Simple, Efficient Routing Schemes for All-Optical Networks*

Michele Flammini†
Department of Mathematics
University of L'Aquila, Via Vetoio loc. Coppito
I-67100 L'Aquila, Italy

Christian Scheideler‡
Heinz Nixdorf Institute and
Dept. of Mathematics & Computer Science
Paderborn University
33095 Paderborn, Germany

Abstract

All-optical networks promise data transmission rates several orders of magnitude higher than current networks. The key to high transmission rates in these networks is to maintain the signal in optical form, thereby avoiding the prohibitive overhead of conversion to and from the electrical form, and to exploit the large bandwidth of optical fibers by sending many signals at different frequencies along the same optical link. Optical technology, however, is not as mature as electronic technology. Hence it is important to understand how efficiently simple routing elements can be used for all-optical communication. In this paper, we consider two types of routing elements. Both types can move messages at different wavelengths to different directions. If in the first type a message wants to use an outgoing link that is already occupied by another message using the same wavelength, the arriving message is eliminated (and therefore has to be rerouted). The second type can evaluate priorities of messages. If more than one message wants to use the same wavelength at the same time then the message with the highest priority wins. We prove nearly matching upper and lower bounds for the runtime of a simple and efficient protocol for both types of routing elements, and apply our results to meshes, butterflies, and node-symmetric networks.

1 Introduction

The subject of this paper is to present and analyze a simple protocol for sending messages in an emerging generation of networks known as *all-optical networks* [6, 10, 16, 20, 30, 33]. These networks promise data transmission rates several orders of magnitudes higher than current networks. The key to high speeds in these networks is to maintain the signal in optical form, thereby avoiding the prohibitive overhead of conversion to and from the electrical form. (Traditional networks use the electrical form to switch signals along routes, and to restore signal strength. Signals can be modulated electronically at a maximum bit rate of about 50 Gbit/s, while the optical fiber bandwidth is about 25 THz [7].) The

* Authors supported in part by the EU ESPRIT Long Term Research Project 20244 (ALCOM-IT).

† email: flammini@univaq.it. Supported in part by the EU TMR Research Training Grant N. ERBFMBICT960861, the Project SLOOP I3S-CNRS/INRIA/Université de Nice–Sophia Antipolis, France, and the Italian MURST 40% project "Algoritmi, Modelli di Calcolo e Strutture Informative".

‡ email: chrsch@uni-paderborn.de. Supported in part by DFG-Sonderforschungsbereich 376 "Massive Parallelität: Algorithmen, Entwurfsmethoden, Anwendungen".

high bandwidth of the optical fiber is utilized through *wavelength-division multiplexing*: two signals connecting different source-destination pairs may share a link, provided they are transmitted on carriers having different frequencies (or wavelengths) of light.

The major applications for such networks are in video conferencing, scientific visualization and real-time medical imaging, high-speed supercomputing and distributed computing [16, 33, 12]. We consider routing elements that are capable of directing messages at different wavelengths to different destinations and detecting collisions of messages. A routing element (or *router* in short) consists of wavelength-selective *switches* and *couplers*.

The task of the switches is to direct different wavelengths to different directions. Several types of optical switches have already been developed [19, 5].

The task of the couplers is to combine the signals from many incoming optical fibers into one outgoing optical fiber. Since we do not want to rely on central control, collisions might occur, that is, two or more signals from different incoming fibers use the same wavelength. In our design of protocols we will consider two different strategies to avoid collisions:

- If a message that arrives at a coupler uses a wavelength already used by another message traversing the coupler, the new message is eliminated. This can be realized with the help of detector arrays that tell the electronic control of the coupler which wavelengths are currently used, and wavelength-selective filters at each incoming fiber.

- If a message that arrives at a coupler uses a wavelength already used by another message traversing the coupler, the message with higher priority is forwarded and the other suspended. There are prototypes of all-optical routing elements in which priorities are implemented by giving different powers to the messages (see, e.g., [21]). However, these routers are more complicated, and therefore it would be worth knowing whether priorities can improve the routing performance.

We call a coupler using the first rule a *serve-first coupler* and we call it a *priority coupler* otherwise. The following picture illustrates how a 2×2 router can be built by switches and couplers.

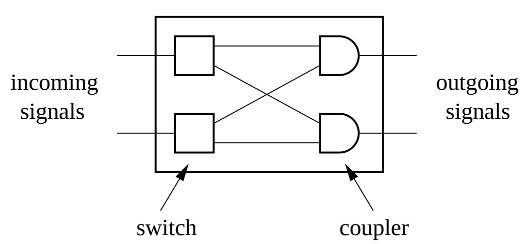

Figure 1: A 2×2 router.

1.1 The Model

We model the topology of an optical network as an undirected graph $G = (V, E)$ where each node in V represents a router (that is connected to a processor) and each edge in E represents two optical links, one in each direction. Each node in V contains an *injection buffer* and a *delivery buffer*. Initially, each message is stored in the injection buffer of its source. Once a message reaches its destination, it is stored in the destination's delivery buffer. During the routing a message cannot be buffered and therefore has to be either moved forward or eliminated. Since we do not want to convert messages to and from the electrical form, we do not require the nodes to operate in discrete, synchronous time steps. Instead, we just need to assume that the nodes are fast enough to operate correctly according to one of the collision rules defined above. Hence one time step within our model is defined as the time some fixed amount of bits (later called *flit*) of the message needs to traverse a link.

In general, a routing scheme consists of two (not necessarily independent) parts: The first part is responsible for selecting a path for each message, and the second part is responsible for sending the messages across their paths. Within our general framework, we assume that some suitable strategy for the path selection is given. (We will later give examples of how paths can be selected in specific networks.) The routing problem will be therefore defined by specifying a path collection $\mathcal{P}$, which is a multiset of paths in G. A path collection is called

- *short-cut free* if there is no subpath of a path that is shortcut by a subpath of another path in $\mathcal{P}$, and
- *leveled* if levels can be assigned to the nodes in $\mathcal{P}$ such that for every path in $\mathcal{P}$ every edge leads from a node in level i to a node in level $i+1$ for some $i \geq 0$.

A path collection is, for instance, always short-cut free if there are no two paths in it that meet, separate and meet again. Since this is mostly the case both in theory and practice, the class of short-cut free path collections is fairly general.

The problem is to route one message along each of the paths in $\mathcal{P}$ in such a way that the time required to route all messages is minimal. We measure the routing performance of our protocols by

- the number n of paths in $\mathcal{P}$,
- the *dilation* D of $\mathcal{P}$, that is, the length of the longest path in $\mathcal{P}$, and
- the *path congestion* $\tilde{C}$ of $\mathcal{P}$, that is, the maximum over all paths p in $\mathcal{P}$ of the number of paths that share an edge with p.

Note that the path congestion should not be mixed up with the commonly used *congestion* of a path collection, which is defined as the maximum over all edges e of the number of paths that contain e.

A major problem in all-optical networks is to interpret the address header of messages arriving at optical switches, since their switching time is still slow compared with the transmission speed in optical fibers. An approach investigated by AT&T [15, 17] and elsewhere employs a low bit-rate header which is read on the fly by a photodiode or a contact on a semiconductor amplifier. These electrical bits are fed to a controller that operates an optical switch that sends the unconverted optical data bits along the proper path.

A message might occupy several links on its way through the network. We therefore model the messages as *worms*, each of which consists of a sequence of fixed size units called *flits*. We assume that it takes one time step to send a flit along a link. The *length* of a worm is defined as the number of flits it contains. The first flit is called the *head* and the remaining flits are called the *body* of the worm. During the routing, a worm occupies a contiguous sequence of links along its path, one flit per link.

The number of wavelengths a router can handle is called the *bandwidth* of the router and denoted by B. As defined for the coupler above, we distinguish between two rules for the router: the *serve-first rule* and the *priority rule*.

1.2 Previous Results

All-optical routing problems have been considered for two basic network models: the *non-reconfigurable* or *switchless* networks, and the *reconfigurable* networks. In the first class of networks, a fixed set of wavelengths is assigned to every connection between any input and output of a router, whereas in the second class switches are allowed, that is, connections between the inputs and the outputs of a router can change. There are basically two possible ways of changing connections: either by simply switching wires or by switching wavelengths (see Figure 2 for an example). The first type of switch is called *elementary* switch and the second type is called *generalized* switch.

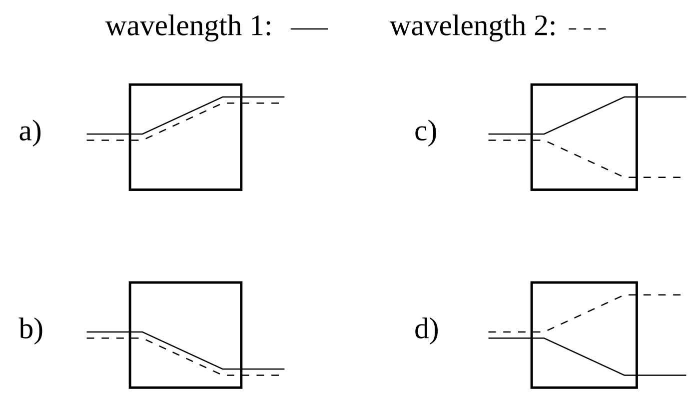

Figure 2: An elementary switch with two outputs only allows configurations a) and b), whereas a generalized switch allows all four configurations.

Clearly, the elementary switch cannot direct different wavelengths arriving at some input to different outputs, whereas the generalized switch can do this. Figure 3 gives an example of a non-reconfigurable and an elementary router.

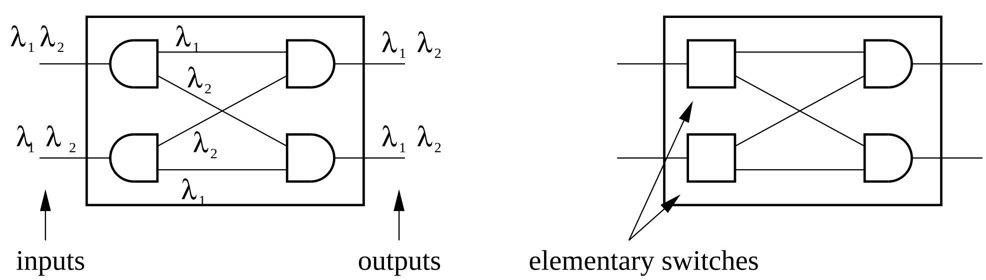

Figure 3: Structure of a switchless (left) and an elementary (right) router.

Almost all papers in the field of all-optical routing deal with the problem of assigning wavelengths to the paths of the messages such that no two paths use the same wavelength at an edge (i.e., conflicts among messages cannot occur). In this situation it is important to find networks and path selection strategies such that the number of wavelengths required is minimal.

The routing strategies developed for the class of non-reconfigurable networks can be separated into two categories: the *single hop* strategies and the *multi hop* strategies. In a single hop strategy, messages are not allowed to change their wavelength somewhere along their routing path, while in a multi hop strategy they are usually allowed to do so for a bounded number of times (each time is referred to as a *hop*).

For the class of single hop strategies, Barry and Humblet prove in [3] that, for any network, permutation routing requires $\Omega(\sqrt{n})$ wavelengths, where n is the number of nodes in the network. They also show that oblivious permutation routing can be done using $\lceil n/2 \rceil + 2$ wavelengths. Awerbuch *et* *al.* prove in [1] the existence of a switchless permutation network using $O(\sqrt{n \log n})$ wavelengths. They also show how to construct a switchless permutation network using $O(\sqrt{n} 2^{(\log n)^{0.8+o(1)}})$ wavelengths.

Multi hop strategies have been considered, e.g., by [8, 9, 37, 13, 14, 22]. Chlamtac *et al.* [8] study the problem of establishing multi hop paths for a given static and dynamic set of circuit demands. In another paper, Chlamtac *et al.* [9] consider the problem of embedding regular networks into the original fiber topology. They present bounds on the number of wavelengths required to simulate a regular topology. Furthermore, algorithms for embedding different regular topologies are described and their performances are evaluated. Zhang and Acampora [37] follow this line by studying two heuristic algorithms for embedding the hypercube into a physical fiber topology. Gerstel and Zaks [13, 14] study layouts for chains, rings, meshes and trees. Kranakis *et al.* [22] give asymptotically tight bounds on the number of hops required for the chain and the mesh given the number of available wavelengths (there expressed as congestion).

Within the field of reconfigurable networks, a number of papers [2, 10, 26] have formulated the routing problem for both elementary and generalized routers as combinatorial optimization problems. For networks with w available wavelengths and elementary routers, Barry and Humblet show in [4] that the number of 2×2 routers required to support permutation routing is $\Omega(n \log(n/w^2))$. Awerbuch *et al.* [1] prove the existence of a permutation network using $O(n \log \frac{n \log w}{w^2})$ routers and constructed a permutation network using $O(n \log(2^{(\log w)^{0.8+o(1)}}/w^2))$ routers. When the transmitters are fixed-tuned and the receivers are tunable, Pieris and Sasaki [31] show that the number of 2×2 routers required for permutation routing is $\Omega(n \log(n/w))$, and constructed such a network using $O(n \log(n/w))$ routers.

Pankaj proves in [29] that, if generalized routers are used, $\Omega(\log n)$ wavelengths are required for permutation routing. He also shows that rearrangeably non-blocking permutation routing* can be done with $O(\log^2 n)$ wavelengths and wide-sense non-blocking permutation routing† can be done with $O(\log^3 n)$ wavelengths in popular interconnection networks such as the shuffle exchange network, the DeBruijn network, and the hypercube. Awerbuch *et al.* prove in [1] a tight bound of $\Theta(\log n)$ for the number of wavelengths required for both rearrangeable and wide-sense non-blocking networks.

Raghavan and Upfal [32] prove results that establish a connection between the expansion of a network and the number of wavelengths required for routing on it, considering both elementary and generalized routers. In [34], Ramaswami and Sivarajan present a lower bound on the blocking probability for any so-called routing and wavelength assignment (RWA) algorithm if requests and terminations of connections arrive at random, and generalized routers are used. They study both the case that wavelength conversion is allowed and not allowed at the routers.

To our knowledge, nothing has been found out so far about the maximum number of trials to send a message to its destination given an arbitrary path collection and a fixed bandwidth, if wavelength conversion is not allowed. In case that wavelength conversion is allowed at every router, Cypher *et al.* [11] presented an online protocol that routes messages of length L along any simple path collection with congestion C and dilation D in time $O((L \cdot C \cdot D^{1/B} + (D+L)\log n)/B)$, w.h.p.‡. However, all-optical devices for wavelength conversion are still a research topic and might significantly increase the cost of a router. Therefore we want to show in this paper how far one can get without wavelength conversion.

* "Rearrangeably" means that for any permutation known completely in advance there is a set of paths and an assignment of wavelengths to the paths such that no two paths crossing an edge use the same wavelength.

† "Wide-sense" means that paths can be set up and wavelengths can be selected for the paths one after another so that it is still possible to connect any unused source with any unused destination in a way that no two paths crossing an edge use the same wavelength.

‡ By "with high probability" (or w.h.p. for short) we mean a probability of at least $1 - 1/n^k$ for any constant $k > 0$.

1.3 New Results

In this paper we investigate how much time is necessary to route messages to their destinations given an arbitrary short-cut free or leveled path collection with some fixed bandwidth in case that wavelength conversion is not allowed.

In order for a protocol to be simple and fast, the processors should avoid any coordination. That is, the processors should be able to decide locally and without any outside information when and how to send out their messages. Since messages cannot be buffered during the routing there are basically two types of local-control strategies for transmitting messages: assigning random initial delays or priorities to the messages. The following simple routing strategy uses both strategies to send worms along a fixed path collection using routers with bandwidth B (the parameter T will be specified later).

Trial-and-Failure Protocol:
all n worms are declared active
for $t = 1$ to T do:

- each active worm is sent out from its source with random startup delay in some suitably chosen range $[\Delta_t]$ using a random wavelength in $[B]$
- for every worm that completely reaches its destination, an acknowledgement is sent back to the source immediately afterwards
- every source that gets back an acknowledgement declares its worm as inactive

Let us call the execution of one for-loop one *round*. Clearly, round t requires at most $\Delta_t + 2(D + L)$ steps to be sure that either an acknowledgement of a successful worm reaches its source, or the worm or its acknowledgement has been (partly) discarded. (Note that if we use priority routers it can happen that worms are only partly discarded.)

Previously, only delay sequence arguments were used to analyze such protocols (see, e.g., [11, 35]). In this paper we use delay tree arguments that yield much more accurate upper bounds on the runtime. In particular, we are able to prove the following three results depending on the contention resolution rule. Their proofs can be found in Section 2 and Section 3. Let $\alpha = \tilde{C} + B(\frac{D}{L} + 1) + 2$ and $\beta = \alpha/\tilde{C} + 2$. The first theorem presents a nearly tight analysis of the protocol above for leveled path collections.

Main Theorem 1.1 *For any leveled path collection of size n with dilation D and path congestion $\tilde{C}$ using serve-first routers with bandwidth B the protocol above routes a worm of length L along each of these paths in time*

$$O\left(\frac{L \cdot \tilde{C}}{B} + \left(\sqrt{\log_\alpha n} + \log\log_\beta n\right)\left(D + L + \frac{L \log n}{B}\right)\right),$$

w.h.p. Furthermore there exists a leveled path collection such that, for any $L \geq 2$, the expected runtime is bounded by

$$\Omega\left(\frac{L \cdot \tilde{C}}{B} + \left(\sqrt{\log_\alpha n} + \log\log_\beta n\right)(D + L)\right).$$

As we will see later, the upper bound results from choosing $T = O(\sqrt{\log_\alpha n} + \log\log_\beta n)$ and $\Delta_t = O((L\tilde{C}/(2^t + \log n) + \log n)/B)$ for all $t \in \{1, \ldots, T\}$. (The upper bounds below decompose in a similar way.)

Since in contrast to leveled path collections it can happen in some short-cut free path collections that worms prevent each other from reaching their destinations, we get a slightly worse result for arbitrary short-cut free path collections.

Main Theorem 1.2 *For any short-cut free path collection of size n with dilation D and path congestion $\tilde{C}$ using serve-first routers with bandwidth B the protocol above routes a worm of length L along each of these paths in time*

$$O\left(\frac{L \cdot \tilde{C}}{B} + (\log_\alpha n + \log\log_\beta n)\left(D + L + \frac{L \log^{3/2} n}{B}\right)\right),$$

w.h.p. Furthermore there exists a short-cut free path collection such that, for any $L \geq 2$, the expected runtime is bounded by

$$\Omega\left(\frac{L \cdot \tilde{C}}{B} + (\log_\alpha n + \log\log_\beta n)(D + L)\right).$$

As will become clear in the proof of Main Theorem 1.2, for the case $L = 1$ or there are no directed loops consisting of less than $\sqrt{\log_\alpha n}$ subpaths, the upper bound in Main Theorem 1.2 can be reduced to the upper bound in Main Theorem 1.1. For any other situation, we also obtain this bound if we replace the serve-first routers by priority routers.

Main Theorem 1.3 *For any collection of n short-cut free paths with dilation D and path congestion $\tilde{C}$ using priority routers with bandwidth B the protocol above routes a worm of length L along each of these paths in time*

$$O\left(\frac{L \cdot \tilde{C}}{B} + \left(\sqrt{\log_\alpha n} + \log\log_\beta n\right)\left(D + L + \frac{L \log n}{B}\right)\right),$$

w.h.p. Furthermore there is a short-cut free path collection and a strategy for assigning priorities to the worms such that, for any $L \geq 2$, the expected runtime is bounded by

$$\Omega\left(\frac{L \cdot \tilde{C}}{B} + \left(\sqrt{\log_\alpha n} + \log\log_\beta n\right)(D + L)\right).$$

We will show in Section 2 that the upper bound holds for *any* assignment of priorities to the worms such that no two worms with the same priority can meet in one round, whether these priorities are changed from round to round, chosen randomly, or deterministically.

The main theorems indicate that for short-cut free path collections the priority rule is more powerful than the serve-first rule. Often, $\Omega(\frac{L \cdot \tilde{C}}{B} + D + L)$ is a lower bound for any protocol using serve-first or priority routers. In this case the runtime of our protocol is (asymptotically) optimal if $\tilde{C}$ is large enough compared to D and L. Note that, for instance, for the butterfly network of size N the average path congestion of permutation routing problems is $\Theta(\log^2 N)$, whereas its diameter is $O(\log N)$.

The upper and lower bounds in Main Theorems 1.1 and 1.3 will be proved in Section 2, and the upper and lower bound in Main Theorem 1.2 will be given in Section 3. In the following, we describe some applications of the trial-and-failure protocol.

1.4 Applications

In this section we demonstrate that our protocol can be efficiently applied to routing messages in some important classes of networks. First let us introduce some terminology. For any $n \in \mathbb{N}$ let $[n]$ denote the set $\{0, \ldots, n-1\}$. Given a network of size n, by "routing a function" we always mean that, given a function $f : [n] \to [n]$, send one message from node i to node $f(i)$ for all $i \in [n]$. Furthermore, "routing a q-function" means routing a function $f : [q] \times [n] \to [n]$, i.e., each node is the source of q messages.

A "random (q-)function" denotes a function that is chosen uniformly at random from the set of all possible (q-)functions. Given a network G, a *path system* of G is defined as a collection of paths that contains a path for every pair of nodes in G.

The results presented in the previous section can be applied, e.g., to node-symmetric networks. This class of networks is defined as follows.

Definition 1.4 *A network $G = (V, E)$ is called* node-symmetric *if for any pair of nodes $u, v \in V$ there exists an isomorphism $\varphi : V \to V$ with $\varphi(u) = v$ such that the graph $G_\varphi = (V, E_\varphi)$ with $E_\varphi = \{\{\varphi(x), \varphi(y)\} \mid \{x, y\} \in E\}$ is equal to G.*

Intuitively, node-symmetry means that a network looks the same from any node. Node-symmetric networks form a very general class and include most of the standard networks such as the d-dimensional torus, the wrap-around butterfly, the hypercube, etc. Furthermore, the best expanders that have an explicit construction are all node-symmetric (see, e.g., [24, 25, 28]). For node-symmetric networks we can show the following result.

Theorem 1.5 *For any bounded degree node-symmetric network of size n with diameter D using priority routers with bandwidth B there is an online protocol for routing a randomly chosen function in time*

$$O\left(\frac{L \cdot D^2}{B} + \left(\sqrt{\log_D n} + \log\log n\right)(D + L)\right),$$

w.h.p.

Proof. In [27] it is shown that for every node-symmetric network with diameter D there exists a shortcut free path system with optimal dilation D and the following property: The expected congestion caused at any edge e if paths are selected from this system according to a randomly chosen function is at most D. Using this property together with Chernoff bounds [18] it is easy to show that w.h.p. a randomly chosen function has a path collection with path congestion $O(D^2 + \log n)$, where n is the size of the network. Using this in the time bound of Main Theorem 1.3 yields the theorem. ∎

The previous best time bound for the case $B = 1$ was $O(L \cdot D^2 + (D + L)\log n)$ [11]. (Note that for $B > 1$ the protocols in [11] allow wavelength conversion which we do not allow here.) The result in Theorem 1.5 can be improved for d-dimensional meshes and tori.

Theorem 1.6 *For any d-dimensional mesh of side length n using serve-first routers with bandwidth B there is an online protocol for routing a randomly chosen function in time*

$$O\left(\frac{L \cdot d \cdot n}{B} + (\sqrt{d} + \log\log n)\left(d \cdot n + L + \frac{L \cdot d \log n}{B}\right)\right),$$

w.h.p.

Proof. Using techniques in [11], it is easy to show that there exists a routing strategy for routing a randomly chosen function that has a path congestion of $O(d \cdot n)$, w.h.p., and in which it cannot happen that some set of worms eliminate each other. Since the size N of a d-dimensional mesh with side length n is equal to n^d, it follows that

$$\sqrt{\log_\alpha N} = O\left(\sqrt{d \log_{dn} n}\right) = O\left(\sqrt{d}\right),$$

where α is chosen as in the main theorems. In case that $\sqrt{d} \leq \log\log N$ we have that $n \geq N^{1/\log\log N}$ and therefore $\log\log N = O(\log\log n)$. This concludes the proof. ∎

The previous best time bound for the case $B = 1$ was $O(L \cdot d \cdot n + (d \cdot n + L)\log n)$ [11], achieved with a similar protocol as our protocol. They can only show that, by using priorities, $O(\log n)$ rounds suffice w.h.p. to route all worms. Our result, however, implies that even without priorities all worms are routed within $O(\log\log n)$ rounds, w.h.p., which is an exponential improvement. In case that we use butterfly networks, we obtain the following result.

Theorem 1.7 *For any $\log n$-dimensional butterfly using serve-first routers with bandwidth B there is a leveled path system such that a randomly chosen q-function can be routed from the inputs to the outputs in time*

$$O\left(\frac{L \cdot q\log n}{B} + \sqrt{\frac{\log n}{\log(q\log n)}}\left(L + \log n + \frac{L\log n}{B}\right)\right),$$

w.h.p.

For $B = 1$, this improves for some cases the previous best time bound of $O(L \cdot q\log n + (L + \log n)\log n)$ [11].

2 Proof of Main Theorems 1.1 and 1.3

In this section we prove upper and lower bounds on the runtime of our protocol using serve-first routers in leveled path collections, or priority routers in short-cut free path collections. In order to simplify the presentation, we will concentrate on serve-first routers in leveled path collections, and note the analogy to routing with priority routers in short-cut free path collections whenever it is necessary.

So suppose we want to route worms of length L along a collection of n leveled paths with path congestion $\tilde{C}$ and dilation D, using serve-first routers with bandwidth B. In order to simplify the analysis we assume that a bandwidth of $2B$ is given, where B wavelengths are reserved for the messages and B wavelengths are reserved for the acknowledgements. (Or we could assume that messages and acknowledgements are sent in separate rounds.) As a further simplification we assume that $\tilde{C}$ covers both messages and acknowledgements. That is, if c many paths share an edge with path p than we define p's path congestion as $2c$. In this case both events are covered that either the message following p collides with other messages or its acknowledgement collides with other acknowledgements. These assumptions allow us to view a round of the trial-and-failure protocol as simply one forward pass (i.e., we can ignore the fact that acknowledgements are sent back).

2.1 The Upper Bound

In this section we want to prove an upper bound for the number T of rounds that is necessary to route all worms using the trial-and-failure protocol with some suitable values of Δ_t. We first want to find a structure that witnesses a long runtime of the protocol.

Assume that a worm w_0 is still active after t rounds. Then there must have been a worm w_1 that prevented it from moving forward in round t. But if w_0 and w_1 have been active at round t there must have been (not necessarily different) worms w_2 and w_3 which prevented w_0 and w_1 from moving forward in round $t-1$. Continuing with this argumentation until round 1 we find:

If worm w_0 is still active after t rounds then a tree of the form as shown in Figure 4 can be constructed such that the nodes represent worms and two nodes with a common father correspond to a collision event.

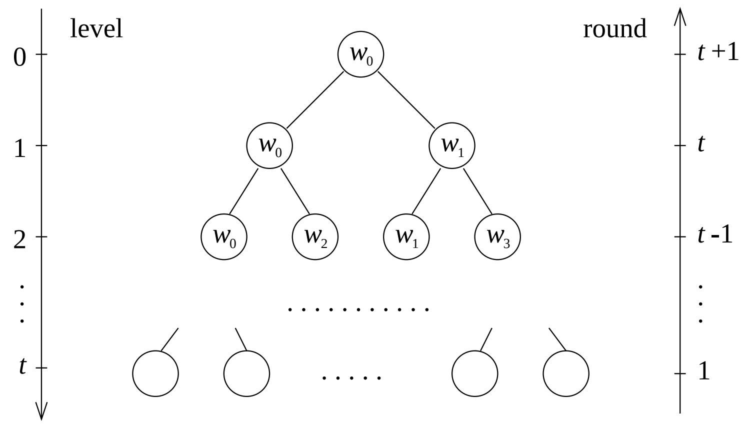

Figure 4: The witness tree of depth t.

Let us call this tree a *witness tree of depth* t, and denote it by $\mathcal{W}(t)$. The following definition formalizes what kind of embeddings of worms into the nodes of $\mathcal{W}(t)$ we only have to consider.

Definition 2.1 *Let φ be an embedding of worms into the nodes of $\mathcal{W}(t)$. A pair of worms (w, w') is called* collision pair *if w is embedded in the left son, and w' is embedded in the right son of a common father in $\mathcal{W}(t)$. We call φ* valid *if for every collision pair (w, w') embedded at level i of $\mathcal{W}(t)$ it holds that*

- $w \neq w'$,
- *w is also embedded in the father of w and w',*
- *there is no collision pair (w, w'') at level i with $w' \neq w''$, and*
- *the paths of w and w' share an edge.*

A valid embedding is called active *if for any collision pair (w, w') embedded at level i of $\mathcal{W}(t)$ it holds that w and w' use the same wavelength and w' prevents w from moving forward in round $t - i + 1$.*

Following the discussion above, we can state the following lemma.

Lemma 2.2 *If worm w_0 is still active after t rounds then there is an active embedding φ of worms into $\mathcal{W}(t)$ that maps w_0 to the root of $\mathcal{W}(t)$.*

The above lemma implies that it suffices to find an upper bound for the probability (w.r.t. random choices for the delays and wavelengths used by the worms) that there is an active embedding φ for any worm w_0 in order to prove the upper bound in Main Theorem 1.1.

In order to count the number of valid embeddings we introduce the following type of graphs.

Definition 2.3 *Let φ be a valid embedding. For each level $i \in \{1, \ldots, t\}$ of $\mathcal{W}(t)$, let $G_i = (V_i, E_i)$ be a directed graph whose nodes represent the set of worms embedded in level i and whose edges (w, w') represent the collision pairs (w, w') in level i. We call the worms in V_{i-1}* old *and the worms in $V_i \setminus V_{i-1}$* new *w.r.t. G_i.*

We assume G_0 to be the graph consisting only of a single node. Let the set of graphs $G_0, \ldots, G_t$ be called *valid* if they represent a valid embedding into $\mathcal{W}(t)$. Clearly, each valid embedding into $\mathcal{W}(t)$ has a unique valid set of graphs $G_0, \ldots, G_t$, and vice versa. Thus we can switch between either considering valid sets of graphs $G_0, \ldots, G_t$ or considering valid embeddings into $\mathcal{W}(t)$ in an arbitrary way.

For any valid embedding φ into the witness tree $\mathcal{W}(t)$, let $m_i = |V_i|$ denote the total number of worms and $\ell_i = m_i - m_{i-1}$ denote the number of new worms at level i. Let $\tilde{C}_j$ be an upper bound for the path congestion that holds at round j w.h.p. using the protocol above for suitably chosen $\Delta_1, \ldots, \Delta_j$ (determined later). Then it holds for the number $V(t,k)$ of valid embeddings in $\mathcal{W}(t)$ using k worms:

$$V(t,k) \leq n \sum_{\substack{\ell_1,\ldots,\ell_t \geq 0,\\ \sum_i \ell_i = k-1}} \prod_{i=1}^{t} \binom{m_{i-1}}{\ell_i} \cdot \tilde{C}_{t-i+1}^{\ell_i} \cdot (\ell_i + m_{i-1})^{m_{i-1}-\ell_i} \ ,$$

w.h.p. This formula is derived as follows.

- There are n ways to choose the worm that is embedded in the root of $\mathcal{W}(t)$.
- For every level i there are $\binom{m_{i-1}}{\ell_i}$ possibilities to choose ℓ_i old worms that collide with (and therefore narrow down the choices for) each of the ℓ_i new worms. Hence afterwards there are at most $\tilde{C}_{t-i+1}^{\ell_i}$ ways w.h.p. to choose the ℓ_i new worms.
- For the remaining $m_{i-1} - \ell_i$ old worms there are at most $\ell_i + m_{i-1}$ possibilities to choose the worm that prevents it from moving forward.

Before we can proceed with our calculation, we need an upper bound that holds for the path congestion after every round w.h.p. (Lemma 2.4), and an upper bound for the probability that any of the embeddings counted in $V(t,k)$ is active (Lemma 2.5).

Lemma 2.4 *For all $t \geq 2$ it holds that, if $\Delta_i \geq 8\mathrm{e}\frac{L\tilde{C}}{B2^{i-1}}$ for all $i \in \{1, \ldots, t-1\}$, then the path congestion $\tilde{C}_t$ at round t is at most $\max\{\frac{\tilde{C}}{2^{t-1}},\ O(\log n)\}$, w.h.p.*

Proof. The proof will be done by induction. Suppose the path congestion at the beginning of round t is bounded by $\frac{\tilde{C}}{2^{t-1}} \geq 2\alpha \log n$ for some arbitrary constant $\alpha > 1$. Let $\Delta_t \geq 8\mathrm{e}\frac{L\tilde{C}}{B2^{t-1}}$ be the delay range in round t. Consider any fixed worm w. Let $w_1, \ldots, w_k$ be the worms participating in round t whose paths share a link with the path of w, $k \leq \frac{\tilde{C}}{2^{t-1}}$. Further let the binary random variable $X_i = 1$ if and only if w_i fails to reach its destination in round t. Then $X = \sum_{i=1}^{k} X_i$ is a random variable denoting the path congestion of w after round t.

Since we only consider short-cut free paths, it holds for every pair of worms w_{i_1} and w_{i_2} that the difference between the time points when their first flits pass an edge remains the same for any commonly used edge. As there are at most $2L$ possibilities for the delays of two worms such that they meet during the routing, and each worm has $B\Delta_t$ possibilities to choose a wavelength and a delay, it holds that

$$\Pr[w_{i_1} \text{ is (partly) discarded by } w_{i_2}] \leq \frac{2L}{B\Delta_t} \ .$$

Therefore,

$$\Pr[X_i = 1] \leq \frac{\tilde{C}_t \cdot 2L}{B\Delta_t} \leq \frac{1}{4\mathrm{e}}$$

and hence $\mathrm{E}[X] \leq \frac{\tilde{C}}{4\mathrm{e}\cdot 2^{t-1}}$. Let $\mu = \frac{\tilde{C}}{4\mathrm{e}\cdot 2^{t-1}}$. Since the worms choose their delays and wavelengths independently at random, we can use Chernoff bounds (see [18]) to prove that, for $\epsilon = 2\mathrm{e} - 1$,

$$\Pr[X \geq (1+\epsilon)\mu] \leq \left(\frac{\mathrm{e}}{1+\epsilon}\right)^{(1+\epsilon)\mu} = \left(\frac{1}{2}\right)^{2\mathrm{e}\frac{\tilde{C}}{4\mathrm{e}\cdot 2^{t-1}}} \leq \left(\frac{1}{2}\right)^{\alpha \log n} = \left(\frac{1}{n}\right)^{\alpha} .$$

For $\alpha > 1$, this yields the lemma. ∎

Hence in the following we can assume that $\tilde{C}_i = \max\{\frac{\tilde{C}}{2^{i-1}}, O(\log n)\}$ for all $i \in \{1, \ldots, t\}$. Next we bound the probability that any of the embeddings counted in $V(t,k)$ is active.

Lemma 2.5 *For every valid embedding into a witness tree is holds that the probability that it is an active embedding for level $i \geq 1$ is at most*

$$\left(\frac{2L}{B\Delta_{t-i+1}}\right)^{m_{i-1}} .$$

Proof. As noted above, the probability that a collision pair (w, w') in level i of $\mathcal{W}(t)$ is active is at most $\frac{2L}{B\Delta_{t-i+1}}$. Let a node in G_i be called a *root* if it has outdegree 0. Then we can prove the following nice property.

Claim 2.6 *For every level i, the connected components in G_i are directed trees with new worms as roots.*

Proof. Every old worm needs a witness for its collision in round i and therefore cannot be a root. In contrast, new worms have no witness since they are just introduced as witnesses in round i. Furthermore a connected component cannot have a cycle since

- in leveled path collections using the serve-first rule this would mean that a worm w_1 is discarded at level ℓ_1 by a worm w_2 that is successful at that level, and w_2 is discarded at level $\ell_2 > \ell_1$ by a worm w_3 that is successful at that level, and so on, until we arrive at a worm w_i that is discarded at level $\ell_i > \ell_{i-1}$ by w_1. Since w_1 already fails at level $\ell_1 < \ell_i$, this cannot happen.
- in short-cut free path collections using the priority rule this would mean that a worm w_1 is discarded by a worm w_2 that has a higher priority than w_1, and w_2 is discarded by a worm w_3 that has a higher priority than w_2, and so on, until we arrive at a worm w_i that is discarded by w_1, since it has a higher priority than w_i. This, however, is not possible as long as no two worms with the same rank can meet in a round.

∎

Since every directed tree in G_i of size s implies a probability of $\leq (\frac{2L}{B\Delta_{t-i+1}})^{s-1}$ that its edges correspond to collisions of worms, and since there are exactly ℓ_i trees in G_i, we obtain a probability of at most

$$\left(\frac{2L}{B\Delta_{t-i+1}}\right)^{(m_{i-1}+\ell_i)-\ell_i} = \left(\frac{2L}{B\Delta_{t-i+1}}\right)^{m_{i-1}}$$

that the collisions in level i are active. This proves the lemma. ∎

Therefore the probability $P(t,k)$ that there exists an active embedding in $\mathcal{W}(t)$ is at most

$$n \sum_{\substack{\ell_1,\ldots,\ell_t \geq 0,\\ \sum_i \ell_i = k-1}} \prod_{i=1}^{t} \binom{m_{i-1}}{\ell_i} \cdot \tilde{C}_{t-i+1}^{\ell_i} \cdot (\ell_i + m_{i-1})^{m_{i-1}-\ell_i} \left(\frac{2L}{B\Delta_{t-i+1}}\right)^{m_{i-1}}$$

We now show how to simplify the formula. In case that $\ell_i \le m_{i-1}/2$, we get

$$\begin{aligned}\binom{m_{i-1}}{\ell_i}(\ell_i + m_{i-1})^{m_{i-1}-\ell_i} &\le \left(\frac{\mathrm{e}m_{i-1}}{m_{i-1}-\ell_i}\right)^{m_{i-1}-\ell_i}\left(\tfrac{3}{2}m_{i-1}\right)^{m_{i-1}-\ell_i}\\ &\le (3\mathrm{e}m_{i-1})^{m_{i-1}-\ell_i} \ ,\end{aligned}$$

and otherwise (that is, $m_{i-1}/2 < \ell_i \le m_{i-1}$)

$$\binom{m_{i-1}}{\ell_i}(\ell_i + m_{i-1})^{m_{i-1}-\ell_i} \le 2^{2\ell_i}(2m_{i-1})^{m_{i-1}-\ell_i} \ .$$

Therefore,

$$\begin{aligned}P(t,k) &\le n \sum_{\substack{\ell_1,\ldots,\ell_t \ge 0,\\ \sum_i \ell_i = k-1}} \prod_{i=1}^{t} 2^{2\ell_i}\cdot(3\mathrm{e}m_{i-1})^{m_{i-1}-\ell_i}\cdot \tilde{C}_{t-i+1}^{\ell_i}\left(\frac{2L}{B\Delta_{t-i+1}}\right)^{m_{i-1}}\\ &\le n\cdot\left(\frac{8L\cdot\tilde{C}}{B\Delta_1}\right)^{k-1} \sum_{\substack{\ell_1,\ldots,\ell_t \ge 0,\\ \sum_i \ell_i = k-1}} \prod_{i=1}^{t}\left(\frac{6\mathrm{e}Lm_{i-1}}{B\Delta_{t-i+1}}\right)^{m_{i-1}-\ell_i}\end{aligned}$$

if all Δ_i are chosen such that $\frac{\tilde{C}}{\Delta_1} \ge \frac{\tilde{C}_i}{\Delta_i}$. Furthermore, the following lemma holds. Its proof can be found in the appendix.

Lemma 2.7 *If $\Delta_i \ge \frac{40\mathrm{e}^2Lk}{B}$ and $\Delta_{i+1} \le \Delta_i$ for all $i \in \{1,\ldots,t-1\}$ then*

$$\max_{\substack{\ell_1,\ldots,\ell_t \ge 0,\\ \sum_i \ell_i = k-1}} \prod_{i=1}^{t}\left(\frac{6\mathrm{e}Lm_{i-1}}{B\Delta_{t-i+1}}\right)^{m_{i-1}-\ell_i} \le \left(\frac{6\mathrm{e}Lt}{B\Delta_t}\right)^{\frac{1}{2}(t-\lceil\log k\rceil)^2} .$$

Clearly, there are $\binom{t+k-1}{t} \le 2^{t+k-1}$ possibilities for choosing the $\ell_1,\ldots,\ell_t$ such that $\sum_{i=1}^{t}\ell_i = k-1$. Thus, if the requirements of Lemma 2.7 are fulfilled then

$$\begin{aligned}P(t,k) &\le n\left(\frac{8L\cdot\tilde{C}}{B\Delta_1}\right)^{k-1} 2^{t+k-1}\left(\frac{6\mathrm{e}Lt}{B\Delta_t}\right)^{\frac{1}{2}(t-\lceil\log k\rceil)^2}\\ &= n\cdot 2^t\left(\frac{16L\cdot\tilde{C}}{B\Delta_1}\right)^{k-1}\left(\frac{6\mathrm{e}Lt}{B\Delta_t}\right)^{\frac{1}{2}(t-\lceil\log k\rceil)^2} .\end{aligned}$$

For any constant $\gamma > 0$, let

$$k_0 = \frac{(2+\gamma)\log n}{\log\left(2 + \frac{B}{16\tilde{C}}\left(\frac{D}{L}+1\right)\right)} + 1$$

and

$$T = \sqrt{\frac{2(2+\gamma)\log n}{\log\left(\frac{1}{\sqrt{2k_0}}\left[\max\left\{\frac{\tilde{C}}{\log n},\ \log n\right\} + \frac{B}{6\mathrm{e}}\left(\frac{D}{L}+1\right)\right]\right)}} + \lceil\log k_0\rceil \ .$$

According to Lemma 2.2 we know that if the routing takes more than T rounds then there is an active embedding of worms into the witness tree $\mathcal{W}(T)$. Since the number of worms embedded in a level can only double from level i to $i+1$, for any such embedding one of the following two cases must be true:

(1) There is a level $t \leq T$ so that from level 0 to level t of $\mathcal{W}(T)$ $k \in \{k_0, \ldots, 2k_0\}$ different worms are embedded.

(2) Only $k \leq k_0$ different worms are embedded in $\mathcal{W}(T)$.

When restricting to these cases, we can set $\Delta_t = \max\{\frac{32L\cdot\tilde{C}_t}{B}, \frac{32L\cdot\tilde{C}}{B\log n}, \frac{40e^2L\cdot\delta\log n}{B}\} + D + L$ for all $t \leq T$, where δ is a sufficiently large constant depending on α in the proof of Lemma 2.4 and γ in the formula of k_0. It is easy to check that in this case the terms $\frac{32L\cdot\tilde{C}_t}{B}$ and $\frac{40e^2L\cdot\delta\log n}{B}$ in the formula above ensure that the Δ_t's fulfill the requirements of Lemma 2.4 and Lemma 2.7 and the requirement above that $\frac{\tilde{C}}{\Delta_1} \geq \frac{\tilde{C}_t}{\Delta_t}$. The term $\frac{32L\cdot\tilde{C}}{B\log n}$ was added to the formula of Δ_t to ensure that the first expression in the formula for T is bounded by $O(\sqrt{\log_{\tilde{C}} n})$. Hence it holds that

$$\begin{aligned}
&\Pr[\text{The routing takes more than } T \text{ rounds}]\\
&\leq \Pr[\text{Case (1) holds}] + \Pr[\text{Case (2) holds}]\\
&\leq \sum_{t=\log k_0}^{T} \sum_{k=k_0}^{2k_0} P(t,k) + \sum_{k=T}^{k_0} P(T,k)\\
&\leq \sum_{t=\log k_0}^{T} \sum_{k=k_0}^{2k_0} n\cdot 2^t \left(\frac{16L\cdot\tilde{C}}{B\Delta_1}\right)^{k-1} + \sum_{k=T}^{k_0} n\cdot 2^T \left(\frac{16L\cdot\tilde{C}}{B\Delta_1}\right)^{k-1} \left(\frac{6eLT}{B\Delta_T}\right)^{\frac{1}{2}(T-\lceil\log k\rceil)^2},
\end{aligned}$$

since $\frac{6eLt}{B\Delta_t} \leq 1$ for all $t \leq T$. Let us assume w.l.o.g. that $\tilde{C} \geq \gamma\log n$, since this does not affect the upper bound we want to prove. In this case $T \leq \sqrt{2k_0}$. Thus we get together with the formulas for Δ_1 and Δ_T that

$$\begin{aligned}
&\Pr[\text{The routing takes more than } T \text{ rounds}]\\
&\leq \sum_{t=\log k_0}^{T} \sum_{k=k_0}^{2k_0} n\cdot 2^t \left(\frac{1}{2+\frac{B}{16\tilde{C}}\left(\frac{D}{L}+1\right)}\right)^{k_0-1+(k-k_0)} +\\
&\quad \sum_{k=T}^{k_0} n\cdot 2^T \left(\frac{1}{2}\right)^{k-1} \left(\frac{\sqrt{2k_0}}{\max\left\{\frac{\tilde{C}}{\log n},\ \log n\right\} + \frac{B}{6e}\left(\frac{D}{L}+1\right)}\right)^{\frac{1}{2}(T-\lceil\log k\rceil)^2}\\
&\leq \sum_{t=\log k_0}^{T} \sum_{k=k_0}^{2k_0} n\cdot 2^t \left(\frac{1}{2}\right)^{(2+\gamma)\log n+(k-k_0)} + \sum_{k=T}^{k_0} n\cdot 2^T \left(\frac{1}{2}\right)^{k-1} \left(\frac{1}{2}\right)^{(2+\gamma)\log n}\\
&\leq \sum_{t=\log k_0}^{T} n\cdot 2^{t+1} \left(\frac{1}{n}\right)^{2+\gamma} + n\cdot 2^T \left(\frac{1}{2}\right)^{T-2} \left(\frac{1}{n}\right)^{2+\gamma}\\
&\leq \frac{1}{2n^\gamma} + \frac{1}{2n^\gamma} \leq n^{-\gamma}
\end{aligned}$$

Therefore the overall runtime is

$$\begin{aligned}
&\sum_{t=1}^{T} (\Delta_t + 2(D+L))\\
&= O\left(\sum_{t=1}^{T}\left(D + L + \frac{L}{B}\left(\frac{\tilde{C}}{2^{t-1}} + \frac{\tilde{C}}{\log n} + \log n\right)\right)\right)\\
&= O\left(\frac{L\tilde{C}}{B} + T\left(D + L + \frac{L\log n}{B}\right)\right),
\end{aligned}$$

w.h.p., which is bounded by

$$O\left(\frac{L\cdot\tilde{C}}{B}+\left(\sqrt{\log_\alpha n}+\log\log_\beta n\right)\left(D+L+\frac{L\log n}{B}\right)\right),$$

where $\alpha=\tilde{C}+B(\frac{D}{L}+1)+2$ and $\beta=2+\frac{B}{\tilde{C}}(\frac{D}{L}+1)$. This completes the proof of the upper bound of Main Theorems 1.1 and 1.3.

2.2 The Lower Bound

In this section we will prove the lower bound in Main Theorems 1.1 and 1.3. We use a path collection that consists of the following two types of subcollections.

- Let $d=\lfloor\frac{L-1}{2}\rfloor+1$. The first type consists of $n/(2\sqrt{\log n})$ structures consisting of $\sqrt{\log n}$ paths of length D that are connected as shown in Figure 5.

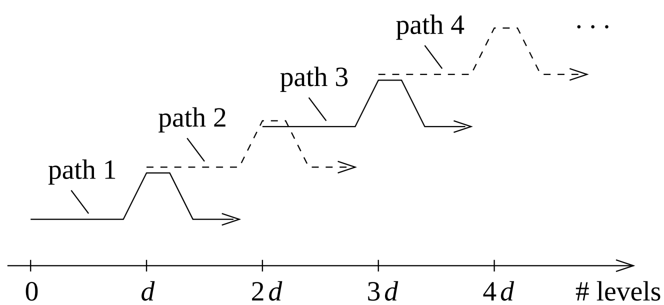

Figure 5: A type-1 structure.

  In general, the ith path starts in level $(i-1)d$ for all $i\geq 0$. Paths i and $i+1$ have a common edge from level $i\cdot d$ to level $i\cdot d+1$.

- The second type consists of $n/(2\tilde{C})$ structures each consisting of $\tilde{C}$ identical paths of length D.

We assume that along each of these paths one worm of length $L\geq 2$ has to be sent. Our aim is to show that the expected runtime of our protocol applied to these structures is at least

$$\Omega\left(\frac{L\tilde{C}}{B}+\left(\sqrt{\log_\alpha n}+\log\log_\beta n\right)(D+L)\right)\quad,$$

where the $\sqrt{\log_\alpha n}$-term is due to the type-1 structures and the $\log\log_\beta n$-term is due to the type-2 structures.

We first want to compute how long it takes to route all worms in a type-1 structure. In case of routing along short-cut free paths using priority routers, we assume that the worm traversing path i has rank i, and in case of conflicts worms with higher ranks are preferred. In order to bound the number of ways to assign delays and wavelengths to the worms such that conflicts occur, we need the following lemma.

Lemma 2.8 *Consider an arbitrary round of the trial-and-failure protocol with delay range $\Delta\geq L$. Suppose that the worms traversing the first $i+1$ paths are still active at the beginning of this round. Then with probability at least $\left(\frac{L-1}{2B\Delta}\right)^i$ the worms traversing the first i paths are discarded.*

Proof. Let Δ be the delay range of the round. Further let us denote the worm traversing path j by w_j, and its delay by δ_j. Clearly, there are $B\Delta$ ways to choose a wavelength and a delay for w_1. In the following we show that for any delay δ_i of worm w_i, there are at least $\frac{L-1}{2}$ ways to assign a delay δ_{i+1} to worm w_{i+1} such that w_{i+1} blocks w_i.

According to the construction of the type-1 structure, w_{i+1} starts $\lfloor\frac{L-1}{2}\rfloor+1$ levels after w_i. Hence, if $\delta_{i+1} \leq \delta_i + \lfloor\frac{L-1}{2}\rfloor$ then w_{i+1} is at least one level ahead of w_i during the routing. On the other hand, if $\delta_{i+1} \geq \delta_i - \lfloor\frac{L-1}{2}\rfloor$ then w_{i+1} is at most $L-1$ levels ahead of w_i during the routing. Since $|\{\delta_i - \lfloor\frac{L-1}{2}\rfloor, \ldots, \delta_i + \lfloor\frac{L-1}{2}\rfloor\} \cap [\Delta]| \geq \lfloor\frac{L-1}{2}\rfloor + 1 \geq \frac{L-1}{2}$ for $\Delta \geq L$, the number of ways to assign delays to w_{i+1} such that w_i is blocked by w_{i+1} is at least $\frac{L-1}{2}$.

Thus altogether there are at least $B\Delta(\frac{L-1}{2})^i$ ways to choose delays and wavelengths for the worms such that the worms traversing the first i paths are discarded. Hence this happens with a probability of at least

$$\frac{B\Delta(\frac{L-1}{2})^i}{(B\Delta)^{i+1}} = \left(\frac{L-1}{2B\Delta}\right)^i .$$

■

Consider now the situation that it takes $t+1$ rounds to route the worms traversing the first $t+1$ paths in a type-1 structure. This could happen, e.g., if in round i only w_{t-i+2} is able to reach its destination, and the worms $w_1, \ldots, w_{t-i+1}$ are discarded. According to the lemma above, for $L \geq 2$ the probability of such an event is at least

$$\prod_{i=1}^{t+1} \frac{B\Delta_i\left(\frac{L-1}{2}\right)^{t-i+1}}{(B(\Delta_i+L))^{t-i+2}} = \prod_{i=1}^{t}\left(\frac{L-1}{2B(\Delta_i+L)}\right)^{t-i+1}, \qquad (1)$$

where $\Delta_i \geq 1$ is the delay range for round i. Clearly, the number of time steps necessary for the t rounds is at least $\Omega(\sum_{i=1}^t (\Delta_i + D + L))$. Given a fixed $\Delta = \sum_{i=1}^t \Delta_i$, the product in (1) is minimal if $\Delta_i + L = (t-i+1)(\Delta + t\cdot L)/\binom{t+1}{2}$ for all $i \in \{1, \ldots, t\}$. This is shown in the following lemma. Its proof can be found in the appendix.

Lemma 2.9 *Consider $x_1, \ldots, x_n \in \mathbb{R}_+$ with $y = \sum_{i=1}^n x_i$. Then, for every $\alpha \in [0, y]$, $\prod_{i=1}^n (x_i+\alpha)^i$ is maximal if $x_i + \alpha = i(y + n\cdot\alpha)/\binom{n+1}{2}$ for all $i \in \{1, \ldots, n\}$.*

Let $\bar{\Delta} = \Delta/t$. Since there are $n/(2\sqrt{\log n})$ type-1 structures, and each structure has a probability of at least

$$\prod_{i=1}^{t}\left(\frac{(L-1)(t+1)}{2B\cdot 2(t-i+1)(\bar{\Delta}+L)}\right)^{t-i+1} \geq \left(\frac{L-1}{4B(\bar{\Delta}+L)}\right)^{t^2}$$

to have active worms after t rounds, the expected number of type-1 structures that have active worms after t rounds is at least

$$\frac{n}{2\sqrt{\log n}}\left(\frac{L-1}{4B(\bar{\Delta}+L)}\right)^{t^2} .$$

Note that,

$$\frac{n}{2\sqrt{\log n}}\left(\frac{L-1}{4B(\bar{\Delta}+L)}\right)^{t^2} < 1 \quad \Leftrightarrow \quad t > \sqrt{\frac{\log\left(\frac{n}{2\sqrt{\log n}}\right)}{\log\left(\frac{4B(\bar{\Delta}+L)}{L-1}\right)}}$$

Hence the expected number of rounds that are needed to route all worms in all type-1 structures is at least

$$\Omega\left(\sqrt{\log_{B(\bar{\Delta}/L+2)} n}\right) .$$

In order to bound the time needed to route all worms in the type-2 structures, we distinguish between the cases $\tilde{C} \geq 2^{\sqrt{\log n}}$ and $\tilde{C} \leq 2^{\sqrt{\log n}}$.

Case $\tilde{C} \leq 2^{\sqrt{\log n}}$:
Note that any routing protocol needs at least $\Omega(\frac{L\tilde{C}}{B} + D + L)$ steps to route all worms in a type-2 structure. Therefore the expected number of steps the protocol needs to route all worms is at least

$$\Omega\left(\frac{L\tilde{C}}{B} + \sqrt{\log_{B(\bar{\Delta}/L+2)} n}(\bar{\Delta} + D + L)\right) .$$

Since the runtime bound is minimal for some $\bar{\Delta}$ chosen in $O(\frac{L\tilde{C}}{B} + D + L)$, the expected runtime of the protocol is at least

$$\Omega\left(\frac{L\tilde{C}}{B} + \sqrt{\log_\alpha n}(D + L)\right) ,$$

where $\alpha = \tilde{C}+B(\frac{D}{L}+1)+2$. Let $\beta = \alpha/\tilde{C}+2$. Since $\tilde{C} \leq 2^{\sqrt{\log n}}$, it holds that $\sqrt{\log_\alpha n} \leq \log\log n$ only if $B(\frac{D}{L}+1) \geq 2^{\log n/(\log\log n)^2} \gg \tilde{C}$. In this case, however, $\log\beta = \Theta(\log\alpha)$, that is, $\sqrt{\log_\alpha n} \geq \log\log_\beta n$. Therefore we arrive at an expected runtime of the protocol of at least

$$\Omega\left(\frac{L\tilde{C}}{B} + \left(\sqrt{\log_\alpha n} + \log\log_\beta n\right)(D + L)\right)$$

time steps.

Case $\tilde{C} \geq 2^{\sqrt{\log n}}$:
Let $\tilde{C}_i$ be the minimum over all type-2 structures P of the number of worms that are still active in P after i rounds. Then the following lemma holds. Its proof can be found in the appendix.

Lemma 2.10 *For every $t \geq 2$ and $L(\frac{\tilde{C}}{B}+2) \leq \Delta_1, \ldots, \Delta_{t-1} \leq \hat{\Delta}$ with $\tilde{C}/(\frac{32B\hat{\Delta}}{(L-1)\tilde{C}})^{2^{t-1}-1} \geq 9\ln n$ it holds that*

$$\tilde{C}_t \geq \frac{\tilde{C}}{\left(\frac{32B\hat{\Delta}}{(L-1)\tilde{C}}\right)^{2^{t-1}-1}}$$

w.h.p.

Thus for any $L \geq 2$ and $\hat{\Delta} \geq 1$ it holds for the expected number t of rounds to route all worms in type-2 structures that

$$\frac{\tilde{C}}{\left(\frac{32B(\hat{\Delta}+L(\tilde{C}/B+2))}{(L-1)\tilde{C}}\right)^{2^t-1}} \leq 9\ln n$$

$$\Leftrightarrow \quad t \geq \log\left(1 + \log_\gamma \frac{\tilde{C}}{9\ln n}\right) ,$$

where $\gamma = \frac{32B(\hat{\Delta}+L(\tilde{C}/B+2))}{(L-1)\tilde{C}}$. Since $\tilde{C} \geq 2^{\sqrt{\log n}}$, the expected runtime of the protocol is at least

$$\Omega\left(\frac{L\tilde{C}}{B} + \left(\sqrt{\log_{\frac{B\bar{\Delta}}{L}+2} n} + \log\log_\gamma n\right)(\bar{\Delta} + D + L)\right) .$$

As is not difficult to see, this bound is minimal for $\bar{\Delta} = O(\frac{L\tilde{C}}{B} + D + L)$. Thus we get an expected runtime of at least

$$\Omega\left(\frac{L\tilde{C}}{B} + \left(\sqrt{\log_\alpha n} + \log\log_\beta n\right)(D+L)\right)$$

time steps, where $\alpha = \tilde{C} + B(\frac{D}{L}+1) + 2$ and $\beta = \alpha/\tilde{C} + 2$.

3 Proof of Main Theorem 1.2

In this section we prove upper and lower bounds on the runtime of our protocol for short-cut free path collections using serve-first routers. Hence suppose we want to route worms of length L along a collection of n short-cut free paths with path congestion $\tilde{C}$ and dilation D, using serve-first routers with bandwidth B. (We again assume that $\tilde{C}$ covers both messages and acknowledgments.)

3.1 The Upper Bound

In this section we want to prove the upper bound in Main Theorem 1.2. Let the witness tree $\mathcal{W}(t)$ be defined as in Section 2. For any valid embedding φ into $\mathcal{W}(t)$, let $m_i = |V_i|$ denote the total number of worms and $\ell_i = m_i - m_{i-1}$ denote the number of new worms at level i. Furthermore let c_i denote the number of old worms that are in a connected component in G_i with a new worm. Let $\tilde{C}_j$ be an upper bound for the path congestion that holds w.h.p. after round j using the trial-and-failure protocol for suitably chosen $\Delta_1, \ldots, \Delta_j$ (determined later). Then it holds for the number $V(t,k)$ of valid embeddings in $\mathcal{W}(t)$ using k worms:

$$V(t,k) \le n \sum_{\substack{\ell_1,\ldots,\ell_t \ge 0,\\ \sum_i \ell_i = k-1}} \prod_{i=1}^{t} \sum_{c_i=\ell_i}^{m_{i-1}} \binom{m_{i-1}}{c_i}\binom{c_i}{\ell_i} \cdot \tilde{C}_{t-i+1}^{\ell_i} \cdot (\ell_i + c_i)^{c_i - \ell_i} \cdot (m_{i-1} - c_i)^{m_{i-1} - c_i} \ ,$$

w.h.p. This formula is derived as follows.

- There are n ways to choose the worm that is embedded in the root of $\mathcal{W}(t)$.
- For each level i, there are $\binom{m_{i-1}}{c_i}$ possibilities to choose c_i old worms that lie in a connected component in G_i with a new worm, and $\binom{c_i}{\ell_i}$ possibilities to choose ℓ_i old worms that collide with (and therefore narrow down the choices for) each of the ℓ_i new worms. Therefore afterwards there are at most $\tilde{C}_{t-i+1}^{\ell_i}$ ways w.h.p. to choose the ℓ_i new worms. For the remaining $c_i - \ell_i$ old worms there are at most $\ell_i + c_i$ possibilities to choose the worm that prevents it from moving forward.
- For each of the remaining $m_{i-1} - c_i$ old worms there are at most $m_{i-1} - c_i$ ways to determine the old worm which prevents it from moving forward.

Before we can proceed with our calculation, we need an upper bound that holds for the path congestion w.h.p., and need an upper bound for the probability that the embeddings counted in $V(t,k)$ are active.

Since the delays and wavelengths are chosen independently and we only consider short-cut free paths, it holds for every pair of worms w_i and w_j at round t that

$$\Pr[w_i \text{ is blocked by } w_j] \le \frac{L}{B\Delta_t} \ .$$

Therefore, analogous to Lemma 2.4 we get that, if $\Delta_i \geq 4\mathrm{e}\frac{L\tilde{C}}{B2^{i-1}}$ for all $i \in \{1,\dots,t-1\}$, then the path congestion $\tilde{C}_t$ at round t is at most $\max\{\frac{\tilde{C}}{2^{t-1}}, O(\log n)\}$, w.h.p.

Next we bound the probability that the embeddings counted in $V(t,k)$ are active. As noted above, the probability that a collision pair (w,w') in level i of $\mathcal{W}(t)$ is active is at most $\frac{L}{B\Delta_{t-i+1}}$. For every level i, each connected component in G_i that contains no new worms has a size of at least three. This is true since we only allow the worms to be routed along short-cut free paths and therefore two worms cannot block each other. Hence there are at most $g_i \leq \frac{m_i - c_i}{3}$ components with no new worms. Since every connected component of size s implies a probability of at most $(\frac{L}{B\Delta_{t-i+1}})^{s-1}$ that its edges represent collisions of worms we obtain a probability of at most

$$\left(\frac{L}{B\Delta_{t-i+1}}\right)^{((m_{i-1}-c_i)-g_i)} \leq \left(\frac{L}{B\Delta_{t-i+1}}\right)^{\frac{2(m_{i-1}-c_i)}{3}}$$

that these components are active. Note that we can improve this bound if we know that at least $k > 3$ subpaths of paths in the collection are needed to obtain a directed cycle in G_i.

According to Definition 2.1, for every level of a valid embedding every old worm can only have one witness. Hence each connected component in G_i that contains a new worm has exactly one edge less than its size and therefore must form a tree. Furthermore each new worm lies in a different connected component. Therefore the probability that the edges of components with new worms represent collisions of worms is at most

$$\left(\frac{L}{B\Delta_{t-i+1}}\right)^{(\ell_i+c_i)-\ell_i} .$$

Altogether the probability that all collision pairs in level i are active given m_{i-1} and c_i is at most

$$\left(\frac{L}{B\Delta_{t-i+1}}\right)^{c_i+\frac{2(m_{i-1}-c_i)}{3}} .$$

Therefore the probability $P(t,k)$ that there exists an active embedding in $\mathcal{W}(t)$ is at most

$$n \sum_{\substack{\ell_1,\dots,\ell_t \geq 0, \\ \sum_i \ell_i = k-1}} \prod_{i=1}^{t} \sum_{c_i=\ell_i}^{m_{i-1}} \binom{m_{i-1}}{c_i}\binom{c_i}{\ell_i} \tilde{C}_{t-i+1}^{\ell_i} (\ell_i+c_i)^{c_i-\ell_i}(m_{i-1}-c_i)^{m_{i-1}-c_i}\left(\frac{L}{B\Delta_{t-i+1}}\right)^{c_i+\frac{2(m_{i-1}-c_i)}{3}}$$

In order to simplify this formula, we have to distinguish between two cases. If $\ell_i \leq c_i/2$ we get

$$\begin{aligned}\binom{c_i}{\ell_i}(\ell_i+c_i)^{c_i-\ell_i} &\leq \left(\frac{\mathrm{e}c_i}{c_i-\ell_i}\right)^{c_i-\ell_i}\left(\tfrac{3}{2}c_i\right)^{c_i-\ell_i} \\ &\leq (3\mathrm{e}c_i)^{c_i-\ell_i} ,\end{aligned}$$

and otherwise

$$\binom{c_i}{\ell_i}(\ell_i+c_i)^{c_i-\ell_i} \leq \tfrac{1}{2}2^{2\ell_i}(2c_i)^{c_i-\ell_i} .$$

Let $\Delta_i \geq \frac{216L}{B}$ for all $i \in \{1,\dots,t\}$. Then it holds

$$\sum_{c_i=\ell_i}^{m_{i-1}} \binom{m_{i-1}}{c_i}\binom{c_i}{\ell_i}(\ell_i+c_i)^{c_i-\ell_i}(m_{i-1}-c_i)^{m_{i-1}-c_i}\left(\frac{L}{B\Delta_{t-i+1}}\right)^{c_i+\frac{2(m_{i-1}-c_i)}{3}}$$

$$
\begin{aligned}
&\le \sum_{c_i=\ell_i}^{m_{i-1}} \left(\frac{\mathrm{e}m_{i-1}}{m_{i-1}-c_i}\right)^{m_{i-1}-c_i} \binom{c_i}{\ell_i} (\ell_i+c_i)^{c_i-\ell_i} (m_{i-1}-c_i)^{m_{i-1}-c_i} \left(\frac{L}{B\Delta_{t-i+1}}\right)^{c_i+\frac{2(m_{i-1}-c_i)}{3}} \\
&\le (\mathrm{e}m_{i-1})^{m_{i-1}-\ell_i} \left(\frac{L}{B\Delta_{t-i+1}}\right)^{\ell_i+\frac{2(m_{i-1}-\ell_i)}{3}} \cdot \\
&\qquad \sum_{c_i=\ell_i}^{m_{i-1}} \binom{c_i}{\ell_i} (\ell_i+c_i)^{c_i-\ell_i} \left(\frac{1}{\mathrm{e}m_{i-1}}\right)^{c_i-\ell_i} \left(\frac{L}{B\Delta_{t-i+1}}\right)^{\frac{c_i-\ell_i}{3}} \\
&\le (\mathrm{e}m_{i-1})^{m_{i-1}-\ell_i} \left(\frac{L}{B\Delta_{t-i+1}}\right)^{\ell_i+\frac{2(m_{i-1}-\ell_i)}{3}} \cdot \\
&\qquad \sum_{c_i=\ell_i}^{m_{i-1}} \tfrac{1}{2} 2^{2\ell_i} (3\mathrm{e}c_i)^{c_i-\ell_i} \left(\frac{1}{\mathrm{e}m_{i-1}}\right)^{c_i-\ell_i} \left(\frac{L}{B\Delta_{t-i+1}}\right)^{\frac{c_i-\ell_i}{3}} \\
&\le (\mathrm{e}m_{i-1})^{m_{i-1}-\ell_i} \left(\frac{L}{B\Delta_{t-i+1}}\right)^{\ell_i+\frac{2(m_{i-1}-\ell_i)}{3}} 4^{\ell_i} \cdot \tfrac{1}{2} \sum_{c_i=\ell_i}^{m_{i-1}} \left(\frac{27L}{B\Delta_{t-i+1}}\right)^{\frac{c_i-\ell_i}{3}} \\
&\le 4^{\ell_i} (\mathrm{e}m_{i-1})^{m_{i-1}-\ell_i} \left(\frac{L}{B\Delta_{t-i+1}}\right)^{\ell_i+\frac{2(m_{i-1}-\ell_i)}{3}}
\end{aligned}
$$

Thus we get

$$
\begin{aligned}
P(t,k) &\le n \sum_{\substack{\ell_1,\dots,\ell_t\ge 0,\\ \sum_i \ell_i=k-1}} \prod_{i=1}^{t} 4^{\ell_i} (\mathrm{e}m_{i-1})^{m_{i-1}-\ell_i} \tilde{C}_{t-i+1}^{\ell_i} \left(\frac{L}{B\Delta_{t-i+1}}\right)^{\ell_i+\frac{2(m_{i-1}-\ell_i)}{3}} \\
&\le n \cdot \left(\frac{4L\cdot\tilde{C}}{B\Delta_1}\right)^{k-1} \sum_{\substack{\ell_1,\dots,\ell_t\ge 0,\\ \sum_i \ell_i=k-1}} \prod_{i=1}^{t} \left(\frac{L(\mathrm{e}m_{i-1})^{3/2}}{B\Delta_{t-i+1}}\right)^{\frac{2(m_{i-1}-\ell_i)}{3}}
\end{aligned}
$$

if all Δ_i are chosen such that $\frac{\tilde{C}}{\Delta_1} \ge \frac{\tilde{C}_i}{\Delta_i}$. Furthermore, the following lemma holds. Its proof can be found in the appendix.

Lemma 3.1 *If $B\Delta_i \ge (3\mathrm{e})^3 Lk^{3/2}$ and $\Delta_{i+1} \le \Delta_i$ for all $i \in \{1,\dots,t-1\}$ then*

$$
\max_{\substack{\ell_1,\dots,\ell_t\ge 0,\\ \sum_i \ell_i=k-1}} \prod_{i=1}^{t} \left(\frac{L(\mathrm{e}m_{i-1})^{3/2}}{B\Delta_{t-i+1}}\right)^{\frac{2(m_{i-1}-\ell_i)}{3}} \le \left(\frac{13L}{B\Delta_t}\right)^{t-\lceil \log k\rceil} .
$$

Clearly, there are $\binom{t+k-1}{t} \le 2^{t+k-1}$ possibilities for choosing the $\ell_1,\dots,\ell_t$ such that $\sum_{i=1}^t \ell_i = k-1$. Thus, if the requirements of Lemma 3.1 are fulfilled,

$$
\begin{aligned}
P(t,k) &\le n \left(\frac{4L\cdot\tilde{C}}{B\Delta_1}\right)^{k-1} 2^{t+k-1} \left(\frac{13L}{B\Delta_t}\right)^{t-\lceil \log k\rceil} \\
&\le n\cdot 2k \left(\frac{8L\cdot\tilde{C}}{B\Delta_1}\right)^{k-1} \left(\frac{26L}{B\Delta_t}\right)^{t-\lceil \log k\rceil} .
\end{aligned}
$$

For any constant $\gamma > 0$, let

$$k_0 = \frac{(2+\gamma)\log n}{\log\left(2 + \frac{B}{8\tilde{C}}\left(\frac{D}{L}+1\right)\right)} + 1$$

and

$$T \geq \frac{(2+\gamma)\log n}{\log\left(\max\left\{\frac{\tilde{C}}{2\log n},\ \log^{3/2} n\right\} + \frac{B}{26}\left(\frac{D}{L}+1\right)\right)} + \lceil \log k_0 \rceil \ .$$

If the routing takes more than T rounds then, analogous to Section 2.1, one of the following two cases must be true:

(1) There must exist a valid reduced embedding into a witness tree $\mathcal{W}(t)$ with $t \leq T$ and $k \in \{k_0, \ldots, 2k_0\}$ different worms.

(2) There must exist a valid reduced embedding into a witness tree $\mathcal{W}(T)$ with $k \leq k_0$ different worms.

It is easy to check (see Section 2.1) that, when restricting to these cases, all requirements above for the Δ_t are fulfilled by setting $\Delta_i = \max\{\frac{16L\cdot\tilde{C}_i}{B}, \frac{16L\cdot\tilde{C}}{B\log n}, \frac{(3e)^3 L\cdot\delta\log^{3/2} n}{B}\} + D + L$, where δ is a sufficiently large constant depending on α in the proof of Lemma 2.4 and γ in the bound of k_0. Then we get:

$$\begin{aligned}
&\Pr[\text{The routing takes more than } T \text{ rounds}]\\
&\leq \Pr[\text{Case (1) holds}] + \Pr[\text{Case (2) holds}]\\
&\leq \sum_{t=\log k_0}^{T}\sum_{k=k_0}^{2k_0} P(t,k) + \sum_{k=2}^{k_0} P(T,k)\\
&\leq \sum_{t=\log k_0}^{T}\sum_{k=k_0}^{2k_0} n\cdot 2k\left(\frac{8L\cdot\tilde{C}}{B\Delta_1}\right)^{k-1} + \sum_{k=2}^{k_0} n\cdot 2k\left(\frac{8L\cdot\tilde{C}}{B\Delta_1}\right)^{k-1}\left(\frac{26L}{B\Delta_T}\right)^{T-\lceil\log k\rceil}\\
&\leq \sum_{t=\log k_0}^{T}\sum_{k=k_0}^{2k_0} n\cdot 2k\left(\frac{1}{2+\frac{B}{8\tilde{C}}\left(\frac{D}{L}+1\right)}\right)^{k_0-1+(k-k_0)} +\\
&\quad \sum_{k=2}^{k_0} n\cdot 2k\left(\frac{1}{2}\right)^{k-1}\left(\frac{26}{\left(\frac{16\tilde{C}}{\log n}+(3e)^3\log^{3/2} n\right)+B\left(\frac{D}{L}+1\right)}\right)^{T-\lceil\log k\rceil}\\
&\leq \sum_{t=\log k_0}^{T}\sum_{k=k_0}^{2k_0} n\cdot 2k\left(\frac{1}{2}\right)^{(2+\gamma)\log n+(k-k_0)} + \sum_{k=2}^{k_0} n\cdot 2k\left(\frac{1}{2}\right)^{k-1}\left(\frac{1}{2}\right)^{(2+\gamma)\log n}\\
&\leq \frac{1}{2n^\gamma} + \frac{1}{2n^\gamma} \leq n^{-\gamma}
\end{aligned}$$

Therefore the overall runtime is

$$\begin{aligned}
&\sum_{t=1}^{T}(\Delta_t + 2(D+L))\\
&\quad = O\left(\sum_{t=1}^{T}\left(D + L + \frac{L}{B}\left(\frac{\tilde{C}}{2^t} + \frac{\tilde{C}}{\log n} + \log^{3/2} n\right)\right)\right)\\
&\quad = O\left(\frac{L\tilde{C}}{B} + T\left(D + L + \frac{L\log^{3/2} n}{B}\right)\right) \ ,
\end{aligned}$$

w.h.p., which is bounded by

$$O\left(\frac{L\cdot\tilde{C}}{B}+(\log_\alpha n+\log\log_\beta n)\left(\frac{L\log^{3/2}n}{B}+D+L\right)\right),$$

where $\alpha=\tilde{C}+B(\frac{D}{L}+1)+2$ and $\beta=\alpha/\tilde{C}+2$.

3.2 The Lower Bound

In this section we will prove the lower bound in Main Theorem 1.2. We use a path collection that consists of the following two types of subcollections.

- The first type consists of $n/6$ structures consisting of three paths of length D that are connected as shown in Figure 6.

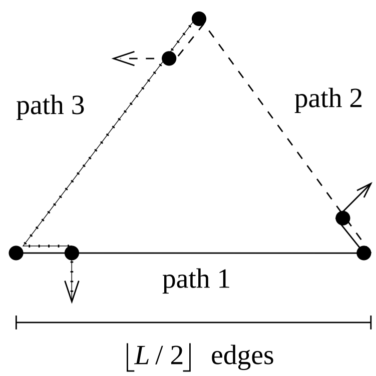

Figure 6: A type-1 structure.

- The second type consists of $n/(2\tilde{C})$ structures each consisting of $\tilde{C}$ identical paths of length D.

We assume that along each of these paths one worm of length $L\geq 2$ has to be sent. (Note that in case of $L=1$ no cycles of colliding worms can occur, that is, we are in a situation of Main Theorems 1.1 and 1.3.) Our aim is to show that the expected runtime of our protocol applied to these structures is at least

$$\Omega\left(\frac{L\tilde{C}}{B}+\left(\log_\alpha n+\log\log_\beta n\right)(D+L)\right) \quad ,$$

where the $\log_\alpha n$-term is due to the type-1 structures and the $\log\log_\beta n$-term is due to the type-2 structures.

We first compute how long it takes to route all worms in a type-1 structure. Consider an arbitrary round i of the trial-and-failure protocol. Suppose that in a given type-1 structure all three worms are still active. Then we want to calculate the probability that these three worms block each other in round i.

Suppose that $\Delta_i\geq L$. Let the worm traveling along path $j\in\{1,2,3\}$ be called w_j. Let δ_1 be the delay chosen by worm w_1. In case that $\delta_1\leq\Delta_i-\lfloor\frac{L}{2}\rfloor$ we get that w_1, w_2, and w_3 collide if w_2 and w_3 choose the same wavelength as w_1 and delays in the range $[\delta_1,\delta_1+\lfloor\frac{L}{2}\rfloor-1]$. In case that $\delta_1>\Delta_i-\lfloor\frac{L}{2}\rfloor$, then w_1, w_2, and w_3 collide if w_2 and w_3 choose the same wavelength as w_1 and delays in the range $[\delta_1-(\lfloor\frac{L}{2}\rfloor-1),\delta_1]$. Hence in both cases there are at least $\lfloor\frac{L}{2}\rfloor$ possibilities for both w_2 and w_3 to choose a wavelength and a delay such that w_1, w_2, and w_3 collide. Thus the probability that w_1, w_2, and w_3 collide at round i is at least $(\lfloor \frac{L}{2} \rfloor / (B\Delta_i))^2$ if $\Delta_i \geq L$. Therefore the probability that w_1, w_2, and w_3 collide for t rounds is at least

$$\prod_{i=1}^{t} \left(\frac{\lfloor L/2 \rfloor}{B(\Delta_i + L)} \right)^2$$

for any choice of $\Delta_1, \ldots, \Delta_t \geq 1$. Given a fixed $\Delta = \sum_{i=1}^{t} \Delta_i$ this product yields the smallest probability if $\Delta_i = \Delta/t$ for all $i \in \{1, \ldots, t\}$. Hence assume that all delay ranges are equal to $\bar{\Delta} = \Delta/t$. Since there are $n/6$ type-1 structures, and each structure has a probability of at least $(\frac{L}{3B(\bar{\Delta}+L)})^{2t}$ to have active worms after t rounds, the expected number of type-1 structures that have active worms after t rounds is at least

$$\frac{n}{6} \left(\frac{L}{3B(\bar{\Delta} + L)} \right)^{2t} < 1 \quad \Leftrightarrow \quad t \geq \frac{\log(n/6)}{2 \log \left(\frac{3B(\bar{\Delta}+L)}{L} \right)} \ .$$

Hence the expected number of rounds that are needed to route all worms is $\Omega(\log_{B(\bar{\Delta}/L+1)} n)$. In order to bound the time needed to route worms in the type-2 structures, we distinguish between the cases $\tilde{C} \geq 2^{\sqrt{\log n}}$ and $\tilde{C} \leq 2^{\sqrt{\log n}}$.

Case $\tilde{C} \leq 2^{\sqrt{\log n}}$:
Note that any routing protocol needs at least $\Omega(\frac{L\tilde{C}}{B} + D + L)$ steps to route all worms in a type-2 structure. Therefore the expected runtime of the protocol is at least

$$\begin{aligned} & \Omega\left(\frac{L\tilde{C}}{B} + \log_{\frac{B\bar{\Delta}}{L}+2} n \cdot (\bar{\Delta} + D + L) \right) \\ = \ & \Omega\left(\frac{L\tilde{C}}{B} + (\log_\alpha n + \log\log_\beta n)(D + L) \right), \end{aligned}$$

where $\alpha = \tilde{C} + B(\frac{D}{L} + 1) + 2$ and $\beta = \alpha/\tilde{C} + 2$.

Case $\tilde{C} \geq 2^{\sqrt{\log n}}$:
This case follows analogous to Section 2.

4 Summary and Open Problems

In case that wavelength conversion is not allowed we presented a very accurate analysis of the performance of a simple routing protocol for arbitrary short-cut free path collections. Important questions that remain are:

- How do the bounds change if arbitrary simple (i.e., loop free) path collections are allowed?
- What is the exact time bound for the runtime of the trial-and-failure protocol if wavelength conversion is allowed? (The bound presented in [11] seems to be too weak compared to the bounds obtained in this paper.)

Furthermore it would be interesting to consider cases in which only a few routers can convert wavelengths (see, e.g., [23]), or worms are allowed a bounded number of hops (i.e., conversions to and from electrical form) in the network.

References

[1] A. Aggarwal, A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, M. Sudan. Efficient routing and scheduling algorithms for optical networks. In *Proc. of the 5th Ann. ACM-SIAM Symp. on Discrete Algorithms*, pp. 412-423, 1994.

[2] K. Bala, T.E. Stern. Algorithms for routing in a linear lightwave network. In *Proc. of INFOCOM*, pp. 1-9, 1991.

[3] R.A. Barry, P.A. Humblet. Bounds on the number of wavelengths needed in WDM networks. In *LEOS'92 Summer Topical Mtg. Digest*, pp. 114-127, 1992.

[4] R.A. Barry, P.A. Humblet. On the number of wavelengths and switches in all-optical networks. In *IEEE Trans. on Communications* **42**, pp. 583-591, 1994.

[5] T.A. Birks, D.O. Culverhouse, S.G. Farwell, P.St.J. Russel. 2×2 single-mode fiber routing switch. *Optics Letters* **21**, pp. 722-724, 1996.

[6] C. Brackett. Dense wavelength division multiplexing networks: Principles and applications. *IEEE J. Selected Areas in Comm.* **8**, pp. 373-380, August 1990.

[7] V.W.S. Chan. All-optical networks. *Scientific American*, pp. 56-59, Sept. 1995.

[8] I. Chlamtac, A. Ganz, G. Karmi. Lightpath communications: A novel approach to high bandwidth optical WANs. *IEEE Trans. on Communications* **40**, July 1992.

[9] I. Chlamtac, A. Ganz, G. Karmi. Lightnets: Topologies for high-speed optical networks. *IEEE/OSA J. Lightwave Techn.* **11** (5/6), pp. 951-961, 1993.

[10] N.K. Chung, K. Nosu, G. Winzer. *IEEE JSAC: Special Issue on Dense WDM Networks* **8**, 1990.

[11] R. Cypher, F. Meyer auf der Heide, C. Scheideler, B. Vöcking. Universal algorithms for store-and-forward and wormhole routing. In *28th Ann. ACM Symp. on Theory of Computing*, pp. 356-365, 1996.

[12] D.H.C. Du, R.J. Vetter. Distributed computing with high-speed optical networks. *IEEE Computer* **26**, pp. 8-18, 1993.

[13] O. Gerstel, S. Zaks. The virtual path layout problem in fast networks. In *Proc. of the 13th ACM Conf. on Principles of Distributed Computing*, pp. 235-243, 1994.

[14] O. Gerstel, S. Zaks. Path layout problem in ATM networks. In *Proc. of the 1st Colloquium on Structure, Information and Communication Complexity*, pp. 151-166, 1994.

[15] R.D. Gitlin, Z. Haas. Field coding: A high-speed 'almost-all' optical interconnect. *25th Ann. Conf. Information Sci. Syst. CISS* (Baltimore, MD), March 20-22, 1991.

[16] P.E. Green. *Fiber-Optic Communication Networks.* Prentice Hall, 1992.

[17] Z. Haas. The 'staggering switch': an 'almost-all' optical packet switch. In *Tech. Dig. Optical Fiber Commun. Conf. OFC'92* (San José, CA), paper WH6.

[18] T. Hagerup, C. Rüb. A guided tour of Chernoff bounds. *Information Processing Letters* **33**, pp. 305-308, 1989/90.

[19] G.R. Hill, P.J. Chidgey et al. A transport network layer based on optical network elements. *J. of Lightwave Techn.* **11**, pp. 667-677, 1993.

[20] H.S. Hinton. Architectural considerations for photonic switching networks. *IEEE J. Selected Areas in Comm.* **6**, pp. 1209-1226, August 1988.

[21] A. Jajszczyk, H.T. Mouftah. Photonic Fast Packet Switching. *IEEE Commun. Magazine* **2**, pp. 65-66, 1993.

[22] E. Kranakis, D. Krizanc, A. Pelc. Hop-congestion trade-offs for high-speed networks. *Int. Jornal of Foundations of Computer Science* **8**, 1997.

[23] K.C. Lee, V.O.K. Li. A wavelength-convertible optical network. *J. of Lightwave Techn.* **11**, pp. 962-970, 1993.

[24] A. Lubotzky, R. Phillips, R. Sarnak. Ramanujan graphs. *Combinatorica* **8**(3), pp. 261-277, 1988.

[25] G.A. Margulis. Explicit group theoretical constructions of combinatorial schemes and their application to the design of expanders and superconcentrators. *Problems Inform. Transmission* **11**, pp. 39-46, 1988.

[26] F. Matera, S. Settembre. All optical implementations of high capacity TDMA networks. *Fiber and Integrated Optics* **12**, pp. 173-186, 1993.

[27] F. Meyer auf der Heide, C. Scheideler. Communication in parallel systems. To appear at *SOFSEM'96.*

[28] M. Morgenstern. Existence and explicit constructions of $q+1$ regular Ramanujan graphs for every prime power q. *Journal of Comb. Theory, Series B* **62**, pp. 44-62, 1994.

[29] R.K. Pankaj. *Architectures for Linear Lightwave Networks.* PhD thesis, MIT, 1992.

[30] S. Personick. Review of fundamentals of optical fiber systems. *IEEE J. Selected Areas in Comm.* **3**, pp. 373-380, April 1983.

[31] G.R.Pieris, G.H. Sasaki. A linear lightwave Benes network. *IEEE/ACM Trans. on Networking* **1**, 1993.

[32] P. Raghavan, E. Upfal. Efficient routing in all-optical networks. In *Proc. of the 26th Ann. ACM Symp. on Theory of Computing*, 1994.

[33] R. Ramaswami. Multi-wavelength lightwave networks for computer communication. *IEEE Communications Magazine* **31**, pp. 78-88, 1993.

[34] R. Ramaswami, K.N. Sivarajan. Routing and wavelength assignment in all-optical networks. *IEEE/ACM Trans. on Networking* **3**(5), pp. 489-500, 1995.

[35] C. Scheideler, B. Vöcking. Universal continuous routing strategies. In *Proc. of the 8th Ann. ACM Symp. on Parallel Algorithms and Architectures*, pp. 142-151, 1996.

[36] J.P. Schmidt, A. Siegel, A. Srinivasan. Chernoff-Hoeffding bounds for applications with limited independence. *SIAM J. Disc. Math.* **8**(2), pp. 223-250, 1995.

[37] Z. Zhang, A.S. Acampora. A heuristic wavelength assignment algorithm for multihop WDM networks with wavelength routing and wavelength re-use. *IEEE/ACM Trans. on Networking* **3**(3), pp. 281-288, 1995.

A Appendix

A.1 Proof of Lemma 2.7

According to Claim 2.6 there must be at least one new worm per level. Thus $\sum_{i=1}^{t} m_i$ is minimized if we set $m_i = i+1$ for all $i \in \{1, \ldots, t - \lceil \log k \rceil + 1\}$, and $m_i = 2m_{i-1}$ for all $i > t - \lceil \log k \rceil + 1$. Then it holds that $\ell_i = 1$ for all $i \in \{1, \ldots, t - \lceil \log k \rceil + 1\}$, and $\ell_i = m_{i-1}$ for all $i > t - \lceil \log k \rceil + 1$. Therefore we get with $\Delta_{i+1} \leq \Delta_i$ for all $i \in \{1, \ldots, t-1\}$:

$$\begin{aligned}\prod_{i=1}^{t} \left(\frac{6\mathrm{e}Lm_{i-1}}{B\Delta_{t-i+1}}\right)^{m_{i-1}-\ell_i} &\leq \prod_{i=2}^{t-\lceil \log k \rceil+1} \left(\frac{6\mathrm{e}Li}{B\Delta_{t-i+1}}\right)^{i-1} \\ \leq \left(\frac{6\mathrm{e}Lt}{B\Delta_t}\right)^{\sum_{i=1}^{t-\lceil \log k \rceil} i} &\leq \left(\frac{6\mathrm{e}Lt}{B\Delta_t}\right)^{\frac{1}{2}(t-\lceil \log k \rceil)^2} .\end{aligned}$$

Next we show that for all other distributions of the m_i it holds that

$$\prod_{i=1}^{t} \left(\frac{6\mathrm{e}Lm_{i-1}}{B\Delta_{t-i+1}}\right)^{m_{i-1}-\ell_i} \leq \left(\frac{6\mathrm{e}Lt}{B\Delta_t}\right)^{\frac{1}{2}(t-\lceil \log k \rceil)^2} \tag{2}$$

if all $\Delta_i \geq \frac{40\mathrm{e}^2 Lk}{B}$.

Consider increasing the number m_j of worms at a stage $j < t$ with $m_j < m_{j+1}$ by 1. Then two terms in the product in (2) change: the $(i = j)$-term and the $(i = j+1)$-term. Before increasing m_j, these terms are

$$\left(\frac{6\mathrm{e}Lm_{j-1}}{B\Delta_{t-j+1}}\right)^{m_{j-1}-\ell_j} \left(\frac{6\mathrm{e}Lm_j}{B\Delta_{t-j}}\right)^{m_j-\ell_{j+1}} , \tag{3}$$

and after increasing m_j by 1, they change to

$$\left(\frac{6\mathrm{e}Lm_{j-1}}{B\Delta_{t-j+1}}\right)^{m_{j-1}-(\ell_j+1)} \left(\frac{6\mathrm{e}L(m_j+1)}{B\Delta_{t-j}}\right)^{(m_j+1)-(\ell_{j+1}-1)} \tag{4}$$

It holds that (3) $\geq$ (4)

$$\begin{aligned}&\Leftrightarrow \quad \frac{6\mathrm{e}Lm_{j-1}}{B\Delta_{t-j+1}} \geq \left(\frac{m_j+1}{m_j}\right)^{m_j-\ell_{j+1}} \left(\frac{6\mathrm{e}L(m_j+1)}{B\Delta_{t-j}}\right)^2 \\ &\Leftarrow \quad \frac{m_{j-1}}{\Delta_{t-j+1}} \geq \mathrm{e} \cdot \frac{6\mathrm{e}L(2m_{j-1}+1)^2}{B\Delta_{t-j}^2} \\ &\Leftarrow \quad \frac{m_{j-1}}{\Delta_{t-j+1}} \geq \mathrm{e} \cdot \frac{40\mathrm{e}Lm_{j-1}^2}{B\Delta_{t-j+1}^2} \\ &\Leftarrow \quad \Delta_{t-j+1} \geq \frac{40\mathrm{e}^2 Lk}{B}\end{aligned}$$

Since any distribution of the m_i can be obtained from the initial distribution above by repeatedly increasing one of the m_i by 1, the lemma follows.

A.2 Proof of Lemma 2.9

For $n = 1$, the lemma is trivially true. We will show by complete induction on n that the assumption above is also true for $n > 1$. Suppose that we have already shown that $f(y,n) = \prod_{i=1}^{n}(i(y+n\alpha)/\binom{n+1}{2})^i$ is the maximal value the product of the $(x_i + \alpha)^i$ can reach. Then we want to find the $x_{n+1} + \alpha$ for which $f(y - x_{n+1}, n) \cdot (x_{n+1} + \alpha)^{n+1}$ is maximal. Clearly,

$$f(y - x_{n+1}, n) \cdot (x_{n+1} + \alpha)^{n+1} = (y - x_{n+1} + n\alpha)^{\binom{n+1}{2}} \cdot (x_{n+1} + \alpha)^{n+1} \cdot g(n) \ , \tag{5}$$

where $g(n)$ is a function that only depends on n. Taking the logarithm yields

$$\begin{aligned} &\log\left((y - x_{n+1} + n\alpha)^{\binom{n+1}{2}} \cdot (x_{n+1} + \alpha)^{n+1} \cdot g(n)\right) \\ =& \binom{n+1}{2} \log(y - x_{n+1} + n\alpha) + (n+1)\log(x_{n+1} + \alpha) + \log g(n) \ . \end{aligned} \tag{6}$$

Since a maximum of this function is also a maximum for the function in (5), it remains to determine the maximum of the function in (6). As (6) is a convex function, this can be done by finding the x_{n+1} for which the derivation of (6) is 0.

$$\begin{aligned} -\binom{n+1}{2} \cdot \frac{1}{y - x_{n+1} + n\alpha} + \frac{n+1}{x_{n+1} + \alpha} &= 0 \\ \Leftrightarrow \quad \frac{x_{n+1} + \alpha}{n+1} &= \frac{y - x_{n+1} + n\alpha}{\binom{n+1}{2}} \\ \Leftrightarrow \quad x_{n+1} + \alpha &= \frac{(n+1)(y + (n+1)\alpha)}{\binom{n+1}{2}} \cdot \frac{1}{1 + \frac{n+1}{\binom{n+1}{2}}} \\ \Leftrightarrow \quad x_{n+1} + \alpha &= \frac{(n+1)(y + (n+1)\alpha)}{\binom{n+2}{2}} \ . \end{aligned}$$

Hence the x_i with $i \le n$ have to be chosen such that

$$\sum_{i}^{n}(x_i + \alpha) = y + (n+1)\alpha - \frac{(n+1)(y + (n+1)\alpha)}{\binom{n+2}{2}} \ .$$

According to the induction hypothesis, $\prod_{i=1}^{n}(x_i + \alpha)^i$ is maximal if, for every $i \in \{1, \ldots, n\}$,

$$\begin{aligned} x_i + \alpha &= \frac{i(y - x_{n+1} + n\alpha)}{\binom{n+1}{2}} \\ &= \frac{i\left(y - \left(\frac{(n+1)(y+(n+1)\alpha)}{\binom{n+2}{2}} - \alpha\right) + n\alpha\right)}{\binom{n+1}{2}} \\ &= \frac{i\left(\frac{\binom{n+2}{2} - (n+1)}{\binom{n+2}{2}} \cdot y + \frac{\binom{n+2}{2} - (n+1)}{\binom{n+2}{2}} \cdot (n+1)\alpha\right)}{\binom{n+1}{2}} \\ &= \frac{i(y + (n+1)\alpha)}{\binom{n+2}{2}} \ . \end{aligned}$$

From this the lemma follows.

A.3 Proof of Lemma 2.10

For $t=1$, the bound on $\tilde{C}_t$ trivially holds. Suppose that the bound above for $\tilde{C}_t$ is true for some $t \geq 1$. Then we want to show that, if $\Delta_t \leq \hat{\Delta}$ and $\tilde{C}/(\frac{32B\hat{\Delta}}{(L-1)\tilde{C}})^{2^t-1} \geq 9\ln n$, then we get

$$\tilde{C}_{t+1} \geq \frac{\tilde{C}}{\left(\frac{32B\hat{\Delta}}{(L-1)\tilde{C}}\right)^{2^t-1}}$$

w.h.p. Assume in the following that $\tilde{C}/(\frac{32B\hat{\Delta}}{(L-1)\tilde{C}})^{2^t-1} \geq 8\alpha \ln n$ for some fixed constant $\alpha > 1$. Consider any fixed type-2 structure P. Let $w_1, \ldots, w_c$ be the worms participating in round t that use this type-2 structure, $c \geq \tilde{C}/(\frac{32B\hat{\Delta}}{(L-1)\tilde{C}})^{2^{t-1}-1}$. Further let the binary random variable $X_i = 1$ if and only if w_i fails to reach its destination in round t. Then $X = \sum_{i=1}^{c} X_i$ is a random variable denoting the path congestion at P after round t. In order to bound X, we introduce a random variable Y that is defined as follows:

For every $i \in \{1, \ldots, c\}$, let δ_i be the delay of worm w_i. Let $W = \{w_{c/2+1}, \ldots, w_c\}$. For all $i \in \{1, \ldots, c/2\}$, let the binary random variable Y_i be one if and only if there is a worm $w_j \in W$ that chooses the same wavelength as worm w_i and a delay $\delta_j \in [\delta_i - (L-1), \delta_i - 1]$. Furthermore let $Y = \sum_{i=1}^{c/2} Y_i$. We assume in the following that in case of priority routers the priorities of the worms are chosen in such a way that $\text{rank}(w_1) < \text{rank}(w_2) < \ldots < \text{rank}(w_c)$. Then the following claim holds for both serve-first and priority routers.

Claim A.1 *For any $\ell \in \{0, \ldots, c/2\}$ it holds that $Y = \ell$ implies that $X \geq \ell$.*

Proof. Consider any event that results in $Y = \ell$. Let us choose any worm w in W that causes some, say a set W', of the worms w_i with $i \in \{0, \ldots, c/2\}$ to have $Y_i = 1$. Let δ be the delay of w. Then either all worms of W' are discarded because of w or w has been discarded and also at least $|W'| - 1$ other worms in W', since only one worm with a delay in $[\delta - (L-1), \delta - 1]$ can survive. In both cases at least $|W'|$ worms are eliminated. Continuing this argument for the worms w_i with $Y_i = 1$ that have not been considered yet yields the claim. ∎

Hence $\Pr[X > \ell] \geq \Pr[Y > \ell]$ or equivalently $\Pr[X \leq \ell] \leq \Pr[Y \leq \ell]$ for all $\ell \in \{0, \ldots, c/2\}$. It therefore remains to find an upper bound for $\Pr[Y \leq \ell]$ in order to have an upper bound for $\Pr[X \leq \ell]$. First, let us bound the probability that $Y_i = 1$ for any i. Since $\Delta_t \geq L(\frac{\tilde{C}}{B} + 2)$ and $c \geq \log n$, we get

$$\begin{aligned}
\Pr[Y_i = 0] &= \sum_{d=0}^{\Delta_t - 1} \Pr[\delta_i = d] \cdot \Pr[Y_i = 0 \mid \delta_i = d] \\
&= \sum_{d=0}^{\Delta_t - 1} \frac{1}{\Delta_t} \prod_{w \in W} (1 - \Pr[w \text{ causes } Y_i = 1 \mid \delta_i = d]) \\
&= \sum_{d=0}^{\Delta_t - 1} \frac{1}{\Delta_t} \left(1 - \frac{\min\{d, L-1\}}{B\Delta_t}\right)^{c/2} \\
&\leq \frac{1}{\Delta_t} + \frac{L-1}{\Delta_t}\left(1 - \frac{1}{B\Delta_t}\right)^{c/2} + \left(1 - \frac{L}{\Delta_t}\right)\left(1 - \frac{L-1}{B\Delta_t}\right)^{c/2} \\
&\leq \frac{1}{\Delta_t} + \frac{L-1}{\Delta_t}\left(1 - \frac{c}{4B\Delta_t}\right) + \left(1 - \frac{L}{\Delta_t}\right)\left(1 - \frac{c(L-1)}{4B\Delta_t}\right)
\end{aligned}$$

$$\begin{aligned}
&= \frac{L}{\Delta_t} - \frac{c(L-1)}{4B\Delta_t^2} + 1 - \frac{c(L-1)}{4B\Delta_t} - \frac{L}{\Delta_t}\left(1 - \frac{c(L-1)}{4B\Delta_t}\right) \\
&= 1 - \frac{c(L-1)}{4B\Delta_t} + \frac{c(L-1)^2}{4B\Delta_t^2} \\
&\leq 1 - \frac{c(L-1)}{8B\Delta_t} \quad .
\end{aligned}$$

Hence,

$$\Pr[Y_i = 1] \geq \frac{c(L-1)}{8B\Delta_t} \quad .$$

Thus we get

$$\begin{aligned}
\mathrm{E}[Y] &= \sum_{i=1}^{c/2} \mathrm{E}[Y_i] \;\geq\; \frac{\tilde{C}_t}{2} \cdot \frac{(L-1)\tilde{C}_t}{8B\hat{\Delta}} \\
&\geq \frac{L-1}{16B\hat{\Delta}} \cdot \left(\frac{\tilde{C}}{\left(\frac{32B\hat{\Delta}}{(L-1)\tilde{C}}\right)^{2^{t-1}-1}} \right)^2 \\
&= 2 \cdot \frac{\tilde{C}}{\left(\frac{32B\hat{\Delta}}{(L-1)\tilde{C}}\right)^{2^{t-1}-1+2^{t-1}-1+1}} \;=\; 2\frac{\tilde{C}}{\left(\frac{32B\hat{\Delta}}{(L-1)\tilde{C}}\right)^{2^t-1}}
\end{aligned}$$

Let $\mu = 2\tilde{C}/(\frac{32B\hat{\Delta}}{(L-1)\tilde{C}})^{2^t-1}$. In order to apply the Chernoff bounds for $\Pr[Y \leq \mu/2]$ we have to show that the Y_i are independent. Let $\mathcal{A}$ be the set of all assignments $A \in ([B]\times[\Delta_t])^{c/2}$ of wavelengths and startup delays to the worms in W. Since the worms choose wavelengths and startup delays independently at random, it clearly holds for any assignment $A \in \mathcal{A}$ and any subset $\{i_1, \ldots, i_\ell\} \subseteq \{1, \ldots, c/2\}$ that

$$\Pr[Y_{i_1} \cdot \ldots \cdot Y_{i_\ell} \mid A] = \prod_{j=1}^{\ell} \Pr[Y_{i_j} \mid A]$$

and therefore

$$\Pr[Y_{i_1} \cdot \ldots \cdot Y_{i_\ell}] = \prod_{j=1}^{\ell} \Pr[Y_{i_j}] \quad .$$

Hence we can use Chernoff bounds (see [18]) to prove that, for $\epsilon = \frac{1}{2}$,

$$\begin{aligned}
\Pr\left[X \leq \frac{\tilde{C}}{\left(\frac{32B\hat{\Delta}}{(L-1)\tilde{C}}\right)^{2^t-1}}\right] &\leq \Pr\left[Y \leq \frac{\tilde{C}}{\left(\frac{32B\hat{\Delta}}{(L-1)\tilde{C}}\right)^{2^t-1}}\right] \leq \Pr\left[Y \leq (1-\epsilon)\mu\right] \\
&\leq \mathrm{e}^{-\epsilon^2\mu/2} \;\leq\; \mathrm{e}^{-\alpha \ln n} \;=\; \left(\frac{1}{n}\right)^{\alpha} \quad .
\end{aligned}$$

Hence for $\alpha > 1$ the path congestion after round t is bounded by $\tilde{C}/(\frac{32B\hat{\Delta}}{(L-1)\tilde{C}})^{2^t-1}$ for all type-2 structures, w.h.p.

A.4 Proof of Lemma 3.1

We start with the valid embedding of k worms in $\mathcal{W}(t)$ that minimizes $\sum_{i=1}^{t} m_i$. Clearly, in each level $i \geq 1$ the number of different worms has to be at least 2. Hence we can choose the distribution $m_i = 2$ for all $i \in \{1, \ldots, t - \lceil \log k \rceil + 1\}$. Then it holds that $\ell_1 = 1$ and $\ell_i = 0$ for all $i \in \{2, \ldots, t - \lceil \log k \rceil + 1\}$. Therefore we get:

$$\prod_{i=1}^{t} \left(\frac{L(\mathrm{e}m_{i-1})^{3/2}}{B\Delta_{t-i+1}} \right)^{\frac{2(m_{i-1}-\ell_i)}{3}} \leq \prod_{i=2}^{t-\lceil \log k \rceil + 1} \left(\frac{L(2\mathrm{e})^{3/2}}{B\Delta_{t-i+1}} \right)^{4/3} \leq \left(\frac{13L}{B\Delta_t} \right)^{\frac{4}{3}(t - \lceil \log k \rceil)} .$$

In the following we show that for all other distributions of the m_i it holds that

$$\prod_{i=1}^{t} \left(\frac{L(\mathrm{e}m_{i-1})^{3/2}}{B\Delta_{t-i+1}} \right)^{\frac{2(m_{i-1}-\ell_i)}{3}} \leq \left(\frac{13L}{B\Delta_t} \right)^{\frac{4}{3}(t - \lceil \log k \rceil)} \tag{7}$$

if $B\Delta_i \geq (3\mathrm{e})^3 L k^{3/2}$.

Consider increasing the number m_j of worms at a stage $j < t$ with $m_j < m_{j+1}$ by 1. Then two terms in the product in (7) change: the $(i = j)$-term and the $(i = j + 1)$-term. Before increasing m_j, these terms are

$$\left(\frac{L(\mathrm{e}m_{j-1})^{3/2}}{B\Delta_{t-j+1}} \right)^{\frac{2(m_{j-1}-\ell_j)}{3}} \left(\frac{L(\mathrm{e}m_j)^{3/2}}{B\Delta_{t-j}} \right)^{\frac{2(m_j-\ell_{j+1})}{3}} \tag{8}$$

and after increasing m_j by 1, they change to

$$\left(\frac{L(\mathrm{e}m_{j-1})^{3/2}}{B\Delta_{t-j+1}} \right)^{\frac{2(m_{j-1}-(\ell_j+1))}{3}} \left(\frac{L(\mathrm{e}(m_j+1))^{3/2}}{B\Delta_{t-j}} \right)^{\frac{2((m_j+1)-(\ell_{j+1}-1))}{3}} . \tag{9}$$

It holds that

$$\begin{aligned}
(8) \geq (9) \quad &\Leftrightarrow \quad \left(\frac{L(\mathrm{e}m_{j-1})^{3/2}}{B\Delta_{t-j+1}} \right)^{2/3} \geq \left(\frac{m_j + 1}{m_j} \right)^{m_j - \ell_{j+1}} \left(\frac{L(\mathrm{e}(m_j+1))^{3/2}}{B\Delta_{t-j}} \right)^{4/3} \\
&\Leftarrow \quad \left(\frac{L(\mathrm{e}m_{j-1})^{3/2}}{B\Delta_{t-j+1}} \right)^{2/3} \geq \mathrm{e} \cdot \left(\frac{L(\mathrm{e}(2m_{j-1}+1))^{3/2}}{B\Delta_{t-j+1}} \right)^{4/3} \\
&\Leftarrow \quad \frac{B\Delta_{t-j+1}}{L m_{j-1}^{3/2}} \geq (3\mathrm{e})^3 \\
&\Leftarrow \quad B\Delta_{t-j+1} \geq (3\mathrm{e})^3 L k^{3/2}
\end{aligned}$$

Since any distribution of m_i can be obtained from the initial distribution above by performing the action described above again and again, the lemma follows.