Routing in Histograms

Man-Kwun Chiu[1], Jonas Cleve[1][0000−0001−8480−1726], Katharina Klost[1], Matias Korman[2], Wolfgang Mulzer[1][0000−0002−1948−5840], André van Renssen[0000−0002−9294−9947], Marcel Roeloffzen, and Max Willert[1]

[1] Institut für Informatik, Freie Universität Berlin, 14195 Berlin, Germany
`chiumk@zedat.fu-berlin.de`,
`[jonascleve,kathklost,mulzer,willerma]@inf.fu-berlin.de`
[2] Department of Computer Science, Tufts University, Medford, MA, USA
`matias.korman@tufts.edu`
[3] School of Computer Science, University of Sydney, Sydney, Australia
`andre.vanrenssen@sydney.edu.au`
[4] Department of Mathematics and Computer Science, TU Eindhoven, Eindhoven, The Netherlands `m.j.m.roeloffzen@tue.nl`

Abstract. Let P be an x-monotone orthogonal polygon with n vertices. We call P a *simple histogram* if its upper boundary is a single edge; and a *double histogram* if it has a horizontal chord from the left boundary to the right boundary. Two points p and q in P are *co-visible* if and only if the (axis-parallel) rectangle spanned by p and q completely lies in P. In the *r-visibility graph* $G(P)$ of P, we connect two vertices of P with an edge if and only if they are co-visible. We consider *routing with preprocessing* in $G(P)$. We may preprocess P to obtain a *label* and a *routing table* for each vertex of P. Then, we must be able to route a packet between any two vertices s and t of P, where each step may use only the label of the target node t, the routing table and the neighborhood of the current node, and the packet header. The routing problem has been studied extensively for general graphs, where truly compact and efficient routing schemes with polylogartihmic routing tables have turned out to be impossible. Thus, special graphs classes are of interest.

We present a routing scheme for double histograms that sends any data packet along a path of length at most twice the (unweighted) shortest path distance between the endpoints. The labels, routing tables, and headers need $O(\log n)$ bits. For the simple histograms, we obtain a routing scheme with optimal routing paths, $O(\log n)$-bit labels, one-bit routing tables, and no headers.

1 Introduction

The *routing problem* is a classic question in distributed graph algorithms [6, 10]. We would like to preprocess a graph G for the following task: given a data packet located at a *source* vertex s, route the packet to a *target* vertex t, identified by its *label*. We strive for three main properties: *locality* (to find the next step of the packet, the scheme should use only information at the current vertex or in the packet header), *efficiency* (the packet should choose a path that is similar to a shortest path between s and t), and *compactness* (the space for labels, routing tables, and packet headers should be as small as possible). The ratio between the length of the *routing path* and a shortest path is called *stretch factor*.

Obviously, we could store at each vertex v of G the complete shortest path tree of v. This routing scheme is local and perfectly efficient: we can send the packet along a shortest path. However, the scheme lacks compactness. Thus, the general challenge is to balance the (potentially) conflicting goals of compactness and efficiency, while maintaining locality.

There are many routing schemes for general graphs (e.g., [11] and the references therein). For example, the scheme by Roditty and Tov [11] stores a poly-logarithmic number of bits in the packet header, and it routes a packet from s to t on a path of length $O\big(k\Delta + m^{1/k}\big)$, where $k > 2$ is any fixed integer, Δ is the shortest path distance between s and t, and m is the number of edges. The routing tables use $mn^{O(1/\sqrt{\log n})}$ total space, where n is the number of vertices. In the late 1980's, Peleg and Upfal [10] proved that in general graphs, any routing scheme with constant stretch factor must store $\Omega(n^c)$ bits per vertex, for some constant $c > 0$. This provides ample motivation to focus on special graph classes to obtain better routing schemes. For instance, trees admit routing schemes that always follow the shortest path and that store $O(\log n)$ bits at each node [5, 13]. Moreover, in planar graphs, for any fixed $\varepsilon > 0$, there is a routing scheme with a poly-logarithmic number of bits in each routing table that always finds a path that is within a factor of $1 + \varepsilon$ from optimal [12]. Similar results are also available for unit disk graphs [9, 14].

Another approach is *geometric routing*: the graph lies in a geometric space, and the routing algorithm must find the next vertex for the packet based on the coordinates of the source and the target vertex, the current vertex, and its neighborhood; e.g., [3] and the references therein. In contrast to compact routing schemes, there are no routing tables, and the routing is purely based on the local geometry (and possibly the packet header). For example, the routing algorithm for triangulations by Bose and Morin [4] uses the line segment between the source and the target for its routing decisions. In a recent result, Bose *et al.* [3] show that if vertices do not store any routing tables, no geometric routing scheme can achieve stretch factor $o(\sqrt{n})$. This holds irrespective of the header size.

Here, we combine the approaches from routing in abstract graphs and from geometric routing. For this, we consider a particularly interesting and practically relevant class of geometric graphs, namely visibility graphs of polygons. Banyassady *et al.* [1] presented a routing scheme for polygonal domains with n vertices and h holes that uses $O(\log n)$ bits for the label, $O((\varepsilon^{-1} + h) \log n)$ bits for the routing tables, and achieves a stretch of $1+\varepsilon$, for any fixed $\varepsilon > 0$. However, their approach is efficient only if the edges of the visibility graph are weighted with their Euclidean lengths. Banyassady *et al.* ask whether there is an efficient routing scheme for visibility graphs with unit weights (the *hop-distance*). This setting seems to be more relevant in practice, and similar results have already been obtained in unit disk graphs for routing schemes [9, 14] and for spanners [2, 8].

To address this problem, we use routing tables at the vertices to represent information about the structure of the graph (as in abstract routing), but we also assume that the labels of all adjacent vertices are directly visible at a node (as in geometric routing). This aligns well with the practical situation, as a node must be aware of its physical neighbors and their labels for meaningful communication to be possible. The size of the neighbor list does not count for the compactness, as it depends on the graph and cannot be influenced during preprocessing. We focus on r-visibility graphs of orthogonal simple and double histograms. At first, this may seem a strong restriction. However, even this case turns out to be quite challenging and reveals the whole richness of the compact routing problem in unweighted, geometrically defined graphs: on the one hand, the problem is still highly nontrivial, while on the other hand, much better results than in general graphs are possible. Histograms constitute a natural starting point, as they are often crucial building blocks in visibility problems. Moreover, r-visibility is a popular concept in orthogonal polygons that enjoys many useful structural properties; see, e.g., [7] and the references therein for more background on histograms and r-visibility.

A simple histogram is a monotone orthogonal polygon whose upper boundary consists of a single edge; a double histogram is a monotone orthogonal polygon that has a horizontal chord that touches the boundary of P only at the left and the right boundary. Let P be a (simple or double) histogram with n vertices. Two vertices v and w in P are connected in the visibility graph $G(P)$ by an unweighted edge if and only if the axis-parallel rectangle spanned by v and w is contained in the (closed) region P. We say that v and w are *co-visible*. We present the first efficient and compact routing schemes for polygonal domains under the hop-distance. In particular, in simple histograms, we can route along a shortest path with no headers, $O(\log n)$-bit labels, and $O(1)$-bit routing tables. In double histograms, we achieve stretch factor 2 and need labels, routing tables, and headers of $O(\log n)$ bits. The precise results are in Theorems 3.4 and 4.12. For space reasons, all proofs are moved to the appendix.

2 Preliminaries

Let $G = (V, E)$ be a *simple, undirected, unweighted, connected graph.* The (closed) *neighborhood* $N(v)$ of a vertex $v \in V$ is the set containing v and its adjacent nodes. Let $v, w \in V$. A sequence $\pi : \langle v = p_0, p_1, \ldots, p_k = w\rangle$ of vertices with $p_{i-1}p_i \in E$, for $i = 1, \ldots, k$, is called a *path* of length k between v and w. The length of π is denoted $|\pi|$. We define $d(v, w) = \min_\pi |\pi|$ as the length of a shortest path between v and w, where π goes over all paths between v and w. Next, we define a *routing scheme*. The algorithm that decides the next step of the packet is modeled by a *routing function.* During preprocessing, every node is assigned a (binary) *label* that identifies it in the network. The routing function uses local information at the current node, the label of the target node, and the *header* stored in the packet. The local information of a node v has two parts: (i) the *link table*, a list of the labels of $N(v)$, and (ii) the *routing* *table*, a bitstring chosen during preprocessing to represent relevant topological properties of G. Formally, a routing scheme of a graph G consists of (a) a label $\text{lab}(v) \in \{0,1\}^+$ for each node $v \in V$; (b) a routing table $\rho(v) \in \{0,1\}^*$ for each node $v \in V$; and (c) a routing function $f\colon \big(\{0,1\}^*\big)^4 \to \big(\{0,1\}^*\big)^2$. The routing function takes the link table and routing table of a current node $s \in V$, the label $\text{lab}(t)$ of a target t, and a header $h \in \{0,1\}^*$. From these, it determines the label $\text{lab}(v)$ of a node adjacent to s and a new header h'. The local information in the packet is updated to h', and it is sent to v. The routing scheme is *correct* if: for any two sites $s, t \in V$, consider the sequence $(\ell_0, h_0) = (\text{lab}(s), \varepsilon)$ and $(\text{lab}(p_{i+1}), h_{i+1}) = f\Big(\text{lab}\left(N(p_i)\right), \rho(p_i), \text{lab}(t), h_i\Big)$, for $i \geq 0$. Then, there is a $k = k(s,t) \geq 0$ with $p_k = t$ and $p_i \neq t$, for $i = 0, \ldots, k-1$. We say the routing scheme *reaches* t in k steps, and $\pi : \langle p_0, \ldots, p_k \rangle$ is the *routing path* from s to t. The *routing distance* is denoted $d_\rho(s,t) = |\pi|$. Let $\mathcal{R}$ be a correct routing scheme for a graph class $\mathcal{G}$, i.e., $\mathcal{R}$ is a correct routing scheme for every graph in $\mathcal{G}$. There are several measures for the quality of $\mathcal{R}$. For one, the various pieces of information used for the routing should be small. This is measured by the *maximum label size* $\text{Lab}(n)$, the *maximum routing table size* $\text{Tab}(n)$, and the *maximum header size* $H(n)$, over all graphs in $\mathcal{G}$ of a certain size. They are defined as $\text{Lab}(n) = \max_{|V|=n} \max_{v \in V} |\text{lab}(v)|$, $\text{Tab}(n) = \max_{|V|=n} \max_{v \in V} |\rho(v)|$, and $H(n) = \max_{|V|=n} \max_{s \neq t \in V} \max_{i=0,\ldots,k(s,t)} |h_i|$. Furthermore, the *stretch* $\zeta(n)$ relates the length of the routing path to the shortest path: $\zeta(n) = \max_{|V|=n} \max_{s \neq t \in V} d_\rho(s,t)/d(s,t)$.

Let P be a *simple orthogonal (axis-aligned) polygon* with n vertices $V(P)$ so that no three vertices in $V(P)$ are on the same vertical or horizontal line. The vertices are indexed counterclockwise from 0 to $n-1$; the lexicographically largest vertex has index $n-1$. For $v \in V(P)$, we write v_x and v_y for the x- and y-coordinate, and v_{id} for the index. We consider *r-visibility*: $p, q \in P$ *see each other* (are *co-visible*) if and only if the axis-aligned rectangle spanned by p and q is inside (the closed set) P. The *visibility graph* $G(P) = \big(V(P), E(P)\big)$ of P has an edge between two vertices $v, w \in V(P)$ if and only if v and w are co-visible. The distance $d(v,w)$ between two vertices $v, w \in V(P)$ is called the *hop distance* of v and w in P.

A *histogram* is an x-monotone orthogonal polygon where the upper boundary consists of exactly one horizontal edge, the *base edge*. By our convention, the endpoints of the base edge have index 0 (left) and $n-1$ (right). They are called the *base vertices*. A *double histogram* is an x-monotone orthogonal polygon P with a *base line*, a horizontal line segment whose relative interior lies in the interior of P and whose left and right endpoints are on the left and right boundary edge of P. We assume that the base line lies on the x-axis. Two vertices v, w in P lie *on the same side* if both are below or above the base line, i.e., if $v_y w_y > 0$. Every histogram is also a double histogram. From now on, we let P denote a (double) histogram.

Next, we classify the vertices of P. A vertex v in P is incident to exactly one horizontal edge h. We call v a *left* vertex if it is the left endpoint of h; otherwise, v is a *right* vertex. Furthermore, v is *convex* if the interior angle at v is $\pi/2$;

otherwise, v is *reflex*. Accordingly, every vertex of P is either *ℓ-convex*, *r-convex*, *ℓ-reflex*, or *r-reflex*.

To understand the shortest paths in P, we associate with each $v \in V(P)$ three landmark points in P (not necessarily vertices); see Fig. 1. The *corresponding*

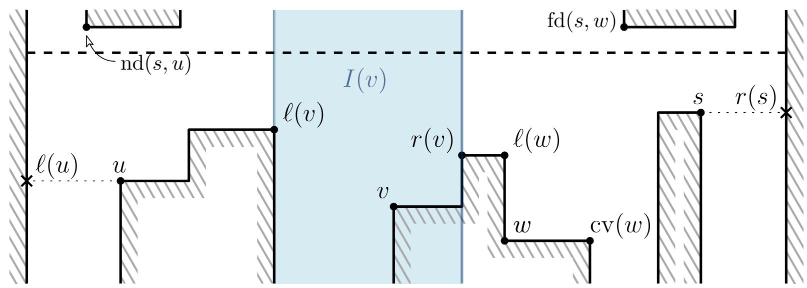

Fig. 1. Left and right points, the corresponding vertex, and the near and far dominators. The interval $I(v)$ of v is the set of vertices between $\ell(v)$ and $r(v)$. The dashed line is the base line.

vertex of v, $\mathrm{cv}(v)$, is the unique vertex with the same horizontal edge as v. To obtain the *left point* $\ell(v)$ of v, we shoot a leftward horizontal ray r from v. Let e be the vertical edge where r first hits the boundary of P. If e is the left boundary of P; then if P is a simple histogram, we let $\ell(v)$ be the left base vertex; and otherwise $\ell(v)$ is the point where r hits e. If e is not the left boundary of P, we let $\ell(v)$ be the endpoint of e closer to the base line. The *right point* $r(v)$ of v is defined analogously, by shooting the horizontal ray to the right.

Let p and q be two points in P. We say that p is *(strictly) to the left of* q, if $p_x \le q_x$ (or $p_x < q_x$). The term *(strictly) to the right of* is defined analogously. The *interval* $[p, q]$ of p and q is the set of vertices in P between p and q, i.e., $[p, q] = \{v \in V(P) \mid p_x \le v_x \le q_x\}$. By general position, this corresponds to index intervals in simple histograms. More precisely, if P is a simple histogram and p is either an r-reflex vertex or the left base vertex and q is either ℓ-reflex or the right base vertex, then $[p, q] = \{v \in V(P) \mid p_{\mathrm{id}} \le v_{\mathrm{id}} \le q_{\mathrm{id}}\}$. The *interval of a vertex* v, $I(v)$, is the interval of the left and right point of v, $I(v) = [\ell(v), r(v)]$. Every vertex visible from v is in $I(v)$, i.e., $N(v) \subseteq I(v)$. This interval will be crucial in our routing schemes and gives a very useful characterization of visibility in double histograms.

Let s and t be two vertices with $t \in I(s) \setminus N(s)$. We define two more landmarks for s and t. Assume that t lies strictly to the right of s, the other case is symmetric. The *near dominator* $\mathrm{nd}(s, t)$ of t with respect to s is the rightmost vertex in $N(s)$ to the left of t. If there is more than one such vertex, $\mathrm{nd}(s, t)$ is the vertex closest to the base line. Since t is not visible from s, the near dominator always exists. The *far dominator* $\mathrm{fd}(s, t)$ of t with respect to s is the leftmost vertex in $N(s)$ to the right of t. If there is no such vertex, we set $\mathrm{fd}(s, t) = r(s)$, the projection of s on the right boundary. The interval $I(s,t) = [\text{nd}(s,t), \text{fd}(s,t)]$ has all vertices between the near and far dominator; see Fig. 1.

3 Simple Histograms

Let P be a simple histogram with n vertices. The idea for our routing scheme is as follows: as long as the target vertex t is not in the interval $I(s)$ of the current vertex s, i.e., as long as there is a higher vertex that blocks visibility between s and t, we have to leave the current pocket as fast as possible. Once $t \in I(s)$, we have to find the pocket containing t. To do this, we must analyse in detail how shortest paths between vertices in P behave.

Paths in a Simple Histogram. We analyze the (shortest) paths in a simple histogram. The following lemma identifies certain "bottleneck" vertices that appear on any path; see. Fig. 2.

Lemma 3.1. *Let $v, w \in V(P)$ be co-visible vertices such that v is either r-reflex or the left base vertex and w is either ℓ-reflex or the right base vertex. Let s and t be two vertices with $s \in [v,w]$ and $t \notin [v,w]$. Then, any path between s and t includes v or w.*

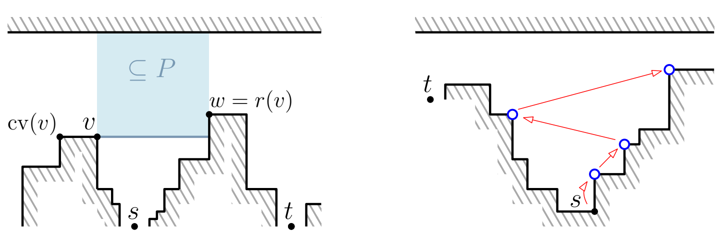

Fig. 2. Left: Any path from s to t includes v or w, since the blue rectangle contains only v and w as vertices. Right: A shortest path from s to t using the highest vertex.

An immediate consequence of Lemma 3.1 is that if $t \notin I(s)$, then any path from s to t uses $\ell(s)$ or $r(s)$. The next lemma shows that if $t \notin I(s)$, there is a shortest path from s to t that uses the higher vertex of $\ell(s)$ and $r(s)$, see Fig. 2.

Lemma 3.2. *Let s and t be two vertices with $t \notin I(s)$. If $\ell(s)_y > r(s)_y$ (resp., $\ell(s)_y < r(s)_y$), then there is a shortest path from s to t using $\ell(s)$ ($r(s)$).*

The next lemma considers the case where t is in $I(s)$. Then, the near and far dominator are the potential vertices that lie on a shortest path from s to t (see also Fig. 3).

Lemma 3.3. *Let s and t be two vertices with $t \in I(s) \setminus N(s)$. Then,* $\text{nd}(s,t)$ *is reflex and either* $\text{fd}(s,t) = \ell(\text{nd}(s,t))$ *or* $\text{fd}(s,t) = r(\text{nd}(s,t))$*.*

The Routing Scheme. We now describe our routing scheme and prove that it gives a shortest path. Let $v \in V(P)$. If v is convex and not a base vertex, it is labeled with its id, i.e., $\text{lab}(v) = v_{\text{id}}$. Otherwise, suppose that v is an r-reflex vertex or the left base vertex. The *breakpoint of* v, $\text{br}(v)$, is defined as the left endpoint of the horizontal edge with the highest y-coordinate to the right of and below v that is visible from v; analogous definitions apply to ℓ-reflex vertices and the right base vertex. The label of v consists of the ids of v and its breakpoint, i.e., $\text{lab}(v) = (v_{\text{id}}, \text{br}(v)_{\text{id}})$. Therefore, $\text{Lab}(n) = 2 \cdot \lceil \log n \rceil$. The routing table of v stores one bit, indicating whether $\ell(v)_y > r(v)_y$. Hence, $\text{Tab}(n) = 1$.

We are given the current vertex s and the label $\text{lab}(t)$ of the target vertex t. The routing function does not need a header, i.e., $H(n) = 0$. If t is visible from s, i.e., if $\text{lab}(t) \in \text{lab}(N(s))$, we directly go from s to t on a shortest path. Thus, assume that t is not visible from s. First, we check if $t \in I(s)$. This is done as follows: we determine the smallest and largest id in the link table $\text{lab}(N(s))$ of s. The corresponding vertices are $\ell(s)$ and $r(s)$. Then, we can check if $t_{\text{id}} \in [\ell(s)_{\text{id}}, r(s)_{\text{id}}]$, which is the case if and only if $t \in I(s)$. Now, there are two cases, illustrated in Fig. 3. First, suppose $t \notin I(s)$. If the bit in the routing table

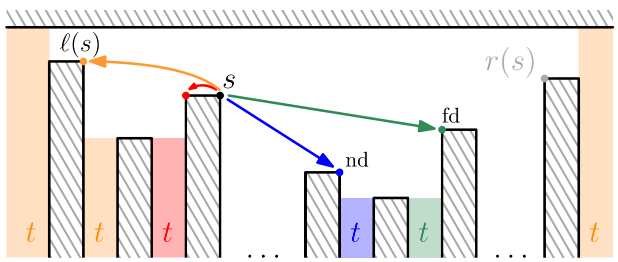

Fig. 3. The cases where the vertex t lies and the corresponding vertices where the data packet is sent to. If $t \in [\ell(s), s]$ we have $\text{nd}(s,t) = \text{cv}(s)$ and $\text{fd}(s,t) = \ell(s)$.

of s indicates that $\ell(s)$ is higher than $r(s)$, we take the hop to $\ell(s)$; otherwise, we hop to $r(s)$. By Lemma 3.2, this hop lies on a shortest path from s to t.

Second, suppose that $t \in I(s) \setminus N(s)$. This case is slightly more involved. We use the link table $\text{lab}(N(s))$ of s and the label $\text{lab}(t)$ of t to determine $\text{fd}(s,t)$ and $\text{nd}(s,t)$. Again, we can do this by comparing the ids. Lemma 3.3 states that either $\text{fd}(s,t) = \ell(\text{nd}(s,t))$ or $\text{fd}(s,t) = r(\text{nd}(s,t))$. We discuss the case that $\text{fd}(s,t) = r(\text{nd}(s,t))$, the other case is symmetric. By Lemma 3.1, any shortest path from s to t includes $\text{fd}(s,t)$ or $\text{nd}(s,t)$. Moreover, due to Lemma 3.3, $\text{nd}(s,t)$ is reflex, and we can use its label to access $b_{\text{id}} = \text{br}(\text{nd}(s,t))_{\text{id}}$. The vertex b splits $I(s,t) = [\text{nd}(s,t), \text{fd}(s,t)]$ into two disjoint subintervals $[\text{nd}(s,t), b]$ and $[\text{cv}(b), \text{fd}(s,t)]$. Also, b and $\text{cv}(b)$ are not visible from s, as they are located strictly between the far and the near dominator. Based on b_{id}, we can now decide on the next hop.

If $t \in [\mathrm{nd}(s,t), b]$, we take the hop to $\mathrm{nd}(s,t)$. If $t = b$, our packet uses a shortest path of length 2. Thus, assume that t lies between $\mathrm{nd}(s,t)$ and b. This is only possible if b is ℓ-reflex, and we can apply Lemma 3.1 to see that any shortest path from s to t includes $\mathrm{nd}(s,t)$ or b. But since $d(s,b) = 2$, our data packet routes along a shortest path.

If $t \in [\mathrm{cv}(b), \mathrm{fd}(s,t)]$, we take the hop to $\mathrm{fd}(s,t)$. If $t = \mathrm{cv}(b)$, our packet uses a shortest path of length 2. Thus, assume that t lies between $\mathrm{cv}(b)$ and $\mathrm{fd}(s,t)$. This is only possible if $\mathrm{cv}(b)$ is r-reflex, so we can apply Lemma 3.1 to see that any shortest path from s to t uses $\mathrm{fd}(s,t)$ or $\mathrm{cv}(b)$. Since $d(s, \mathrm{cv}(b)) = 2$, our packet routes along a shortest path. Thus:

Theorem 3.4. *Let P be a simple histogram with n vertices. There is a routing scheme for $G(P)$ with 1-bit routing tables, no header, and label size $2 \cdot \lceil \log n \rceil$, such that we can route between any two vertices on a shortest path.*

4 Double Histograms

Let P be a double histogram with n vertices. Similar to the simple histogram case, we first focus on the structure of shortest paths in P. Again, if the target vertex t is not in the interval $I(s)$ of the current vertex s, we should widen the interval as fast as possible. However, in contrast to simple histograms, we can now change sides arbitrarily often. Nevertheless, we can guarantee that in each step, the interval comes closer to t. Once we have reached the case that t is in the interval of the current vertex, we again have to find the right pocket. Unlike in simple histograms, this case is now simpler to describe.

Paths in a Double Histogram. To understand shortest paths in double histograms, we distinguish three cases, depending on where t lies relative to s. First, if t is close, i.e., if $t \in I(s)$, we focus on the near and far dominators. Second, if $t \notin I(s)$ but there is a vertex v visible from s with $t \in I(v)$, then we can find a vertex on a shortest path from s to t. Third, if there is no visible vertex v from s such that $t \in I(v)$, we can apply our intuition from simple histograms: go as fast as possible towards the base line.

Let s, t be two vertices with $t \in I(s) \setminus N(s)$. In contrast to simple histograms, $\mathrm{fd}(s,t)$ now might not be a vertex. Furthermore, $\mathrm{fd}(s,t)$ and $\mathrm{nd}(s,t)$ might be on different sides of the base line. In this case, Lemma 3.3 no longer holds. However, the next lemma establishes a visibility relation between them.

Lemma 4.1. *Let $s, t \in V(P)$ with $t \in I(s) \setminus N(s)$. Then, $\mathrm{nd}(s,t)$ and $\mathrm{fd}(s,t)$ are co-visible.*

The proof of the next lemma uses Lemma 4.1 to find a shortest path vertex.

Lemma 4.2. *One of $\mathrm{nd}(s,t)$ or $\mathrm{fd}(s,t)$ is on a shortest path from s to t. If $\mathrm{fd}(s,t)$ is not a vertex, then $\mathrm{nd}(s,t)$ is on a shortest path from s to t.*

Next, we consider the case where $\mathrm{fd}(s,t)$ is a vertex but not on a shortest path from s to t. Then, $\mathrm{fd}(s,t)$ cannot see t, and we define $\mathrm{fd}^2(s,t) = \mathrm{fd}(\mathrm{fd}(s,t),t)$. By Lemma 4.1, $\mathrm{nd}(s,t)$ and $\mathrm{fd}(s,t)$ are co-visible, so $\mathrm{fd}^2(s,t)$ has to be in the interval $[\mathrm{nd}(s,t),t]$, and therefore it is a vertex. The following lemma states that $\mathrm{fd}^2(s,t)$ is strictly closer to t than s.

Lemma 4.3. *If $\mathrm{fd}(s,t)$ is a vertex but not on a shortest path from s to t, then we have $d(\mathrm{fd}^2(s,t),t) = d(s,t) - 1$.*

Let s,t be two vertices so that $t \notin I(s)$ but there is a vertex $v \in N(s)$ with $t \in I(v)$. For clarity of presentation, we will always assume that s is below the base line. The crux of this case is this: there might be many vertices visible from s that have t in their interval. However, we can find a best vertex as follows: once t is in the interval of a vertex, the goal is to shrink the interval as fast as possible. Therefore, we must find a vertex $v \in N(s)$ whose left or right interval boundary is closest to t among all vertices in $N(s)$. This leads to the following inductive definition of two sequences $a^i(s)$ and $b^i(s)$ of vertices in $N(s)$. For $i = 0$, we let $a^0(s) = b^0(s) = s$. For $i > 0$, if the set $A^i(s) = \{v \in N(s) \mid \ell(v)_x < \ell(a^{i-1}(s))_x\}$ is nonempty, we define $a^i(s) = \mathrm{argmin}\{v_x \mid v \in A^i(s)\}$; and $a^i(s) = a^{i-1}(s)$, otherwise. If the set $B^i(s) = \{v \in N(s) \mid r(v)_x > r(b^{i-1}(s))_x\}$ is nonempty, we define $b^i(s) = \mathrm{argmax}\{v_x \mid v \in B^i(s)\}$; and $b^i(s) = b^{i-1}(s)$, otherwise. We force unambiguity by choosing the vertex closer to the base line. Let $a^*(s)$ be the vertex with $a^*(s) = a^i(s) = a^{i-1}(s)$, for an $i > 0$, and $b^*(s)$ the vertex with $b^*(s) = b^i(s) = b^{i-1}(s)$, for an $i > 0$. If the context is clear, we write a^i instead of $a^i(s)$ and b^i instead of $b^i(s)$.

Let us try to understand this definition. For $i \geq 0$, we write ℓ^i for $\ell(a^i)$; and we write ℓ^* for $\ell(a^*)$. Then, we have $a^0 = s$ and $\ell^0 = \ell(s)$. Now, if $\ell(s)$ is not a vertex, then $a^* = s$, because there is no vertex whose left point is strictly to the left of the left boundary of P. On the other hand, if ℓ^0 is a vertex in P, we have $a^1 = \ell^0 = \ell(s)$, and $[\ell^1, a^1]$ is an interval between points on the lower side of P. Then comes a (possibly empty) sequence of intervals $[\ell^2, a^2], [\ell^3, a^3], \dots, [\ell^k, a^k]$ between points on the upper side of P; possibly followed by the interval $[\ell(r(s)), r(s)]$. There are four possibilities for a^*: it could be s, $\ell(s)$, a vertex a^i on the upper side of P, or $r(s)$. If $a^* \neq s$, then the intervals $[\ell^1, a^1] \subset [\ell^2, a^2] \subset \cdots \subset [\ell^*, a^*]$ are strictly increasing: ℓ^i is strictly to the left of ℓ^{i-1} and a^i is strictly to the right of a^{i-1}; see Fig. 4. Symmetric observations apply for the b^i; we write r^i for $r(b^i)$ and r^* for $r(b^*)$.

Lemma 4.4. *For $i \geq 1$, the vertices ℓ^{i-1}, a^i as well as r^{i-1}, b^i are co-visible.*

Finally, the next lemma tells us the following: if $t \in [\ell^*, r^*]$ we find a vertex $v \in N(s)$ with $t \in I(v)$. Its quite technical proof needs Lemma 4.4.

Lemma 4.5. *If $t \in [\ell^i, \ell^{i-1}]$, for some $i \geq 1$, then a^i is on a shortest path from s to t. If $t \in [r^{i-1}, r^i]$, for some $i \geq 1$, then b^i is on a shortest path from s to t.*

Finally, we consider the case that there is no vertex $v \in N(s)$ with $t \in I(v)$, i.e., $t \notin [\ell^*, r^*]$. The intuition now is as follows: to widen the interval, we should

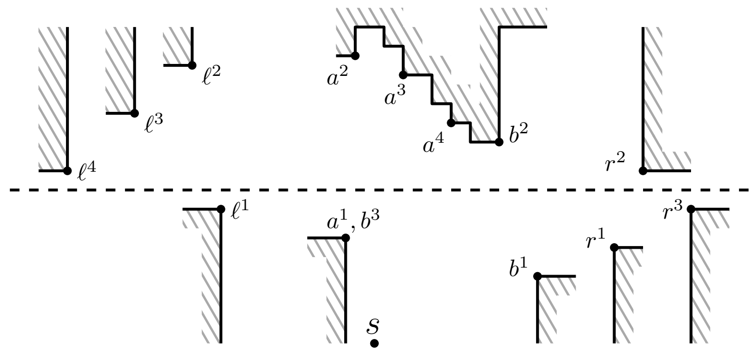

Fig. 4. The vertices a^i and b^i are illustrated. Observe that $\ell(s) = a^1 = b^3$ and $r(s) = b^1$.

go to a vertex that is visible from s, but closest to the base line. In simple histograms, there was only one such vertex, but in double histograms there might be a second one on the other side. These two vertices are the *dominators of* s. These two dominators might have their own dominators, and so on. This leads to the following inductive definition.

For $k \geq 0$, we define the k-th *bottom dominator* $\mathrm{bd}^k(s)$, the k-th *top dominator* $\mathrm{td}^k(s)$, and the k-th *interval* $I^k(s)$ of s. For any set $Q \subset V(P)$, we write Q^- (resp. Q^+) for all points in Q below (resp. above) the base line. We set $\mathrm{bd}^0(s) = \mathrm{td}^0(s) = s$ and $I^0(s) = \{s\}$. For $k > 0$, we set $I^k(s) = I(\mathrm{bd}^{k-1}(s)) \cup I(\mathrm{td}^{k-1}(s))$. If $I^k(s)^-$ is nonempty, we let $\mathrm{bd}^k(s)$ be the leftmost vertex inside $I^k(s)^-$ that minimizes the distance to the base line. If $I^k(s)^+$ is nonempty, we let $\mathrm{td}^k(v)$ be the leftmost vertex inside $I^k(s)^+$ that minimizes the distance to the base line. If one of the two sets is empty, the other one has to be nonempty, since $s \in I^k(s)$. In this case, we let $\mathrm{td}^k(s) = \mathrm{bd}^k(s)$. We write $\mathrm{bd}(s)$ for $\mathrm{bd}^1(s)$ and $\mathrm{td}(s)$ for $\mathrm{td}^1(s)$. Observe, that $I^1(s) = I(s)$ and $I^2(s) = [\ell^*, r^*]$. If $I(\mathrm{bd}^{k-1}(s)) = V(P)$, we have $\mathrm{bd}^k(s) = \mathrm{bd}^{k-1}(s)$. The same holds for the top dominator. We provide a few technical properties concerning the k-th interval as well as the k-th dominators.

Lemma 4.6. *For any* $s \in V(P)$ *and* $k \geq 0$*, we have* $I^k(s) \subseteq I(\mathrm{bd}^k(s)) \cap I(\mathrm{td}^k(s))$ *and* $\mathrm{bd}^k(s)$*,* $\mathrm{td}^k(s)$ *are co-visible.*

The following lemma seems rather specific, but will be needed later to deal with short paths.

Lemma 4.7. *For any* $s \in V(P)$*, we have* $I^3(s) = I^2\big(\mathrm{bd}(s)\big) \cup I^2\big(\mathrm{td}(s)\big)$.

Intuitively, the meaning of $I^k(s)$ is as follows: let ℓ be the leftmost and r be the rightmost vertex with hop distance exactly k from s, then, $I^k(s) = [\ell, r]$. We do not really need this property, so we leave it as an exercise for the reader to find a proof for this. Instead, we prove the following weaker statement. For this, recall that due to its definition, $\mathrm{bd}^k(s)$ might not be on the lower side of the histogram (and $\mathrm{td}^k(s)$ might not be on the upper side).

Lemma 4.8. *Let $k \geq 0$ and let $s, t \in V(P)$ with $d(s,t) \leq k$. Then, $t \in I^k(s)$.*

Let $k \geq 0$ and $s \in V(P)$. For $i = 1, \dots, k$, by Lemma 4.6, $\mathrm{bd}^{i-1}(s), \mathrm{td}^{i-1}(s) \in I(\mathrm{bd}^i(s)) \cap I(\mathrm{td}^i(s))$. Moreover, by definition, $\mathrm{bd}^i(s), \mathrm{td}^i(s) \in I(\mathrm{bd}^{i-1}(s)) \cup I(\mathrm{td}^{i-1}(s))$. As we show in the appendix (Observation C.1), now both $\mathrm{bd}^i(s)$ and $\mathrm{td}^i(s)$ can see at least one of $\mathrm{bd}^{i-1}(s)$ or $\mathrm{td}^{i-1}(s)$. Therefore, there is a path $\pi_b(s,k) : \langle s = p_0, \dots, p_k = \mathrm{bd}^k(s) \rangle$ from s to $\mathrm{bd}^k(s)$ and a path $\pi_t(s,k) : \langle s = q_0, \dots, q_k = \mathrm{td}^k(s) \rangle$ from s to $\mathrm{td}^k(s)$ with $p_i, q_i \in \{\mathrm{bd}^i(s), \mathrm{td}^i(s)\}$, for $i = 0, \dots, k$. We call $\pi_b(s,k)$ and $\pi_t(s,k)$ the *canonical path* from s to $\mathrm{bd}^k(s)$ and from s to $\mathrm{td}^k(s)$, respectively. The following two lemmas show that for every $t \notin I^{k+1}(s)$ one of the canonical paths is the prefix of a shortest path from s to t. To show Lemma 4.10 we need Lemmas 4.8 and 4.9.

Lemma 4.9. *Let $k \geq 1$ and $s \in V(P)$. If $I(\mathrm{bd}^{k-1}(s)) \neq V(P)$, we have that $d(s, \mathrm{bd}^k(s)) = k$. If $I(\mathrm{td}^{k-1}(s)) \neq V(P)$ we have $d(s, \mathrm{td}^k(s)) = k$.*

Lemma 4.10. *Let s and t be vertices and $k \geq 1$ an integer such that $t \notin I^{k+1}(s)$. Then $\mathrm{bd}^k(s)$ or $\mathrm{td}^k(s)$ is on a shortest path from s to t.*

Routing Scheme. Let v be a vertex. The label of v consists of its x- and y-coordinate as well as the bounding x-coordinates of $I(v)$. We do not need v_{id} since (v_x, v_y) identifies the vertex in the network. Thus, $\mathrm{Lab}(n) = 4 \cdot \lceil \log n \rceil$ since we can assume that $v_x, v_y \in \{0, \dots, n-1\}$. In the routing table of v, we store the bounding x-coordinates of $I^2(\mathrm{bd}(v))$ as well as the bounding x-coordinates of $I^2(\mathrm{td}(v))$. Furthermore, we store $(\mathrm{bd}^2(v)_x, \mathrm{bd}^2(v)_y, bit)$ where *bit* indicates whether $\mathrm{td}(v)$ or $\mathrm{bd}(v)$ is on the path $\pi_b(v,2)$. Thus, $\mathrm{Tab}(n) = 6 \cdot \lceil \log n \rceil + 1$.

We are given a current vertex s together with its routing table and link table, the label of a target vertex t, and a header. If $t \in N(s)$, then $\mathrm{lab}(t)$ is in the link table of s, and we send the data packet directly to t. If the header is non-empty, it will contain the coordinates of exactly one vertex visible from s. We clear the header and go to this respective vertex. The remaining discussion assumes that the header is empty and that $t \notin N(s)$. The routing function now distinguishes four cases depending on whether $t \in I(s)$, $t \in I^2(s)$ or $t \in I^3(s)$. We can check the first and the second condition locally, using the link table of s as well as the label of t (note that from the link table of s, we can deduce $a^*(s)$ and $b^*(s)$, and their interval boundaries). To check the third condition locally, we use Lemma 4.7 which shows that $I^3(s) = I^2(\mathrm{bd}(s)) \cup I^2(\mathrm{td}(s))$. Since we stored the bounding x-coordinates of these two intervals in the routing table of s, we can check $t \in I^3(s)$ easily.

Case 1 ($t \in I(s) \setminus N(s)$): if $\mathrm{fd}(s,t)$ is a vertex, we can determine it by using the link table and the label of t. The packet is sent to $\mathrm{fd}(s,t)$. If $\mathrm{fd}(s,t)$ is not a vertex, we determine $\mathrm{nd}(s,t)$ and send the packet there. The header remains empty.

Case 2 ($t \in I^2(s) \setminus I(s)$): there is an $i \geq 1$ with $t \in [\ell^i, \ell^{i-1}]$ or $t \in [r^{i-1}, r^i]$. We find i using the link table and $\mathrm{lab}(t)$. The packet is sent to a^i or b^i. The header remains empty.

Case 3 ($t \in I^3(s) \setminus I^2(s)$): if $t \in I^2(\mathrm{bd}(s))$, we send the packet to $\mathrm{bd}(s)$. Otherwise, $t \in I^2(\mathrm{td}(s))$, and we send the packet to $\mathrm{td}(s)$. In both cases, the header remains empty.

Case 4 ($t \notin I^3(s)$): the routing table has the entry $(\mathrm{bd}^2(s)_x, \mathrm{bd}^2(s)_y, b)$. We store $(\mathrm{bd}^2(s)_x, \mathrm{bd}^2(s)_y)$ in the header and send the packet to $\mathrm{bd}(s)$ or $\mathrm{td}(s)$, as indicated by b.

Obviously, $H(n) = 2 \cdot \lceil \log n \rceil$. It remains to analyze the stretch factor. For this, we use the following lemma:

Lemma 4.11. *Let $s, t \in V(P)$. After at most two steps of the routing scheme from s with target label $\mathrm{lab}(t)$, we reach a vertex v with $d(v,t) \leq d(s,t) - 1$.*

This immediately gives a stretch factor of 2 and our main theorem.

Theorem 4.12. *Let P be a double histogram with n vertices. There is a routing scheme for $G(P)$ with routing table, label and header size $O(\log n)$, such that we can route between any two vertices with stretch at most 2.*

References

1. B. Banyassady, M.-K. Chiu, M. Korman, W. Mulzer, A. van Renssen, M. Roeloffzen, P. Seiferth, Y. Stein, B. Vogtenhuber, and M. Willert. Routing in polygonal domains. In *28th ISAAC*, pages 10:1–10:13, 2017.
2. A. Biniaz. Plane hop spanners for unit disk graphs: Simpler and better. `arXiv:1902.10051`, 2019.
3. P. Bose, R. Fagerberg, A. van Renssen, and S. Verdonschot. Competitive local routing with constraints. *JoCG*, 8:125–152, 2017.
4. P. Bose and P. Morin. Competitive online routing in geometric graphs. *TCS*, 324:273–288, 2004.
5. P. Fraigniaud and C. Gavoille. Routing in trees. In *28th ICALP*, pages 757–772, 2001.
6. S. Giordano and I. Stojmenovic. Position based routing algorithms for ad hoc networks: A taxonomy. In *Ad hoc wireless networking*, pages 103–136. Springer-Verlag, 2004.
7. F. Hoffmann, K. Kriegel, S. Suri, K. Verbeek, and M. Willert. Tight bounds for conflict-free chromatic guarding of orthogonal art galleries. *CGTA*, 73:24–34, 2018.
8. P. Jacquet and L. Viennot. Remote-spanners: What to know beyond neighbors. In *23rd IPDPS*, pages 1–10, 2009.
9. H. Kaplan, W. Mulzer, L. Roditty, and P. Seiferth. Routing in unit disk graphs. *Algo.*, 80:830–848, 2018.
10. D. Peleg and E. Upfal. A trade-off between space and efficiency for routing tables. *JACM*, 36:510–530, 1989.
11. L. Roditty and R. Tov. Close to linear space routing schemes. *Dist. Comp.*, 29:65–74, 2016.
12. M. Thorup. Compact oracles for reachability and approximate distances in planar digraphs. *JACM*, 51:993–1024, 2004.
13. M. Thorup and U. Zwick. Compact routing schemes. In *13th SPAA*, pages 1–10, 2001.
14. C. Yan, Y. Xiang, and F. F. Dragan. Compact and low delay routing labeling scheme for unit disk graphs. *CGTA*, 45:305–325, 2012.

A Acknowledgments

This work was initiated during a visit to Tohoku University in February 2018. We would like to thank Takeshi Tokuyama and all the members of his group for their hospitality and for creating a conducive research environment.

M.-K.C., J.C. and W.M. partially supported by ERC STG 757609. J.C. partially supported by DFG grant MU 3501/1-2. M.K. partially supported by MEXT KAKENHI No. 17K12635 and the NSF award CCF-1422311. A.v.R. partially supported by by JST ERATO Grant Number JPMJER1201, Japan.

B Additional Material for Section 3

B.1 Visibility in Simple Histograms

We begin with some observations on the visibility in P. As P is a simple histogram, for all $v \in V(P)$, the points $\ell(v)$ and $r(v)$ are vertices of P. Hence, the far dominators also must be vertices. The next observations are now immediate.

Observation B.1. *Let $v \in V(P)$ be r-reflex or the left base vertex, and let $u \in [v, r(v)]$ be a vertex distinct from v and $r(v)$. Then, $I(u) \subseteq [v, r(v)]$.*

Proof. Assume that $\ell(u)$ or $r(u)$ is outside of $[v, r(v)]$. Then, u must have a larger y-coordinate than v. It follows that v cannot see $r(v)$. This contradicts the definition of $r(v)$. ☐

Observation B.2. *Let $v \in V(P)$ be a left (right) vertex distinct from the base vertex. Then, v can see exactly two vertices to its right (left), namely $\mathrm{cv}(v)$ and $r(v)$ ($\ell(v)$).*

Proof. Suppose that v is a left vertex; the other case is symmetric. Any vertex visible from v to the right of v lies in in $[\mathrm{cv}(v), r(v)]$. If $\mathrm{cv}(v)$ is convex, the observation is immediate, since then $[\mathrm{cv}(v), r(v)] = \{\mathrm{cv}(v), r(v)\}$. Otherwise, $\mathrm{cv}(v)$ is r-reflex and $r(v) = r(\mathrm{cv}(v))$. By Observation B.1, we get that for all $u \in [\mathrm{cv}(v), r(v)] \setminus \{\mathrm{cv}(v), r(v)\}$, we have $I(u) \subseteq [\mathrm{cv}(v), r(v)]$. Thus, $v \notin I(u)$ for any such u, and since $N(u) \subseteq I(u)$, v cannot see u. ☐

B.2 Missing Proofs

Proof (of Lemma 3.1). Let $I = [v, w]$. Since $t \notin I$, not both v, w are base vertices. Thus, suppose without loss of generality that $v_y < w_y$. Then, $\mathrm{cv}(v)$ is a left vertex and can see w. Hence, Observation B.2 implies that $r(v) = r(\mathrm{cv}(v)) = w$. By Observation B.1, we get $N(u) \subseteq I(u) \subseteq I$, for any $u \in I \setminus \{v, w\}$. Thus, any path between s and t must include v or w. ☐

Proof (of Lemma 3.2). Assume $\ell(s)_y > r(s)_y$, the other case is symmetric. Let $\pi : \langle s = p_0, \ldots, p_k = t\rangle$ be a shortest path from s to t. If π contains $\ell(s)$, we are done. Otherwise, by Lemma 3.1, there is a $0 < j < k$ with $p_j = r(s)$ and

$p_i \neq r(s)$, for $i > j$. Thus, $p_{j+1} \notin I(s)$. Since we assumed $\ell(s)_y > r(s)_y$, it follows that $\ell(p_j) = \ell(r(s)) = \ell(s)$, so p_{j+1} must be to the right of p_j. Therefore, by Observation B.2, we can conclude that $p_{j+1} \in \{\mathrm{cv}(p_j), r(p_j)\}$. Now, since $\ell(s)$ is higher than $r(s)$, it can also see $\mathrm{cv}(p_j)$ and $r(p_j)$, in particular, it can see p_{j+1}. Hence, $\langle s, \ell(s), p_{j+1}, \ldots, p_k\rangle$ is a valid path of length at most $|\pi|$, so there exists a shortest path from s to t through $\ell(s)$. □

Proof (of Lemma 3.3). Without loss of generality, t lies strictly to the right of s. First, assume that $\mathrm{nd}(s,t)$ is ℓ-convex. Since s can see $\mathrm{nd}(s,t)$ and since $\mathrm{nd}(s,t)$ is to the right of s, it follows that s and $\mathrm{nd}(s,t)$ share the same vertical edge. Then, $\mathrm{cv}(\mathrm{nd}(s,t))$ is also visible from s and its horizontal distance to t is smaller. This contradicts the definition of $\mathrm{nd}(s,t)$.

Next, assume that $\mathrm{nd}(s,t)$ is r-convex. Let v be the reflex vertex sharing a vertical edge with $\mathrm{nd}(s,t)$. Then, $N(\mathrm{nd}(s,t)) \subseteq N(v)$ and $v \in N(s)$. Furthermore, since t is strictly to the right of v but still inside $I(s)$, the vertices v and $r(s)$ must be distinct. Thus, $v_y < s_y$, so that $\mathrm{cv}(v)$ is also visible from s. Moreover, the horizontal distance of $\mathrm{cv}(v)$ and t is smaller than the horizontal distance of $\mathrm{nd}(s,t)$ and t. This again contradicts the definition of $\mathrm{nd}(s,t)$. The first part of the lemma follows.

It remains to show that $\mathrm{fd}(s,t) = r(\mathrm{nd}(s,t))$. First of all, $\mathrm{fd}(s,t)$ is higher than $\mathrm{nd}(s,t)$, since otherwise $\mathrm{fd}(s,t)$ would not be visible from s. Moreover, if $\mathrm{nd}(s,t)$ and $\mathrm{fd}(s,t)$ are not co-visible, there must be a vertex v strictly between $\mathrm{nd}(s,t)$ and $\mathrm{fd}(s,t)$ that is visible from s and higher than $\mathrm{nd}(s,t)$. Now, either $t \in [\mathrm{nd}(s,t), v]$ or $t \in [v, \mathrm{fd}(s,t)]$. In the first case, the horizontal distance between v and t is smaller than between t and $\mathrm{fd}(s,t)$, and in the second case, the horizontal distance between v and t is smaller than between t and $\mathrm{nd}(s,t)$. Either case leads to a contradiction. Therefore, $\mathrm{fd}(s,t)$ is higher than $\mathrm{nd}(s,t)$, strictly to the right of $\mathrm{nd}(s,t)$ and visible from $\mathrm{nd}(s,t)$. Thus, Observation B.2 gives $\mathrm{fd}(s,t) = r(\mathrm{nd}(s,t))$. □

C Additional Material for Section 4

C.1 Visibility in Double Histograms

The structure of the shortest paths in double histograms can be much more involved than in simple histograms; in particular, Lemma 3.1 does not hold anymore. However, the following observations provide some structural insight that can be used for an efficient routing scheme.

Observation C.1. *Two vertices v, w are co-visible if and only if $v \in I(w)$ and $w \in I(v)$.*

Proof. The forward direction is immediate, as co-visibility implies $v \in N(w) \subseteq I(w)$ and $w \in N(v) \subseteq I(v)$. For the backward direction, let Q be the rectangle spanned by v and w. Since $v \in I(w)$ and $w \in I(v)$, the upper and lower boundary of Q do not contain a point outside P. As P is a double histogram, this implies that the left and right boundary of Q also do not contain any point outside P. The claim follows since P has no holes. □

Observation C.2. *Let a, b, c, and d be vertices in P with $a_x \leq b_x \leq c_x \leq d_x$. If $a \in I(c)$ and $d \in I(b)$, then b and c are co-visible.*

Proof. This follows immediately from Observation C.1. □

Observation C.3. *The intervals on one side of P form a laminar family, i.e., for any two vertices v and w on the same side of the base line, we have (i) $I(v) \cap I(w) = \emptyset$, (ii) $I(v) \subseteq I(w)$, or (iii) $I(w) \subseteq I(v)$.*

Proof. Suppose there are two vertices v and w on the same side of P with $\ell(v)_x < \ell(w)_x \leq r(v)_x < r(w)_x$. By Observation C.2, $\ell(w)$ and $r(v)$ are co-visible. Since $\ell(w)$ and $r(v)$ are on the same side of P, either $r(v)$ cannot see any vertex to the left of $\ell(w)$ or $\ell(w)$ cannot see any vertex to the right of $r(v)$. This contradicts the fact that the $\ell(v)$ and $r(v)$ as well as $\ell(w)$ and $r(w)$ must be co-visible. □

C.2 Missing Proofs

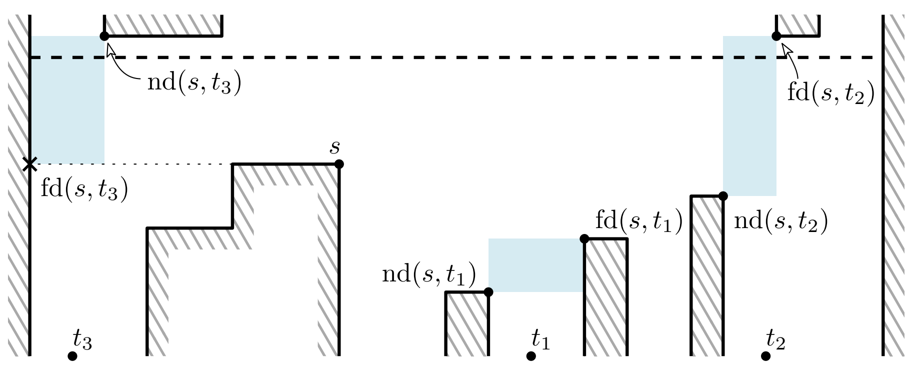

Fig. 5. The far and the near dominator can see each other.

Proof (of Lemma 4.1). Without loss of generality, t is strictly to the right of s, see Fig. 5. Suppose for a contradiction that $r(\text{nd}(s,t))$ is strictly left of $\text{fd}(s,t)$. Then, we get $r(\text{nd}(s,t)) \in I(s)$. Also, $s \in I(\text{nd}(s,t)) \subseteq I(r(\text{nd}(s,t)))$. Hence, by Observation C.1, s can see $r(\text{nd}(s,t))$. But then $r(\text{nd}(s,t))$ is a vertex strictly between the near and far dominator visible from s, contradicting the choice of the dominators. Thus, $s_x \leq \text{nd}(s,t)_x \leq \text{fd}(s,t)_x \leq r(\text{nd}(s,t))_x$, and Observation C.2 gives the result. □

Proof (of Lemma 4.2). Without loss of generality, t is to the right of s. Let $\pi : \langle s = p_0, \ldots, p_k = t\rangle$ be a shortest path from s to t, and let p_j be the last vertex outside of $I(s,t)$. If $j = 0$, then p_{j+1} must be one of the dominators, since by definition they are the only vertices in $I(s,t)$ visible from s. Now, assume $j \geq 1$. If p_j is to the left of $\mathrm{nd}(s,t)$, we apply Lemmas 4.1 and C.2 on the four points p_j, $\mathrm{nd}(s,t)$, p_{j+1}, and $\mathrm{fd}(s,t)$ to conclude that $\mathrm{nd}(s,t)$ can see p_{j+1}. Symmetrically, if p_j is to the right of $\mathrm{fd}(s,t)$, the same argument shows that the far dominator can see p_{j+1}. Thus, depending on the position of p_j we can exchange the subpath $p_1, \ldots, p_j$ in π by $\mathrm{nd}(s,t)$ or $\mathrm{fd}(s,t)$ and get a valid path of length $k-j+1 \leq k$. The second part of the lemma follows because p_j cannot be to the right of $\mathrm{fd}(s,t)$, if $\mathrm{fd}(s,t)$ is not a vertex but a point on the right boundary. □

Proof (of Lemma 4.3). Without loss of generality, t is to the right of s; see Fig. 6. By Lemma 4.2, $\mathrm{nd}(s,t)$ lies on a shortest path from s to t. Let $\langle s = p_0, \mathrm{nd}(s,t) = p_1, p_2, \ldots, p_k = t\rangle$ be such a shortest path. We claim that $\mathrm{fd}^2(s,t)$ can see p_2. Then, $\langle \mathrm{fd}^2(s,t), p_2, \ldots, p_k = t\rangle$ is a valid path of length $k-1 = d(s,t)-1$. To prove that $\mathrm{fd}^2(s,t)$ can indeed see p_2, we show that $p_2 \in I\big(\mathrm{fd}^2(s,t)\big)$ and $\mathrm{fd}^2(s,t) \in I(p_2)$ and then apply Observation C.1.

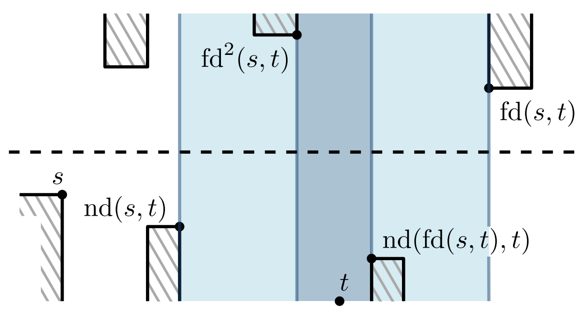

Fig. 6. $\mathrm{fd}^2(s,t)$ lies between $\mathrm{nd}(s,t)$ and $\mathrm{fd}(s,t)$ and is closer to t than s. The darker region is $I(\mathrm{fd}(s,t),t)$ and a subset of $I(s,t)$, the brighter region.

First, we show $p_2 \in I(\mathrm{fd}(s,t),t)$ by contradiction. Thus, suppose that $p_2 \notin I(\mathrm{fd}(s,t),t)$. Since $t \in I(\mathrm{fd}(s,t),t)$, there is a $j \geq 2$ with $p_{j+1} \in I(\mathrm{fd}(s,t),t)$ and $p_j \notin I(\mathrm{fd}(s,t),t)$. First, if $p_{j,x} < \mathrm{fd}^2(s,t)_x$, then $p_{j,x} < \mathrm{fd}^2(s,t)_x \leq p_{j+1,x} \leq \mathrm{nd}(\mathrm{fd}(s,t),t)_x$. By Lemma 4.1, $\mathrm{fd}^2(s,t)$ and $\mathrm{nd}\,\big(\mathrm{fd}(s,t),t\big)$ are co-visible, so Observation C.2 implies that $\mathrm{fd}^2(s,t)$ and p_{j+1} are co-visible. Then it follows that $\langle s, \mathrm{fd}(s,t), \mathrm{fd}^2(s,t), p_{j+1}, \ldots, p_k = t\rangle$ is a valid path of length $k-j+2 \leq k$, contradicting the assumption that $\mathrm{fd}(s,t)$ is not on a shortest path. If $\mathrm{nd}(\mathrm{fd}(s,t),t)_x < p_{j,x}$, it follows with the same reasoning that $\mathrm{nd}(\mathrm{fd}(s,t),t)$ and p_{j+1} are co-visible then $\mathrm{fd}(s,t)$ is on $\langle s, \mathrm{fd}(s,t), \mathrm{nd}(\mathrm{fd}(s,t),t), p_{j+1}, \ldots, p_k = t\rangle$ which is a valid path of length $k-j+2 \leq k$. This again contradicts the assumption. Now, since $I(\mathrm{fd}(s,t),t) = \big[\mathrm{fd}^2(s,t), \mathrm{nd}(\mathrm{fd}(s,t),t)\big] \subseteq I(\mathrm{fd}^2(s,t))$, we get $p_2 \in I(\mathrm{fd}^2(s,t))$. Since p_2 sees $\mathrm{nd}(s,t)$ which is to the left of $\mathrm{fd}^2(s,t)$ and since p_2 is in $I(\mathrm{fd}(s,t),t)$, and thus to the right of $\mathrm{fd}^2(s,t)$, it follows that $\mathrm{fd}^2(s,t) \in I(p_2)$. □

Proof (of Lemma 4.4). We focus on ℓ^{i-1} and a^i. We show that $\ell^{i-1} \in I(a^i)$ and $a^i \in I(\ell^{i-1})$; the lemma follows from Observation C.1. The claim $\ell^{i-1} \in I(a^i)$ is due to the facts that $[\ell^i, a^i] \subseteq I(a^i)$ and $\ell^{i-1} \in [\ell^i, a^i]$ (this holds also for $i = 1$, as then $\ell^{i-1} = a^i$). Next, since $a^{i-1} \in I(\ell^{i-1})$, the vertex a^{i-1} is to the left of $r(\ell^{i-1})$; and since $\ell^{i-1} \in I(r(\ell^{i-1}))$, the point $\ell(r(\ell^{i-1}))$ is to the left of ℓ^{i-1}. Thus, if $r(\ell^{i-1})$ is visible from s, we have $a^i = r(\ell^{i-1})$, by the definition of a^i. On the other hand, if $r(\ell^{i-1})$ is not visible from s, the visibility must be blocked by $r(s)$, and then $a^i = r(s)$. In either case, we have $a^i \in [a^{i-1}, r(\ell^{i-1})] \subseteq I(\ell^{i-1})$, as desired. □

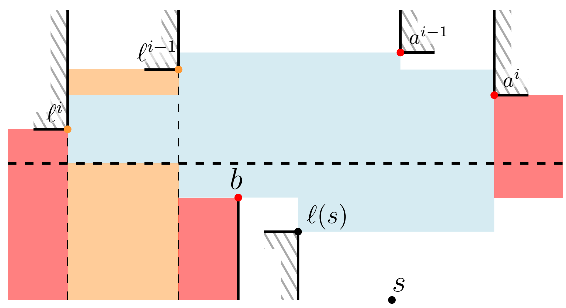

Fig. 7. The vertex a^i is on a shortest path. The vertex p_j can lie in the red regions, the vertex p_{j+1} can lie in the orange region, and the blue region cannot contain any point outside of P.

Proof (of Lemma 4.5). We focus on the first statement; see Fig. 7. Let $\pi : \langle s = p_0, \dots, p_k = t\rangle$ be a shortest path from s to t, and let p_j be the last vertex on π outside of $[\ell^i, \ell^{i-1}]$. If $j = 0$, then p_{j+1} must be $\ell^0 = \ell(s)$, because this is the only vertex ℓ^{i-1} visible from s. Then, $i = 1$ and $a^i = \ell(s)$ is on π. From now on, we assume that $j \geq 1$.

First, suppose that p_{j+1} and a^i are co-visible. Then $\langle s, a^i, p_{j+1}, \dots, p_k\rangle$ is a path from s to t that uses a^i and has length $k - j + 1 \leq k$. Second, suppose that p_{j+1} and a^i are not co-visible. Then, the contrapositive of Observation C.2 applied to the four points ℓ^i, p_{j+1}, a^i, and p_j shows that p_j is strictly to the left of a^i. There are two subcases, depending on whether p_j is strictly to the left of ℓ^i or strictly to the right of ℓ^{i-1}.

If p_j is strictly to the left of ℓ^i, then $j \geq 2$, since ℓ^i is to the left of $\ell^0 = \ell(s)$ and we need at least two hops to reach a point strictly to the left of $\ell(s)$ from s. We apply Observation C.2 on the four points p_j, ℓ^i, p_{j+1}, and a^i, and get that ℓ^i and p_{j+1} are co-visible. Hence, $\langle s, a^i, \ell^i, p_{j+1}, \dots, p_k\rangle$ is a path that uses a^i and has length $k - j + 2 \leq k$.

Finally, assume that p_j is strictly to the right of ℓ^{i-1}. By Lemma 4.4, a^i can see ℓ^{i-1}. Thus, $p_{j+1} \neq \ell^{i-1}$ and there is no vertex strictly between ℓ^{i-1} and a^i

on the same side as ℓ^{i-1} that can see a vertex strictly to the left of ℓ^{i-1}. Thus, p_j and a^{i-1} are on different sides of the base line. Let b be the rightmost vertex that (i) lies on the same side of P as p_j; (ii) is strictly between ℓ^{i-1} and a^i; (iii) is closest to the base line. The vertex b exists (since p^j is a candidate), is not visible from s (because b is strictly left of a^i and can see strictly left of ℓ^{i-1}); and thus strictly left of $\ell(s)$. The vertex p_j cannot be strictly to the right of b, as otherwise b would obstruct visiblity between p_j and p_{j+1}. We conclude that $j \geq 2$, since we need at least two hops to reach a point strictly to the left of $\ell(s)$ from s. If $p_j \in \{b, \mathrm{cv}(b)\}$, a^i can see p_j and thus, $\langle s, a^i, p_j, p_{j+1}, \ldots, p_k\rangle$ is a path of length $k - j + 2 \leq k$ using a^i. If $p_j \notin \{b, \mathrm{cv}(b)\}$, then b is strictly closer to the base line than p_j. Then, we have $j \geq 3$, because we need two hops to cross the vertical line through $\ell(s)$ and one more hop to cross the horizontal line through b. We apply Observation C.2 on the four points p_{j+1}, ℓ^{i-1}, p_j, and b to conclude that ℓ^{i-1} can see p_j. Hence, $\langle s, a^i, \ell^{i-1}, p_j, p_{j+1}, \ldots, p_k\rangle$ is a path of length $k - j + 3 \leq k$ using a^i. □

Proof (of Lemma 4.6). We have $I^k(s) \subseteq I(\mathrm{bd}^k(s))$, since by definition, interval $I^k(s)$ contains no vertex that is on the same side as $\mathrm{bd}^k(s)$ and strictly closer to the base line, so no vertex can obstruct horizontal visibility of $\mathrm{bd}^k(s)$ in $I^k(s)$. Analogously, $I^k(s) \subseteq I(\mathrm{td}^k(s))$, as desired.

By definition and the first part, $\mathrm{td}^k(s) \in I^k(s) \subseteq I(\mathrm{bd}^k(s))$ and $\mathrm{bd}^k(s) \in I^k(s) \subseteq I(\mathrm{td}^k(s))$. The claim now follows from Observation C.1. □

Proof (of Lemma 4.7). We begin by showing that

$$I\big(\mathrm{bd}^2(s)\big) = I\big(\mathrm{bd}(\mathrm{td}(s))\big) \cup I\big(\mathrm{bd}(\mathrm{bd}(s))\big). \tag{1}$$

If $\mathrm{bd}(s)$ is above the base line, then $\mathrm{bd}(s) = \mathrm{td}(s)$ and $I^2(s) = I(\mathrm{bd}(s)) \cup I(\mathrm{td}(s)) = I(\mathrm{td}(s))$. The definition of $\mathrm{bd}^2(s)$ then gives $\mathrm{bd}^2(s) = \mathrm{bd}(\mathrm{td}(s))$, and (1) follows.

If $\mathrm{bd}(s)$ is below the base line, the vertex $b_1 = \mathrm{bd}(\mathrm{bd}(s))$ is below the base line. Let $b_2 = \mathrm{bd}(\mathrm{td}(s))$. By Lemma 4.6, $\mathrm{bd}(s)$ and $\mathrm{td}(s)$ are co-visible, so $\mathrm{bd}(s) \in I(\mathrm{td}(s))^-$. Therefore, b_2 is below the base line. Since $I(b_1)$ and $I(b_2)$ are not disjoint (both contain s) and since b_1 and b_2 are on the same side of the base line, Observation C.3 gives $I(b_1) \subseteq I(b_2)$ or $I(b_2) \subseteq I(b_1)$. Because $\mathrm{bd}^2(s)$ is the highest vertex in $\big(I(\mathrm{bd}(s)) \cup I(\mathrm{td}(s))\big)^-$, we get that $\mathrm{bd}^2(s)$ is b_1 or b_2, and (1) follows also in this case. Symmetrically, we have

$$I(\mathrm{td}^2(s)) = I(\mathrm{td}(\mathrm{td}(s))) \cup I(\mathrm{td}(\mathrm{bd}(s))). \tag{2}$$

We use the definitions and (1,2) to get

$$\begin{aligned} I^3(s) &= I\big(\mathrm{bd}^2(s)\big) \cup I\big(\mathrm{td}^2(s)\big) \\ &= I\big(\mathrm{bd}(\mathrm{td}(s))\big) \cup I\big(\mathrm{bd}(\mathrm{bd}(s))\big) \cup I\big(\mathrm{td}(\mathrm{td}(s))\big) \cup I\big(\mathrm{td}(\mathrm{bd}(s))\big) \\ &= I^2\big(\mathrm{bd}(s)\big) \cup I^2\big(\mathrm{td}(s)\big), \end{aligned}$$

as desired. □

Proof (of Lemma 4.8). We show that for any $j \geq 0$ and any vertex $v \in I^j(s)$, we have $N(v) \subseteq I^{j+1}(s)$. The lemma then follows by induction. If $v \in I^j(s)^-$, then $\mathrm{bd}^j(s)$ is on the lower side, and by definition, $I(v) \subseteq I(\mathrm{bd}^j(s))$. If $v \in I^j(s)^+$, by a similar argument $I(v) \subseteq I(\mathrm{td}^j(s))$. Thus, $N(v) \subseteq I(v) \subseteq I(\mathrm{bd}^j(s)) \cup I(\mathrm{td}^j(s)) = I^{j+1}(s)$, as desired. □

Proof (of Lemma 4.9). On the one hand, $d(s, \mathrm{bd}^k(s)) \leq |\pi_b(s,k)| = k$ and $d(s, \mathrm{td}^k(s)) \leq |\pi_t(s,k)| = k$. On the other hand, we show that $\mathrm{bd}^k(s) \notin I^{k-1}(s)$ and $\mathrm{td}^k(s) \notin I^{k-1}(s)$. The claim then follows by the contrapositive of Lemma 4.8.

Case 1: First, assume that $\mathrm{bd}^{k-1}(s) \in I^{k-1}(s)^-$. Since $I(\mathrm{bd}^{k-1}(s)) \neq P$, at least one of its bounding points is a vertex v contained in $I^k(s)$. Then, v is strictly closer to the base line than $\mathrm{bd}^{k-1}(s)$, and since v is a candidate for $\mathrm{bd}^k(s)$, the same applies to $\mathrm{bd}^k(s)$. It follows that $\mathrm{bd}^k(s) \notin I^{k-1}(s)$. Similarly, we get that if $\mathrm{td}^{k-1}(s) \in I^{k-1}(s)^+$, the vertex $\mathrm{td}^k(s)$ is not in $I^{k-1}(s)$.

Case 2: Second, assume that $\mathrm{bd}^{k-1}(s) \in I^{k-1}(s)^+$. Then, we have that $\mathrm{bd}^k(s) \notin I^{k-1}(s)^-$, since this set is empty. Thus, suppose for a contradiction that $\mathrm{bd}^k(s) \in I^{k-1}(s)^+$. This can only be the case if $\mathrm{bd}^{k-1}(s) = \mathrm{td}^{k-1}(s)$ and $\mathrm{bd}^k(s) = \mathrm{td}^k(s)$. However, in Case 1 we showed that $\mathrm{td}^k(s) \notin I^{k-1}(s)^+$ if $\mathrm{td}^{k-1}(s) \in I^{k-1}(s)^+$. Hence, $\mathrm{bd}^k(s) \notin I^{k-1}(s)^+$, as desired. □

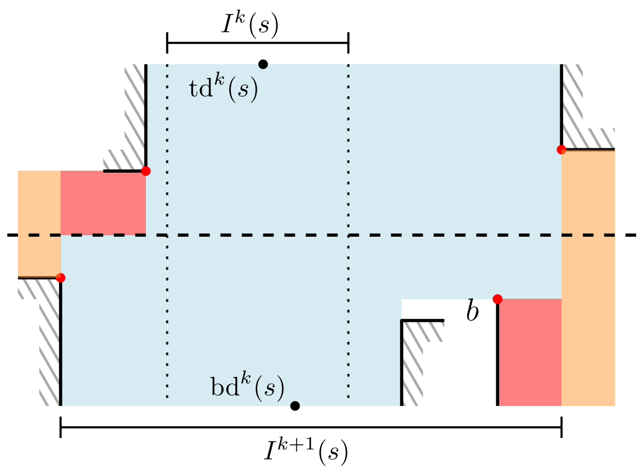

Fig. 8. $\mathrm{bd}^k(s)$ or $\mathrm{td}^k(s)$ is on a shortest path. The vertex p_j lies in one of the red regions, p_{j+1} lies in one of the orange regions, and the blue region cannot contain any point outside of P.

Proof (of Lemma 4.10). First, observe that $I^{k+1}(s) = I(\mathrm{bd}^k(s)) \cup I(\mathrm{td}^k(s)) \neq P$, as $t \notin I^{k+1}(s)$. Let $\pi : \langle s = p_0, \ldots, p_m = t \rangle$ be a shortest path from s to t, and p_j the last vertex in $I^{k+1}(s)$. Without loss of generality, p_{j+1} is strictly to the right of s. By Lemma 4.8, we get that p_j is not in $I^k(s)$ and thus, again by Lemma 4.8, we have $j \geq d(s, p_j) \geq k+1$.

First, suppose that p_j and $\mathrm{bd}^k(s)$ (resp. $\mathrm{td}^k(s)$) are co-visible. Then, $\pi_b(s,k) \circ \langle p_j, \dots, t\rangle$ (resp. $\pi_t(s,k) \circ \langle p_j, \dots, t\rangle$) is a valid path of length $k+1+(m-j) \leq m$. Here, $\circ$ concatenates two paths. Second, suppose p_j be visible from neither $\mathrm{bd}^k(s)$ nor $\mathrm{td}^k(s)$. First, we claim that p_j is strictly to the right of $\mathrm{bd}^k(s)$ and $\mathrm{td}^k(s)$. Otherwise, since p_{j+1} is strictly to the right of both dominators, we would get $\mathrm{bd}^k(s), \mathrm{td}^k(s) \in I(p_j)$. Moreover, $p_j \in I^{k+1}(s) = I(\mathrm{bd}^k(s)) \cup I(\mathrm{td}^k(s))$. Observation C.1 now would imply that p_j can see $\mathrm{bd}^k(s)$ or $\mathrm{td}^k(s)$—a contradiction. The claim follows. Next, we claim $\mathrm{bd}^k(s) \neq \mathrm{td}^k(s)$. If not, $p_j \in I(\mathrm{bd}^k(s)) = I^{k+1}(s)$, and since p_j can see a point outside of $I^{k+1}(s)$, we would get $p_j \in \{\ell(\mathrm{bd}^k(s)), r(\mathrm{bd}^k(s))\}$, which again contradicts our assumption that p_j cannot see $\mathrm{bd}^k(s)$. The claim follows. There are two cases, depending on which dominator sees further to the right.

Case 1: $r(\mathrm{bd}^k(s))_x < r(\mathrm{td}^k(s))_x$; see Fig. 8. Let b be the leftmost vertex in $[r(\mathrm{bd}^k(s)), r(\mathrm{td}^k(s))]^-$ closest to the base line. Observe that b is strictly to the right of $I^k(s)$, because $r(\mathrm{bd}^k(s))$ is strictly to the right of $I^k(s)$. Since p_j is not visible from $\mathrm{td}^k(s)$, it has to be strictly to the right of and strictly below b. Next, we claim that no vertex $v \in I^k(s)$ can see p_j. If one could, by Observation C.1, we would have $v \in I(p_j)$. But since p_j is strictly to the right of and strictly below b, then v would be to the right of b, which is impossible. This shows the claim. Thus, by Lemma 4.8, $j \geq d(s,p_j) \geq k+2$. We apply Observation C.2 to $\mathrm{td}^k(s)$, p_j, $r(\mathrm{td}^k(s))$ and p_{j+1} and get that $r(\mathrm{td}^k(s))$ can see p_j. Therefore, $\pi_t(s,k) \circ \langle r(\mathrm{td}^k(s)), p_j, \dots, t\rangle$ is a valid path of length $k+1+(1+m-j) \leq m$.

Case 2: $r(\mathrm{td}^k(s))_x < r(\mathrm{bd}^k(s))_x$. Let b be the leftmost vertex in the interval $[r(\mathrm{td}^k(s)), r(\mathrm{bd}^k(s))]^+$ closest to the base line. Observe that b is strictly to the right of $I^k(s)$, because $r(\mathrm{td}^k(s))$ is strictly to the right of $I^k(s)$. Since p_j is not visible from $\mathrm{bd}^k(s)$, it has to be strictly to the right of and strictly above b. Next, we claim that no vertex $v \in I^k(s)$ can see p_j. If one could, by Observation C.1, we would have $v \in I(p_j)$. But since p_j is strictly to the right of and strictly above b, then v would be to the right of b, which is impossible. This shows the claim. Thus, by Lemma 4.8, $j \geq d(s,p_j) \geq k+2$. We apply Observation C.2 to $\mathrm{bd}^k(s)$, p_j, $r(\mathrm{bd}^k(s))$ and p_{j+1} and get that $r(\mathrm{bd}^k(s))$ can see p_j. Therefore, $\pi_b(s,k) \circ \langle r(\mathrm{bd}^k(s)), p_j, \dots, t\rangle$ is a valid path of length $k+1+(1+m-j) \leq m$, as desired. ☐

Proof (of Lemma 4.11). First, if $t \in N(s)$, then we take one hop and decrease the distance to 0. Second, suppose that $t \in I(s) \setminus N(s)$. If $\mathrm{fd}(s,t)$ is not a vertex, the next vertex is $\mathrm{nd}(s,t)$, which is on a shortest path from s to t due to Lemma 4.2. Otherwise, $\mathrm{fd}(s,t)$ is the next vertex. If $\mathrm{fd}(s,t)$ is on a shortest path from s to t, we are done. Otherwise, t is not visible from $\mathrm{fd}(s,t)$, so $\mathrm{fd}^2(s,t)$ has to be the second vertex on the routed path. By Lemma 4.3, we have $d(\mathrm{fd}^2(s,t),t) = d(s,t)-1$. Third, if $t \in I^2(s) \setminus I(s)$, there is an $i \geq 1$ such that the next vertex is either a^i or b^i. By Lemma 4.5, this vertex is on a shortest path. Fourth, assume $t \in I^3(s) \setminus I^2(s)$. Let v_1 and v_2 be the next two vertices on the routing path. We use Lemma 4.6 and Lemma 4.10 to conclude $d(v_1, t) \leq d(s, t)$, as v_1 is either $\mathrm{td}(s)$ or $\mathrm{bd}(s)$. Due to the construction of the routing function, we have $t \in I^2(v_1) \setminus I(v_1)$. Thus, there is an $i \geq 1$, such that $v_2 = a^i(v_1)$ or $v_2 = b^i(v_1)$. By Lemma 4.5, the vertex v_2 is on a shortest path from v_1 to t and we can conclude $d(v_2, t) = d(v_1, t) - 1 \leq d(s, t) - 1$. Last, assume $t \notin I^3(s)$. Then, the packet is routed to a vertex $p \in \{\mathrm{bd}(s), \mathrm{td}(s)\}$, whichever is on a shortest path to $\mathrm{bd}^2(s)$, and then $\mathrm{bd}^2(s)$. Lemmas 4.6 and 4.10 give $d(\mathrm{bd}^2(s), t) \leq d(s, t) - 1$, as claimed. $\square$