

---

# The Algebraic Path Problem for Graph Metrics

---

Enrique Fita Sanmartín<sup>1</sup> Sebastian Damrich<sup>1</sup> Fred A. Hamprecht<sup>1</sup>

## Abstract

Finding paths with optimal properties is a foundational problem in computer science. The notions of shortest paths (minimal sum of edge costs), minimax paths (minimal maximum edge weight), reliability of a path and many others all arise as special cases of the “algebraic path problem” (APP). Indeed, the APP formalizes the relation between different semirings such as min-plus, min-max and the distances they induce. We here clarify, for the first time, the relation between the potential distance and the log-semiring. We also define a new unifying family of algebraic structures that include all above-mentioned path problems as well as the commute cost and others as special or limiting cases. The family comprises not only semirings but also strong bimonoids (that is, semirings without distributivity). We call this new and very general distance the “log-norm distance”. Finally, we derive some sufficient conditions which ensure that the APP associated with a semiring defines a metric over an arbitrary graph.

## 1. Introduction

Graphs are a versatile abstraction permitting the modeling and analysis of an extremely broad range of problems from vision, NLP and learning with structured data. Measuring the similarity between the nodes in a graph is, in turn, a task of fundamental importance that often decides on success or failure of an application. Consequently, the study and development of graph node metrics with different properties is a problem of high interest.

The shortest path distance is arguably the most popular graph node metric. Given two nodes in a graph with edge costs, the shortest path problem aims to find the path with minimum cost, i.e., the path for which the sum over the cost of its constituent edges is minimized. This problem

is determined by the min and + operations: the + to indicate that edge costs are summed along any path; and the min to state that the overall cost is given by the smallest cost of any single path. Together, these form the “min-plus” semiring. A semiring is an algebraic structure with two operations that relaxes the concept of a ring by dropping the requirement for inverses under the “addition” operation. The algebraic path problem (APP) generalizes the notion of shortest path by replacing the min-plus semiring by an alternative semiring. Different semirings result in dramatically different preferences of paths, see the toy example in Table 1. This generalization encompasses a great variety of problems and applications in diverse research areas like NLP (Cortes et al., 2004; Schwartz et al., 2018) or routing protocols (Griffin & Sobrinho, 2005). For more applications see (Gondran & Minoux, 2008; Baras & Theodorakopoulos, 2010).

Other common graph metrics that raise from the APP framework are the commute cost distance (CCD) (Klein & Randić, 1993; Chandra et al., 1996; Eisner, 2001), which calculates the average first passage cost between two nodes in both directions; and the minimax or bottleneck shortest path distance (Maggs & Plotkin, 1988; Challa et al., 2019), which computes the path with minimal maximum edge cost among its edges.

In some situations, these metrics may fail to take the global structure of the graph into consideration. On the one hand, the minimax and shortest path distances are determined by a single path. Thus, they may ignore the topology of the surrounding graph since the degree of connectivity between the nodes is not reflected by the metrics. On the other hand, though CCD weighs all paths, it is known that for large graphs, it only takes the degree of the source and target nodes into account (Nadler et al., 2009; Luxburg et al., 2010).

We aim to combine the advantages and compensate for the deficiencies of these metrics. To do so, we propose a novel parametrized family of distances, dubbed log-norm distances, that interpolate between the shortest path, CCD and the minimax distances up to a constant factor. We base our interpolation on the algebraic path problem. First, we present a family of semirings whose associated APP yields a metric which interpolates between the shortest and min-

---

<sup>1</sup>IWR at Heidelberg University, 69120 Heidelberg, Germany. Correspondence to: <>.

Table 1. Family of log-norm distances and limit cases (**bold**) with their associated algebraic structure (*red*). The family of log-norm distances interpolates between the commute cost, shortest path and minimax distances. The algebraic structures associated to the cells in cyan ( $r > 1$  and  $0^+ < \mu < \infty$ ) do not form a semiring, but a strong bimonoid. The distances derived from these bimonoids are presented here for the first time. The graphs below represent the contribution of individual edges/paths to the  $s, t$ -distance in a toy problem. The cost of an edge is given by its length and wider edges have higher relevance. The  $r$  parameter regulates the impact of the edge costs in the paths, with higher  $r$  favoring paths with shorter edges. The parameter  $\mu$  regulates the distribution of the contribution of the paths; higher  $\mu$  favors the concentration of the distribution into a smaller number of lowest cost paths. Different distances differ radically in what part of a graph they emphasize

$r \backslash \mu$	$0^+$	$(0, \infty)$	$\infty$
1	<p><i>Eisner semiring</i>  <b>Commute cost distance</b>  <math>\mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [c(\pi)] + \mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} [c(\pi)]</math>                      (Klein &amp; Randić, 1993)</p> 	<p><i>Log-semiring</i>  <b>Potential distance</b>  <math>-\frac{1}{\mu} \left( \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ e^{-\mu c(\pi)} \right] \right) + \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} \left[ e^{-\mu c(\pi)} \right] \right) \right)</math>                      (Kivimäki et al., 2014; Françoisse et al., 2017)</p> 	<p><i>Min-plus semiring</i>  <b>Shortest path distance</b>  <math>\min_{\pi \in \mathcal{P}_{st}} c(\pi)</math></p> 
$(1, \infty)$	<p><i>Exp-norm bimonoid</i>  <b>Exp-norm distance</b>  <math>\mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [\ c(\pi)\ _r] + \mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} [\ c(\pi)\ _r]</math></p> 	<p><i>Log-norm bimonoid</i>  <b>Log-norm distance</b>  <math>-\frac{1}{\mu} \left( \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ e^{-\mu \ c(\pi)\ _r} \right] \right) + \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} \left[ e^{-\mu \ c(\pi)\ _r} \right] \right) \right)</math></p> 	<p><i>Min-norm semiring</i>  <b>Min-norm distance</b>  <math>\min_{\pi \in \mathcal{P}_{st}} \ c(\pi)\ _r</math>                      (Mckenzie &amp; Damein, 2019)</p> 
$\infty$	<p><i>Exp-max bimonoid</i>  <b>Exp-max distance</b>  <math>\mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ \max_{e \in \pi} c(e) \right] + \mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} \left[ \max_{e \in \pi} c(e) \right]</math></p> 	<p><i>Log-max bimonoid</i>  <b>Log-max distance</b>  <math>-\frac{1}{\mu} \left( \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ e^{-\mu \max_{e \in \pi} c(e)} \right] \right) + \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} \left[ e^{-\mu \max_{e \in \pi} c(e)} \right] \right) \right)</math></p> 	<p><i>Minimax semiring</i>  <b>Minimax distance</b>  <math>\min_{\pi \in \mathcal{P}_{st}} \max_{e \in \pi} c(e)</math>                      (Maggs &amp; Plotkin, 1988)</p> 

imax distance. Moreover, we study the potential distance (Kivimäki et al., 2014; Françoisse et al., 2017), which interpolates between the CCD and the shortest path distance. We redefine this distance via the well-known log-semiring and its associated APP. As far as we know, we are the first to formalize the interpolations of these metrics from the APP point of view. Finally, we introduce a greater parametrized set of algebraic structures. Though not all the members of this family of algebraic structures define a semiring, but strong bimonoids (Droste et al., 2010), their associated APP define the log-norm family of metrics. These distances are parametrized by a parameter  $r$ , which controls the relevance of higher cost edges in the paths; and a parameter  $\mu$ , which regulates how much individual paths contribute to the distance, favoring paths with lower cost. Table 1 summarizes these relations and highlights, on a toy example, how drastically different distances vary. Clearly, these properties greatly impact the machine learning on graphs.

Intrigued by the fact that so many metrics can be retrieved from the APP framework, we finally study under which circumstances the APP associated with a semiring defines a metric. In concrete, we focus on the setting where only hitting paths (paths whose last node appears only once)

are considered. We find that one of the key factors lies in the function that maps elements of the semiring to the non-negative real numbers. Under natural conditions, we find that it is necessary, but not sufficient, that this function is subadditive with respect to the product operation of the semiring. We also provide sufficient conditions based on the monotonicity behavior of the function.

In summary we: 1) review the APP and how some of the most common graph metrics can be recovered by defining appropriate semirings, demonstrating this relation for the first time for some of them; 2) introduce a novel unifying family of graph metrics, dubbed log-norm distances, which interpolate between the CCD, minimax and shortest path metrics up to a constant factor; 3) study under which conditions the APP associated with a semiring defines a graph metric.

### 1.1. Related Work

There has been much work on interpolations between the shortest path distance and the CCD. In (Yen et al., 2008), these distances are interpolated by a family of dissimilarities inspired by the randomized shortest paths framework (RSP)

Table 2. Semiring examples.  $\mathcal{P}(A)$  is the power set of a set  $A$ .

Semiring	$S$	$\oplus$	$\otimes$	$\bar{0}$	$\bar{1}$	class
Min-plus	$\mathbb{R}^+$	min	+	$\infty$	0	selective
Minimax	$\mathbb{R}^+$	min	max	$\infty$	0	selective
Power set	$\mathcal{P}(A)$	$\cup$	$\cap$	$\emptyset$	$A$	idempotent

(Saerens et al., 2009). These dissimilarities do not fulfill the triangle inequality, ergo they are not metrics. Also, based on the RSP, (Françoisse et al., 2017) and (Kivimäki et al., 2014) define the same interpolating family of distances, which they call potential and free energy distances, respectively. We study this distance as an instance of the APP framework. The logarithmic forest distances (Chebotarev, 2011) are a family of distances, based on the matrix forest theorem, which also interpolate between the SP and the CCD. The walk distances (Chebotarev, 2012) are a broader set of distances which include the logarithmic forest, the CCD and the SP distances. The set of  $p$ -resistance distances (Alamgir, 2012) generalize the definition of the effective resistance distance (Klein & Randić, 1993), which is proportional to the CCD (Chandra et al., 1996). For  $p = 1$ , the SP distance is retrieved.

Closely related to our work, (Kim & Choi, 2013) define a family of similarity measures whose behavior resembles the one observed for the log-norm distances. Our metric adapts their similarities and transforms them into a proper distance while building a connection between the CCD, shortest path and minimax distances. As far as we know, the family of distances proposed in (Gurvich, 2010) is the only one that interpolates between the three aforementioned distances (CCD, minimax and shortest path distances) besides our proposal. Similar to the  $p$ -resistance approach, Gurvich generalizes the effective resistance concept by introducing two parameters. One difference between this proposal and ours lies in the fourth limit case that emerges from the approaches taken. While our limit computes the expected maximum edge cost of a path, its limit retrieves the inverse value of maximum flow between two nodes.

## 2. Preliminaries

### 2.1. Semirings

To fix notation and give necessary background, in this section we review the semiring algebraic structure, which constitutes the primary tool to understand the algebraic path problem. For a more extensive analysis of semirings, we refer the interested reader to (Gondran & Minoux, 2008).

**Definition 2.1.** A **semiring** is an algebraic structure  $(S, \oplus, \otimes, \bar{0}, \bar{1})$  formed by a set  $S$  and two binary closed operations,  $\oplus$  and  $\otimes$ , with the following properties  $\forall a, b, c \in S$ :

- $\oplus$  **commutativity**:  $a \oplus b = b \oplus a$

- $\oplus$  **associativity**:  $(a \oplus b) \oplus c = a \oplus (b \oplus c)$
- $\bar{0}$  **neutral element of  $\oplus$** :  $a \oplus \bar{0} = \bar{0} \oplus a = a$
- $\otimes$  **associativity**:  $(a \otimes b) \otimes c = a \otimes (b \otimes c)$
- $\bar{1}$  **neutral element of  $\otimes$** :  $a \otimes \bar{1} = \bar{1} \otimes a = a$
- **distributivity of  $\otimes$  relative to  $\oplus$** :

$$(a \oplus b) \otimes c = a \otimes c \oplus b \otimes c, \quad c \otimes (a \oplus b) = c \otimes a \oplus c \otimes b$$

- $\bar{0}$  **absorbing for  $\otimes$** :  $a \otimes \bar{0} = \bar{0} \otimes a = \bar{0}$

Individually,  $\oplus$  and  $\otimes$  define a monoid over  $S$ . A semiring is idempotent if for all  $a \in S$ ,  $a \oplus a = a$ . Furthermore, a semiring is called selective if  $a \oplus b \in \{a, b\}$ . If we drop the distributivity property from the list of requirements of a semiring, then the algebraic structure is called **strong bimonoid** (Droste et al., 2010). Table 2 summarizes some common semirings.

A semiring has a canonical preorder relation<sup>1</sup>,  $\preceq$ , given by

$$a \preceq b \iff \exists c \in S : a \oplus c = b. \quad (1)$$

As for the usual addition and product operations, we can extend the  $\oplus$  and  $\otimes$  to the matrix domain. Let  $A, B$  be two matrices in  $S^{n \times m}$ , then the following operations define a semiring over the matrices on  $S$ :

$$[A \oplus B]_{ij} = A_{ij} \oplus B_{ij}, \quad [A \otimes B]_{ij} = \bigoplus_k A_{ik} \otimes B_{kj}.$$

### 2.2. Graph notation

Let  $G = (V, E)$  be a directed graph where  $V$  and  $E$  represent the sets of vertices and edges respectively. A path  $\pi = (v_0, \dots, v_k)$  from  $s$  to  $t$  is defined as a sequence of adjacent nodes, i.e.  $(v_i, v_{i+1}) \in E$  with  $v_0 = s$  and  $v_k = t$ . Note that a node can appear multiple times in a path. A hitting path is a path whose last node,  $t$ , appears only once. The set of all paths from  $s$  to  $t$  will be represented by  $\mathcal{P}_{st}$ . The subset of paths with exactly  $k$  edges is denoted by  $\mathcal{P}_{st}[k]$ . Analogously, we define the variant  $\mathcal{P}_{st}^h$  for the set of hitting paths.

Let  $(S, \oplus, \otimes, \bar{0}, \bar{1})$  be a semiring. We say that a graph  $G$  is  $S$ -valued or  $S$ -graph if there is a cost function  $c : V \times V \rightarrow S$  that assigns a cost  $c(e) \in S$  to each edge. We set  $c(e) = \bar{0}$  if and only if  $e \notin E$ . Additionally, the cost of a path is defined as the product of the cost of its edges,  $c(\pi) = \bigotimes_{e \in \pi} c(e)$ . For  $S$ -valued graphs, the entry  $A_{ij}$  of the adjacency matrix is equal to the cost of the edge  $(i, j)$ .

### 2.3. Algebraic Path Problem

Given a graph with  $c(e) \in \mathbb{R}^+$ , the shortest path problem (SPP) computes

$$\min_{\pi \in \mathcal{P}_{st}} \sum_{e \in \pi} c(e). \quad (2)$$

<sup>1</sup>Reflexive ( $a \preceq a$ ) and transitive ( $a \preceq b$  and  $b \preceq c \rightarrow a \preceq c$ ) properties are satisfied, but antisymmetry ( $a \preceq b$  and  $b \preceq a \rightarrow a = b$ ) may not hold.

As previously mentioned, the min and  $+$  operations characterize the min-plus semiring (Simon, 1978; Pin, 1998). The algebraic path problem (APP) extends the SPP through the use of general binary operations  $\oplus$  and  $\otimes$  that jointly form a semiring. On the one hand, the  $\otimes$  operation ( $+$  in the SPP) acts over the cost of the edges. It can be interpreted as an edge concatenation operator which constructs the path by “multiplying” the cost of the edges. On the other hand, the  $\oplus$  operation (min in the SPP) acts over paths and behaves like a path aggregation operator which condenses the cost of different paths. When the semiring is selective (e.g. min-plus semiring),  $\oplus$  can also be interpreted as a choice operator where a single path is being selected. Formally, the APP generalizes (2) by calculating

$$\text{APP}(s, t) := \bigoplus_{\pi \in \mathcal{P}_{st}} \bigotimes_{e \in \pi} c(e) = \bigoplus_{\pi \in \mathcal{P}_{st}} c(\pi). \quad (3)$$

Let  $A$  be the adjacency matrix of the graph  $G$ . It can be verified that  $[A^k]_{st} = \bigoplus_{\pi \in \mathcal{P}_{st}[k]} c(\pi)$ . Since  $\bigcup_k \mathcal{P}_{st}[k] = \mathcal{P}_{st}$  and  $\mathcal{P}_{st}[k] \cap \mathcal{P}_{st}[k'] = \emptyset$  if  $k \neq k'$ , we obtain

$$\begin{aligned} \lim_{k \rightarrow \infty} [I \oplus A \oplus \dots \oplus A^k]_{st} &= \bigoplus_{k=0}^{\infty} \bigoplus_{\pi \in \mathcal{P}_{st}[k]} c(\pi) \\ &= \bigoplus_{\pi \in \mathcal{P}_{st}} c(\pi), \end{aligned} \quad (4)$$

where  $I$  is the diagonal matrix with  $\bar{1}$ 's in the diagonal. The limit in equation (4) is called the closure of  $A$  and it is denoted by  $A^*$ . Note that  $A^*$ , and consequently  $\text{APP}(s, t)$ , may not always exist. An interesting property of  $A^*$  is that it is the minimal solution of  $X = A \otimes X \oplus I$  (see Proposition 6.2.2, Ch. 3 (Gondran & Minoux, 2008)). In the semirings considered in this paper, these limits will be well defined. Alternatively, the closure of an element  $a$  in a semiring may be defined as the solution of equation  $x = a \otimes x \oplus \bar{1}$  instead of as a limit (Lehmann, 1977).

Specialized algorithms for the SPP have been generalized to the APP. The Dijkstra algorithm (Dijkstra, 1959), which solves the Single Source Shortest Path Problem, has been generalized for some specific semirings (Mohri, 2002; Huang, 2008). Analogously, the Floyd-Warshall algorithm (Floyd, 1962), which solves the All Pairs of Shortest Paths Problem, can be generalized to solve the APP for any  $S$ -valued graph, for which  $A^*$  exists. This algorithm generalizes the Gauss–Jordan Method and solves  $X = A \otimes X \oplus I$  (Carre, 1971).

## 2.4. Semiring Distances

If the edge costs of a graph are strictly positive, the shortest path distance defines a metric over the nodes of the graph. In this section, we will review other common graph metrics (minimax and CCD) and present how these can be expressed in terms of the APP.

### 2.4.1. MINIMAX DISTANCE

An alternative popular distance in the literature is the minimax distance (Maggs & Plotkin, 1988; Kim & Choi, 2013; Challa et al., 2019). As its name indicates, this metric can be retrieved from the APP framework with the underlying minimax semiring (see Table 2)

$$\bigoplus_{\pi \in \mathcal{P}_{st}} \bigotimes_{e \in \pi} c(e) = \min_{\pi \in \mathcal{P}_{st}} \max_{e \in \pi} c(e). \quad (5)$$

This semiring calculates the so-called minimax path, i.e., the path which minimizes the most expensive edge of a path between two nodes. The minimax semiring can be used to calculate a minimum spanning tree (mST) since every simple path (path where each node appears once) between two nodes in a mST is a minimax path (Maggs & Plotkin, 1988).

### 2.4.2. COMMUTE COST DISTANCE

A prominent metric used for graphs is the commute cost distance (CCD) (Fouss et al., 2016). To compute the CCD between two nodes, it is necessary that each edge  $(i, j)$  has two values,  $p_{ij}$  and  $c_{ij}$ , associated to it. The first represents the probability that a random walker located at node  $i$  will transition from node  $i$  to node  $j$  through edge  $(i, j)$ . Usually,  $p_{ij}$  is set proportional to some weight  $w_{ij}$  that measures the affinity between the nodes  $i$  and  $j$ . The value  $c_{ij}$  is a positive value which indicates some kind of cost associated to the traversal of the edge  $(i, j)$ . The first hitting cost dissimilarity between two nodes  $s$  and  $t$ ,  $\mathcal{H}(s, t)$ , is the expected cost that it takes a random walker starting at  $s$  to reach  $t$  for the first time. The CCD symmetrizes this dissimilarity. Formally,

$$\text{CCD}(s, t) = \underbrace{\mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [c(\pi)]}_{\mathcal{H}(s, t)} + \underbrace{\mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} [c(\pi)]}_{\mathcal{H}(t, s)}. \quad (6)$$

The term  $\mathcal{H}(s, t)$  can be expressed in the framework of the APP if one uses the so-called expectation semirings (Baras & Theodorakopoulos, 2010). In concrete, the Eisner semiring (Eisner, 2001), defined over the ground set  $\mathbb{R}^+ \times \mathbb{R}^+$  with operations:

$(a, b) \oplus (c, d) = (a+c, b+d)$ ,  $(a, b) \otimes (c, d) = (ac, cb+ad)$ , recovers  $\mathcal{H}(s, t)$ . If we set the edge costs of the graph as  $(p_e, p_e c_e)$ , it can be verified that (see appendix B)

$$\mathcal{H}(s, t) = \left[ \bigoplus_{\pi \in \mathcal{P}_{st}^h} \bigotimes_{e \in \pi} (p_e, p_e c_e) \right]_2, \quad (7)$$

where the subindex 2 indicates the second entry. Therefore, CCD can be expressed in terms of APP. Note that (7) considers only hitting paths, in contrast to the general APP. The set  $\mathcal{P}_{st}^h$  over an arbitrary graph  $G$  is equal to the set  $\mathcal{P}_{st}$  over the graph  $G^h[t]$ , where the out-going edges of node  $t$

have been removed. Hence,  $t$  becomes an absorbing node and any path to  $t$  is therefore a hitting path. Consequently, CCD requires two APP on the graphs  $G^h[t]$  and  $G^h[s]$  to compute  $\mathcal{H}(s, t)$  and  $\mathcal{H}(t, s)$  respectively. The shortest path problem can also be constrained to hitting paths, since, as far as the costs are positive, each node appears only once in the optimal path.

*Remark 2.2.* It is known (Kivimäki et al., 2014, appendix) that given some fixed random walker probabilities,  $p_e$ , CCD is proportional to the commute time distance (CTD), i.e. the expected length of a path (expected number of edges, that is  $c_e = 1 \forall e \in E$ ). The  $c_e$  values only determine the proportionality constant between CCD and CTD.

### 3. Log-Norm Distances

In this section, we propose the novel family of log-norm distances, which interpolate between the above mentioned distances up to a constant factor. To do so, we introduce an interpolating family of distances between the minimax and shortest path. Moreover, we study the potential distance (Françoisse et al., 2017), which interpolates between the CCD and the shortest path distance. We prove that both metrics are instances of the APP framework once the appropriate semiring has been defined. Finally, we combine these semirings to define a greater family of strong bimonoids that will define the log-norm family of distances (Table 1).

#### 3.1. Shortest Path and Minimax Distance Interpolation

The min-plus and min-max semiring can be continuously interpolated by the semiring  $S = (R^+, \min, \otimes_r, \infty, 0)$  where  $a \otimes_r b = \sqrt[r]{a^r + b^r}$ . We are not aware that this algebraic structure has been acknowledged in the literature as a semiring, though it has been used in earlier works (Kim & Choi, 2013). Therefore, we dub it min-norm semiring. In appendix A, we demonstrate that it is indeed a semiring. The APP derived from the min-norm semiring is defined as

$$\bigoplus_{\pi \in \mathcal{P}_{st}} \bigotimes_{e \in \pi} c(e) = \min_{\pi \in \mathcal{P}_{st}} \sqrt[r]{\sum_{e \in \pi} (c(e))^r}. \quad (8)$$

Clearly, when  $r = 1$ ,  $\otimes_1$  is equal to the regular sum, and consequently we recover the min-plus semiring. On the other extreme, when  $r \rightarrow \infty$ ,  $\otimes_\infty$  is reduced to the max operation, hence the minimax semiring is retrieved. In appendix F.5, we show that (8) defines a metric over the nodes of a graph, which in turn also interpolates between the shortest path and minimax distances. This distance is also known in the literature as the power weighted shortest path metric (Mckenzie & Damelin, 2019).

The  $r$  parameter, which characterizes  $\otimes_r$ , regulates the impact of the edge costs in the paths. On the one hand, for high  $r$ , the min-norm path tends to have edges with low

cost, though it may contain a higher number of edges (more similar to the minimax distance). On the other hand, for lower  $r$ , the min-norm path is more dominated by the total additive cost of the edges, which must be low overall (closer to the shortest path distance) but may contain edges whose cost are relatively high. Last column of Figure 1 illustrates this pattern, which was already pointed out in (Kim & Choi, 2013). The shortest path ( $r = 1$ ) contains only three, but one long edge in contrast to the minimax path ( $r = \infty$ ) which contains many short edges. The min-norm path ( $1 < r < \infty$ ) interpolates between both path patterns.

#### 3.2. Commute Cost and Shortest Path Distance Interpolation

As mentioned in the introduction, CCD has some inconveniences if the graph is large (Luxburg et al., 2010). Many node distances have been proposed that interpolate between the shortest path and the CCD distances in order to exploit the benefits of both metrics. Among them, we call attention to the potential distance (PD) (Kivimäki et al., 2014; Françoisse et al., 2017). This distance is based on the randomized shortest paths (RSP) framework (Saerens et al., 2009). The PD can be interpreted as the logarithm of the expected reward  $\exp(-\mu c(\pi))$  of the paths. Formally,

$$\begin{aligned} \text{PD}(s, t) = & -\frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [\exp(-\mu c(\pi))] \right) \\ & -\frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} [\exp(-\mu c(\pi))] \right), \end{aligned} \quad (9)$$

where the parameter  $\mu$  regulates implicitly the entropy of the distribution defined by the RSP framework. Kivimäki et al. showed that when  $\mu \rightarrow 0$ , PD tends to CCD, and when  $\mu \rightarrow \infty$ , it tends to the shortest path distance up to a constant factor (see appendix E).

This distance also fits in the APP framework if one uses the log-semiring (Lotito et al., 2005; Litvinov, 2005) defined by  $a \oplus_\mu b = -\frac{1}{\mu} \log(\exp(-\mu a) + \exp(-\mu b))$ ,  $a \otimes b = a + b$ . Though the distance and the semiring were already known, we show the relation between them for the first time. Setting the edge costs of the graph equal to  $-\log(p_e \exp(-\mu c_e)) / \mu$ , it can be verified that (see appendix C)

$$\begin{aligned} \text{PD}(s, t) = & \bigoplus_{\pi \in \mathcal{P}_{st}^h}^\mu \bigotimes_{e \in \pi}^\mu -\frac{1}{\mu} \log(p_e \exp(-\mu c_e)) \\ & + \bigoplus_{\pi \in \mathcal{P}_{ts}^h}^\mu \bigotimes_{e \in \pi}^\mu -\frac{1}{\mu} \log(p_e \exp(-\mu c_e)) \end{aligned} \quad (10)$$

*Remark 3.1.* In (Françoisse et al., 2017) it was noted that the PD could also be computed by applying a generalization of the Bellman-Ford formula (Bellman, 1958). This is a direct consequence of the APP framework, since the PD can be retrieved by the log-semiring.

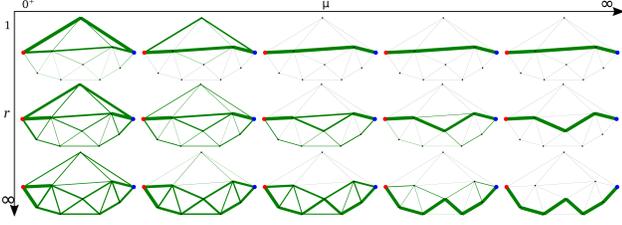


Figure 1. Schematic log-norm distance relevance of paths for different values of  $\mu$  and  $r$  in a graph. We computed the 1000 shortest paths according to the costs  $\|c(\pi)\|_r$  for different  $r$  and  $\mu$  values. The edge width is proportional to  $\sum_{\pi: e \in \pi} \Pr(\pi) \exp(-\mu \|c(\pi)\|_r)$ . It visualizes how much each edge contributes to the value of the APP: the wider the edge, the more significant. The random walker probabilities are uniform at each node. The parameter  $\mu$  regulates how the importance of the paths is distributed conditioned by their cost, while  $r$  regulates the cost of the paths.

*Remark 3.2.* It is worth to mention that although the PD interpolates between the shortest path and CCD distances, the log-semiring does not interpolate between the min-plus and the Eisner semiring. When  $\mu \rightarrow \infty$ ,  $\oplus_\infty$  becomes the min operator and  $\lim_{\mu \rightarrow \infty} -\log(p_e \exp(-\mu c_e))/\mu = c_e$ . Thus, the min-plus semiring arises. Nonetheless, when  $\mu \rightarrow 0^+$ ,  $\oplus_{0^+}$  and the costs themselves are not well defined.

In contrast to the min-norm semirings, the log-semiring is not selective, i.e. it does not make a choice over the paths, but aggregates their costs. The operation  $\oplus_\mu$  weighs all the paths by their probability and cost. On the one hand, when  $\mu$  is close to  $0^+$ , the metric considers all paths based on their probability. In this case, the larger the number of paths with high probability exist between two nodes (the effect of the costs is negligible due to remark 2.2), the lower will be the distance. Intuitively, its more likely that the random walker is absorbed earlier if there are more connections between  $s$  and  $t$ . On the other hand, when  $\mu$  is high, thanks to the RSP framework, the paths with lower cost have higher probability. Therefore, the connecting paths with low cost become relevant. In the limit case  $\mu \rightarrow \infty$ , only the shortest paths are taken into account. See first row of Figure 1.

### 3.3. The family of Log-Norm Distances

In the previous sections we have shown how popular node metrics can be posed as particular instances of the APP. In concrete, we presented a semiring that interpolates between the shortest path and minimax distances via the parameter  $r$ , which conditions the  $\otimes_r$  operation. Additionally, we have discussed the log-semiring, whose APP interpolates between the CCD and the shortest path distances via a parameter  $\mu$  that determines the  $\oplus_\mu$  operation. A natural question arises: *is there any semiring that defines a distance interpolating between the CCD and the minimax distance?*

In this section we aim to answer this question by proposing a family of strong bimonoids, whose associated APP interpolates between those distances.

The key operators that allow to interpolate between the distances are  $\oplus_\mu$  and  $\otimes_r$ . To relate all above mentioned distances, we propose to define an algebraic structure that combines the Eisner, log- and min-norm semirings. We define the following operations over  $\mathbb{R}^+ \times \mathbb{R}^+$

$$(a, b) \oplus_\mu (c, d) = \left( 1, -\frac{1}{\mu} \log \left( ae^{-\mu b} + ce^{-\mu d} \right) \right) \quad (11)$$

$$(a, b) \otimes_r (c, d) = \left( ac, \sqrt[r]{b^r + d^r} \right).$$

Unfortunately, the distributive property of  $\oplus_\mu$  with respect to  $\otimes_r$  does not hold for arbitrary  $r$  and  $\mu$  values. Consequently, these operations do not define a semiring, but a strong bimonoid. Nonetheless, its APP, with edge costs equal to  $(p_e, c_e)$  with  $p_e, c_e$  as defined in Section 2.4.2, still defines a distance which we name log-norm distance (LN):

$$\begin{aligned} \text{LN}(s, t) &= \left[ \bigoplus_{\pi \in \mathcal{P}_{st}^h} \bigotimes_{e \in \pi} (p_e, c_e) + \bigoplus_{\pi \in \mathcal{P}_{ts}^h} \bigotimes_{e \in \pi} (p_e, c_e) \right]_2 \\ &= -\frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [\exp(-\mu \|c(\pi)\|_r)] \right) \\ &\quad -\frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} [\exp(-\mu \|c(\pi)\|_r)] \right), \quad (12) \end{aligned}$$

where  $\|c(\pi)\|_r = \sqrt[r]{\sum_{e \in \pi} (c_e)^r}$ . Table 1 summarizes all the graph metrics that the log-norm family includes along with the algebraic structure, whose APP defines the metric. Note that the algebraic structures do not interpolate between them for the same reason that was explained in remark 3.2.

The log-norm distances are closely related with the similarities presented in (Kim & Choi, 2013), defined as  $\sum_{\pi \in \mathcal{P}_{ts}^h} \exp(-\mu \|c(\pi)\|_r)$ . These similarities are not defined when  $\mu \rightarrow 0^+$ . In our setting, we transform these similarities into a metric by computing the expected similarity between two nodes followed by the  $-\log(\cdot)/\mu$  function. This way, we are able to study the limit  $\mu \rightarrow 0^+$ , which retrieves the CCD when  $r = 1$ .

The log-norm family can also be related with the Helmholtz free energy, following the same reasoning that related the PD with the free energy in (Kivimäki et al., 2014). We prove in appendix H, that the log-norm distance between two nodes  $s$  and  $t$  is equal to  $\Phi_r(\text{Pr}_{st}) + \Phi_r(\text{Pr}_{ts})$ , where  $\Phi_r(\text{Pr}_{st}) = \sum_{\pi \in \mathcal{P}_{st}^h} \text{Pr}_{st}(\pi) \|c(\pi)\|_r + \frac{1}{\mu} \text{KL}(\text{Pr}_{st}, \text{Pr}^{\text{ref}})$  is the free energy of a thermodynamical system and  $\text{Pr}_{st}$  is a probability distribution over the hitting paths between  $s$  and  $t$  that minimizes the free energy at a certain temperature  $1/\mu$ . KL denotes the KL-divergence, and  $\text{Pr}^{\text{ref}}$  denotes

a reference probability distribution over the hitting paths given by the random walker.

In Figure 1, we plot schematically the path relevance determined by the log-norm metric for different  $r$  and  $\mu$  values. We computed the 1000 shortest paths according to the costs  $\|c(\pi)\|_r$  for different  $r$  values. The edge width is proportional to  $\sum_{\pi: e \in \pi} \Pr(\pi) \exp(-\mu \|c(\pi)\|_r)$ . The width of the edges visualizes how much each edge contributes to the value of the APP. The wider the edges, the more significant. The random walker probabilities are set uniform at each node. We observe that for lower  $\mu$ , the influence among the edges of different paths is more distributed. As a consequence of remark 2.2, for the CCD ( $r = 1, \mu \rightarrow 0^+$ ), the cost of a path does not determine the influence of the path. Consequently, the more influential paths are the ones with higher probability. In our example, the high probability paths coincide with the shorter length paths (low number of edges) because we assume a uniform transition probability at each node. For  $r$  and  $\mu$  values close to 1 and 0 respectively, we expect similar behaviour. Indeed, this pattern is present in the top left corner graph. Nonetheless, when the value  $r$  is higher, the influence shifts to the paths situated in the lower part of the graph, which have a lower cost with respect to  $\|\cdot\|_r$ . This pattern indicates that the cost becomes more dominant. However, the mass is still distributed among all the edges of all the paths and the probabilities of the paths are still significant.

Contrarily, for higher  $\mu$ , the contribution of the paths shifts from paths with higher probability to paths with lower cost  $\|c(\pi)\|_r$ . In the limit  $\mu \rightarrow \infty$ , only the path with minimum cost is considered. This convergence is more pronounced for lower  $r$  values since  $\Pr(\pi) \exp(-\mu \|c(\pi)\|_r) \leq \Pr(\pi) \exp(-\mu \|c(\pi)\|_{r'})$  if  $r' \leq r$ . Thus, for higher  $r$  values the relevance of the factor  $\Pr(\pi)$  in  $\Pr(\pi) \exp(-\mu \|c(\pi)\|_r)$  decreases at a slower rate, with respect to the increase of  $\mu$ , than it does for lower  $r$  values. For higher  $r$ , paths which contain low cost edges are favored. In summary,  $r$  and  $\mu$  have the same effect that was shown for the min-norm and potential distances, but can be combined in our framework. Note that Figure 1 is an approximation since only a finite number of paths has been considered. The approximation is more reliable for higher  $\mu$  values, where the influence of high cost paths is negligible.

In appendix G, we show that LN defines a distance. Though PD and LN are similar in form, the same strategy that was followed in (Françoise et al., 2017) to prove that PD is a distance does not apply for LN due to the absence of distributivity. Instead, we expose an step by step derivation to show the triangle inequality. The proof builds on the factorization of the set of paths from  $s$  to  $t$  into those that cross a third node  $q$  and those which do not. Furthermore, the absence of distributivity makes any attempt to

compute the LN distance in finite time impractical, since one can not factor out common terms for different paths. General algorithms proposed for the APP can not be applied here either. Though there have been papers that have tackled the non-distributivity question, they are not applicable here. For instance, in (Daggitt et al., 2018) only selective semirings are considered. Alternatively, the proposed approach in (Lengauer & Theune, 1991) does not simplify the computation of our algebraic structure. Nonetheless, in appendix I we sketch an algorithm to compute the Exp-max and Log-max distances (last row Table 1). The analysis and implementation of this algorithm is out of the scope of the current paper and, therefore, is left for future work.

#### 4. When Does a Semiring Define a Distance?

In the previous sections, we have expressed some of the most common graph metrics in terms of the APP. Now, we wonder which properties must a semiring satisfy such that its associated APP defines a metric. Since all the metrics we have analyzed could be expressed in terms of the APP framework which considers only hitting paths, we will focus on the hitting case, denoted here by  $\text{APP}^h$ . We hope that our results will allow researchers to more easily define semirings that yield new, potentially useful graph distances and reveal some of the underlying structure of semiring based graph distances.

A metric maps a pair of points to a non-negative real number. Since a semiring can be defined over an arbitrary set, we need a function  $g: S \mapsto \mathbb{R}^+ \cup \{\infty\}$  that maps an element of the semiring to the non-negative real numbers. We assume  $g(\bar{1}) = 0$  and  $g(\bar{0}) = \infty$ , since  $\text{APP}^h(s, s) = \bar{1}$  and  $\bar{0}$  represents the cost of the non-existing edges/paths. Let  $G = (V, E)$  be a S-graph. We assume that the graph is strongly connected such that there exists a path connecting two arbitrary nodes and that  $\text{APP}^h(\cdot, \cdot)$  is defined for each pair of nodes. Given  $s, t \in V$  we define the following dissimilarity function

$$d(s, t) = g\left(\text{APP}^h(s, t)\right) + g\left(\text{APP}^h(t, s)\right). \quad (13)$$

We aim to find out which properties  $g$  and the semiring  $S$  must satisfy such that (13) determines a metric function. The non-negativity of  $d$  follows trivially from the definition. We have  $d(s, s) = 0$ , since  $\text{APP}^h(s, s) = \bar{1}$  and  $g(\bar{1}) = 0$ . The opposite direction of the identity of indiscernibles property is more delicate. The next lemma states some conditions that ensure  $s = t$  if  $d(s, t) = 0$ . Concretely, it requires  $\bar{1}$  to be the unique element mapped to 0 and that all paths have cost greater than  $\bar{1}$ . The results stated in this section are proven in appendix F.

**Lemma 4.1.** *Let  $d$  be defined as in (13). If*

- $a \preceq \bar{1} \iff a = \bar{1} \text{ or } a = \bar{0}$ ,

- $g(a) = 0$  if and only if  $a = \bar{1}$ ,
- none of the edge costs is invertible with respect to  $\otimes$ ,

then  $d(s, t) = 0$  implies  $s = t$ .

In the following we will focus on the triangle inequality property. We will show under which circumstances, the left summand of (13),  $d_L(s, t) := g(\text{APP}^h(s, t))$ , satisfies the triangle inequality. If it holds for  $d_L$  it will also hold for  $d$  due to its symmetry. The foundation of our arguments lies in the fact that the set of paths between  $s$  and  $t$  can be partitioned into the set of paths that pass through a node  $q$  and the ones that do not. Consequently, it can be shown that

$$\text{APP}^h(s, t) = \alpha^h \oplus \beta^h \otimes \text{APP}^h(q, t) \quad (14)$$

for certain  $\alpha^h, \beta^h \in S$ , where  $\beta^h \preceq \text{APP}^h(s, q)$  (see appendix F.2). Equation (14) is the starting point of our proofs. The next lemma asserts that a necessary, but not sufficient, condition to satisfy the triangle inequality of  $d_L$  on arbitrary  $S$ -valued graphs is the  $\otimes$ -subadditivity of  $g$ .

**Lemma 4.2.** *If the function  $d_L$ , satisfies the triangle inequality over an arbitrary graph and  $\text{APP}^h(\cdot, \cdot)$  can take any value in the semiring  $S$ , then  $g$  is  $\otimes$ -subadditive, i.e.,*

$$g(a \otimes b) \leq g(a) + g(b), \quad \forall a, b \in S. \quad (15)$$

*Sketch proof.* We consider a graph where all paths from  $s$  to  $t$  pass through  $q$ , such that  $\text{APP}^h(s, q)$  and  $\text{APP}^h(q, t)$  can take any possible value in  $S$ . In that case, equation (14) becomes  $\text{APP}^h(s, t) = \text{APP}^h(s, q) \otimes \text{APP}^h(q, t)$ . The triangle inequality implies that  $g(\text{APP}^h(s, t)) = g(\text{APP}^h(s, q) \otimes \text{APP}^h(q, t)) \leq g(\text{APP}^h(s, q)) + g(\text{APP}^h(q, t))$ . Thus,  $g$  must be subadditive on  $S$ .  $\square$

Now we will present some conditions that are sufficient to ensure that (13) satisfies the triangle inequality over an arbitrary graph. In particular, we analyze the triangle inequality when  $g$  is monotone with respect to  $\preceq$ .

**Theorem 4.3.** *Let  $G = (V, E)$  be an  $S$ -graph. If*

1.  $g$  is subadditive, i.e.,  $g(a \otimes b) \leq g(a) + g(b) \forall a, b \in S$ ,
2.  $g$  is decreasing, i.e.,  $a \preceq b \rightarrow g(b) \leq g(a) \forall a, b \in S$ ,
3.  $a \otimes \text{APP}^h(t, q) \otimes \text{APP}^h(q, t) \preceq a, \forall a \in S, q, t \in V$ .

then  $d$  satisfies the triangle inequality over the nodes of  $G$ .

The third assumption states that, if  $c(\pi) = a \in S$  is the cost of a path,  $\pi$ , and this path is concatenated ( $\otimes$  operation) with all cycles starting at a node  $t$  and traversing an arbitrary node  $q$ , then the aggregation of all these new costs is not greater (according to (1)) than the original cost of the path  $c(\pi)$ . Consequently, since  $g$  is decreasing, the concatenation (and aggregation) of the cycles does not increase the  $g$ -value of the path. The proof of Theorem 4.3 (see appendix F.3)

focuses on demonstrating that

$$\text{APP}^h(s, q) \otimes \text{APP}^h(q, t) \preceq \alpha^h \oplus \beta^h \otimes \text{APP}^h(q, t).$$

If  $\beta^h$  was equal to  $\text{APP}^h(s, q)$ , requiring  $g$  to be decreasing and subadditive would suffice to prove the triangle inequality. Since  $\beta^h \preceq \text{APP}^h(s, q)$ , assumption 3 is needed to ensure the triangle inequality.

One can derive as special cases of Theorem 4.3 that the min-norm distances (including the shortest path and minimax distances) and the PD define metrics over arbitrary graphs (see Corollary F.5 and Corollary F.6). However, this theorem does not apply to CCD. The next theorem mirrors Theorem 4.3, but considers  $g$  to be increasing instead of decreasing. As a corollary, we can prove the fact that the CCD defines a metric (see Corollary F.9).

**Theorem 4.4.** *Let  $G = (V, E)$  be an  $S$ -graph. If*

1.  $g$  is subadditive, i.e.,  $g(a \otimes b) \leq g(a) + g(b) \forall a, b \in S$ ,
2.  $g$  is increasing in  $S \setminus \{\bar{0}\}$ , i.e.,  $a \preceq b \rightarrow g(a) \leq g(b) \forall a, b \in S \setminus \{\bar{0}\}$ ,
3.  $a \preceq a \otimes \text{APP}^h(t, q) \otimes \text{APP}^h(q, t), \forall a \in S, q, t \in V$ .

then  $d$  satisfies the triangle inequality over the nodes of  $G$ .

In the appendix we provide, in addition to simplified proofs of the validity of existing metrics, a case example of a new metric, which can be easily verified to be a metric thanks to our theorems (see appendix H). Note that, none of the results stated in this section apply to the log-norm distance, since it is not defined by a semiring but a strong bimonoid. Distributivity is an essential part of the proofs.

## 5. Conclusion

We have revisited some of the most common graph metrics (shortest path, CCD and minimax distances) and have presented them in terms of the algebraic path problem. We reviewed semirings whose associated algebraic path problems retrieve these metrics. We also discussed the potential distance (which interpolates between the CCD and the shortest path distance) and the min-norm distance (which interpolates between the shortest path and minimax distances) from a new perspective. We showed that these metrics can be expressed as instances of the APP framework.

Moreover, we have proposed a novel unifying family of metrics which includes and relates all the aforementioned distances. This family of metrics is parameterized by a parameter  $r$  that regulates the impact of edges with high cost, and another parameter  $\mu$  which regulates the influence of the paths based on their  $\|\cdot\|_r$ -cost. Moreover, inspired by (Kivimäki et al., 2014), we have proven that the log-norm distance between two nodes coincides with the symmetrized minimum Helmholtz free energy between the nodes. Unfortunately, this distance cannot be obtained as the APP of a

semiring but of a strong bimonoid, and its exact computation remains infeasible. An efficient approximate computation is left for future work.

Finally, we have provided sufficient conditions which ensure that the APP constrained to hitting paths associated with a semiring,  $S$ , defines a metric over the nodes of a graph. In addition, we showed that the function that maps elements from  $S$  to  $\mathbb{R}^+$  must be  $\otimes$ -subadditive if  $g(\text{APP}^h(\cdot, \cdot))$  is to satisfy the triangle inequality. We hope that these results can help in the design of new metrics, and as such help enrich the toolbox available to graph-centric machine learning.

## Acknowledgements

This work is supported, in part, by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2181/1 - 390900948 (the Heidelberg STRUCTURES Excellence Cluster); and in part by Informatics for Life, funded by the Klaus Tschira Foundation.

## References

- Alamgir, M. Shortest path distance in random  $k$ -nearest neighbor graphs. *Proceedings of the 29th International Conference on Machine Learning*, pp. 8, 2012.
- Baras, J. S. and Theodorakopoulos, G. Path Problems in Networks. *Synthesis Lectures on Communication Networks*, 3(1):1–77, January 2010. ISSN 1935-4185, 1935-4193. doi: 10.2200/S00245ED1V01Y201001CNT003. URL <http://www.morganclaypool.com/doi/abs/10.2200/S00245ED1V01Y201001CNT003>.
- Bellman, R. On a routing problem. *Quarterly of Applied Mathematics*, 16(1):87–90, 1958. doi: <https://doi.org/10.1090/qam/102435>.
- Carre, B. A. An Algebra for Network Routing Problems. *IMA Journal of Applied Mathematics*, 7:273–294, 1971. ISSN 0272-4960. doi: 10.1093/imamat/7.3.273.
- Challa, A., Danda, S., Sagar, B. S. D., and Najman, L. Watersheds for Semi-Supervised Classification. *IEEE Signal Processing Letters*, 26(5):720–724, May 2019. ISSN 1558-2361. doi: 10.1109/LSP.2019.2905155.
- Chandra, A. K., Raghavan, P., Ruzzo, W. L., Smolensky, R., and Tiwari, P. The electrical resistance of a graph captures its commute and cover times. *computational complexity*, 6(4):312–340, 1996. doi: 10.1145/73007.73062. Publisher: Springer.
- Chebotarev, P. A Class of Graph-Geodesic Distances Generalizing the Shortest-Path and the Resistance Distances. *Discrete Applied Mathematics*, 159(5):295–302, March 2011. ISSN 0166218X. doi: 10.1016/j.dam.2010.11.017. URL <http://arxiv.org/abs/0810.2717>. arXiv: 0810.2717.
- Chebotarev, P. The Walk Distances in Graphs. *Discrete Applied Mathematics*, 160(10-11):1484–1500, July 2012. ISSN 0166218X. doi: 10.1016/j.dam.2012.02.015. URL <http://arxiv.org/abs/1103.2059>. arXiv: 1103.2059.
- Cortes, C., Haffner, P., and Mohri, M. Rational Kernels: Theory and Algorithms. *Journal of Machine Learning Research (JMLR)*, 5:1035–1062, 2004. URL <http://www.cs.nyu.edu/mohri/postscript/jmlr.pdf>.
- Daggitt, M. L., Gurney, A. J. T., and Griffin, T. G. Asynchronous convergence of policy-rich distributed bellmanford routing protocols. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*, pp. 103–116, Budapest Hungary, August 2018. ACM. ISBN 978-1-4503-5567-4. doi: 10.1145/3230543.3230561. URL <https://dl.acm.org/doi/10.1145/3230543.3230561>.
- Dijkstra, E. W. A note on two problems in connexion with graphs. *Numer. Math.*, 1(1):269–271, December 1959. ISSN 0029-599X, 0945-3245. doi: 10.1007/BF01386390. URL <http://link.springer.com/10.1007/BF01386390>.
- Droste, M., Stüber, T., and Vogler, H. Weighted finite automata over strong bimonoids. *Information Sciences*, 180(1):156–166, January 2010. ISSN 00200255. doi: 10.1016/j.ins.2009.09.003. URL <https://linkinghub.elsevier.com/retrieve/pii/S0020025509003867>.
- Eisner, J. Parameter estimation for probabilistic finite-state transducers. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02*, pp. 1, Philadelphia, Pennsylvania, 2001. Association for Computational Linguistics. doi: 10.3115/1073083.1073085. URL <http://portal.acm.org/citation.cfm?doid=1073083.1073085>.
- Floyd, R. W. Algorithm 97: Shortest path. *Commun. ACM*, 5(6):345, June 1962. ISSN 0001-0782, 1557-7317. doi: 10.1145/367766.368168. URL <https://dl.acm.org/doi/10.1145/367766.368168>.
- Fouss, F., Saerens, M., and Shimbo, M. *Algorithms and Models for Network Data and Link Analysis*. Cambridge University Press, Cambridge, 2016. ISBN 978-1-316-41832-1. doi: 10.1017/CBO9781316418321. URL <http://ebooks.cambridge.org/ref/id/CBO9781316418321>.

- Françoise, K., Kivimäki, I., Mantrach, A., Rossi, F., and Saerens, M. A bag-of-paths framework for network data analysis. *Neural Networks*, 90:90–111, June 2017. ISSN 08936080. doi: 10.1016/j.neunet.2017.03.010. URL <http://arxiv.org/abs/1302.6766>. arXiv: 1302.6766.
- Gondran, M. and Minoux, M. *Graphs, Dioids and Semirings: New Models and Algorithms*. Operations Research/Computer Science Interfaces Series. Springer US, 2008. ISBN 978-0-387-75449-9. doi: 10.1007/978-0-387-75450-5. URL <https://www.springer.com/gp/book/9780387754499>.
- Griffin, T. G. and Sobrinho, J. L. Metarouting. In *Proceedings of the 2005 conference on Applications, technologies, architectures, and protocols for computer communications - SIGCOMM '05*, pp. 1, Philadelphia, Pennsylvania, USA, 2005. ACM Press. ISBN 978-1-59593-009-5. doi: 10.1145/1080091.1080094. URL <http://portal.acm.org/citation.cfm?doid=1080091.1080094>.
- Gurvich, V. Metric and ultrametric spaces of resistances. *Discrete Applied Mathematics*, pp. 10, 2010.
- Huang, L. Advanced Dynamic Programming in Semiring and Hypergraph Frameworks. In *Coling 2008: Advanced Dynamic Programming in Computational Linguistics: Theory, Algorithms and Applications - Tutorial notes*, pp. 1–18, Manchester, UK, August 2008. Coling 2008 Organizing Committee. URL <https://aclanthology.org/C08-5001>.
- Kim, K.-H. and Choi, S. Walking on Minimax Paths for k-NN Search. *Proceedings of the AAAI Conference on Artificial Intelligence*, 27(1):8, 2013.
- Kivimäki, I., Shimbo, M., and Saerens, M. Developments in the theory of randomized shortest paths with a comparison of graph node distances. *Physica A: Statistical Mechanics and its Applications*, 393:600–616, January 2014. ISSN 03784371. doi: 10.1016/j.physa.2013.09.016. URL <http://arxiv.org/abs/1212.1666>. arXiv: 1212.1666.
- Klein, D. J. and Randić, M. Resistance distance. *J Math Chem*, 12(1):81–95, December 1993. ISSN 1572-8897. doi: 10.1007/BF01164627. URL <https://doi.org/10.1007/BF01164627>.
- Lehmann, D. J. Algebraic structures for transitive closure. *Theoretical Computer Science*, 4(1):59–76, February 1977. ISSN 0304-3975. doi: 10.1016/0304-3975(77)90056-1. URL <https://www.sciencedirect.com/science/article/pii/0304397577900561>.
- Lengauer, T. and Theune, D. Unstructured path problems and the making of semirings. In Dehne, F., Sack, J.-R., and Santoro, N. (eds.), *Algorithms and Data Structures*, volume 519, pp. 189–200. Springer-Verlag, Berlin/Heidelberg, 1991. ISBN 978-3-540-54343-5. doi: 10.1007/BFb0028261. URL <http://link.springer.com/10.1007/BFb0028261>. Series Title: Lecture Notes in Computer Science.
- Litvinov, G. L. The Maslov dequantization, idempotent and tropical mathematics: A brief introduction. *arXiv:math/0507014*, July 2005. URL <http://arxiv.org/abs/math/0507014>. arXiv: math/0507014.
- Lotito, P., Quadrat, J.-P., and Mancinelli, E. Traffic assignment & Gibbs-Maslov semirings. In Litvinov, G. L. and Maslov, V. P. (eds.), *Contemporary Mathematics*, volume 377, pp. 209–219. American Mathematical Society, Providence, Rhode Island, 2005. ISBN 978-0-8218-3538-8 978-0-8218-7967-2. doi: 10.1090/conm/377/06993. URL <http://www.ams.org/conm/377/>.
- Luxburg, U., Radl, A., and Hein, M. Getting lost in space: Large sample analysis of the resistance distance. In *Advances in Neural Information Processing Systems*, volume 23. Curran Associates, Inc., 2010.
- Maggs, B. M. and Plotkin, S. A. Minimum-cost spanning tree as a path-finding problem. *Information Processing Letters*, 26(6):291–293, January 1988. ISSN 00200190. doi: 10.1016/0020-0190(88)90185-8. URL <https://linkinghub.elsevier.com/retrieve/pii/0020019088901858>.
- Mckenzie, D. and Damelin, S. Power weighted shortest paths for clustering Euclidean data. *Foundations of Data Science*, 2019. doi: 10.3934/fods.2019014.
- Mohri, M. Semiring Frameworks and Algorithms for Shortest-Distance Problems. *Journal of Automata, Languages and Combinatorics*, 7(3):321–350, 2002.
- Nadler, B., Srebro, N., and Zhou, X. Semi-supervised learning with the graph Laplacian: the limit of infinite unlabelled data. In *Proceedings of the 22nd International Conference on Neural Information Processing Systems, NIPS'09*, pp. 1330–1338, Red Hook, NY, USA, December 2009. Curran Associates Inc. ISBN 978-1-61567-911-9.
- Pin, J.-E. Tropical Semirings. In Gunawardena, J. (ed.), *Idempotency (Bristol, 1994)*, Publ. Newton Inst. 11, pp. 50–69. Cambridge Univ. Press, Cambridge, 1998. URL <https://hal.archives-ouvertes.fr/hal-00113779>.

- Saerens, M., Achbany, Y., Fouss, F., and Yen, L. Randomized Shortest-Path Problems: Two Related Models. *Neural computation*, 21:2363–404, April 2009. doi: 10.1162/neco.2009.11-07-643.
- Schwartz, R., Thomson, S., and Smith, N. A. Bridging CNNs, RNNs, and Weighted Finite-State Machines. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 295–305, Melbourne, Australia, 2018. Association for Computational Linguistics. doi: 10.18653/v1/P18-1028. URL <http://aclweb.org/anthology/P18-1028>.
- Simon, I. Limited subsets of a free monoid. In *19th Annual Symposium on Foundations of Computer Science (sfcs 1978)*, pp. 143–150, October 1978. doi: 10.1109/SFCS.1978.21. ISSN: 0272-5428.
- Yen, L., Saerens, M., Mantrach, A., and Shimbo, M. A family of dissimilarity measures between nodes generalizing both the shortest-path and the commute-time distances. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '08*, pp. 785–793, New York, NY, USA, August 2008. Association for Computing Machinery. ISBN 978-1-60558-193-4. doi: 10.1145/1401890.1401984. URL <https://doi.org/10.1145/1401890.1401984>.

## A. Min-Norm Semiring

**Lemma A.1.** Given  $r > 0$ , the min-norm semiring  $(S, \oplus, \otimes_r, \infty, 0)$ , where  $S = \mathbb{R}^+ \cup \{\infty\}$  and

$$x \oplus y = \min(x, y), \quad x \otimes_r y = \sqrt[r]{x^r + y^r}. \quad (16)$$

is a commutative selective semiring with  $\bar{0} = \infty$  and  $\bar{1} = 0$  as its neutral elements.

*Proof.* We will only show the the associativity of  $\otimes_r$  and the distributivity of  $\otimes_r$  over  $\oplus$ .

- Associativity  $\otimes_r$

$$\begin{aligned} (x \otimes_r y) \otimes_r z &= \sqrt[r]{x^r + y^r} \otimes_r z = \sqrt[r]{(\sqrt[r]{x^r + y^r})^r + z^r} = \sqrt[r]{x^r + y^r + z^r} \\ &= \sqrt[r]{x^r + (\sqrt[r]{y^r + z^r})^r} = x \otimes_r \sqrt[r]{y^r + z^r} = x \otimes_r (y \otimes_r z) \end{aligned} \quad (17)$$

- Distributivity:

$$\begin{aligned} (x \oplus y) \otimes_r z &= \sqrt[r]{(\min(x, y))^r + z^r} = \sqrt[r]{\min(x^r, y^r) + z^r} = \sqrt[r]{\min(x^r + z^r, y^r + z^r)} \\ &= \min(\sqrt[r]{x^r + z^r}, \sqrt[r]{y^r + z^r}) = x \otimes_r z \oplus y \otimes_r z \end{aligned} \quad (18)$$

The left-distributivity is a consequence of the commutativity of  $\otimes$ .

□

**Lemma A.2.** Let  $x, y \geq 0$ . Then

$$\lim_{r \rightarrow \infty} (x^r + y^r)^{1/r} = \max(x, y).$$

*Proof.* If  $x = 0$  or  $y = 0$ , then the limit is trivial. Without loss of generality, we will assume  $x \geq y > 0$ . First note that

$$\sqrt[r]{x^r + y^r} = (x^r + y^r)^{1/r} = \exp\left(\frac{1}{r} \log(x^r + y^r)\right).$$

Then

$$\lim_{r \rightarrow \infty} (x^r + y^r)^{1/r} = \lim_{r \rightarrow \infty} \exp\left(\frac{1}{r} \log(x^r + y^r)\right) = \exp\left(\lim_{r \rightarrow \infty} \frac{1}{r} \log(x^r + y^r)\right).$$

Applying L'Hôpital's rule we obtain

$$\lim_{r \rightarrow \infty} \frac{1}{r} \log(x^r + y^r) = \lim_{r \rightarrow \infty} \frac{x^r \log(x) + y^r \log(y)}{x^r + y^r} = \lim_{r \rightarrow \infty} \frac{\log(x) + \left(\frac{y}{x}\right)^r \log(y)}{1 + \left(\frac{y}{x}\right)^r} \underbrace{=}_{\frac{y}{x} \leq 1} \log(x). \quad (19)$$

Therefore  $\exp\left(\lim_{r \rightarrow \infty} \frac{1}{r} \log(x^r + y^r)\right) = \exp(\log(x)) = x = \max(x, y)$ . □

As a consequence of Lemma A.2,  $(S, \oplus, \otimes_r, \infty, 0)$  interpolates between the min-plus and minimax semiring. Therefore, the min-norm distance also interpolates between the shortest path and minimax distance.

## B. APP of the Eisner Semiring Recovers the First Hitting Cost

We show that the second entry of the APP associated with the Eisner semiring restricted to hitting paths is equal to the first hitting cost  $\mathcal{H}(s, t)$ :

$$\begin{aligned}
 \bigoplus_{\pi \in \mathcal{P}_{st}^h} \bigotimes_{e \in \pi} (p_e, p_e c_e) &= \bigoplus_{\pi \in \mathcal{P}_{st}^h} \left( \prod_{e \in \pi} p_e, \left( \prod_{e \in \pi} p_e \right) \left( \sum_{e \in \pi} c_e \right) \right) \\
 &= \bigoplus_{\pi \in \mathcal{P}_{st}^h} \left( \Pr(\pi), \Pr(\pi) c(\pi) \right) = \left( \sum_{\pi \in \mathcal{P}_{st}^h} \Pr(\pi), \sum_{\pi \in \mathcal{P}_{st}^h} \Pr(\pi) c(\pi) \right) \\
 &= \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [1], \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [c(\pi)] \right) = \left( 1, \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [c(\pi)] \right) \\
 &= (1, \mathcal{H}(s, t))
 \end{aligned} \tag{20}$$

## C. APP of the Log-Semiring Recovers the Potential Distance

We show that the APP associated with the log-semiring restricted to hitting paths retrieves the first summand of the potential distance (Kivimäki et al., 2014; Françoisse et al., 2017).

$$\begin{aligned}
 \bigoplus_{\pi \in \mathcal{P}_{st}^h} \bigotimes_{e \in \pi} -\frac{1}{\mu} \log (p_e \exp(-\mu c_e)) &= \bigoplus_{\pi \in \mathcal{P}_{st}^h} \sum_{e \in \pi} \frac{1}{\mu} \log (p_e \exp(-\mu c_e)) \\
 &= \bigoplus_{\pi \in \mathcal{P}_{st}^h} -\frac{1}{\mu} \log \left( \prod_{e \in \pi} p_e \exp(-\mu \sum_{e \in \pi} c_e) \right) \\
 &= \bigoplus_{\pi \in \mathcal{P}_{st}^h} -\frac{1}{\mu} \log \left( \Pr(\pi) \exp(-\mu c(\pi)) \right) \\
 &= -\frac{1}{\mu} \log \left( \sum_{\pi \in \mathcal{P}_{st}^h} \Pr(\pi) \exp(-\mu c(\pi)) \right) \\
 &= -\frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ \exp(-\mu c(\pi)) \right] \right)
 \end{aligned} \tag{21}$$

## D. Log-Norm Strong Bimonoid

In this section, we prove that the operations defined in equation (11), and restated here in (22), form a strong bimonoid over  $\mathbb{R}^+ \times \mathbb{R}^+ \cup \{\bar{0}\}$ , where  $\bar{0}$  represents the neutral element of the operation  $\oplus_\mu$ . Since there is not a natural neutral element in  $\mathbb{R}^+ \times \mathbb{R}^+$  for  $\oplus_\mu$ , we explicitly need to define an ad hoc neutral element.

**Lemma D.1.** *Let  $r > 1, \mu > 0$ . The log-norm algebraic structure  $(\mathbb{R}^+ \times \mathbb{R}^+ \cup \{\bar{0}\}, \oplus_\mu, \otimes_r, \bar{0}, \bar{1} = (1, 0))$ , where*

$$\begin{aligned}
 (a, b) \oplus_\mu (c, d) &= \left( 1, -\frac{1}{\mu} \log (a e^{-\mu b} + c e^{-\mu d}) \right) \\
 (a, b) \oplus_\mu \bar{0} &= \bar{0} \oplus_\mu (a, b) = (a, b) \\
 (a, b) \otimes_r (c, d) &= (ac, \sqrt[r]{c^r + d^r}) \\
 (a, b) \otimes_r \bar{0} &= \bar{0} \otimes_r (a, b) = \bar{0},
 \end{aligned} \tag{22}$$

defines a strong bimonoid.

*Proof.* Note that in (22) we define  $\bar{0}$  to be absorbing. Thus, we just have left to show that  $\oplus_\mu$  and  $\otimes_r$  define monoids over  $\mathbb{R}^+ \times \mathbb{R}^+$ . The associativity and commutativity of  $\otimes_r$  follow from the associativity and commutativity of the usual product

operation in  $\mathbb{R}^+$  and the product operation of the min-norm semiring (Appendix B, equation (17)). It is also trivial to show that the neutral element of  $\otimes_r$  is  $\bar{1} = (1, 0)$ .

Next we show the associativity of  $\oplus_\mu$ .

$$\begin{aligned}
 ((a_1, a_2) \oplus_\mu (b_1, b_2)) \oplus_\mu (c_1, c_2) &= \left(1, -\frac{1}{\mu} \log \left(a_1 e^{-\mu a_2} + b_1 e^{-\mu b_2}\right)\right) \oplus_\mu (c_1, c_2) \\
 &= \left(1, -\frac{1}{\mu} \log \left(a_1 e^{-\mu a_2} + b_1 e^{-\mu b_2} + c_1 e^{-\mu c_2}\right)\right) \\
 &= (a_1, a_2) \oplus_\mu \left(1, -\frac{1}{\mu} \log \left(b_1 e^{-\mu b_2} + c_1 e^{-\mu c_2}\right)\right) \\
 &= (a_1, a_2) \oplus_\mu ((b_1, b_2) \oplus_\mu (c_1, c_2)).
 \end{aligned} \tag{23}$$

The commutativity of  $\oplus_\mu$  follows from the commutativity of the common sum.  $\square$

**Lemma D.2.** *Let  $r > 1, \mu > 0$ . The associated APP of the log-norm strong bimonoid defines the log-norm distance when the costs are equal to  $(p_e, c_e)$ , i.e.*

$$\begin{aligned}
 \text{LN}(s, t) &= \left[ \bigoplus_{\pi \in \mathcal{P}_{st}^h} \bigotimes_{e \in \pi}^{\mu} (p_e, c_e) + \bigoplus_{\pi \in \mathcal{P}_{ts}^h} \bigotimes_{e \in \pi}^{\mu} (p_e, p_e c_e) \right]_2 \\
 &= -\frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [\exp(-\mu \|c(\pi)\|_r)] \right) - \frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} [\exp(-\mu \|c(\pi)\|_r)] \right).
 \end{aligned} \tag{24}$$

*Proof.*

$$\begin{aligned}
 \bigoplus_{\pi \in \mathcal{P}_{st}^h} \bigotimes_{e \in \pi} (p_e, p_e c_e) &= \bigoplus_{\pi \in \mathcal{P}_{st}^h} \left( \prod_{e \in \pi} p_e, \sqrt[r]{\sum_{e \in \pi} c_e^r} \right) \\
 &= \bigoplus_{\pi \in \mathcal{P}_{st}^h} \left( \Pr(\pi), \|c(\pi)\|_r \right) = \left(1, -\frac{1}{\mu} \log \left( \sum_{\pi \in \mathcal{P}_{st}^h} \Pr(\pi) \exp(-\mu \|c(\pi)\|_r) \right)\right) \\
 &= \left(1, -\frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [\exp(-\mu \|c(\pi)\|_r)] \right)\right)
 \end{aligned} \tag{25}$$

$\square$

**Corollary D.3.** *Let  $\mu > 0$ .*

- The log-max algebraic structure  $(\mathbb{R}^+ \times \mathbb{R}^+ \cup \{\bar{0}\}, \oplus_\mu, \otimes_\infty, \bar{0}, \bar{1} = (1, 0))$ , where

$$\begin{aligned}
 (a, b) \oplus_\mu (c, d) &= \left(1, -\frac{1}{\mu} \log \left(ae^{-\mu b} + ce^{-\mu d}\right)\right) \\
 (a, b) \oplus_\mu \bar{0} &= \bar{0} \oplus_\mu (a, b) = (a, b) \\
 (a, b) \otimes_\infty (c, d) &= (ac, \max(c, d)) \\
 (a, b) \otimes_\infty \bar{0} &= \bar{0} \otimes_r (a, b) = \bar{0},
 \end{aligned} \tag{26}$$

*defines a strong bimonoid.*

- The associated APP of the Log-max strong bimonoid defines the log-norm distance when the costs are equal to  $(p_e, c_e)$ , i.e.

$$\begin{aligned}
 \text{LM}(s, t) &= \left[ \bigoplus_{\pi \in \mathcal{P}_{st}^h} \bigotimes_{e \in \pi}^{\mu} (p_e, c_e) + \bigoplus_{\pi \in \mathcal{P}_{ts}^h} \bigotimes_{e \in \pi}^{\mu} (p_e, p_e c_e) \right]_2 \\
 &= -\frac{1}{\mu} \left( \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ e^{-\mu \max_{e \in \pi} c(e)} \right] \right) + \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} \left[ e^{-\mu \max_{e \in \pi} c(e)} \right] \right) \right).
 \end{aligned} \tag{27}$$

*Proof.* Consequence of Lemma D.1 and Lemma D.2 and the fact that  $\lim_{r \rightarrow \infty} \otimes_r = (\times, \max)$  (see Lemma A.2).  $\square$

Analogously, we define the strong bimonoids that would define the exp-norm and exp-max distances presented in Table 1.

**Lemma D.4.** 1. Let  $r > 0$ . The exp-norm algebraic structure  $(\mathbb{R}^+ \times \mathbb{R}^+ \oplus, \otimes_r, (0, 0), (1, 0))$  defines an strong bimonoid, where

$$\begin{aligned} (a, b) \oplus_\mu (c, d) &= (a + c, b + d) \\ (a, b) \otimes_r (c, d) &= \left( ac, ac \sqrt[r]{\left(\frac{b}{a}\right)^r + \left(\frac{d}{c}\right)^r} \right). \end{aligned} \quad (28)$$

2. The exp-max algebraic structure  $(\mathbb{R}^+ \times \mathbb{R}^+ \oplus, \otimes_\infty, (0, 0), (1, 0))$  defines an strong bimonoid, where

$$\begin{aligned} (a, b) \oplus_\mu (c, d) &= (a + c, b + d) \\ (a, b) \otimes_r (c, d) &= \left( ac, ac \max\left(\frac{b}{a}, \frac{d}{c}\right) \right). \end{aligned} \quad (29)$$

Moreover the associated APP of the exp-norm and exp-max strong bimonoids define the exp-norm and exp-max distances respectively when the costs are equal to  $(p_e, p_e c_e)$ , i.e.

$$\left[ \bigoplus_{\pi \in \mathcal{P}_{st}^h}^\mu \bigotimes_{e \in \pi}^\mu (p_e, p_e c_e) + \bigoplus_{\pi \in \mathcal{P}_{ts}^h}^\mu \bigotimes_{e \in \pi}^\mu (p_e, p_e c_e) \right]_2 = \mathbb{E}_{\pi \sim \mathcal{P}_{st}} [||c(\pi)||_r] + \mathbb{E}_{\pi \sim \mathcal{P}_{ts}} [||c(\pi)||_r], \quad (30)$$

and

$$\left[ \bigoplus_{\pi \in \mathcal{P}_{st}^h}^\mu \bigotimes_{e \in \pi}^\mu (p_e, p_e c_e) + \bigoplus_{\pi \in \mathcal{P}_{ts}^h}^\mu \bigotimes_{e \in \pi}^\mu (p_e, p_e c_e) \right]_2 = \mathbb{E}_{\pi \sim \mathcal{P}_{st}} \left[ \max_{e \in \pi} c(e) \right] + \mathbb{E}_{\pi \sim \mathcal{P}_{ts}} \left[ \max_{e \in \pi} c(e) \right]. \quad (31)$$

*Proof.* We will only proof the exp-norm case, since the exp-max case is analogous. To prove that the exp-norm is an strong bimonoid we will just show that the  $\otimes_r$  is associative. The rest of properties are trivial.

$$\begin{aligned} ((a_1, a_2) \otimes_r (b_1, b_2)) \otimes_r (c_1, c_2) &= \left( a_1 b_1, a_1 b_1 \sqrt[r]{\left(\frac{a_2}{a_1}\right)^r + \left(\frac{b_2}{b_1}\right)^r} \right) \otimes_r (c_1, c_2) \\ &= \left( a_1 b_1 c_1, a_1 b_1 c_1 \sqrt[r]{\left(\frac{a_1 b_1 \sqrt[r]{\left(\frac{a_2}{a_1}\right)^r + \left(\frac{b_2}{b_1}\right)^r}}{a_1 b_1}\right)^r + \left(\frac{c_2}{c_1}\right)^r} \right) \\ &= \left( a_1 b_1 c_1, a_1 b_1 c_1 \sqrt[r]{\left(\frac{a_2}{a_1}\right)^r + \left(\frac{b_2}{b_1}\right)^r + \left(\frac{c_2}{c_1}\right)^r} \right) \\ &= (a_1, a_2) \otimes_r \left( b_1 c_1, b_1 c_1 \sqrt[r]{\left(\frac{b_2}{b_1}\right)^r + \left(\frac{c_2}{c_1}\right)^r} \right) \\ &= (a_1, a_2) \otimes_r ((b_1, b_2) \otimes_r (c_1, c_2)). \end{aligned} \quad (32)$$

The APP of the exp-norm strong bimonoid follows from

$$\begin{aligned}
 \bigoplus_{\pi \in \mathcal{P}_{st}^h}^{\mu} \bigotimes_{e \in \pi}^r (p_e, p_e c_e) &= \bigoplus_{\pi \in \mathcal{P}_{st}^h}^{\mu} \left( \prod_{e \in \pi} p_e, \left( \prod_{e \in \pi} p_e \right) \sqrt[r]{\sum_{e \in \pi} c_e^r} \right) \\
 &= \bigoplus_{\pi \in \mathcal{P}_{st}^h}^{\mu} \left( \Pr(\pi), \Pr(\pi) \|c(\pi)\|_r \right) = \left( \sum_{\pi \in \mathcal{P}_{st}^h} \Pr(\pi), \sum_{\pi \in \mathcal{P}_{st}^h} \Pr(\pi) \|c(\pi)\|_r \right) \\
 &= \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [1], \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [\|c(\pi)\|_r] \right) = \left( 1, \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [\|c(\pi)\|_r] \right).
 \end{aligned} \tag{33}$$

□

## E. Log-Norm Metric Limits

In this section, we prove the limits of the log-norm distance shown in Table 1. First, in Lemma E.1 we prove the limits when  $\mu \rightarrow 0^+$  and  $\mu \rightarrow \infty$  for a finite  $r$ .

**Lemma E.1.**

1.

$$\begin{aligned}
 &\lim_{\mu \rightarrow 0^+} -\frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] \right) - \frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] \right) \\
 &= \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [\|c(\pi)\|_r] + \mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} [\|c(\pi)\|_r]
 \end{aligned} \tag{34}$$

2.

$$\lim_{\mu \rightarrow \infty} -\frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] \right) - \frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] \right) = 2 \min_{\pi \in \mathcal{P}_{st}^h} \|c(\pi)\|_r \tag{35}$$

*Proof.* By showing the limit of the first summands we can derive the limit of the second in an analogous way.

1.

$$\begin{aligned}
 &\lim_{\mu \rightarrow 0^+} -\frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] \right) = \\
 &\lim_{\mu \rightarrow 0^+} -\frac{1}{\mu} \log \left( \sum_{\pi \in \mathcal{P}_{st}^h} (\Pr(\pi) \exp(-\mu \|c(\pi)\|_r)) \right) \stackrel{\text{L'Hôpital's rule}}{=} \\
 &\lim_{\mu \rightarrow 0^+} \frac{\sum_{\pi \in \mathcal{P}_{st}^h} (\Pr(\pi) \|c(\pi)\|_r \exp(-\mu \|c(\pi)\|_r))}{\sum_{\pi \in \mathcal{P}_{st}^h} (\Pr(\pi) \exp(-\mu \|c(\pi)\|_r))} = \sum_{\pi \in \mathcal{P}_{st}^h} \Pr(\pi) \|c(\pi)\|_r = \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} [\|c(\pi)\|_r]
 \end{aligned} \tag{36}$$

2.

$$\begin{aligned}
 &\lim_{\mu \rightarrow \infty} = -\frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] \right) = \\
 &\lim_{\mu \rightarrow \infty} -\frac{1}{\mu} \log \left( \sum_{\pi \in \mathcal{P}_{st}^h} \left( \Pr(\pi) \exp \left( -\mu \sum_{e \in \pi} c_e \right) \right) \right) \stackrel{\text{L'Hôpital's rule}}{=} \\
 &\lim_{\mu \rightarrow \infty} \frac{\sum_{\pi \in \mathcal{P}_{st}^h} (\Pr(\pi) \|c(\pi)\|_r \exp(-\mu \|c(\pi)\|_r))}{\sum_{\pi \in \mathcal{P}_{st}^h} (\Pr(\pi) \exp(-\mu \|c(\pi)\|_r))} = \min_{\pi \in \mathcal{P}_{st}^h} \|c(\pi)\|_r
 \end{aligned} \tag{37}$$

□

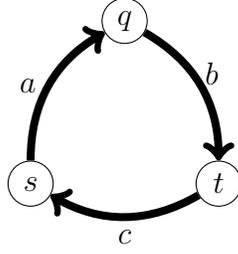


Figure 2. Graph where all paths connecting  $s$  and  $t$  pass through  $q$ . The terms  $a, b \in S$  indicate the cost of the corresponding edges. The triangle inequality can only be satisfied if  $g$  is subadditive such that

$$d_L(s, t) = g(\text{APP}^h(s, q) \otimes \text{APP}^h(q, t)) = g(a \otimes b) \leq g(a) + g(b) = g(\text{APP}^h(s, q)) + g(\text{APP}^h(q, t)) = d(s, q) + d(q, t).$$

Since  $\lim_{r \rightarrow \infty} x \otimes_r y = \lim_{r \rightarrow \infty} \sqrt[r]{x^r + y^r} = \max(x, y)$  as a consequence of Lemma A.2, we obtain all the limits exposed in Table 1. Note that for  $r = 1$  we retrieve the potential distance (Kivimäki et al., 2014; Françoisse et al., 2017) and their limits.

## F. When Does a Semiring Define a Distance?

This section contains the proofs of all the results stated in section 4, which provide sufficient conditions to ensure that  $d(s, t)$  defines a proper metric. We will focus only on the left summand of (13), since for the right term the same properties will follow:

$$d_L(s, t) := g(\text{APP}^h(s, t)) = g\left(\bigoplus_{\pi \in \mathcal{P}_{st}^h} \bigotimes_{e \in \pi} c(e)\right). \quad (38)$$

Through the whole section we assume that distributivity commutativity and associativity of the semiring operations also hold for infinite sums and products.

### F.1. Proof Lemma 4.1

**Lemma F.1** (Lemma 4.1). *Let  $d$  be defined as in (13). If*

1.  $a \preceq \bar{1} \iff a = \bar{1}$  or  $a = \bar{0}$ , where  $\preceq$  is the canonical preorder relation defined in (1),
2.  $g(a) = 0$  if and only if  $a = \bar{1}$ ,
3. none of the edge costs is invertible with respect to  $\otimes$ ,

then  $d(s, t) = 0$  if and only if  $s = t$ .

*Proof.* According to assumption 2, we just need to prove that  $\text{APP}^h(s, t) \neq \bar{1}$  for arbitrary distinct vertices  $s$  and  $t$ . First we recall the definition of the canonical order of a semiring which was stated in equation (1):

$$a \preceq b \iff \exists c \in S \text{ such that } a \oplus c = b.$$

As a consequence of the definition of  $\preceq$  and the first assumption, there do not exist any  $a$  and  $b$  distinct of  $\bar{1}$  such that  $a \oplus b = \bar{1}$ . Therefore,

$$\text{APP}^h(s, t) = \bigoplus_{\pi \in \mathcal{P}_{st}^h} \bigotimes_{e \in \pi} c(e) = \bar{1} \Rightarrow \exists \pi \in \mathcal{P}_{st}^h \text{ such that } \bigotimes_{e \in \pi} c(e) = \bar{1}.$$

Thanks to assumption 3,  $\bigotimes_{e \in \pi} c(e) \neq \bar{1}$ , otherwise the costs  $c(e)$  would have inverse elements.  $\square$

### F.2. The Mapping $g$ is Subadditive

We will assume that the edge costs, and also  $\text{APP}^h(\cdot, \cdot)$ , can take arbitrary values in the semiring  $S$ . To prove that the subadditivity of  $g$  is a necessary condition if  $d_L$  satisfies the triangle inequality on arbitrary graphs, we will define a particular strongly connected graph where the subadditivity is necessary (see Figure 2).

First, we will show two equalities that will prove to be useful.

**Lemma F.2.** *Given a graph  $G$  and arbitrary nodes  $s, t$  and  $q$  then*

1.

$$\text{APP}^h(s, t) = \alpha^h \oplus \beta^h \otimes \text{APP}^h(q, t). \quad (39)$$

2.

$$\text{APP}^h(s, q) = \beta^h \oplus \alpha^h \otimes \text{APP}^h(t, q), \quad (40)$$

where

$$\alpha^h := \bigoplus_{\substack{\pi \in \mathcal{P}_{st}^h \\ q \notin \pi}} \bigotimes_{e \in \pi} c(e), \quad \beta^h := \bigoplus_{\substack{\pi_1 \in \mathcal{P}_{sq}^h \\ t \notin \pi_1}} \bigotimes_{e \in \pi_1} c(e).$$

*Proof.* 1.

$$\begin{aligned} \text{APP}^h(s, t) &= \bigoplus_{\pi \in \mathcal{P}_{st}^h} \bigotimes_{e \in \pi} c(e) \\ &= \bigoplus_{\substack{\pi \in \mathcal{P}_{st}^h \\ q \notin \pi}} \bigotimes_{e \in \pi} c(e) \oplus \bigoplus_{\substack{\pi \in \mathcal{P}_{st}^h \\ q \in \pi}} \bigotimes_{e \in \pi} c(e) \\ &= \bigoplus_{\substack{\pi \in \mathcal{P}_{st}^h \\ q \notin \pi}} \bigotimes_{e \in \pi} c(e) \oplus \bigoplus_{\substack{\pi_1 \in \mathcal{P}_{sq}^h \\ t \notin \pi_1 \\ \pi_2 \in \mathcal{P}_{qt}^h}} \left( \bigotimes_{e \in \pi_1} c(e) \right) \otimes \left( \bigotimes_{e \in \pi_2} c(e) \right) \\ &= \underbrace{\bigoplus_{\substack{\pi \in \mathcal{P}_{st}^h \\ q \notin \pi}} \bigotimes_{e \in \pi} c(e)}_{\alpha^h} \oplus \underbrace{\left( \bigoplus_{\substack{\pi_1 \in \mathcal{P}_{sq}^h \\ t \notin \pi_1}} \bigotimes_{e \in \pi_1} c(e) \right)}_{\beta^h} \otimes \underbrace{\left( \bigoplus_{\pi_2 \in \mathcal{P}_{qt}^h} \bigotimes_{e \in \pi_2} c(e) \right)}_{\text{APP}^h(q, t)} \\ &= \alpha^h \oplus \beta^h \otimes \text{APP}^h(q, t). \end{aligned} \quad (41)$$

2. The second equality is proven analogously to the previous one once the following permutation is done:

$$s \rightarrow s, q \rightarrow t, t \rightarrow q$$

□

The first equality, (39), decomposes the cost of the paths from  $s$  to  $t$  in two terms: one that depends on the paths that pass through a third node  $q$  and a second term where the paths do not pass through  $q$ . Indeed, the term  $\alpha^h$  aggregates all the hitting paths which do not cross node  $q$ , while the term  $\beta^h \otimes \text{APP}^h(q, t)$  considers all hitting paths that pass through  $q$ . The second equality, (40), performs the same decomposition as the first equality but considering the paths from  $s$  to  $q$ .

Let us prove now that  $g$  must be subadditive, if the triangle inequality holds for the function  $d_L$ , (38). It follows from equation (39) that

$$\begin{aligned} g(\alpha^h \oplus \beta^h \otimes \text{APP}^h(q, t)) &= g(\text{APP}^h(s, t)) = d_L^h(s, t) \\ &\leq d_L^h(s, q) + d_L^h(q, t) = g(\text{APP}^h(s, q)) + g(\text{APP}^h(q, t)). \end{aligned} \quad (42)$$

Lemma 4.2 claims that the subadditivity of  $g$  with respect to  $\otimes$  is a necessary condition to ensure the triangle inequality of  $d_L$  in any graph. Indeed, for a graph where all paths from  $s$  to  $t$  cross node  $q$  (e.g. Figure 2) we have  $\alpha^h = 0$  and therefore  $\beta^h = \text{APP}^h(s, q)$ . Hence

$$g(\text{APP}^h(s, t)) = g(\text{APP}^h(s, q) \otimes \text{APP}^h(q, t)) = d_L(s, t) \leq d_L(s, q) + d_L(q, t) = g(\text{APP}^h(s, q)) + g(\text{APP}^h(q, t)). \quad (43)$$

Note that in the graph of Figure 2 the cost of the edge from  $s$  to  $q$ ,  $a$ , is equal to  $\text{APP}^h(s, q)$ . Analogously,  $\text{APP}^h(q, t) = b$ . Thus, for arbitrary values  $a, b \in S$  the subadditivity of  $g$  in  $S$  follows:

$$g(a \otimes b) \leq g(a) + g(b).$$

We have proven that  $g$  must be subadditive such that  $d_L$  satisfies the triangle inequality on the graph in Figure 2. Thus,  $g$  must be subadditive such that  $d_L$  satisfies the triangle inequality on all graphs.

*Remark F.3.* Note that  $\text{APP}^h(s, q)$  and  $\text{APP}^h(q, t)$  may not take any possible value in the semiring. For instance, in the Eisner semiring, which characterizes the commute cost distance, the possible values of these variables lie in the set  $\{1\} \times \mathbb{R}^+$ . In the Eisner semiring, the first entry of  $\text{APP}(s, t)$  is always equal to  $\sum_{\pi \in \mathcal{P}_{st}^h} \text{Pr}(\pi) = 1$ , since it is the sum of the probabilities of the hitting paths from  $s$  to  $t$ .<sup>2</sup> In this concrete case, the subadditivity should be constrained to this set.

### F.3. Proof Theorem 4.3

**Theorem F.4** (Theorem 4.3). *Let  $G = (V, E)$  be an  $S$ -graph. If*

1.  $g$  is  $\otimes$ -subadditive, i.e.  $g(a \otimes b) \leq g(a) + g(b)$ ,  $\forall a, b \in S$
2.  $g$  is decreasing with respect to the order defined in (1), i.e.,  $a \preceq b \rightarrow g(b) \leq g(a)$   $\forall a, b \in S$ ,
3.  $a \otimes \text{APP}^h(t, q) \otimes \text{APP}^h(q, t) \preceq a$ ,  $\forall a \in S$ ,  $q, t \in V$ , i.e., aggregating the cost of the cycles starting at an arbitrary node  $q$  and traversing a node  $t$ , does not increase (according to (1)) the cost  $a$  of a path.<sup>3</sup>

then  $d$ , as defined in (13), satisfies the triangle inequality over the nodes of  $G$ .

*Proof.* We need to prove that the triangle inequality (42) holds. Let  $s, q$  and  $t$  be arbitrary nodes of  $G$ . Due to the subadditivity of  $g$ , we have

$$g(\text{APP}^h(s, q) \otimes \text{APP}^h(q, t)) \leq g(\text{APP}^h(s, q)) + g(\text{APP}^h(q, t)).$$

Therefore, as a consequence of (39), it will be enough to show

$$d_L(s, t) = g(\text{APP}^h(s, t)) = g(\alpha^h \oplus \beta^h \otimes \text{APP}^h(q, t)) \leq g(\text{APP}^h(s, q) \otimes \text{APP}^h(q, t)) = d_L(s, q) + d_L(q, t),$$

which will follow from

$$\text{APP}^h(s, q) \otimes \text{APP}^h(q, t) \preceq \alpha^h \oplus \beta^h \otimes \text{APP}^h(q, t), \quad (44)$$

since  $g$  is decreasing. As a consequence of (40), it suffices to prove the following inequality

$$\begin{aligned} \text{APP}^h(s, q) \otimes \text{APP}^h(q, t) &= (\alpha^h \otimes \text{APP}^h(t, q) \oplus \beta^h) \otimes \text{APP}^h(q, t) \\ &= \alpha^h \otimes \text{APP}^h(t, q) \otimes \text{APP}^h(q, t) \oplus \beta^h \otimes \text{APP}^h(q, t) \\ &\preceq \alpha^h \oplus \beta^h \otimes \text{APP}^h(q, t), \end{aligned} \quad (45)$$

which holds if

$$\alpha^h \otimes \text{APP}^h(t, q) \otimes \text{APP}^h(q, t) \preceq \alpha^h. \quad (46)$$

Indeed, (46) holds thanks to our third assumption. □

**Corollary F.5.** *Min-norm distances, including the shortest path and minimax distances, are graph node metrics.*

*Proof.* We will apply the previous theorem. The subadditivity follows from the subadditivity of  $r$ -th roots:

$$a \otimes_r b = \sqrt[r]{a^r + b^r} \leq a + b, \quad \forall r \geq 1, \quad (47)$$

since  $g$  is equal to the identity function. Trivially, the subadditivity also holds for the max operation.

<sup>2</sup>Appendix A (François et al., 2017) proves that the sum of the path likelihoods is equal to 1 for hitting paths.

<sup>3</sup>Note that, although the cost of a path  $a$  is not increased when one aggregates the cost of the mentioned cycles, the distance does not decrease because  $g$  is decreasing. Thus, aggregating cycles leaves equal or increases the distance.

Furthermore, the identity function  $g$  is a decreasing function since

$$a \preceq b \iff \exists c \in S \text{ s.t. } \min(a, c) = a \oplus c = b \iff b \leq a. \quad (48)$$

Finally, the third assumption is a consequence of the increasing nature of the  $\otimes_r$  and  $\max$  operations: for  $a, b \in \mathbb{R}^+$  we have

$$a \otimes_r b = \sqrt[r]{a^r + b^r} \geq a \Rightarrow a \otimes_r b \preceq a; \quad \max(a, b) \geq a \Rightarrow a \otimes_\infty b \preceq a$$

Therefore,  $a \otimes \text{APP}^h(q, t) \otimes \text{APP}^h(t, q) \preceq a$ .

□

**Corollary F.6.** *The potential distance (Kivimäki et al., 2014; Françoisse et al., 2017) defines a metric.*

*Proof.* We will apply Theorem F.4. The potential distance can be retrieved by the APP associated with the log-semiring. In this case  $\otimes$  coincides with  $+$ . Therefore,  $g$  is the identity  $\otimes$ -homomorphism and the subadditivity follows trivially.

The function  $g$  is a decreasing function since

$$\begin{aligned} a \preceq b &\iff \exists c \in S \text{ s.t. } a \oplus_\mu c = -\frac{1}{\mu} \log(e^{-\mu a} + e^{-\mu c}) = b \\ &\iff \exists c \text{ s.t. } c = -\frac{1}{\mu} \log(e^{-\mu b} - e^{-\mu a}) \iff b \leq a \end{aligned} \quad (49)$$

The third assumption of Theorem 4.3 follows from the fact that the cost of a path is strictly positive and that  $a \otimes b = a + b \preceq a$  since  $a \leq a + b$  for  $b \geq 0$ . □

#### F.4. Proof Theorem 4.4

**Theorem F.7** (Theorem 4.4). *Let  $G = (V, E)$  be an  $S$ -graph. If*

1.  $g$  is  $\otimes$ -subadditive, i.e.  $g(a \otimes b) \leq g(a) + g(b)$ ,  $\forall a, b \in S$ ,
2.  $g$  is increasing in  $S \setminus \{\bar{0}\}$  with respect to the order defined in (1), i.e.,  $a \preceq b \rightarrow g(a) \leq g(b)$   $\forall a, b \in S \setminus \{\bar{0}\}$ ,
3.  $a \preceq a \otimes \text{APP}^h(t, q) \otimes \text{APP}^h(q, t) \forall a \in S, q, t \in V$ , i.e., aggregating the cost of the cycles starting at an arbitrary node  $q$  and traversing a node  $t$ , does not decrease the cost.<sup>4</sup>

then  $d$ , as defined in (13), satisfies the triangle inequality over the nodes of  $G$ .

*Remark F.8.* We need to consider that  $g$  is increasing in  $S \setminus \{\bar{0}\}$  because  $\bar{0} \preceq s$ ,  $\forall s \in S$  since  $s \oplus \bar{0} = s$ . However, by assumption  $g(\bar{0}) = \infty$ , thus  $g(\bar{0}) \geq g(s)$ . If we did not exclude  $\bar{0}$ ,  $g$  would map all elements of  $S$  to  $\infty$ .

*Proof.* We need to prove that the triangle inequality (42) holds. Let  $s, q$  and  $t$  be arbitrary nodes of  $G$ . Due to the subadditivity of  $g$ , we have

$$g(\text{APP}^h(s, q) \otimes \text{APP}^h(q, t)) \leq g(\text{APP}^h(s, q)) + g(\text{APP}^h(q, t)).$$

Therefore, as a consequence of (39), it will be enough to show

$$d_L(s, t) = g(\text{APP}^h(s, t)) = g(\alpha^h \oplus \beta^h \otimes \text{APP}^h(q, t)) \leq g(\text{APP}^h(s, q) \otimes \text{APP}^h(q, t)) = d_L(s, q) + d_L(q, t),$$

which will follow from

$$\alpha^h \oplus \beta^h \otimes \text{APP}^h(q, t) \preceq \text{APP}^h(s, q) \otimes \text{APP}^h(q, t). \quad (50)$$

since  $g$  is increasing. As a consequence of Lemma F.2, it suffices to show the following inequality

$$\begin{aligned} \alpha^h \oplus \beta^h \otimes \text{APP}^h(q, t) &\preceq \text{APP}^h(s, q) \otimes \text{APP}^h(q, t) \\ &= (\alpha^h \otimes \text{APP}^h(t, q) \oplus \beta^h) \otimes \text{APP}^h(q, t) \\ &= \alpha^h \otimes \text{APP}^h(t, q) \otimes \text{APP}^h(q, t) \oplus \beta^h \otimes \text{APP}^h(q, t), \end{aligned} \quad (51)$$

<sup>4</sup>And also the distance, since  $g$  is increasing.

which holds if

$$\alpha^h \preceq \alpha^h \otimes \text{APP}^h(t, q) \otimes \text{APP}^h(q, t). \quad (52)$$

Indeed, (52) holds thanks to our third assumption.  $\square$

**Corollary F.9.** *The Commute Cost Distance defines a metric.*

*Proof.* Note that the values of the semiring that we consider lie in  $\{1\} \times \mathbb{R}^+$ , since the first entry of  $\text{APP}(s, t)^h$  is equal to  $\sum_{\pi \in \mathcal{P}_{st}} \text{Pr}(\pi) = 1$  (see Remark F.3). Thus we just need to show the properties of Theorem 4.4 for this subset of elements.

In this case  $g$  is the projection of the second entry of  $\mathbb{R}^+ \times \mathbb{R}^+$ . Hence, the subadditivity follows from

$$g((1, c_1) \otimes (1, c_2)) = g((1, c_1 + c_2)) = c_1 + c_2 = g((1, c_1)) + g((1, c_2)). \quad (53)$$

The function  $g$  is an increasing function since

$$\begin{aligned} (p_1, c_1) \preceq (p_2, c_2) &\iff \\ \exists (p_3, c_3) \in S \text{ s.t. } (p_1, c_1) \oplus_{\mu} (p_3, c_3) &= (p_1 + p_3, c_1 + c_3) = (p_2, c_2) \\ \iff \exists p_3, c_3 > 0 \text{ s.t. } p_1 + p_3 &= p_2 \ \& \ c_1 + c_3 = c_2 \\ \Rightarrow c_1 &\leq c_2. \end{aligned} \quad (54)$$

The third assumption of Theorem 4.4 follows from

$$(p_1, c_1) \preceq (p_1, c_1) \otimes \underbrace{\text{APP}^h(t, q)}_{=(1, c_2)} \otimes \underbrace{\text{APP}^h(q, t)}_{=(1, c_3)} = (p_1, c_1) \otimes (1, c_2) \otimes (1, c_3) = (p_1, c_1 + p_1(c_2 + c_3)),$$

since  $c_1 \leq c_1 + p_1(c_2 + c_3)$  because  $p_1, c_2, c_3 \geq 0$ .  $\square$

### F.5. Use Case of the Results in Section 4

In order to illustrate the application of the results exposed in section 4, we will define a graph distance that can be easily verified to be a metric thanks to Theorem 4.3.

Let  $p_{ij} \in [0, 1)$  be the transition probabilities of a vanishing random walker, i.e. a random walker for which there exist at least one node  $k$  where  $\sum_j p_{kj} < 1$ . That is, there is a non-zero probability that the random walker "vanishes". Alternatively, one could interpret that there exists an absorbing node,  $i^*$ , to which every node,  $i$ , can transition with probability  $p_{ii^*} = 1 - \sum_{j \neq i^*} p_{ij}$ . In such case, there exists at least one node  $k$  such that  $p_{ki^*} > 0$ .

Let

$$d(s, t) := \begin{cases} \overbrace{-\frac{1}{r} \log \left( \sum_{\pi \in \mathcal{P}_{st}^h} \prod_{(i,j) \in E} p_{ij}^r \right)}^{d_L(s,t)} \overbrace{-\frac{1}{r} \log \left( \sum_{\pi \in \mathcal{P}_{ts}^h} \prod_{(i,j) \in E} p_{ij}^r \right)}^{d_R(s,t)} & \text{if } s \neq t \\ 0 & \text{otherwise} \end{cases} \quad (55)$$

We claim that (55) defines a metric. To prove it we will apply Theorem 4.3.

First, consider the semiring  $S = \{\mathbb{R}^+, \oplus_r, \cdot, 0, 1\}$  with  $x \oplus_r y = \sqrt[r]{x^r + y^r}$ ,  $r \geq 1$  and  $g(x) = -\log(x)$ . Note that if  $s \neq t$ , then

$$\bigoplus_{\pi \in \mathcal{P}_{st}^h} \bigotimes_{e \in \pi} p_e = \prod_{\pi \in \mathcal{P}_{st}^h} \sqrt[r]{\sum_{e \in \pi} p_e^r} = \sqrt[r]{\prod_{\pi \in \mathcal{P}_{st}^h} \sum_{e \in \pi} p_e^r}. \quad (56)$$

If we apply the function  $g(x) = -\log(x)$  to (56), we retrieve (55):

$$g \left( \bigoplus_{\pi \in \mathcal{P}_{st}^h} \bigotimes_{e \in \pi} p_e \right) = -\log \left( \sqrt[r]{\prod_{\pi \in \mathcal{P}_{st}^h} \sum_{e \in \pi} p_e^r} \right) = -\frac{1}{r} \log \left( \sum_{\pi \in \mathcal{P}_{st}^h} \prod_{(i,j) \in E} p_{ij}^r \right) = d_L(s, t). \quad (57)$$

From the definition of (55), it is clear that  $d(s, t)$  satisfies the symmetry and indiscernible properties. It is only left to prove the triangle inequality to show that (55) is indeed a metric. This property follows easily from Theorem 4.3, since:

- $g(x) = -\log(x)$  is  $\otimes$ -subadditive:

$$\begin{aligned} g(a \otimes b) &= -\log(a \otimes b) = \log\left(\sqrt[r]{a^r + b^r}\right) = -\frac{1}{r} \log(a^r + b^r) \\ &= -\frac{1}{r} \log(a^r) - \frac{1}{r} \log(b^r) = -\log(a) - \log(b) = g(a) + g(b) \end{aligned}$$

- $g(x) = -\log(x)$  is decreasing: Note that  $a \preceq b \iff \exists c \in \mathbb{R}^+$  such that  $a^r + c^r = b^r$ . Consequently,

$$a \preceq b \iff a \leq b. \quad (58)$$

Thus, if  $a \preceq b$ , then

$$g(b) = -\log(b) = -\log\left(\sqrt[r]{a^r + c^r}\right) = -\frac{1}{r} \log(a^r + c^r) \leq -\frac{1}{r} \log(a^r) = -\log(a) = g(a).$$

- A similar argument as was used in (Françoisse et al., 2017) (see Remark F.3) can be used to show that  $\text{APP}^h(t, q) = \bigoplus_{\pi \in \mathcal{P}_{st}^h} \bigotimes_{e \in \pi} p_e \leq 1$ . Hence,  $a \otimes \text{APP}^h(t, q) \otimes \text{APP}^h(q, t) \preceq a$ ,  $\forall a \in S$ ,  $q, t \in V$  follows from this fact together with (58).

We have just proven the following corollary.

**Corollary F.10.** *Given the transition probabilities of a vanishing random walker,  $p_{ij} \in [0, 1)$ , over an arbitrary graph  $G$ , the function  $d(s, t)$  defines a metric.*

$$d(s, t) := \begin{cases} -\frac{1}{r} \log\left(\sum_{\pi \in \mathcal{P}_{st}^h} \prod_{(i,j) \in E} p_{ij}^r\right) - \frac{1}{r} \log\left(\sum_{\pi \in \mathcal{P}_{ts}^h} \prod_{(i,j) \in E} p_{ij}^r\right) & \text{if } s \neq t \\ 0 & \text{otherwise} \end{cases}. \quad (59)$$

We provide a bit of intuition about this metric: when  $r = 1$ , the expression  $\sum_{\pi \in \mathcal{P}_{st}^h} \prod_{(i,j) \in E} p_{ij}^r$  is equal to the absorbing probability of  $t$  from  $s$  before the random walker vanishes.

$$\lim_{r \rightarrow 1} \sum_{\pi \in \mathcal{P}_{st}^h} \prod_{(i,j) \in E} p_{ij}^r = \sum_{\pi \in \mathcal{P}_{st}^h} \prod_{(i,j) \in E} p_{ij} = \sum_{\pi \in \mathcal{P}_{st}^h} \text{Pr}(\pi)$$

Thus, when  $r \rightarrow 1$ , nodes are closer if they have higher absorbing probability before the RW vanishes. On the other extreme, when  $r \rightarrow \infty$ , then the distance focuses on the path of maximum probability between two nodes. If we consider  $p_{ij} = \exp(-c_{ij})$ ,<sup>5</sup> the limit case  $r \rightarrow \infty$  would be equivalent to twice the shortest path cost with edge costs equal to  $c_{ij}$ :

$$\begin{aligned} \lim_{r \rightarrow \infty} d_L(s, t) &= \lim_{r \rightarrow \infty} -\frac{1}{r} \log\left(\sum_{\pi \in \mathcal{P}_{st}^h} \prod_{(i,j) \in E} \exp(-r \cdot c_{ij})\right) = \lim_{r \rightarrow \infty} -\log\left(\sqrt[r]{\sum_{\pi \in \mathcal{P}_{st}^h} \exp(-r \cdot c(\pi))}\right) \\ &= -\log\left(\max_{\pi \in \mathcal{P}_{st}^h} \exp(-c(\pi))\right) = \min_{\pi \in \mathcal{P}_{st}^h} c(\pi). \end{aligned}$$

## G. Log-Norm Distance

In this section we will prove that the log-norm distance defines a metric over the nodes of a graph with positive edge-costs. Since the log-norm operations  $\otimes_r$  and  $\oplus_\mu$  do not define a semiring, we can not use the results developed in Section 4. Therefore, we present an additional proof for the log-norm distance.

**Lemma G.1.** *Given  $r \geq 1$ ,  $\mu > 0$  and  $c : E \mapsto \mathbb{R}^+$  then*

$$\text{LN}(s, t) = -\frac{1}{\mu} \left( \log\left(\mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ \exp\left(-\mu c(\|\pi\|_r)\right)\right]\right) + \log\left(\mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} \left[ \exp\left(-\mu c(\|\pi\|_r)\right)\right]\right) \right) \quad (60)$$

defines a distance over the vertices of the graph.

<sup>5</sup>In this case we are assuming that  $c_{ij}$  are high enough for all  $i, j$  such that  $\sum_k c_{ik} < 1$  for all nodes  $i$ .

*Proof.* The symmetry, non-negativity and the equality  $\text{LN}(s, s) = 0$  are trivial consequences of the definition. Moreover, if we assign to each  $e \in E$  a cost  $c(e) > 0$ , then  $c(\|\pi\|_r) > 0$  for any path  $\pi$ . Consequently,  $\text{LN}(s, t) > 0$  for  $s \neq t$ .

From now on, we focus on the triangle inequality. We will show the triangle inequality for the terms with the paths in  $\mathcal{P}_{st}^h$  since the case for  $\mathcal{P}_{ts}^h$  is analogous. Hence, we claim

$$\begin{aligned} & -\frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] \right) \\ & \leq -\frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{sq}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] \right) - \frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{qt}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] \right) \end{aligned} \quad (61)$$

Equation (61) is equivalent to

$$\mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] \geq \mathbb{E}_{\pi \sim \mathcal{P}_{sq}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] \mathbb{E}_{\pi \sim \mathcal{P}_{qt}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] \quad (62)$$

after applying  $-\frac{1}{\mu} \log(\cdot)$  to both sides of (62). Next, in order to isolate the terms on the right-hand side of (62), we separate the set of hitting paths from  $s$  to  $t$  into those that cross the third node  $q$  and those that do not.

$$\begin{aligned}
 & \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] = \sum_{\pi \in \mathcal{P}_{st}^h} \Pr(\pi) \exp \left( -\mu \|c(\pi)\|_r \right) \\
 & \stackrel{*}{\geq} \sum_{\substack{\pi_1 \in \mathcal{P}_{sq}^h \\ t \notin \pi_1}} \sum_{\pi_2 \in \mathcal{P}_{qt}^h} \Pr(\pi_1) \Pr(\pi_2) \exp \left( -\mu \|c(\pi_1)\|_r \right) \exp \left( -\mu \|c(\pi_2)\|_r \right) \\
 & + \sum_{\substack{\pi \in \mathcal{P}_{st}^h \\ q \notin \pi}} \Pr(\pi) \exp \left( -\mu \|c(\pi)\|_r \right) \\
 & = \left( \sum_{\substack{\pi_1 \in \mathcal{P}_{sq}^h \\ t \notin \pi_1}} \Pr(\pi_1) \exp \left( -\mu \|c(\pi_1)\|_r \right) \right) \left( \sum_{\pi_2 \in \mathcal{P}_{qt}^h} \Pr(\pi_2) \exp \left( -\mu \|c(\pi_2)\|_r \right) \right) \\
 & + \sum_{\substack{\pi \in \mathcal{P}_{st}^h \\ q \notin \pi}} \Pr(\pi) \exp \left( -\mu \|c(\pi)\|_r \right) \\
 & = \left( \sum_{\pi_1 \in \mathcal{P}_{sq}^h} \Pr(\pi_1) \exp \left( -\mu \|c(\pi_1)\|_r \right) - \sum_{\substack{\pi_1 \in \mathcal{P}_{sq}^h \\ t \in \pi_1}} \Pr(\pi_1) \exp \left( -\mu \|c(\pi_1)\|_r \right) \right) \\
 & \times \left( \sum_{\pi_2 \in \mathcal{P}_{qt}^h} \Pr(\pi_2) \exp \left( -\mu \|c(\pi_2)\|_r \right) \right) + \sum_{\substack{\pi \in \mathcal{P}_{st}^h \\ q \notin \pi}} \Pr(\pi) \exp \left( -\mu \|c(\pi)\|_r \right) \tag{63} \\
 & = \left( \sum_{\pi_1 \in \mathcal{P}_{sq}^h} \Pr(\pi_1) \exp \left( -\mu \|c(\pi_1)\|_r \right) \right) \left( \sum_{\pi_2 \in \mathcal{P}_{qt}^h} \Pr(\pi_2) \exp \left( -\mu \|c(\pi_2)\|_r \right) \right) \\
 & - \left( \sum_{\substack{\pi_1 \in \mathcal{P}_{sq}^h \\ t \in \pi_1}} \Pr(\pi_1) \exp \left( -\mu \|c(\pi_1)\|_r \right) \right) \left( \sum_{\pi_2 \in \mathcal{P}_{qt}^h} \Pr(\pi_2) \exp \left( -\mu \|c(\pi_2)\|_r \right) \right) \\
 & + \sum_{\substack{\pi \in \mathcal{P}_{st}^h \\ q \notin \pi}} \Pr(\pi) \exp \left( -\mu \|c(\pi)\|_r \right) \\
 & = \mathbb{E}_{\pi \sim \mathcal{P}_{sq}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] \mathbb{E}_{\pi \sim \mathcal{P}_{qt}^h} \left[ \exp \left( -\mu \|c(\pi)\|_r \right) \right] \\
 & - \left( \sum_{\substack{\pi_1 \in \mathcal{P}_{sq}^h \\ t \in \pi_1}} \Pr(\pi_1) \exp \left( -\mu \|c(\pi_1)\|_r \right) \right) \left( \sum_{\pi_2 \in \mathcal{P}_{qt}^h} \Pr(\pi_2) \exp \left( -\mu \|c(\pi_2)\|_r \right) \right) \\
 & + \sum_{\substack{\pi \in \mathcal{P}_{st}^h \\ q \notin \pi}} \Pr(\pi) \exp \left( -\mu \|c(\pi)\|_r \right)
 \end{aligned}$$

In \* we have used the fact that

$$\sqrt[r]{\sum_{i \in I_1} x_i^r + \sum_{i \in I_2} y_i^r} \leq \sqrt[r]{\sum_{i \in I_1} x_i^r} + \sqrt[r]{\sum_{i \in I_2} y_i^r}.$$

and that  $\exp(-x)$  is a decreasing function. Note that for  $r = 1$  equality holds.

In order to show (62) from (63) we need to prove

$$\begin{aligned} & \sum_{\substack{\pi \in \mathcal{P}_{st}^h \\ q \notin \pi}} \Pr(\pi) \exp\left(-\mu \|c(\pi)\|_r\right) \\ & \geq \left( \sum_{\substack{\pi \in \mathcal{P}_{sq}^h \\ t \in \pi}} \Pr(\pi) \exp\left(-\mu \|c(\pi)\|_r\right) \right) \left( \sum_{\pi \in \mathcal{P}_{qt}^h} \Pr(\pi) \exp\left(-\mu \|c(\pi)\|_r\right) \right) \end{aligned} \quad (64)$$

Let  $\pi_1 \odot \pi_2$  denote the concatenation of paths. Since  $c(e) > 0$  we deduce

$$c(\pi_1 \odot \pi_2) = \sqrt[r]{\sum_{e \in \pi_1 \odot \pi_2} (c(e))^r} \geq \sqrt[r]{\sum_{e \in \pi_1} (c(e))^r} = c(\pi_1).$$

Thus,

$$\begin{aligned} & \sum_{\substack{\pi \in \mathcal{P}_{sq}^h \\ t \in \pi}} \Pr(\pi) \exp\left(-\mu \|c(\pi)\|_r\right) = \left( \sum_{\substack{\pi_1 \in \mathcal{P}_{st}^h \\ q \notin \pi_1}} \sum_{\pi_2 \in \mathcal{P}_{tq}^h} \Pr(\pi_1) \Pr(\pi_2) \exp\left(-\mu c(\pi_1 \odot \pi_2)\right) \right) \\ & \leq \left( \sum_{\substack{\pi_1 \in \mathcal{P}_{st}^h \\ q \notin \pi_1}} \sum_{\pi_2 \in \mathcal{P}_{tq}^h} \Pr(\pi_1) \Pr(\pi_2) \exp\left(-\mu c(\pi_1)\right) \right) = \sum_{\substack{\pi_1 \in \mathcal{P}_{st}^h \\ q \notin \pi_1}} \Pr(\pi_1) \exp\left(-\mu c(\pi_1)\right) \end{aligned} \quad (65)$$

Since  $\|c(\pi)\|_r > 0$  for any  $\pi \in \mathcal{P}_{ij}^h$  for any vertices  $i \neq j$ . Therefore,

$$\sum_{\pi \in \mathcal{P}_{st}^h} \Pr(\pi) \exp\left(-\mu \|c(\pi)\|_r\right) < \sum_{\pi \in \mathcal{P}_{st}^h} \Pr(\pi) = 1. \quad (66)$$

Then equation (64) follows from (65) and (66).  $\square$

Note that, as a consequence of the previous theorem, the triangle inequality of the exp-max, log-max and exp-norm limits exposed in Table 1 is also satisfied. Indeed, by taking the corresponding limits in both sides of the log-norm triangle inequality, the inequality will still hold.

## H. Log-Norm Distance and the Randomized Shortest Paths

In this section we will relate the log-norm distance with the Helmholtz free energy, following the same reasoning that related the potential distance with the free energy in (Kivimäki et al., 2014). Let  $s$  and  $t$  two arbitrary but fixed nodes in the graph  $G$ . As defined in (Kivimäki et al., 2014), the free energy of a thermodynamical system modelled by the probability distribution  $\Pr_{st}$  and with temperature  $T = 1/\mu$  is given by

$$\Phi(\Pr_{st}) = \sum_{\pi \in \mathcal{P}_{st}^h} \Pr_{st}(\pi) c(\pi) + \frac{1}{\mu} \text{KL}(\Pr_{st}, \Pr^{\text{ref}}), \quad (67)$$

where  $\Pr^{\text{ref}}(\pi) = \prod_{e \in \pi} p_e$  is the probability that a path  $\pi \in \mathcal{P}_{st}^h$  is generated by a random walker and KL is the Kullback-Leibler divergence. In our setting, the cost of a path will be given by  $\|c(\pi)\|_r$ . Therefore, we use the following expression for the free energy

$$\Phi_r(\Pr_{st}) = \sum_{\pi \in \mathcal{P}_{st}^h} \Pr_{st}(\pi) \|c(\pi)\|_r + \frac{1}{\mu} \text{KL}(\Pr_{st}, \Pr^{\text{ref}}). \quad (68)$$

We will show that the symmetrized minimum free energy between two nodes  $s$  and  $t$  (free energy distance in (Kivimäki et al., 2014)) coincides with the log-norm distance, i.e.

$$\text{LN}(s, t) = \Phi_r(\text{Pr}_{st}) + \Phi_r(\text{Pr}_{ts}).$$

First, we define the probability distribution over the hitting paths from  $s$  to  $t$  as the one that minimizes the free energy

$$\text{Pr}_{st}^*(\cdot) := \arg \min_{\text{Pr}(\cdot)} \sum_{\pi \in \mathcal{P}_{st}^h} \text{Pr}(\pi) \|c(\pi)\|_r + \frac{1}{\mu} \text{KL}(\text{Pr}_{st}, \text{Pr}^{\text{ref}}). \quad (69)$$

It can be easily checked that the minimizer is given by the following Gibbs probability distribution (Kivimäki et al., 2014)

$$\text{Pr}_{st}^*(\pi) = \frac{\text{Pr}^{\text{ref}}(\pi) \exp(-\mu \|c(\pi)\|_r)}{\sum_{\hat{\pi} \in \mathcal{P}_{st}^h} \text{Pr}^{\text{ref}}(\hat{\pi}) \exp(-\mu \|c(\hat{\pi})\|_r)}. \quad (70)$$

If we now compute the KL-divergence between  $\text{Pr}_{st}^*$  and  $\text{Pr}^{\text{ref}}$  we obtain

$$\begin{aligned} \text{KL}(\text{Pr}_{st}^*, \text{Pr}^{\text{ref}}) &= \sum_{\pi \in \mathcal{P}_{st}^h} \text{Pr}_{st}^*(\pi) \log \left( \frac{\text{Pr}_{st}^*(\pi)}{\text{Pr}^{\text{ref}}(\pi)} \right) \\ &= \sum_{\pi \in \mathcal{P}_{st}^h} \text{Pr}_{st}^*(\pi) \log \left( \frac{\text{Pr}^{\text{ref}}(\pi) \exp(-\mu \|c(\pi)\|_r)}{\sum_{\hat{\pi} \in \mathcal{P}_{st}^h} \text{Pr}^{\text{ref}}(\hat{\pi}) \exp(-\mu \|c(\hat{\pi})\|_r)} \right) - \sum_{\pi \in \mathcal{P}_{st}^h} \text{Pr}_{st}^*(\pi) \log(\text{Pr}^{\text{ref}}(\pi)) \\ &= \sum_{\pi \in \mathcal{P}_{st}^h} \text{Pr}_{st}^*(\pi) \log(\text{Pr}^{\text{ref}}(\pi)) - \mu \sum_{\pi \in \mathcal{P}_{st}^h} \text{Pr}_{st}^*(\pi) \|c(\pi)\|_r \\ &\quad - \log \left( \sum_{\hat{\pi} \in \mathcal{P}_{st}^h} \text{Pr}^{\text{ref}}(\hat{\pi}) \exp(-\mu \|c(\hat{\pi})\|_r) \right) - \sum_{\pi \in \mathcal{P}_{st}^h} \text{Pr}_{st}^*(\pi) \log(\text{Pr}^{\text{ref}}(\pi)) \\ &= -\mu \sum_{\pi \in \mathcal{P}_{st}^h} \text{Pr}_{st}^*(\pi) \|c(\pi)\|_r - \log \left( \sum_{\pi \in \mathcal{P}_{st}^h} \text{Pr}^{\text{ref}}(\pi) \exp(-\mu \|c(\pi)\|_r) \right) \end{aligned} \quad (71)$$

Combining this result with (68) it follows that

$$\Phi_r(\text{Pr}_{st}^*) = -\frac{1}{\mu} \log \left( \sum_{\pi \in \mathcal{P}_{st}^h} \text{Pr}^{\text{ref}}(\pi) \exp(-\mu \|c(\pi)\|_r) \right).$$

Finally, symmetrizing this expression we obtain the log-norm distance.

## I. Exp-Max and Log-Max Metric Computation

Currently, there does not exist any efficient algorithm to compute the log-norm distance, LN, in its general form. Nonetheless, we briefly sketch here a possible algorithm to compute the novel exp-max distance (EM) that arises as a limit case of LN (see Table 1).

$$\text{EM}(s, t) = \underbrace{\mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ \max_{e \in \pi} c(e) \right]}_{\text{EM}_L(s, t)} + \underbrace{\mathbb{E}_{\pi \sim \mathcal{P}_{ts}^h} \left[ \max_{e \in \pi} c(e) \right]}_{\text{EM}_R(s, t)} \quad (72)$$

Let  $G = (V, E)$  be a graph,  $l(E)$  be the set of edge costs instantiated by the graph  $G$ , and  $\mathcal{P}_{st}^h(c)$  be the set of paths with maximum cost equal to  $c$ :

$$l(E) := \{c(e) : e \in E\} \quad (73)$$

$$\mathcal{P}_{st}^h(c) := \{\pi \in \mathcal{P}_{st}^h : c = \max_{e \in \pi} c(e)\} \quad (74)$$

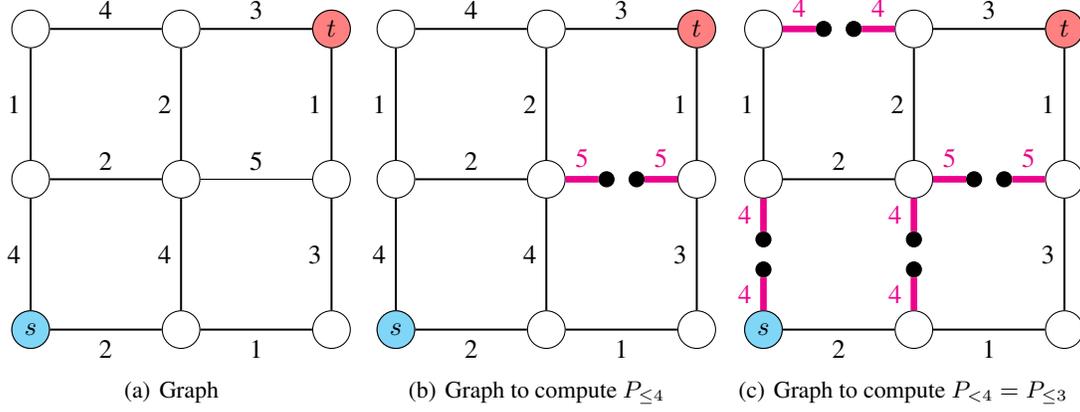


Figure 3. 3(a) Graph with edge costs and two marked nodes  $s$  and  $t$ . 3(b) Graph with two absorbing nodes (black nodes). Each path connecting  $s$  and  $t$  that has a cost  $\|c(\pi)\|_\infty = \max_{e \in \pi} c_e$  higher than 4 must contain the edge with cost equal to 5. By adding artificial absorbing nodes to this edge,<sup>7</sup> any random walker that crosses this edge will be absorbed. Thus, the node  $t$  absorption probability by a random walker starting at node  $s$  in this graph is equal to the probability of sampling a path with cost lower or equal than 4, i.e.,  $P_{\leq 4}$ . 3(c) Analogously we can compute  $P_{<4}$  if we add absorbing nodes in all edges with cost higher or equal than 4. Note that since there are no edges with cost in between 3 and 4, we have that  $P_{<4} = P_{\leq 3}$ .

We can decompose the left summand of EM as

$$\text{EM}_L(s, t) = \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ \max_{e \in \pi} c(e) \right] = \sum_{\pi \in \mathcal{P}_{st}^h} \Pr(\pi) \max_{e \in \pi} c(e) = \sum_{c \in l(E)} \sum_{\pi \in \mathcal{P}_{st}^h(c)} \Pr(\pi) \max_{e \in \pi} c(e) = \sum_{c \in l(E)} c \Pr(\pi \in \mathcal{P}_{st}^h(c)).$$

Let  $P_{\square c} := \Pr(\pi \in \cup_{c' \square c} \mathcal{P}_{st}^h(c'))$  with  $\square \in \{<, \leq\}$ . Thus,

$$\Pr(\pi \in \mathcal{P}_{st}^h(c)) = P_{\leq c} - P_{< c} \quad (75)$$

can be computed in closed form, since  $P_{< c}$  ( $P_{\leq c}$ ) is the probability of reaching  $t$  from  $s$  without traversing an edge with lower (or equal) cost than  $c$ . This is equal to the absorption probability of  $t$ , which can be computed analytically by solving a linear system (see 3.7.2 (Fouss et al., 2016)), once extra absorbing nodes have been set on the edges with higher (or equal) cost than  $c$  (see Figure 3). The computational cost of this algorithm scales with  $|l(c)|$ . To reduce the computational cost, we suggest to bin the edge costs coarsely.

Analogously, one can decompose the left summand of the log-max distance (LM) (see Table 1) and approximate it in a similar form:

$$\text{LM}_L(s, t) = -\frac{1}{\mu} \log \left( \mathbb{E}_{\pi \sim \mathcal{P}_{st}^h} \left[ e^{-\mu \max_{e \in \pi} c(e)} \right] \right) = -\frac{1}{\mu} \log \left( \sum_{c \in l(E)} e^{-\mu c} \Pr(\pi \in \mathcal{P}_{st}^h(c)) \right). \quad (76)$$

We have shown that one can compute these particular limit instances of the log-norm distance. The analysis and implementation of this algorithm is out of the scope of the current paper and, therefore, is left for future work.

<sup>7</sup>We add two absorbing nodes per edge to account for the directions of the edges, which could have different costs. Since in this concrete case the graph is undirected, adding one absorbing node would have been enough.