1$ and $P_2$. In this case we count those lattice points as part of $P_1$ and not $P_2$, so that each token belongs in exactly one of the halves, and without loss of generality, we may assume that there are at least as many lattice points in $P_1$ as in $P_2$. In any configuration, we say that a given token is ***improper*** if it is in the opposite half from where it belongs. The number of improper tokens in $P_1$ is always equal to the number of improper tokens in $P_2$, which is at most the total number of lattice points in $P_2$.

We use the intersection magnitude theorem (Theorem 14) to find subgraphs $G_1$ and $G_2$ in $P_1$ and $P_2$ that are reflections over a vertical line and have size at least $\frac{1}{20}|P_2|$. Then we can move up to $|G_1| = |G_2|$ improper tokens into their home halves, using the following sequence of phases:

1. Use the ramp-like bound (Theorem 11) to route within $P_1$ and $P_2$ separately so that as many improper tokens as possible are in $G_1$ and $G_2$. If there are at least $|G_1|$ improper tokens on each side, then $G_1$ and $G_2$ become completely filled with improper tokens.
2. In the case where $G_1$ and $G_2$ do not become completely filled with improper tokens, continue to route within $P_2$ so that the locations of the improper tokens in $G_2$ are exactly the mirror image of the locations of the improper tokens in $G_1$.

3. Route each row of $P_1 \cup P_2$ simultaneously so that the improper tokens in $G_1$ exchange places with their mirror-image improper tokens in $G_2$, leaving no more improper tokens in either $G_1$ or $G_2$.

Each of these three phases takes at most $C(w(P_1 \cup P_2) + h(P_1 \cup P_2))$ routing steps, for some constant $C$. Repeating up to 20 times if necessary, we can move every token into its home half so that no improper tokens remain. Applying the ramp-like bound (Theorem 11) once more to route within each half, we move all tokens to their home lattice points in at most $C(w(P_1 \cup P_2) + h(P_1 \cup P_2))$ routing steps, for some constant $C$.

We now consider the case where the widths of $P_1$ and $P_2$ and the length of $E$ are not all at least 41. Suppose without loss of generality that it is $P_2$ that has height or width less than 41. We apply Lemma 5 with $G = P_1 \cup P_2$, $K = P_1$, the constant $c_1$ is 41, and $S$ is the rightmost column of $P_1$. Because we have a bound on the routing number of $P_1$, Lemma 5 states that the routing number of $P_1 \cup P_2$ is at most $C(w(P_1 \cup P_2) + h(P_1 \cup P_2))$ for some constant $C$. □

Using the bound for a pair of ramp-like polygons, we can follow a similar sequence of steps again to finish proving the bound on routing number of burger bun polygons.

**Theorem 18** (Burger bun bound)**.** *There exists a constant $C > 0$ such that for any burger bun polygon $P$, the routing number of $P$ satisfies the bound* $\mathrm{rt}(P) \leq C \cdot (w(P) + h(P))$.

*Proof.* Let $E$ be the spine of $P$. Without loss of generality we may assume that $E$ is vertical, so $E$ divides $P$ into a left half $P_1$ and a right half $P_2$. Each of $P_1$ and $P_2$, if it is not ramp-like already, is the union of two ramp-like pieces sharing a horizontal spine. Thus, Theorem 17 gives a bound on the routing number of $P_1$ and $P_2$ separately. In the present proof, we follow the proof of Theorem 17, but instead of using the ramp-like bound (Theorem 11) to route the two ramp-like halves, we use Theorem 17 itself to route $P_1$ and $P_2$.

We still need to prove an analogue of the intersection magnitude theorem

(Theorem 14) that applies to the present $P_1$ and $P_2$, which are not necessarily ramp-like. To do this, we construct two right triangles $T_1$ and $T_2$ that are ramp-like with common spine $E$, such that the intersection of $T_1$ with the reflection of $T_2$ is contained in the intersection of $P_1$ with the reflection of $P_2$. Then we apply the intersection magnitude theorem (Theorem 14) to $T_1$ and $T_2$.

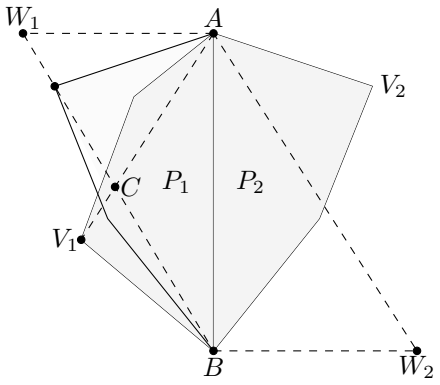



Figure 16: We use the intersection magnitude bound for ramp-like polygons to show a similar bound for burger bun halves, by applying it to ramp-like triangles $T_1 = ABW_1$ and $T_2 = ABW_2$, which are sufficiently large but still have (reflected) intersection inside our burger bun halves.

We construct $T_1$ and $T_2$ as follows, shown in Figure 16. Let $V_1$ and $V_2$ be the points on $P_1$ and $P_2$ (respectively) farthest from edge $E$, and let $A$ and $B$ be the top and bottom points of $E$. Then the triangle $ABV_1$ intersects the reflection of $ABV_2$ in some triangle $ABC$. We construct the third vertex $W_1$ of $T_1$ (that is, the vertex other than $A$ and $B$) by continuing the segment $BC$ upward until it intersects the horizontal line containing $A$. Similarly, we construct the third vertex $W_2$ of $T_2$ by continuing the reflection of $AC$ downward until it intersects the horizontal line containing $B$.

By applying the intersection magnitude theorem (Theorem 14) to $T_1$ and $T_2$, we find subgraphs $G_1$ and $G_2$ in $T_1$ and $T_2$, consisting of at least $\frac{1}{20}\min\{|T_1|, |T_2|\}$ vertices each. The construction of $T_1$ and $T_2$ guarantees that $G_1$ is also in $P_1$ and that $G_2$ is also in $P_2$. We compare $|T_1|$ and $|T_2|$ to $|P_1|$ and $|P_2|$ by observing that $T_1$ and $T_2$ each have at least half as many vertices as their bounding boxes,

and that those bounding boxes contain the bounding boxes of $P_1$ and $P_2$. Thus we have $|T_1| \geq \frac{1}{2}|P_1|$ and $|T_2| \geq \frac{1}{2}|P_2|$. The result is the lower bound

$$|G_1| = |G_2| \geq \frac{1}{40} \min\{|P_1|, |P_2|\},$$

in the case where the width and height of $P_1$ and $P_2$ are all at least 41.

Thus, following the proof of the routing between ramp-like theorem (Theorem 17), in the case where the height and width of $P_1$ and $P_2$ are all at least 41, we can move at least $\frac{1}{40}$ of the improper tokens into their home halves by first routing each half to put the improper tokens into $G_1$ and $G_2$, and then routing each row simultaneously to swap the improper tokens in $G_1$ with the improper tokens in $G_2$. Repeating this process at most 40 times puts every token into its home half, and then one more instance of routing within the two halves moves every token to its home vertex.

In the case where the height or width of (without loss of generality) $P_2$ is less than 41, as in the proof of Theorem 17 we may apply Lemma 5 with $G = P_1 \cup P_2$, $K = P_1$, $c_1 = 41$, and $S$ is the rightmost column of $P_1$. □

## 5. Proof of main theorem

In this section we prove the main theorem. For convenience we reproduce the statement of the main theorem here.

**Theorem 1.** *Let $P$ be a connected convex grid piece. Then the routing number of $P$ satisfies the bound* $\mathrm{rt}(P) \leq C(w(P) + h(P))$ *for some universal constant $C$.*

The idea of the proof is much simpler than the details. We show that our arbitrary convex polygon $P$ is related to a burger bun polygon by a shear transformation of at most 45°. Roughly, this transformation corresponds to a map between the sets of enclosed lattice points that stretches distances by at most a fixed factor. We show that if two graphs are related by a map that stretches by at most a fixed factor, then their routing numbers are also related by at most

a fixed factor. Thus, if we shear the original polygon to get a burger bun, then routing the resulting burger bun polygon helps us to route the original polygon.

The actual proof becomes more complicated to account for how the shear transformation does not respect the integer lattice—in particular, it does not necessarily preserve the number of lattice points inside the polygon. Lemmas 19 and 20 describe how we cut off part of $P$ to form $P_1 \subseteq P$, and Lemma 21 describes how we cut off a little more to form $P_2 \subseteq P_1$. Lemmas 22 and 23 describe how we shear $P_2$ to get a burger bun polygon $P_3$, which we know how to route. Lemma 24 describes how to map the lattice points inside $P_3$ into $P_1$ to form a subgraph $\psi(P_3)$ of $P_1$, and the bounded stretch theorem (Theorem 25) implies that the routing number of $\psi(P_3)$ is at most a constant factor greater than that of $P_3$. Then Lemma 26 checks the hypotheses of Lemmas 5 and 6, which will show that the routing number of $P$ is not much greater than that of $\psi(P_3)$. Figure 17 shows the relationship between these polygons and their associated graphs.

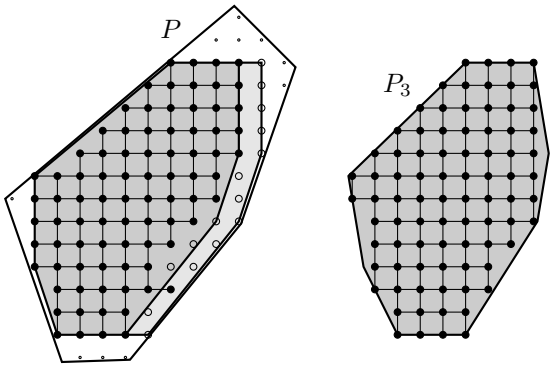



Figure 17: We cut off any short rows or columns of the original polygon $P$ to form $P_1$ (left, shaded), then cut off the rightmost vertex in each row to form $P_2$ (left, darkly shaded) before shearing $P_2$ to get a burger bun polygon $P_3$. The graph $\psi(P_3)$ inside $P_1$ has the same number of vertices per row as $P_3$ has.

Before shearing, we want to know that the shear does not affect whether the enclosed lattice graph is connected. To do this, in the next two lemmas we trim off the short rows and columns of $P$ that would be at risk of being pulled apart

by the shear.

**Lemma 19** (Constructing $P_1$)**.** *Let $P$ be a convex polygon enclosing strictly more than $4(w(P) + h(P))$ lattice points. Then there exists a polygon $P_1 \subseteq P$ such that*

- *$P_1$ is the convex hull of its enclosed lattice points;*
- *The subgraph $P \setminus P_1$ has at most $4(w(P) + h(P))$ vertices; and*
- *Every row and column of $P_1$ has at least 4 lattice points.*

*Proof.* We consider the top and bottom rows of $P$ and the leftmost and rightmost columns of $P$. If each of these has at least 5 lattice points, we set $P_1$ to be the convex hull of the lattice points in $P$. Otherwise, we iteratively remove one row or column at a time from the graph, choosing either the top row, the bottom row, the leftmost column, or the rightmost column, whichever has at most 4 lattice points. Once the top and bottom rows and the leftmost and rightmost columns of the remaining graph all have at least 5 lattice points, we set $P_1$ to be the convex hull of all the lattice points remaining.

Each deletion step reduces the number of lattice points by at most 4, while also reducing either the width or the height (possibly both) by 1. Thus $P \setminus P_1$ has at most $4(w(P) + h(P))$ lattice points, and in particular $P_1$ is nonempty.

To show that every row and column has at least 4 lattice points, because of the symmetry it suffices to show that every row has at least 4 lattice points. Consider the parallelogram formed by any choice of 5 consecutive lattice points from the top row of $P_1$ and 5 consecutive lattice points from the bottom row of $P_1$. Because $P_1$ is convex, this parallelogram is contained in $P_1$. Every horizontal cross-section of the parallelogram is an interval of length 4, so each cross-section at integer height must contain either 4 or 5 lattice points. Thus every row has at least 4 lattice points. □

The following lemma shows that consecutive rows of $P_1$ are well connected to each other, as are consecutive columns.

**Lemma 20** (Property of $P_1$)**.** *Let $P_1$ be a convex polygon with at least 4 lattice points in every row and column. Then every two consecutive rows of $P_1$ have at least 3 columns in common, and every two consecutive columns of $P_1$ have at least 3 rows in common.*

*Proof.* Because of the symmetry, it suffices to show that every two consecutive rows of $P_1$ have at least 3 columns in common. Suppose for the sake of contradiction that there are two consecutive rows with at most two columns in common; without loss of generality, suppose that it is the two (or more) rightmost vertices in the upper row that are not adjacent to vertices in the lower row. Let $u$ be the vertex in the upper row just to the right of the shared columns, and let $v$ be the vertex in the lower row just to the left of the shared columns, as in Figure 18.

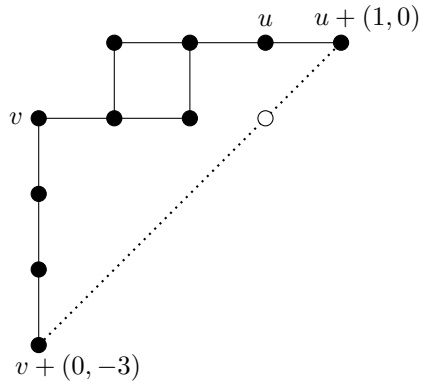



Figure 18: If the rows of $u$ and $v$ have fewer than 3 columns in common, we can contradict the convexity of $P_1$.

Then $v$ has no vertex immediately above it, so it must have at least three vertices below it in the same column, which means that the lattice point $v + (0, -3)$ is a vertex in $P_1$. We also know that $u + (1, 0)$ is a vertex in $P_1$. The segment with endpoints $v + (0, -3)$ and $u + (1, 0)$ has slope at least 1, because the $x$-coordinates of $u$ and $v$ differ by at most 3. The lattice point $u + (0, -1)$ lies on or to the left of this segment, but we have assumed that it is not in $P_1$, giving a contradiction. Thus we may conclude that each pair of consecutive

rows has more than two columns in common. □

It turns out that we want to trim off the right side of $P_1$ to form $P_2$ before shearing. This ensures that later when we construct $\psi(P_3)$, it fits inside $P_1$.

**Lemma 21** (Constructing $P_2$)**.** *Let $P_1$ be a convex polygon. Then there is a convex polygon $P_2$ that encloses all of the lattice points inside $P_1$ except the rightmost lattice point of each row.*

*Proof.* Let $P_{1\text{ymax}}$ and $P_{1\text{ymin}}$ be points on $P_1$ with the maximum and minimum $y$-coordinate, respectively. Consider the portion of the boundary of $P_1$ that is between the points $P_{1\text{ymax}}$ and $P_{1\text{ymin}}$ when moving clockwise from $P_{1\text{ymax}}$ to $P_{1\text{ymin}}$. Translate this piecewise-linear curve to the left by 1, and define the new polygon $P_2$ to be the subset of $P_1$ that is to the left of this translated side. Then $P_2$ has exactly one less vertex per row than $P_1$. □

We construct a shear that transforms $P_2$ into a burger bun polygon, and label this burger bun polygon $P_3$, in the following lemma.

**Lemma 22** (Constructing $P_3$)**.** *Let $P_2$ be a convex polygon with $w(P_2) \leq h(P_2)$. Then there exists a shear*

$$S = \begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix}$$

*with $|m| \leq 1$ such that the resulting polygon $P_3 = SP_2$ is burger bun, and we have $w(P_3) \leq 2w(P_2)$.*

*Proof.* Let $p = P_{\text{ymax}}$ and $q = P_{\text{ymin}}$. Then $\left|\frac{q_x - p_x}{p_y - q_y}\right| \leq 1$ because $|q_x - p_x| \leq w(P) \leq h(P) = |p_y - q_y|$. Define the horizontal shear

$$S = \begin{bmatrix} 1 & \frac{q_x - p_x}{p_y - q_y} \\ 0 & 1 \end{bmatrix}.$$

The region $SP = P'$ is a convex polygon since $S$ is linear. Moreover, $P'$ is burger bun because $S$ fixes the $y$-coordinate of each point and we have con-

structed the matrix $S$ so that $x(P'_{\text{ymax}}) = x(P'_{\text{ymin}})$, as follows:

$$x(P'_{\text{ymax}}) = x(Sp) = p_x + p_y \frac{q_x - p_x}{p_y - q_y} = \frac{-p_x q_y + p_y q_x}{p_y - q_y} =$$
$$= q_x + q_y \frac{q_x - p_x}{p_y - q_y} = x(Sq) = x(P'_{\text{ymin}}).$$

To show that $w(P_3) \leq 2w(P_2)$, let $a$ and $b$ be two arbitrary points in $P_2$. Then the $x$-coordinates of their images in $P_3$ are $a_x + ma_y$ and $b_x + mb_y$, which have absolute difference at most

$$|a_x - b_x| + |m| \cdot |a_y - b_y| \leq w(P_2) + \frac{w(P_2)}{h(P_2)} \cdot h(P_2) = 2w(P_2).$$

□

Before we can apply our burger-bun routing theorem to $P_3$, we need to check in the following lemma that $P_3$ is connected.

**Lemma 23** (Property of $P_3$)**.** *Let $P_3$ be a burger bun polygon with vertical spine. Suppose that $P_3$ contains at least $3 \cdot w(P_3)$ vertices in total and at least 2 vertices in each row. Then the graph $P_3$ is connected.*

*Proof.* If the spine is at an integer $x$-coordinate, then $P_3$ is connected because every vertex is connected by a horizontal path to the spine. Otherwise, we consider the two columns surrounding the spine. Every vertex in $P_3$ can be connected by a horizontal path to one of these two columns, so it suffices to show that the two columns have a row in common. Suppose to the contrary that they do not; without loss of generality, the $y$-coordinates of the subgraph of $P_3$ to the left of the spine are all greater than the $y$-coordinates of the subgraph to the right.

Because $P_3$ has at least $3 \cdot w(P_3)$ vertices, it must have at least three rows; without loss of generality, the left subgraph has at least one row, and the right subgraph has at least two rows. Consider the second-to-right vertex $u$ in the bottom row of the left subgraph, and the second-to-top vertex $v$ in the leftmost column of the right subgraph, as in Figure 19. The midpoint of $u$ and $v$ would connect the left subgraph to the right subgraph, so our assumptions imply that

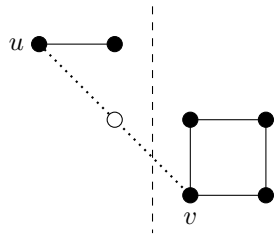



Figure 19: If a burger bun polygon has at least 3 rows, with at least 2 vertices in each row, then the subgraphs on either side of the spine must connect to each other, or else contradict the convexity of the polygon.

it is not a vertex in $P_3$; however, this contradicts the convexity of $P_3$. Thus, it is impossible for the two columns surrounding the spine not to have a row in common, and so $P_3$ is connected. □

The burger bun bound (Theorem 18) shows that we can route $P_3$. To show that this helps us route $P$, we start by finding a distorted copy of $P_3$ inside $P_1$.

**Lemma 24** (Constructing $\psi(P_3)$)**.** *Let $P_1$, $P_2$ and $P_3$ be convex polygons constructed in Lemmas 19, 21, and 22. Then there is an injective map $\psi$ from the vertex set of $P_3$ to the vertex set of $P_1$ with the following property. Consider the graph $\psi(P_3)$, the induced subgraph of $P_1$ whose vertices are the image of $\psi$. Then the map $\psi$ sends adjacent vertices in $P_3$ to vertices no more than 3 edges away from each other in $\psi(P_3)$. In particular, the image $\psi(P_3)$ is connected.*

*Proof.* First we construct the map $\psi$. Figure 20 depicts how $\psi$ is defined as a discrete version of reversing the shear that produced $P_3$ from $P_2$. We start by defining new coordinates for the vertices in $P_1$ and $P_3$. Without loss of generality suppose that the least $y$ coordinate of vertices in each polygon is 1. Let $r_i$ be the set of vertices in $P_1$ with $y$ coordinate $i$, so $r_i$ is the $i$th row of $P_1$. Let $r_{i,j}$ be the $j$th vertex in row $r_i$, counting from left to right. That is, $r_{i,1}$ is the vertex with least $x$ coordinate in row $i$, and $r_{2,i}$ has the second least $x$ coordinate and so on. Similarly, label the vertices of $P_3$ as $r'_{i,j}$. Define an injection $\psi$ from the vertex set of $P_3$ to $P_1$, given by $\psi(r'_{i,j}) = r_{i,j}$. That is, $\psi$ sends the $j$th vertex in row $i$ of $P_3$ to the $j$th vertex in row $i$ of $P_1$. We

have constructed the polygons so that each row of $P_3$ has no more vertices than the corresponding row of $P_1$: every horizontal cross-section of $P_3$ has the same length as the corresponding cross-section of $P_2$, so the number of lattice points in each row of $P_3$ differs by at most 1 from the number of lattice points in the corresponding row of $P_2$, and we know that $P_1$ has one more lattice point per row than $P_2$ has. Thus, our map $\psi$ is well-defined.

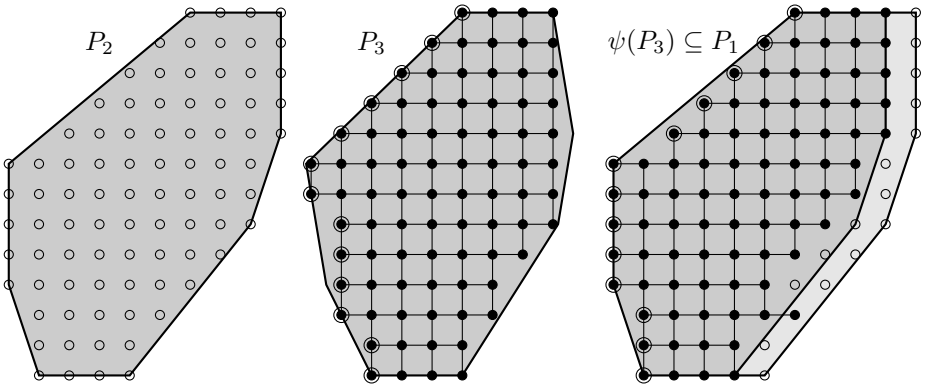



Figure 20: The map $\psi$ sends the first vertex in the $i$th row of $P_3$, denoted by $r'_{i,1}$, to the first vertex in the $i$th row of $P_1$, denoted by $r_{i,1}$. It sends the other vertices $r'_{i,j}$ in the $i$th row of $P_3$ to consecutive vertices $r_{i,j}$ in the $i$th row of $P_1$. The numbers of vertices in corresponding rows of $P_2$ and $P_3$ may differ by up to 1 in either direction, so the image $\psi(P_3)$ may extend past $P_2$ in some rows, and may miss vertices of $P_2$ in other rows.

Before proving that $\psi$ stretches distances by at most a factor of 3, as a preliminary step we show that every two consecutive rows of $\psi(P_3)$ have at least one column in common. By Lemma 20 we know that every two consecutive rows of $P_1$ have at least three columns in common. Then $P_2$ is like $P_1$, but with the rightmost vertex deleted from each row, so every two consecutive rows of $P_2$ have at least two columns in common. Then because corresponding rows of $P_2$ and $P_3$ differ in length by at most 1, we know that $\psi(P_3)$ is like $P_1$, but with up to two of the rightmost vertices deleted from each row, so every two consecutive rows of $\psi(P_3)$ have at least one column in common.

Using these common columns of pairs of consecutive rows, we can prove that $\psi$ sends adjacent vertices to vertices no more than 3 edges away from each other.

Adjacent vertices in a row of $P_3$ are sent to adjacent vertices, so we only need to check what $\psi$ does to adjacent vertices in a column of $P_3$. Thus, it suffices to show that each pair of adjacent rows do not shift in relation to each other by more than 2 edges under $\psi$. That is, for all rows $r_j$ and $r_{j+1}$ in $P_3$, we must show that the difference $|[x(r_{j,1}) - x(r_{j+1,\,1})] - [x(r'_{j,1}) - x(r'_{j+1,\,1})]|$ is no more than 2.

Let $(x_1, j)$ be the leftmost point in $P_3$ with $y$-coordinate $j$, and $(x_2, j+1)$ be the leftmost point of $P_3$ with $y$-coordinate $j+1$. Note that $r_{j,1}$ and $r_{j+1,\,1}$ are the first vertices in each of these rows, so $r_{j,1} = (\lceil x_1 \rceil, j)$ and $r_{j+1,\,1} = (\lceil x_2 \rceil, j+1)$. After the shear by the matrix

$$M = \begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix},$$

the points $(x_1, j)$ and $(x_2, j+1)$ are sent to $(x_1+mj, j)$ and $(x_2+m(j+1), j+1)$. Therefore, the first vertices in these rows of $P_1$, namely $r'_{j,1}$ and $r'_{j+1,\,1}$, have coordinates $(\lceil x_1 + mj \rceil, j)$ and $(\lceil x_2 + m(j+1) \rceil, j+1)$ respectively. Bounding the distance between these two points in $P_1$, we have

$$(x_1 - x_2) - |m| - 1 < \lceil x_1 + mj \rceil - \lceil x_2 + m(j+1) \rceil < (x_1 - x_2) + |m| + 1. \tag{2}$$

And, bounding the distance between the original points in $P_3$, we have

$$(x_1 - x_2) - 1 < \lceil x_1 \rceil - \lceil x_2 \rceil < (x_1 - x_2) + 1. \tag{3}$$

Combining these inequalities, we find that the absolute difference between the quantities $\lceil x_1 \rceil - \lceil x_2 \rceil$ and $\lceil x_1 + mj \rceil - \lceil x_2 + m(j+1) \rceil$ is strictly less than $|m| + 2$, and thus is strictly less than 3. Because the difference is an integer, it must be at most 2; in other words, the rows shift by no more than 2 vertices away from each other. Thus, the map $\psi$ sends adjacent vertices to vertices no more than 3 edges away from each other. □

Applying the next theorem to $\psi$ gives us a way to route $\psi(P_3)$ using our knowledge of how to route $P_3$.

**Theorem 25** (Bounded stretch)**.** *Let $A$ and $B$ be lattice graphs in $\mathbb{Z} \times \mathbb{Z}$, and let $\psi$ be a bijection between the vertices of $A$ and the vertices of $B$, such that for any pair of adjacent vertices $v_1, v_2 \in A$, the path-length distance between $\psi(v_1)$ and $\psi(v_2)$ in $B$ is at most a constant $c$. Then $\mathrm{rt}(B) \leq C(c) \cdot \mathrm{rt}(A)$, where $C(c)$ is a constant depending only on $c$.*

*Proof.* We would like to color the edges of $A$, such that if $(v_1, v_2)$ and $(v_1', v_2')$ are two edges of the same color, and we draw shortest paths between $w_1 = \psi(v_1)$ and $w_2 = \psi(v_2)$ and between $w_1' = \psi(v_1')$ and $w_2' = \psi(v_2')$ in $B$, then these two paths are disjoint. To do this, we would like to assign a color to each pair of vertices in $B$ that are within distance $c$ of each other, such that pairs of the same color are more than distance $c$ apart. We do this by first coloring the vertices of $B$, and then coloring the distance $c$ pairs by the color pairs of their vertices.

Our first coloring assigns colors to the vertices of $B$, such that if two vertices have the same color, they have distance greater than $2c$ in $B$. To do this, we construct a graph $B'$ with the same vertex set as $B$, with an edge between vertices $w_1$ and $w_2$ whenever their distance is at most $2c$. The maximal degree of any vertex in $B'$ is at most $4c(2c+1)$, since there are $4i$ lattice points with distance exactly $i$ away in $\mathbb{Z} \times \mathbb{Z}$ for each $1 \leq i \leq 2c$, and if a pair of vertices have distance at most $2c$ in $B$, they also have distance at most $2c$ in $\mathbb{Z} \times \mathbb{Z}$. Therefore, using a greedy strategy we can color the vertices of $B'$ with no more than $4c(2c+1)+1$ colors, so that no two vertices of the same color are adjacent in $B'$.

Our second coloring has one color for each pair of colors in the first coloring, for a total of $\binom{4c(2c+1)+1}{2} = (4c(2c+1)+1) \cdot 2c(2c+1)$ colors. It assigns one color to each pair of vertices $w_1, w_2$ in $B$ that have distance at most $c$, given by the pair of colors of $w_1$ and $w_2$ in the first coloring. Suppose that $(w_1, w_2)$ and $(w_1', w_2')$ are two pairs of vertices at distance at most $c$ in $B$, and they have the same color in the second coloring; without loss of generality this means that $w_1$ and $w_1'$ have the same color in the first coloring, as do $w_2$ and $w_2'$.

We claim that if we choose any shortest path from $w_1$ to $w_2$ in $B$, and any shortest path from $w_1'$ to $w_2'$ in $B$, then these two paths are disjoint. Suppose to the contrary that these two paths cross at a point. Then either $w_1$ or $w_2$ is within $\frac{1}{2}c$ of the common point, as is either $w_1'$ or $w_2'$. This implies that one of the pairs $(w_1, w_1')$, $(w_1, w_2')$, $(w_2, w_1')$, or $(w_2, w_2')$ has distance at most $c$, and so either $w_1$ and $w_1'$ have distance at most $2c$, or $w_2$ and $w_2'$ have distance at most $2c$. This contradicts their having been colored the same color in the first coloring. Therefore, any shortest path from $w_1$ to $w_2$ in $B$, and any shortest path from $w_1'$ to $w_2'$ in $B$ are disjoint.

For each edge in $A$, we color it by taking the corresponding pair of vertices in $B$, and finding the color of that pair in the second coloring. Then, given a pair of configurations in $B$, we use the following method to route between them. Take the corresponding pair of configurations in $A$, and consider a shortest sequence of steps to route between them. For each step in $A$, a set of disjoint swaps, all swaps along edges of the same color in $A$ can be carried out in parallel in $B$, because the corresponding paths are disjoint. There are at most $(4c(2c+1)+1)\cdot 2c(2c+1)$ colors of edges in $A$, and each color may take $c+1$ steps to route in $B$, since a path of length $c$ can be routed in $c+1$ steps. Therefore, it will take at most

$$C(c) = (4c(2c+1)+1) \cdot 2c(2c+1) \cdot (c+1)$$

steps in total to route a single step in $A$, a set of disjoint swaps. Therefore, $\mathrm{rt}(B) \leq C(c) \cdot \mathrm{rt}(A)$. □

At this point, we have obtained a bound on the routing number of $\psi(P_3)$, using the bound on the routing number of $P_3$. To extend this result to route all of $P$, we have to check the hypotheses of Lemmas 5 and 6, and then applying these lemmas will show that because $P$ is not too much bigger than $\psi(P_3)$, its routing number is also not too much bigger.

**Lemma 26** (Hair and skin of $\psi(P_3)$)**.** *Let $P$ be a convex polygon enclosing strictly more than $4(w(P)+h(P))$ lattice points, let $P_1$, $P_2$, and $P_3$ be convex*

*polygons constructed in Lemmas 19, 21, and 22, and let* $\psi\colon P_3 \to P_1$ *be the injection constructed in Lemma 24. Then the subgraph* $\psi(P_3)$ *of* $P$ *has the following properties:*

1. $P \setminus \psi(P_3)$ *has at most* $6(w(P) + h(P))$ *vertices; and*
2. *There is a subset* $S$ *of vertices in* $\psi(P_3)$ *such that* $S$ *has at most* $4(w(P) + h(P))$ *vertices and the induced subgraph* $S \cup (P \setminus \psi(P_3))$ *of* $P$ *is connected.*

*Proof.* Let $P_4$ be the convex polygon obtained from $P_2$ by removing the rightmost vertex of each row, as in Lemma 21. Then $\psi(P_3)$ contains $P_4$. We know that $P_1 \setminus P_4$ contains exactly $2h(P_1)$ vertices and that $P \setminus P_1$ contains at most $4(w(P) + h(P))$ vertices, so in total $P \setminus P_4$, and therefore $P \setminus \psi(P_3)$, contains at most $4w(P) + 6h(P)$ vertices.

Lemma 6 implies that $P_4$ has at most $2(w(P_4) + h(P_4)) \leq 2(w(P) + h(P))$ vertices in its boundary. We choose $S$ to contain the boundary of $P_4$, as well as all of $\psi(P_3) \setminus P_4$. Because $\psi(P_3)$ has at most two more vertices in each row than $P_4$, the number of vertices in $S$ is at most $2w(P) + 4h(P)$.

To show that $S \cup (P \setminus \psi(P_3))$ is connected, we observe that it is the same induced subgraph of $P$ as the union of the boundary of $P_4$ with $P \setminus P_4$; Lemma 6 states that because $P$ is connected and contains $P_4$, this induced subgraph is also connected. □

Finally we are ready to finish proving the bound on routing number of arbitrary convex polygons.

*Proof of Theorem 1.* Let $P$ be a convex polygon such that the grid piece contained in $P$ is connected. If $P$ has at most $4(w(P)+h(P))$ vertices, then the tree bound (Theorem 3) implies that $\mathrm{rt}(P) \leq 12(w(P)+h(P))$, and there is nothing more to prove. Thus, we may assume that $P$ has more than $4(w(P) + h(P))$ vertices.

We apply Lemma 19 to find $P_1$ inside $P$ with at least 4 vertices per row and column. Without loss of generality, we may assume that $h(P_1) \geq w(P_1)$. Then we apply Lemma 21 to find $P_2$ inside $P_1$ by removing the rightmost vertex of

each row. Then $P_2$ has the same height as $P_1$, and $w(P_2) = w(P_1) - 1$. We apply Lemma 22 to shear $P_2$ to get a burger bun polygon $P_3$.

We would like to apply Lemma 23 to check that $P_3$ is connected, so we need to estimate the number of vertices in $P_3$. We know that $P_1$ has more than $4(w(P_1) + h(P_1))$ vertices, so because $P_2$ is missing one vertex from each row, we see that $P_2$ has more than $4w(P_1) + 3h(P_1)$ vertices. Then every horizontal cross-section of $P_3$ has the same length as the corresponding cross-section of $P_2$, so the number of vertices in each row of $P_3$ differs by at most 1 from the number of vertices in the corresponding row of $P_2$. This implies that $P_3$ has more than $4w(P_1) + 2h(P_1)$ vertices. Lemma 22 tells us that $w(P_3) \leq 2w(P_2)$, and we know that $w(P_2) \leq w(P_1) \leq h(P_1)$, so the number of vertices in $P_3$ is more than $4w(P_2) + 2w(P_2) \geq 3w(P_3)$. Thus we may apply Lemma 23 to conclude that $P_3$ is connected.

The burger bun bound (Theorem 18) implies that because $P_3$ is connected and burger bun, we have $\mathrm{rt}(P_3) \leq C(w(P_3)+h(P_3))$ for some constant $C$. Then Lemma 24 and the bounded stretch theorem (Theorem 25) together imply that $\mathrm{rt}(\psi(P_3)) \leq C(w(P_3) + h(P_3))$ for some larger constant $C$, and so $\mathrm{rt}(\psi(P_3)) \leq C(2w(P_1) + h(P_1))$. Lemma 26 implies that Lemma 5 applies to $G = P$ and $K = \psi(P_3)$, so we may conclude that $\mathrm{rt}(P) \leq C(w(P)+h(P))$ for some constant $C$, as desired. □

## 6. Further questions

Our main theorem suggests several extensions to explore. The first direction is the question of bounded stretch.

**Question 1.** *Suppose that two token configurations on a connected convex grid piece $P$ have the property that the distance between each token's positions in the two configurations is at most $\delta$. Is there a sequence of $C \cdot \delta$ routing steps that transform the first configuration into the second, where $C$ is a constant not depending on the configurations or on $P$?*

An affirmative answer would strengthen our main theorem. Such an answer when $P$ is a rectangular grid graph is proved in [6] with a slightly different routing setup.

A second direction for further research is to ask for the appropriate formulation when the convex polygon $P$ is replaced by a nonconvex polygon. In this case the bound on routing number should depend on the presence of bottlenecks in the polygon. One way to quantify the idea of bottlenecks is as follows. Given a polygon $P$ in the plane, we define a ***chord*** of $P$ to be any line segment such that its two endpoints are on the boundary of $P$, and such that the interior of the segment is in the interior of $P$. Each chord $\gamma$ divides $P$ into two smaller polygons $P_1$ and $P_2$. We define the chord Cheeger constant of $P$, denoted by $\mathrm{ch}(P)$, to be

$$\mathrm{ch}(P) = \inf_{\gamma} \frac{\mathrm{length}(\gamma)}{\min(\mathrm{area}(P_1), \mathrm{area}(P_2))}.$$

A small value of $\mathrm{ch}(P)$ indicates that $P$ has a severe bottleneck. We should expect that $\mathrm{rt}(P)$ is at least proportional to $\frac{1}{\mathrm{ch}(P)}$, because in each routing step, the maximum number of tokens that can cross $\gamma$ is at most proportional to $\mathrm{length}(\gamma)$, and we may need enough routing steps for all the tokens belonging in the smaller of $P_1$ and $P_2$ to cross $\gamma$.

**Question 2.** *Is there a constant $C$ such that given a connected, possibly nonconvex grid piece $P$, we have*

$$\mathrm{rt}(P) \leq \frac{C}{\mathrm{ch}(P)}?$$

An affirmative answer would generalize our main theorem, because when $P$ is convex, the quantity $\frac{1}{\mathrm{ch}(P)}$ is within a constant factor of $w(P) + h(P)$.

A third way to try to extend our main result is by replacing the $\mathbb{Z}^2$ grid and 2-dimensional polygons by the $\mathbb{Z}^d$ grid (again with edges between vertices at distance 1) and $d$-dimensional polytopes. The following theorem on rectangular $d$-dimensional grid graphs is a consequence of the theorem on Cartesian product graphs from [1].

**Theorem 27** ([1])**.** *Let $R$ be a $d$-dimensional rectangular grid graph, of size $p_1$ by $p_2$ by... by $p_d$. Then we have*

$$\mathrm{rt}(P) \leq C_d \cdot (p_1 + p_2 + \cdots + p_d),$$

*where the constant $C_d$ depends only on the dimension.*

We note that the quantity $p_1+\cdots+p_d$ is within a constant factor of the graph diameter of $R$, which is within a dimensional constant factor of the Euclidean diameter of $R$. An affirmative answer to the following question would generalize the theorem above to non-rectangular convex shapes and would generalize our main theorem to all dimensions.

**Question 3.** *Are there dimensional constants $C_d$, such that if $P$ is a connected induced subgraph of the infinite grid graph $\mathbb{Z}^d$ consisting of all vertices contained in a given convex subset of $\mathbb{R}^d$, we have*

$$\mathrm{rt}(P) \leq C_d \cdot \mathrm{diam}(P)?$$

## References


[1] N. Alon, F. R. K. Chung, R. L. Graham, Routing permutations on graphs via matchings, SIAM J. Discrete Math. 7 (3) (1994) 513–530.

[2] Y. Baryshnikov, P. Bubenik, M. Kahle, Min-type Morse theory for configuration spaces of hard spheres, Int. Math. Res. Not. IMRN (9) (2014) 2577–2592.

[3] H. Löwen, Fun with hard spheres, in: Statistical physics and spatial statistics (Wuppertal, 1999), Vol. 554 of Lecture Notes in Phys., Springer, Berlin, 2000, pp. 295–331.

[4] P. Diaconis, The Markov chain Monte Carlo revolution, Bull. Amer. Math. Soc. (N.S.) 46 (2) (2009) 179–205.

[5] M. Farber, Invitation to topological robotics, Zurich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, 2008.

[6] E. D. Demaine, S. P. Fekete, P. Keldenich, C. Scheffer, H. Meijer, Coordinated Motion Planning: Reconfiguring a Swarm of Labeled Robots with Bounded Stretch, in: 34th International Symposium on Computational Geometry (SoCG 2018), Vol. 99 of Leibniz International Proceedings in Informatics (LIPIcs), 2018, pp. 29:1–29:15.

[7] R. Chinta, S. D. Han, J. Yu, Coordinating the motion of labeled discs with optimality guarantees under extreme density, in: M. Morales, L. Tapia, G. Sánchez-Ante, S. Hutchinson (Eds.), Algorithmic Foundations of Robotics XIII, Springer International Publishing, Cham, 2020, pp. 817–834.

[8] H. Alpert, Discrete configuration spaces of squares and hexagons, J. Appl. Comput. Topol. 4 (2) (2020) 263–280.