Lossy gossip and composition of metrics

LOSSY GOSSIP AND COMPOSITION OF METRICS

ANDRIES E. BROUWER, JAN DRAISMA, AND BART J. FRENK

1. Introduction and results

Imagine travelling between three locations such as Eindhoven (E, a medium-sized town in the Netherlands), a parking lot P on the border of the Dutch capital Amsterdam, and the city center A of Amsterdam. In Figure 1 the travel times by car between these locations are depicted by the leftmost triangle, while the travel times by bike are depicted by the second triangle. The large distances between E and either P or A are covered much faster by car than by bike. On the other hand, because of crowded streets, the short distance between P and A is covered considerably faster by bike than by car. As a consequence, an attractive alternative for travelling from E to A by car is to travel by car from E to P and continue by bike to A. In other words, to get from E to A we first do a step in the car metric and then a step in the bike metric, where we optimise the sum of the two travel times. Computing this *car-bike metric* for the remaining ordered pairs leads to the picture on the right in Figure 1. The corresponding matrix computation is

$$\begin{bmatrix} 0 & 90 & 140 \\ 90 & 0 & 60 \\ 140 & 60 & 0 \end{bmatrix} \odot \begin{bmatrix} 0 & 630 & 640 \\ 630 & 0 & 20 \\ 640 & 20 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 90 & 110 \\ 90 & 0 & 20 \\ 140 & 20 & 0 \end{bmatrix},$$

where $\odot$ is *tropical* or *min-plus* matrix multiplication, obtained from usual matrix multiplication by changing plus into minimum and times into plus. Note that the resulting matrix is not symmetric (the transpose corresponds to the "first bike, then car" metric), and that it does not satisfy the triangle inequality either. Observe also that if we vary the travel times in the two metric matrices slightly, their min-plus product moves in a three-dimensional space, where the entry corresponding to 110 remains the sum of the entries corresponding to 90 and 20, and smaller than the entry corresponding to 140. This preservation of dimension when tropically multiplying cones of distance matrices is one of the key results of this paper.

While keeping this min-plus product in the back of our minds, we next contemplate the following different setting. Three gossipers, Eve, Patricia, and Adam, each have an individual piece of gossip, which they can share through one-to-one

JD is supported by a Vidi grant from the Netherlands Organisation for Scientific Research (NWO), and BJF by an NWO free competition grant.

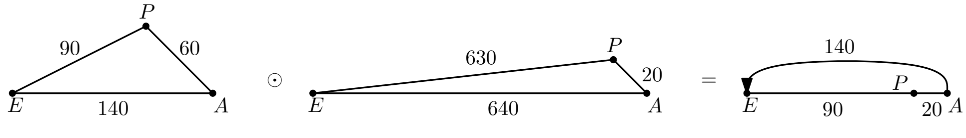


Figure 1. Composing the car metric with the bike metric.

phone calls in which both callers update each other on all the gossip they know. Record the knowledge of E, P, A in a three-by-three uncertainty matrix with entries 0 (for "i's gossip is known by j") and ∞ (for the other entries). Then initially that matrix is the *tropical identity matrix*, with zeroes along the diagonal and ∞ outside the diagonal. A phone call beteen E and P, for example, corresponds to tropically right-multiplying that tropical identity matrix matrix with

$$\begin{bmatrix} 0 & 0 & \infty \\ 0 & 0 & \infty \\ \infty & \infty & 0 \end{bmatrix},$$

resulting in this very same matrix. A second phone call between P and A leads to

$$\begin{bmatrix} 0 & 0 & \infty \\ 0 & 0 & \infty \\ \infty & \infty & 0 \end{bmatrix} \odot \begin{bmatrix} 0 & \infty & \infty \\ \infty & 0 & 0 \\ \infty & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \infty & 0 & 0 \end{bmatrix}.$$

Note the resemblance of this computation with the car-bike metric computation above. This resemblance can be made more explicit by passing from gossip to *lossy gossip*, where each phone call between gossipers k and l comes with a parameter $q \in [0, 1]$ to be interpreted as the fraction of information that gets broadcast correctly through the phone line, and where each gossiper j knows a fraction $p_{ij} \in [0, 1]$ of i's gossip. Assume the (admittedly simplistic) procedure where k updates his knowledge of gossip i to $q \cdot p_{il}$ if this is larger than p_{ik} and retains his knowledge p_{ik} of gossip i otherwise, and similarly for gossiper l. In this manner, the fractions p_{ij} are updated through a series of lossy phone-calls. Passing from p_{ij} to the *uncertainty* $u_{ij} := -\log p_{ij} \in [0, \infty]$ of gossiper j about gossip i and from q to the *loss* $a := -\log q \in [0, \infty]$ of the phone line in the call between k and l, the update rule changes u_{ik} into the minimum of u_{ik} and $u_{il} + a$, and similarly for u_{il}. This is just tropical right-multiplication with the matrix $C_{kl}(a)$ having 0's on the diagonal, ∞'s everywhere else, except an a on positions (k, l) and (l, k). So *lossy gossip is tropical matrix multiplication*. Note that lossy gossip is different from *gossip over faulty telephone lines* discussed in [BH86, HRS87].

This paper concerns the entirety of such uncertainty matrices, or compositions of finite metrics. Our main result uses the following notation: fixing a number n (of gossipers or vertices), let $D = D_n$ be the set of all *metric* $n \times n$ matrices, i.e., matrices with entries $a_{ij} \in \mathbb{R}_{\geq 0}$ satisfying $a_{ii} = 0$ and $a_{ij} = a_{ji}$ and $a_{ij} + a_{jk} \geq a_{ik}$.

Theorem 1.1. *The set* $\{A_1 \odot \cdots \odot A_k \mid k \in \mathbb{N},\ A_1, \ldots, A_k \in D_n\}$ *is the support of a (finite) polyhedral complex of dimension* $\binom{n}{2}$*, whose topological closure in* $[0, \infty]^{n \times n}$ *is the monoid generated by the matrices* $C_{kl}(a)$ *with* $k, l \in [n] := \{1, \ldots, n\}$ *and* $a \in [0, \infty]$.

Theorem 1.2. *For* $n \leq 5$ *the complex in the previous theorem is pure and connected in codimension* 1*. Moreover, for* $n \leq 4$*, there is a unique coarsest such complex. This coarsest complex has* D_2, D_3, D_4 *among the* $1, 7, 289$ *full-dimensional cones; and in total it has* $1, 2, 16$ *orbits of full-dimensional cones under the groups* $\mathrm{Sym}(2), \mathrm{Sym}(3), \mathrm{Sym}(4)$*, respectively.*

For some statistics for $n = 5$ we refer to Section 6. We conjecture that the pureness and connectedness in codimension 1 carry through to arbitrary n.

About the length of products we can say the following.

Theorem 1.3. *For $n \leq 5$ every element of G_n is the tropical product of at most $\binom{n}{2}$ lossy phone call matrices $C_{kl}(a)$, but not every element is the tropical product of fewer factors.*

We conjecture that the restriction $n \leq 5$ can be omitted.

Our next result concerns "pessimal" ordinary gossip.

Theorem 1.4. *Any sequence of phone calls among n gossiping parties such that in each phone call both participants exchange all they know, and at least one of the parties learns something new, has length at most $\binom{n}{2}$, and equality occurs.*

Corollary 1.5. *In the monoid generated by the matrices $C_{kl}(0)$, $k, l \in [n]$ every irredundant product of such matrices has at most $\binom{n}{2}$ factors.*

Our motivation for this paper is twofold. First, it establishes a connection between gossip networks and composition of metrics that seems worth pursuing further. Second, the *lossy gossip monoid* defined below is a beautiful example of a submonoid of $(\mathbb{R} \cup \{\infty\})^{n\times n}$; a general theory of such submonoids also seems very worthwile. Note that sub*groups* of this semigroup (but with identity element an arbitrary idempotent matrix) have been investigated in [IJK12].

The remainder of this paper is organised as follows. Sections 2 and 3 contain observations that pave the way for the analysis for $n = 3, 4$ in Sections 4 and 5. In Section 6 we report on extensive computations for $n = 5$. In Section 7 we use tropical algebraic geometry to prove the first statement of Theorem 1.1. Interestingly, no polyhedral-combinatorial proof is known. In Section 8 we study the monoid generated by the ordinary gossip matrices $C_{kl}(0)$, $k, l \in [n]$: for small values of n we determine its order, and we prove Theorem 1.4. We conclude with a number of open questions in Section 9.

2. Preliminaries

Fixing a natural number n, we define $\overline{D}_n$ to be the topological closure of D_n in $[0, \infty]^{n\times n}$, and we denote by G_n the monoid generated by $\overline{D}_n$ under min-plus matrix multiplication. We call G_n the *lossy gossip monoid* with n gossipers. This terminology is justified by the following lemma.

Lemma 2.1. *The lossy gossip monoid G_n is generated by the* lossy phone call matrices $C_{kl}(a)$ $(k, l \in [n], a \in [0, \infty])$ *having zeroes on the diagonal and ∞ everywhere else except for values a on positions (k, l) and (l, k).*

Proof. Lossy phone call matrices lie in $\overline{D}_n$, so the monoid that they generate is contained in G_n. For the converse it suffices to show that every element A of $\overline{D}_n$ is the product of lossy phone call matrices. We claim that, in fact, $A = \prod_{k<l} C_{kl}(a_{kl}) =: B$, where the a_{kl} are the entries of A and the product is taken in any order. Indeed, the (i, j)-entry of B is the minimum of expressions of the form $a_{i_0,i_1} + a_{i_1,i_2} + \ldots + a_{i_{s-1},i_s}$ where $s \leq \binom{n}{2}$, $i_0 = i$, $i_s = j$, and where the $C_{i_0,i_1}, \ldots, C_{i_{s-1},i_s}$ (with $s \leq \binom{n}{2}$) appear in that order (though typically interspersed with other factors) in the product expression for B. By the triangle inequalities among the entries of A, the minimum of these expressions equals $a_{i,j}$. □

Although elements of G_n need not be symmetric, they have a symmetric core.

Lemma 2.2. *Each element A of G_n satisfies $a_{i,j} = a_{j,i}$ for at least $n-1$ pairs of distinct indices i, j. If we view such pairs of indices as edges in the complete graph with vertex set $[n]$, then these edges form a connected spanning subgraph (for $n > 1$).*

Proof. Consider a partition of $[n]$ into two nonempty parts K and L. Let a_{ij} be the smallest among the values a_{kl} with $k \in K$, $l \in L$. Then we see from a representation of A as product of (symmetric) lossy phone call matrices that $a_{ij} = a_{ji}$. □

Every connected graph on $[n]$ occurs as symmetric core of some element of G_n.

Observe that $C_{kl}(a) \odot C_{kl}(b) = C_{kl}(a \oplus b)$, where $\oplus$ denotes tropical addition (defined by $a \oplus b = \min(a, b)$). Thus Lemma 2.1 exhibits G_n as a monoid generated by certain *one-parameter submonoids*, reminiscent of the generation of algebraic groups by one-parameter subgroups. This resemblance will be exploited in Section 7.

We define the *length* of an element X of G_n as the minimal number of factors in any expression of X as a tropical product of lossy phone call matrices. The following lemma exhibits a uniform upper bound on the length of elements of G_n.

Lemma 2.3. *The length of an element of G_n is at most $n(n-1)^2$.*

Proof. Let A be an element of G_n and write

$$A = C_{I_1}(a_1) \odot \cdots \odot C_{I_k}(a_k)$$

where the a_j are non-negative real numbers and the I_j are unordered pairs of distinct numbers in $[n]$. The entry at position (i, j) of A is the minimum of expressions $a_{k_1} + \ldots + a_{k_s}$, where $(I_{k_1}, \ldots, I_{k_s})$ is a path from i to j in the complete graph on $[n]$ and $k_1 < \cdots < k_s$. Since revisiting a site is never cheaper, without loss of generality $s \leq n-1$. So for each of the (at most) $n(n-1)$ non-zero entries of A only (at most) $n-1$ of the factors above are essential. Thus in total at most $n(n-1)^2$ of the factors above are essential. Leaving out all other factors yields a product that still equals A. □

One may wonder how long an irredundant product of factors can be when all factors are lossy phone call matrices $C_{kl}(a)$. The above proof shows that $n(n-1)^2$ is an upper bound. A slight sharpening of the proof yields the upper bound $n^2(n-1)/2$. A construction shows that $\binom{n+1}{3}$ is a lower bound. For $n < 5$ this latter value is the actual maximum length.

Lemma 2.4. *The length of an expression that is an irredundant tropical product of lossy phone call matrices in G_n is at most $n^2(n-1)/2$.*

Proof. Let A be an element of G_n and write

$$A = C_{I_1}(a_1) \odot \cdots \odot C_{I_k}(a_k)$$

where the a_j are non-negative real numbers and the I_j are unordered pairs of distinct numbers in $[n]$. The entry at position (h, i) of A is the minimum of expressions $a_{k_1} + \ldots + a_{k_s}$, where $(I_{k_1}, \ldots, I_{k_s})$ is a path from h to i in the complete graph on $[n]$ and $k_1 < \cdots < k_s$. Suppose (for $i \neq j$) that a cheapest path from h to i passes through j. Then we see a path from h to j, and there may be cheaper paths. Hence $A_{hj} \leq A_{hi}$ and there is a cheapest path from h to j that does not pass through i. We see that the entries A_{hi} with fixed h involve at most $n(n-1)/2$ factors. □

Lemma 2.5. *There exists an expression that is an irredundant tropical product of $\binom{n+1}{3}$ lossy phone call matrices in G_n.*

Proof. Induction on n. For $n = 1$ there are no factors. Let W_{n-1} be an irredundant expression over G_n of length $\binom{n}{3}$ not involving the index 1. Let P_h be the product

$$P_h = C_{12}(b_{h1}) \odot C_{23}(b_{h2}) \odot \cdots \odot C_{h,h+1}(b_{hh})$$

(of length h) and put

$$W_n = W_{n-1} \odot P_{n-1} \odot P_{n-2} \odot ... \odot P_1.$$

Then the expression for W_n has length $\binom{n+1}{3}$. Order the constants involved such that those in W_{n-1} are small, those in P_1 (just b_{11}) much larger, those in P_2 larger again, and those in P_{n-1} the largest. The matrix that is the result of multiplying out the expression W_n has (i, j)-entry as found for W_{n-1} when $i, j \neq 1$, but $(1, h+1)$-entry as found for P_h (since 1 is not found in W_{n-1}, $h+1$ is not found later than in P_h, and earlier P_j are too expensive). It follows that no factor of P_h is redundant. □

Proposition 2.6. *The closure of D_n under tropical matrix multiplication is the support of some finite polyhedral complex in $\mathbb{R}_{\geq 0}^{n\times n}$ and equals $G_n \cap \mathbb{R}_{\geq 0}^{n\times n}$. Its topological closure in $[0,\infty]^{n\times n}$ equals G_n.*

Note that this is Theorem 1.1 minus the (deepest) claim that the dimension of that complex is (not more than) $\binom{n}{2}$.

Proof. By Lemma 2.1 and the proof of Lemma 2.3 the closure of D_n under tropical matrix multiplication is a finite union of images of orthants $\mathbb{R}_{\geq 0}^{k}$ with $k \leq n(n-1)^2$ under piecewise linear maps. Such an image is the support of some polyhedral complex. The remaining two statements are straightforward. □

From now on, we will sometimes use the term "polyhedral fan" for the topological closure in $[0,\infty]^N$ of a polyhedral fan in $\mathbb{R}_{\geq 0}^{N}$. Thus G_n itself is a polyhedral fan in $[0,\infty]^{n\times n}$.

Recall that the *Kleene star* of $A \in [0,\infty]^{n\times n}$ is defined as

$$A^* := I \oplus A \oplus A^{\odot 2} \oplus \cdots = I \oplus A \oplus A^{\odot 2} \oplus \cdots \oplus A^{\odot(n-1)} = (I \oplus A)^{\odot(n-1)}$$

where I is the tropical identity matrix [But10, p. 21]. The (i, j)-entry of A^* records the length of the shortest path from i to j in the directed graph on $[n]$ with edge lengths a_{ij}. From this interpretation it follows readily that for $A_1, \ldots, A_s \in [0,\infty]^{n\times n}$ and $\pi \in \mathrm{Sym}(s)$ we have $(A_1 \odot \cdots \odot A_s)^* = (A_{\pi(1)} \odot \cdots \odot A_{\pi(s)})^*$.

Lemma 2.7. *The Kleene star maps G_n into its subset $\overline{D}_n$.*

Proof. Let $A \in G_n$ be the tropical product of lossy phone call matrices $C_1, \ldots, C_k$. Note that $C_i^\top = C_i$. We have

$$\begin{aligned} A^* &= (C_1 \odot \cdots \odot C_k)^* = (C_k \odot \cdots \odot C_1)^* = (C_k^\top \odot \cdots \odot C_1^\top)^* \\ &= ((C_1 \odot \cdots \odot C_k)^\top)^* = ((C_1 \odot \cdots \odot C_k)^*)^\top = (A^*)^\top, \end{aligned}$$

where we have used the remark above, the fact that transposition reverses multiplication order, and the fact that Kleene star commutes with transposition. Thus A^* is a symmetric Kleene star and hence a metric matrix. □

3. Graphs with detours

In the next two sections we will visualise elements of the lossy gossip monoids G_3 and G_4, as well as the polyhedral structures on these monoids. We will do this through combinatorial gadgets that we dub *graphs with detours*. We first recall realisations of ordinary metrics, i.e., elements of D_n (see, e.g., [Dre84, ISoPZ84]).

Let $\Gamma = (V, E)$ be a finite, undirected graph and $w : E \to \mathbb{R}_{\geq 0}$ be a function assigning lengths to the edges of Γ. The weight of a path in (Γ, w) is the sum of the weights of the individual edges in the path. A map $\ell : [n] \to V$ is called a *labelling*, or $[n]$*-labelling*, if we need to be precise, and the pair (Γ, ℓ) is referred to as a labelled graph, or an $[n]$-labelled graph.

A weighted $[n]$-labelled graph gives rise to a matrix $A(\Gamma, w, \ell)$ in D_n whose entry at position (i, j) is the minimal weight of a path between $\ell(i)$ and $\ell(j)$. We say that the weighted labelled graph (Γ, w, ℓ) *realises* the matrix $A(\Gamma, w, \ell)$. Any matrix $X \in D_n$ has a realisation by some weighted, $[n]$-labelled graph, e.g., the graph with vertex set $[n]$, the entries of X as weights, and ℓ equal to the identity. However, typically more efficient realisations exist, in the following sense. A weighted, $[n]$-labelled graph $(\Gamma = (V, E), w, \ell)$ is called an *optimal realisation* of X if the sum $\sum_e w(e)$ is minimal among all realisations [ISoPZ84]. We will, moreover, require of an optimal realisation that no edges get weight 0 (since such edges can be removed and their endpoints identified), and that no vertices in $V \setminus \ell([n])$ have valency 2 (since such vertices can be removed and their incident edges glued together). Optimal realisations of any $X \in D_n$ exist [ISoPZ84], and there is an interesting question concerning the uniqueness of optimal realisations for generic X [Dre84, Conjecture 3.20].

Our first step in describing the cones of G_3 and G_4 is to find weighted labelled graphs that realise the elements of D_3, D_4, as follows (for much more about this see [Dre84, DHLM06]). We write J_0 for the matrix of the appropriate size with all entries 0.

Example 3.1. We give optimal realisations of the elements of D_n, for $n = 2, 3, 4$. For the cases $n = 5, 6$ see [KLM09] and [SY04].

(1) An element of $D_2 \setminus \{J_0\}$ is optimally realised by the graph on two vertices having one edge with the right weight. The choice of labelling is inconsequential as long as it is injective. The matrix J_0 is optimally realised by the graph on one vertex.
(2) Any matrix in D_3 is realised by the top labelled graph of the poset depicted in Figure 2 with suitable edge weights (note that we allow these to be zero), but only the matrices in the relative interior of the cone D_3 are optimally realised by it. Matrices on the boundary are optimally realised by some graph further down the poset, depending on the smallest face of D_3 in which the matrix lies.
(3) The case of D_4 is similar to that of D_3 in the sense that there exists a single graph Γ which, appropriately labelled and weighted, realises any $X \in D_4$. However, unlike for D_3, three distinct labellings are required. The labelled graphs are depicted in Figure 3. For graphs in the relative interior of D_4, the given realisation is optimal (and in fact the unique optimal realisation).

We now extend realisation of metric matrices by graphs to realisations of arbitrary matrices in $\mathbb{R}_{\geq 0}^{n \times n}$ with zeroes on the diagonal. For this we need an extension of the concept of a labelled weighted graph. Let i and j be distinct elements of $[n]$. A *detour* from i to j in an $[n]$-labelled weighted graph is simply a walk p starting at $\ell(i)$ and ending at $\ell(j)$ that has larger total weight than the path of minimal weight between $\ell(i)$ and $\ell(j)$. Such a walk is allowed to traverse the same edge more than once, and even to turn around partway through an edge. The data specifying

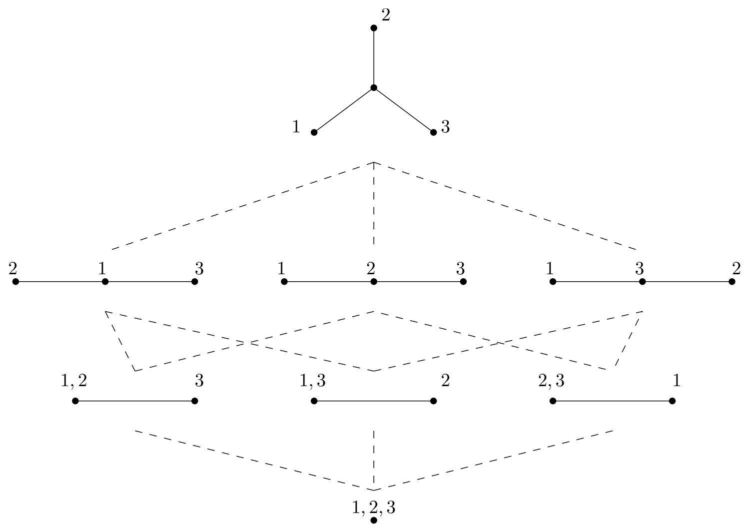


FIGURE 2. Minimal realisations of three-point metrics.

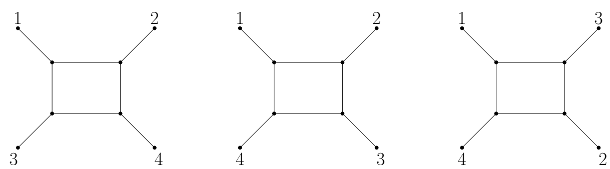


FIGURE 3. Minimal realisations of four-point metrics. The parallel sides of the middle rectangle have equal weight.

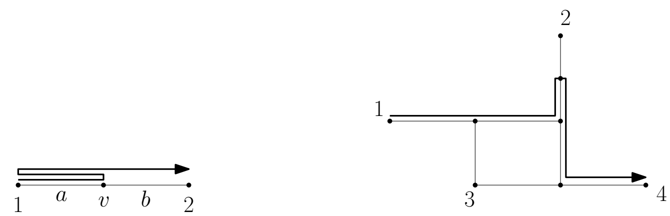
the detour is the triple (i, j, p). A *labelled weighted graph with detours* is a tuple consisting of a labelled weighted graph and a finite set of detours between distinct ordered pairs (i, j).

Let $(\Gamma, w, \ell, \mathcal{D})$ be an $[n]$-labelled weighted graph with set of detours $\mathcal{D}$. It gives rise to a matrix $A(\Gamma, w, \ell, \mathcal{D})$ whose entry at position (i, j) equals the weight of the detour from i to j, if there is any, or the weight of a path of minimal weight between i and j, if there is no detour between i and j in $\mathcal{D}$. In particular, $A(\Gamma, w, \ell, \mathcal{D})$ need not be symmetric, but its diagonal entries are 0. Again, if $X \in \mathbb{R}_{\geq 0}^{n\times n}$ and $X = A(\Gamma, w, \ell, \mathcal{D})$, then $(\Gamma, w, \ell, \mathcal{D})$ is said to realise X. Any non-negative matrix with zeroes on the diagonal is realised by some labelled weighted graph with detours. Observe also that replacing all detours (i, j, p) by the detours (j, i, p'), where p' is the opposite of p, corresponds to transposing the realised matrix.

Example 3.2. We give two examples of labelled weighted graphs with detours. First, the graph in Figure 4(a) has a single detour from 1 to 2, and realises the matrix

$$\begin{bmatrix} 0 & 3a+b \\ a+b & 0 \end{bmatrix}.$$

Except when $a = 0$, this matrix is not in G_2. Second, the graph in Figure 4(b) has a single detour from 1 to 4. By varying the edge lengths in $\mathbb{R}^6_{\geq 0}$, giving parallel sides of



(a) Path with a single detour from 1 to 2.

(b) Graph with a single detour from 1 to 4.

FIGURE 4. Examples of labelled weighted graphs with detours.

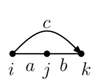
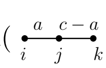
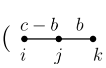
the rectangle the same length, this realises all matrices A in a six-dimensional cone, one of whose supporting equations is $a_{41} = a_{43} + a_{31}$ and one of whose bounding inequalities is $a_{14} \geq a_{41}$. This cone will turn out to be one of the maximal cones in G_4.

By Lemma 2.7, the Kleene star of a matrix A in G_n lies in $\overline{D}_n$. Thus it makes sense to look for a realisation of A by a labelled weighted graph with detours that, when forgetting the detours, realises A^*. This is what we will do in the next two sections for $n = 3$ and $n = 4$.

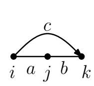
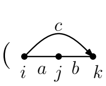
4. THREE GOSSIPERS

Since G_3 is a pointed fan, no combinatorial information is lost by intersecting that fan with a sphere centered around the all-zero metrix. The resulting spherical polyhedral complex is depicted in Figure 5. Detour graphs realising the maximal cones can be constructed by realising the arrows in an arbitrary manner as detours in the undirected graph. The middle cone is (the topological closure of) D_3, with its three codimension-one faces corresponding to the second layer in Figure 2 and its three codimension-two faces corresponding to the third layer.

The computations to show that Figure 5 gives all of G_3 are elementary and can be done by hand. We use pictorial notation and write $A(\Gamma)$ for the matrix realised by a labelled weighted graph with detours Γ. First, to prove that the matrices $A(\Gamma)$ with Γ as in the figure are indeed in G_3 we observe that

(1) $A($  $) = C_{jk}(b) \odot A($  $) = A($  $) \odot C_{ij}(a)$,

for any $c \geq a + b$ (and $a, b \geq 0$ as always). Together with the fact that $C_{ij}(a) \odot C_{ij}(d) = C_{ij}(a \oplus d)$ this implies that

$A($  $) \odot C_{ij}(d)$ and $C_{jk}(d) \odot A($  $)$

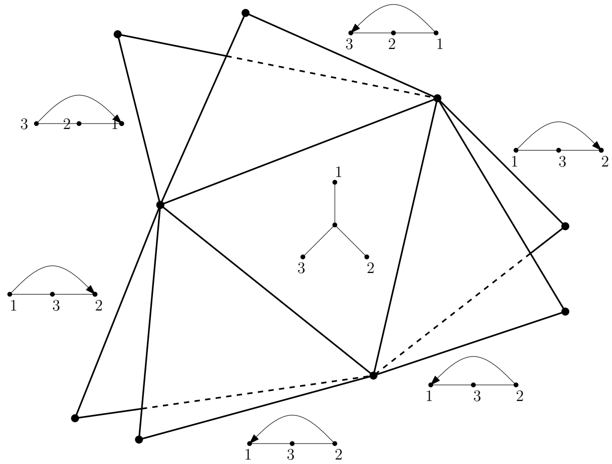


FIGURE 5. Representation of the spherical complex of G_3. The labelled graphs with detours corresponding to the maximal cells are indicated. The middle triangle represents the cone of distance matrices.

are contained in the complex of Figure 5 for all choices of a, b, c and d with $c \geq a+b$. Next we compute

$$C_{ij}(d) \odot A(\underset{i\;a\;j\;b\;k}{\overset{c}{\frown}}), \quad = \begin{cases} A(\underset{i\;a\;j\;b\;k}{\overset{c}{\frown}}), & c-b \leq d, \\ A(\underset{i\;a\;j\;b\;k}{\overset{b+d}{\frown}}), & a \leq d \leq c-b, \text{ and} \\ A(\underset{i\;d\;j\;b\;k}{\overset{a+b}{\frown}}), & 0 \leq d \leq a; \end{cases}$$

and, for $m := \max(a-b, b-a)$,

$$C_{ik}(d) \odot A(\underset{i\;a\;j\;b\;k}{\overset{c}{\frown}}) \quad = \begin{cases} A(\underset{i\;a\;j\;b\;k}{\overset{c}{\frown}}), & c \leq d, \\ A(\underset{i\;a\;j\;b\;k}{\overset{d}{\frown}}), & a+b \leq d \leq c, \\ A(\text{star: } i \tfrac{a+d-b}{2},\ k \tfrac{b+d-a}{2},\ j \tfrac{a+b-d}{2}), & m \leq d \leq a+b, \text{ and} \\ A(\underset{j\;a\;i\;d\;k}{\overset{b}{\frown}}), & 0 \leq d \leq m. \end{cases}$$

It follows by transposition that the products

$$A(\underset{i\;a\;j\;b\;k}{\overset{c}{\frown}}) \odot C_{ik}(d), \text{ and } A(\underset{i\;a\;j\;b\;k}{\overset{c}{\frown}}) \odot C_{jk}(d)$$

are also contained in one of the cones of Figure 5. This concludes the proof of Theorem 1.2 for $n = 3$.

5. Four gossipers

The computations for G_4 are too cumbersome to do by hand. Instead we used `Mathematica` to compute a fan structure on G_4. Figure 6 gives realising graphs with detours of all the cones of G_4, up to transposition and the action of Sym(4). The *surplus length* of a detour from i to j is defined as the difference between the length of the detour and the minimal distance between i and j in the graph. Two detours between i to j and k to l have the same color if their surplus lengths are equal.

These graphs were obtained as follows. First, generate all 6^6 possible piecewise linear affine maps $[0, \infty]^6 \to G_4$ of the form

$$(a_1, \ldots, a_6) \to C_{I_1}(a_1) \odot C_{I_2}(a_2) \odot \ldots \odot C_{I_6}(a_6),$$

where $I_1, \ldots, I_6$ are unordered pairs of distinct indices. Among the image cones, select only the six-dimensional ones, and compute their linear spans. There are 289 different linear spans. Compute the Sym(4)-orbits on these spans; this yields 16 orbits. Choose a representative for each of these orbits on spans, and for each representative select all cones with that span. It turns out that, for each representative span, one of the cones contains all other cones. To show that the orbits of these 16 maximal cones give all of G_4, left-multiply each of these 16 cones with all possible lossy phone call matrices and show that the resulting unions of cones are contained in the union of the 289 maximal cones; this is facilitated by the fact that each of these cones is the intersection of G_4 with (the topological closure in $[0, \infty]^{n \times n}$ of) a 6-dimensional subspace. This yields the statement about the unique coarsest fan structure in Theorem 1.2, as well as the numbers 16 and 289.

Next, the group $\mathbb{Z}/2\mathbb{Z}$ acts on G_4 by transposition. Taking orbit representatives under the larger group $\mathrm{Sym}(4) \times (\mathbb{Z}/2\mathbb{Z})$ from among the 16 yields 11 cones. Among these, 9 are simplicial (have six facets), the cone D_4 has 12 facets, and the remaining cone has 9 facets. The cone D_4 is the union of three simplicial cones (see Figure 3), which are permuted by Sym(4), so we need only one. This is C_5 in Figure 6. The cone with 9 facets turns out to be the union of two simplicial cones. Splitting this up yields C_{11} and C_{12} in the figure. It turns out that each C_i is the image of $\mathbb{R}^6_{\geq 0}$ under a linear map into $\mathbb{R}^{4\times 4}_{\geq 0}$ with non-negative integral entries with respect to the standard bases, and that these maps can be realised using weighted, labelled graphs with detours. These are the graphs in the picture. The graphs without the detours realise the Kleene star A^* with $A \in C_i$.

Finally, connectivity in codimension 1 is proved by Figure 7. It shows that any maximal cone is connected in codimension 1 to D_4; note the specified labelling. Most intersections in Figure 7 are of a simple type, where one of the edge weights becomes zero to go from one cone to the neighbouring cone; these contracted edges are then marked with an asterix on both sides. The only exception is the connection from C_7 to C_{12}. Although (suitable elements in the Sym(4)-orbits of) these cones intersect in a five-dimensional boundary cone, the boundary cone is obtained from the parametrizations specified by the graphs with detours by restricting the parametrization to a hyperplane where two of the weights are equal. This leads to the following theorem.

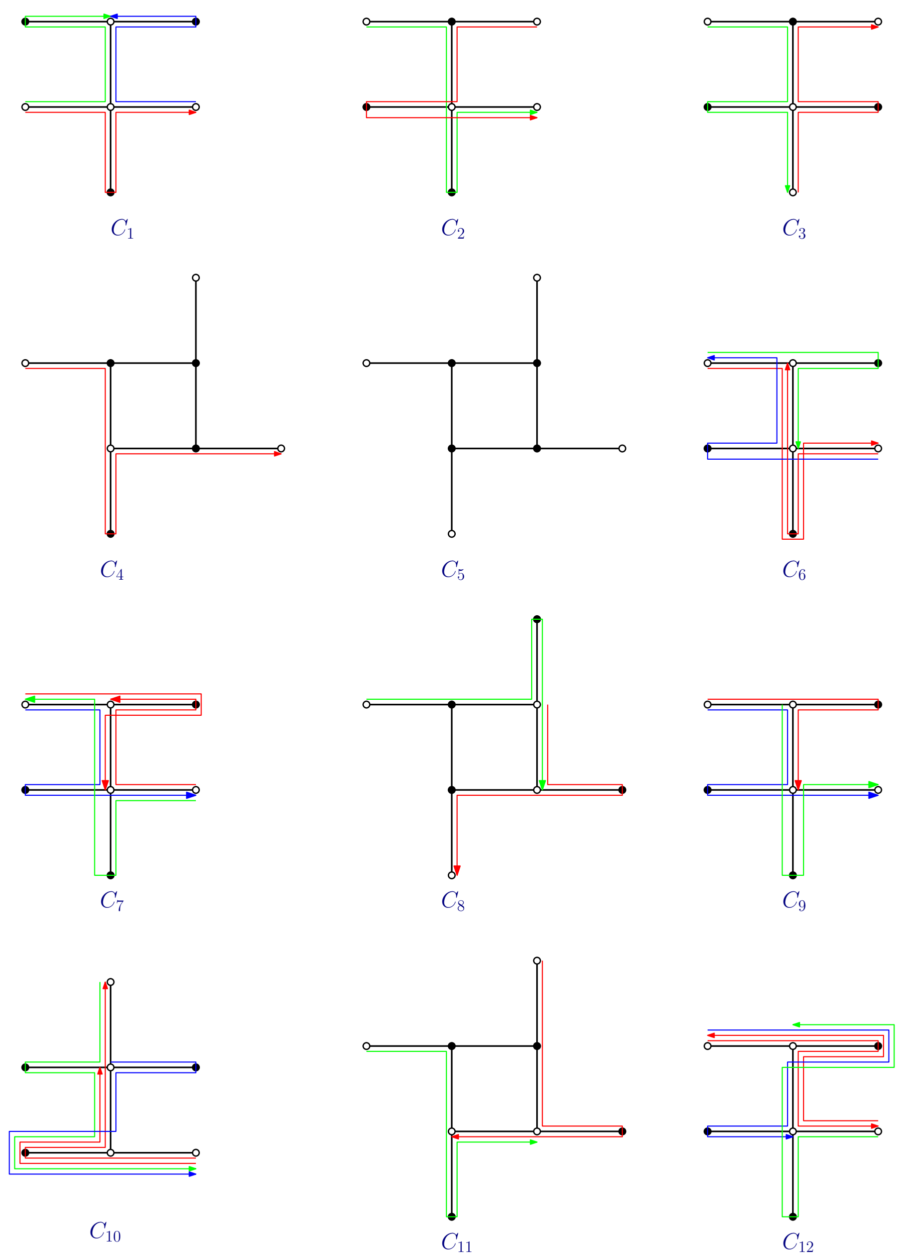


FIGURE 6. Orbit representatives of labelled weighted graphs with detours realising a polyhedral fan structure on G_4 with simplicial cones. The white vertices are the labelled vertices.

Theorem 5.1. *The cones realised by the graphs of Figure 6 give a polyhedral fan structure on* G_4*. This polyhedral fan is pure of dimension* 6 *and connected in codimension* 1*. Its intersection with a sphere around the origin is a simplicial*

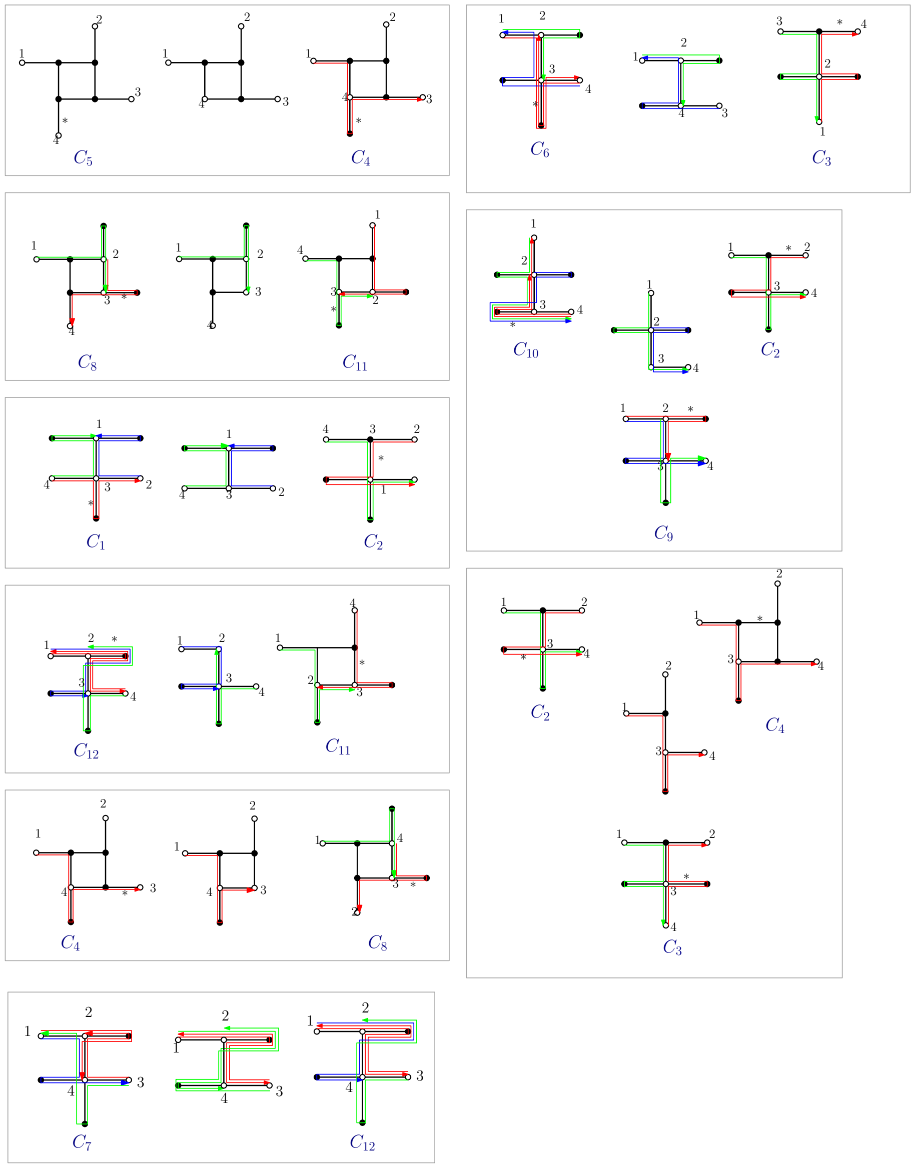


FIGURE 7. Walking from maximal cones to maximal cones by edge contraction, except in case C_7–C_{12}. The edge to be contracted is indicated by an asterix $*$. This shows that the cones in the grey boxes intersect in a cone of dimension 5. The intersection between C_7 and C_{12} is obtained by setting equal certain surplus lengths in the graphs representing C_7 and C_{12}.

spherical complex. Moreover, every element of G_4 is the product of (at most) 6 *lossy phone call matrices.*

n	# spans	# orbits	orbit size distribution
2	1	1	1×1
3	7	2	1×1, 1×6
4	289	16	1×1, 6×12, 9×24
5	91151	787	1×1, 2×20, 1×30, 48×60, 735×120

TABLE 1. Numbers of subspaces spanned by full-dimensional cones, and their numbers of orbits under $\mathrm{Sym}(n)$.

6. Five gossipers

Similar but somewhat more extensive computations establish the claimed facts about G_5. Every element has an expression as a tropical product of at most 10 lossy phone call matrices. The set G_5 has a polyhedral fan structure which is pure of dimension 10, and connected in codimension 1. The situation for $n = 5$ is more complicated than that for smaller n in that it is no longer true that the subspace spanned by a maximal polyhedral cone contains a unique maximal polyhedral cone. Some statistics is given in Table 1. Of course the single orbit of size 1 is that of D_n.

7. Dimension of the lossy gossip monoid

In the previous sections we have established Theorem 1.2 through explicit computations. We do not know of any systematic, combinatorial description of a polyhedral structure on G_n for larger n. However, we will now establish that G_n, which is the support set of *some* finite polyhedral fan by Lemma 2.3, has dimension $\binom{n}{2}$. Clearly, since G_n contains D_n, we have $\dim G_n \geq \dim D_n = \binom{n}{2}$. So the difficulty of Theorem 1.1 is in proving that its dimension does not exceed $\binom{n}{2}$.

Motivated by the analogy between the lossy phone call matrices $C_{ij}(a), a \in \mathbb{R}_{\geq 0}$ and one-parameter subgroups of algebraic groups (see Section 2), we introduce the one-parameter subgroups $g_{ij}(x)$ by

$$g_{ij}(x) := \begin{bmatrix} 1 & & & & \\ & \cos(x) & \cdots & -\sin(x) & \\ & \vdots & 1 & \vdots & \\ & \sin(x) & \cdots & \cos(x) & \\ & & & & 1 \end{bmatrix},$$

where the 1s stand for identity matrices, the cosines and sines are in the $\{i,j\} \times \{i,j\}$-submatrix, and the empty entries are 0. For any choice of x in the field $K = \mathbb{C}\{\{t\}\}$ of Puiseux series in t such that the order $\mathrm{val}(x)$ of x at zero is positive, the matrix $g_{ij}(x)$ is a well-defined matrix in the orthogonal group $\mathrm{O}_n(K)$, an algebraic subvariety of $K^{n\times n}$. Recall that the *tropicalisation* of $\mathrm{O}_n(K)$ is the topological closure of the image of $\mathrm{O}_n(K)$ in $(-\infty,\infty]^{n\times n}$ under the component-wise valuation map $\mathrm{val} : K^{n\times n} \to (\mathbb{Q} \cup \{\infty\})^{n\times n}$.

Proposition 7.1. *The lossy gossip monoid G_n is contained in the tropicalisation of $\mathrm{O}_n(K)$.*

Proof. First note that $\mathrm{val}(g_{ij}(x)) = C_{ij}(\mathrm{val}(x))$, so the statement would be immediate if we knew that the tropicalisation of $\mathrm{O}_n(K)$ were closed under tropical

n	$\lvert G_n(\{0,\infty\})\rvert$	max. length
1	1	0
2	2	1
3	11	3
4	189	4
5	9152	6
6	1,092,473	10
7	293,656,554	13
8	166,244,338,221	16
9	188,620,758,836,916	19

TABLE 2. Sizes and maximal lengths of $G_n(\{0,\infty\})$, for $n = 1, \ldots, 9$.

matrix multiplication. We do not know whether this is the case, but in fact we need only something weaker. Let $a_1, \ldots, a_k$ be strictly positive rational numbers and let $(i_1, j_1), \ldots, (i_k, j_k)$ be pairs of distinct indices. Then for a vector $(c_1, \ldots, c_k) \in \mathbb{C}^k$ outside some proper hypersurface, no cancellation takes place in the expression

$$g_{i_1,j_1}(c_1 t^{a_1}) \cdots g_{i_k,j_k}(c_k t^{a_k}),$$

in the sense that

$$\mathrm{val}[g_{i_1,j_1}(c_1 t^{a_1}) \cdots g_{i_k,j_k}(c_k t^{a_k})] = \mathrm{val}[g_{i_1,j_1}(c_1 t^{a_1})] \odot \cdots \odot \mathrm{val}[g_{i_k,j_k}(c_k t^{a_k})],$$

which equals $C_{i_1,j_1}(a_1) \odot \cdots \odot C_{i_k,j_k}(a_k)$. This shows that the latter expression is contained in the tropicalisation of the orthogonal group. Since that tropicalisation is closed in the Euclidean topology, all of G_n is contained in it. □

The dimension claim in Theorem 1.1 now follows from the Bieri-Groves theorem [BG84], which says that the tropicalisation of a variety has dimension equal to that of the variety—indeed, the dimension of the orthogonal group is $\binom{n}{2}$.

8. Ordinary gossip

In this section we study the *ordinary gossip monoid* $G_n(\{0,\infty\})$, which is the submonoid of G_n of matrices with entries in $\{0,\infty\}$. Note that there is a surjective homorphism $G_n \to G_n(\{0,\infty\})$ mapping non-∞ entries to 0 and ∞ to ∞, which shows that the length of an element of $G_n(\{0,\infty\})$ inside G_n is the same as the minimal number of non-lossy phone calls $C_{ij}(0)$ needed to express it. A classical result says that length of the all-zero matrix is exactly 1 for $n = 2$, 3 for $n = 3$, and $2n - 4$ for $n \geq 4$ [BS72, Bum81, HMS72, Tij71], and this result spurred a lot of further activity on gossip networks. But the all-zero matrix does not necessarily have the largest possible length—see Table 2, which records sizes and maximal element lengths for $G_n(\{0,\infty\})$ with $n \leq 9$. The first 8 rows were computed by former Eindhoven Master's student Jochem Berndsen [Ber12].

Proof of Theorem 1.4. Consider n ladies, each with a different gossip item. They communicate by telephone, and whenever two ladies talk, each tells the other all she knows. We will determine the maximal length of a sequence of calls, when in each call at least one participant learns something new. The answer turns out to be $n(n-1)/2$.

That $n(n-1)/2$ is a lower bound, is shown by the following scenario: Let the participants be $A_1, \ldots, A_n$. All calls involve A_1. For $i = 2, \ldots, n$ she calls $A_i, A_{i-1}, \ldots, A_2$, for a total of $1 + 2 + \cdots + (n-1) = n(n-1)/2$ calls. Of course there are many other scenarios.

We now argue that $n(n-1)/2$ is an upper bound. Since each of the n participants must learn $n-1$ items, and at least one item is learned on each call, there are at most $n(n-1)$ calls. If during a call $1+e$ items of gossip are transmitted, we say that there were e extra items. We will show that in all calls together at least $n(n-1)/2$ extra items occur, by pointing at (at least) one extra item for each unordered pair A, B of participants. Then the total number of calls is at most $n(n-1) - n(n-1)/2 = n(n-1)/2$.

We shall assign a call to each ordered pair (A, B) of participants in such a way that a call with e extra items is assigned to (A, B) for at most e unordered pairs A, B. If A, B start out knowing a, b, respectively, the call assigned to (A, B) will be one during which a and b are exchanged, or be the one where B first learns a. If a and b are exchanged during the call where B first learns a, this will be the assigned call. If a and b are not exchanged during the call assigned to (A, B), then a different call is assigned to (B, A), and we find a sharper bound.

Number the calls in time order. There is a canonical call where a and b are exchanged: At the end B knows both a and b. Trace calls backward from B to uniquely find a strictly increasing series of calls numbered $i_1, ..., i_h$ and a maximal series of ladies X_0, X_1, ..., $X_h = B$, such that call i_j is between ladies X_{j-1} and X_j, where X_j knows both a and b at the end of the call, but did not know both at the start of the call. Now at call i_1 the ladies X_0 and X_1 exchanged a and b. The call assigned to the pair (A, B) will be either i_1 or i_h. The series of calls $i_1, ..., i_h$ will be called the canonical path of a to B.

We shall compare paths $i_1, ..., i_h$ in reverse lexicographic order (revlex), that is lexicographic order on the reversed paths $i_h, ..., i_1$. For two paths where one is a tail of the other, the smaller path is the one that starts with the smallest element, i.e., the one that is longest.

Let R be the set of items that B learned during call number i. Let a_0 be some element of R (to be fixed later) that revlex minimizes the canonical path to B. Regard all items $a \in R \setminus \{a_0\}$ as extra on call i. This assigns an extra item to the pair A, B whenever the canonical path of a to B is not revlex minimal among the paths with the same final element.

Suppose $i_1, ..., i_h = i$ is the revlex minimal one among the canonical paths to B ending in i, where this is the canonical path of a to B, and let i_1 be a call between P and Q, where P is the participant learning b. If $p \in R$, then $i_1, ..., i_h$ is the canonical path of p to B. Now pick $a_0 = p$, and assign the pair (P, B) to call i_1 as promised. If $p \notin R$ then pick $a_0 \in R$ arbitrarily, and assign the pair (A_0, B) to call i_1. (In this case the canonical path for (P, B) is revlex smaller than $i_1, ..., i_h$, and does not use the call i_1. The pair (P, B) will be assigned somewhere on that path.)

Let Q, P learn the sets of items S, T of sizes s, t (respectively) during call i_1. Then $s + t$ items were transmitted, so $s + t - 1$ extra items, and $s, t \geq 1$. We are allowed to assign $s + t - 1$ pairs to call i_1. If that call exchanges p and q, then a possible assignment is (P, B) for all $b \in T$ and (A, Q) for all $a \in S$. If we assign (A, B) then we do not assign (P, B), so that the total remains ok. If call i_1 does not exchange p and q, for example because $p \notin S$, then for every $b \in T$ we pick a unique $a \in S$, and at most t pairs are assigned to call i_1, again ok. □

We do not know whether there are elements in the ordinary gossip monoid of length $\binom{n}{2}$.

9. Open questions

In view of the extensive computations in Sections 4–6 and the rather indirect dimension argument in Section 7, the most urgent challenge concerning the lossy gossip monoid is the following.

Question 9.1. *Find a purely combinatorial description of a polyhedral fan structure with support G_n. Use this description to prove or disprove the pureness of dimension $\binom{n}{2}$ and the connectedness in codimension one.*

A related question, motivated on the one hand by the fact that G_n has dimension $\binom{n}{2}$ and on the other hand by Theorem 1.4, is the following.

Question 9.2. *Is the length of any element of G_n at most $\binom{n}{2}$?*

Finally, once a satisfactory polyhedral complex for G_n is found, the somewhat ad-hoc graphs in Sections 4 and 5 lead to the following challenge.

Question 9.3. *Find a useful notion of optimal realisations of elements of G_n by graphs with detours, and a notion of* tight spans *of such elements.*

References

[Ber12] Jochem Berndsen, *Three problems in algebraic combinatorics*, Master's thesis, Eindhoven University of Technology, 2012. `http://alexandria.tue.nl/extra1/afstversl/wsk-i/berndsen2012.pdf` .

[BG84] Robert Bieri and John R. J. Groves, *The geometry of the set of characters induced by valuations*, J. reine angew. Math. **347** (1984) 168–195.

[BH86] Kenneth A. Berman and Michael Hawrylycz, *Telephone problems with failures*, SIAM J. Algebraic Discrete Methods **7** (1986) 13–17.

[BS72] Brenda Baker and Robert Shostak, *Gossips and telephones*, Discrete Math. **2** (1972) 191–193.

[Bum81] Richard T. Bumby, *A problem with telephones*, SIAM J. Algebraic Discrete Methods **2** (1981) 13–18.

[But10] Peter Butkovič, *Max-linear systems, Theory and algorithms*, Springer Monographs in Mathematics, Springer, London, 2010.

[DHLM06] Andreas Dress, Katharina T. Huber, Alice Lesser, and Vincent Moulton. *Hereditarily optimal realizations of consistent metrics*, Ann. Comb. **10** (2006) 63–76.

[Dre84] Andreas W. M. Dress, *Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: A note on combinatorial properties of metric spaces*, Adv. Math. **53** (1984) 321–402.

[HMS72] András Hajnal, Eric C. Milner, and Endre Szemerédi, *A cure for the telephone disease*, Can. Math. Bull. **15** (1972) 447–450.

[HRS87] Ramsey W. Haddad, Shaibal Roy, and Alejandro A. Schäffer, *On gossiping with faulty telephone lines*, SIAM J. Algebraic Discrete Methods **8** (1987) 439–445.

[IJK12] Zur Izhakian, Marianne Johnson, and Mark Kambites, *Tropical matrix groups*, preprint, 2012. `http://arxiv.org/abs/1203.2449` .

[ISoPZ84] Wilfried Imrich, J. M. S. Simões Pereira, and Christina M. Zamfirescu, *On optimal embeddings of metrics in graphs*, J. Combin. Theory Ser. B **36** (1984) 1–15.

[KLM09] Jack Koolen, Alice Lesser, and Vincent Moulton, *Optimal realizations of generic five-point metrics*, Eur. J. Comb. **30** (2009) 1164–1171.

[SY04] Bernd Sturmfels and Josephine Yu, *Classification of six-point metrics*, Electron. J. Comb. **11** (2004) R44.

[Tij71] R. Tijdeman, *On a telephone problem*, Nieuw Arch. Wiskd., III. Ser. **19** (1971) 188–192.

Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

E-mail address: aeb@cwi.nl, j.draisma@tue.nl, b.j.frenk@tue.nl