Dilworth rate: a generalization of Witsenhausen's zero-error rate for directed graphs

Gábor Simonyi and Ágnes Tóth

***Abstract*—We investigate a communication setup where a source output is sent through a free noisy channel first and an additional codeword is sent through a noiseless but expensive channel later. With the help of the second message the decoder should be able to decide with zero-error whether its decoding of the first message was error-free. This scenario leads to the definition of a digraph parameter that generalizes Witsenhausen's zero-error rate for directed graphs. We investigate this new parameter for some specific directed graphs and explore its relations to other digraph parameters like Sperner capacity and dichromatic number.**

We also look at the natural variant of the above problem, where the decoder should decode the first message with zero-error, not only decide whether its earlier decoding was correct. In this case the Witsenhausen rate of an appropriately defined undirected graph turns out to be the relevant parameter.

***Index Terms*—zero-error, graph products, Sperner capacity, dichromatic number, Witsenhausen rate**

I. Introduction

CONSIDER the following situation. Alice writes a message to Bob consisting of the numbers of several bank accounts to which Bob has to send some money. She writes in a hurry (she just got to know that the transfers are urgent if they do not want to pay delay punishment, but currently she has little time). Therefore her characters are not very well legible, so Bob may misread some numbers. However, there are some rules for the possible mistakes, e.g., a 7 may be thought to be a 1 but never a 6. This relation between the possible digits need not be symmetric: it is possible that a 0 is sometimes read as a 6 but a 6 may not be decoded as a 0. These rules of possible confusions are known both by Alice and by Bob.

As Alice is aware of the possibility that Bob misread her message, later in the day she sends another message to Bob, the goal of which is to make Bob certain whether he read (decoded) the first message correctly or not. If he did he can transfer the money with complete confidence that he sends it to the right accounts. If he did not he will know that he does not know the account numbers correctly and so he better wait and pay the punishment than transfer the money to the wrong place.

G. Simonyi and Á. Tóth are with the Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, Hungary, e-mail: simonyi.gabor@renyi.mta.hu, toth.agnes@renyi.mta.hu;
Research of the first author is partially supported by the Hungarian Foundation for Scientific Research, Grants K104343 and K105840; Research of the second author is partially supported by the Hungarian Foundation for Scientific Research, Grants K104343 and K108947. This paper was presented in part at the 2014 IEEE International Symposium on Information Theory.

The second message will be received by Bob correctly for certain, but it uses an expensive device, e.g., Alice sends it as an sms from another country after she has arrived there. (Now we understand why she was in a hurry: she had to arrive to the airport in time.) For some reason, every character sent from this foreign country costs a significant amount of money for her. So she wants to send the shortest possible message that makes it sure (here we insist on zero-error) that Bob will know whether his decoding of the original handwritten message was error-free or not. The problem is to determine the best rate of communication over the second channel as the length of the original message received tends to infinity.

In Section II we describe the abstract communication model for this scenario and show that the best achievable rate is a parameter of an appropriate directed graph. We will see that this parameter of a directed graph is a generalization of the parameter (of an undirected graph) called Witsenhausen rate.

In Section III we investigate the relationship with other graph parameters. These include Sperner capacity and the dichromatic number. The former is a generalization of Shannon's graph capacity [33] to directed graphs. Though originally defined to give a general framework for some problems in extremal set theory (see [16], [17]), Sperner capacity also has its own information theoretic relevance, see [12], [30], [7]. The dichromatic number is a generalization of the chromatic number to directed graphs introduced in [29]. Using the above mentioned relations we determine our new parameter for some specific directed graphs.

In Section IV we consider a compound channel type version of the problem parallel to [30], [36].

In Section V some connections to extremal set theory are pointed out.

In Section VI we will consider the setup where the requirement is more ambitious and we want that Alice's second (the error-free but expensive) message make Bob able to decode the original message with zero-error. (That is, he will know the message itself not only the correctness or incorrectness of his original decoding of Alice's handwriting.) We will see that this setting leads to the Witsenhausen rate of an undirected graph related to the problem. This gives a new interpretation of Witsenhausen rate.

The natural logarithm of a number x will be denoted by $\ln x$. All other logarithms are meant to be of base 2.

II. The Dilworth rate of a directed graph

A. The communication model

The abstract setting for our communication scenario is the following. We have a source whose output is sent through a noisy channel. (This belongs to Alice's handwriting.) The input and output alphabets of this channel are identical and they coincide with the output alphabet of the source. It is known how the noisy channel can deform the input, in particular we know what (input) letters can become a certain, possibly different (output) letter on the other side. (We always assume though that every letter can result in itself, that is get through the noisy channel without alteration. At the same time, there is no other limitation than the length on how many characters of a sequence can get changed.) Later another message is sent (by the same sender) to the same receiver. This second message is sent via a noiseless channel and its goal is to make zero-undetected-error decoding possible, i.e., after having received this second message the receiver should be able to decide whether it decoded the first message correctly. The use of the noiseless channel is expensive, so the second message should be kept as short as possible.

Let the shortest possible message that satisfies the criteria have length $h(t)$ when t characters of the source output are encoded together. (Though we will let t going to infinity, for finite t we always assume that its value is known to Alice and Bob, and thus may affect their coding strategy.) Let H denote the noisy channel. The efficiency of the communication is measured by the quantity

$$R_{\mathrm{D}}(H) := \liminf_{t\to\infty} \frac{h(t)}{t}$$

that we call the Dilworth rate of the noisy channel H. (For an explanation of the name see Remark 5 in Subsection II-B.)

Remark 1: Note the special feature of the problem that we characterize a channel by a rate, that is with a parameter that, unlike channel capacity, we want to be as small as possible. The reason is that we measure the reliability of a channel not by the amount of information it can safely transfer but with the amount of information needed *to be added* for making the communication reliable. □

B. Dilworth rate and Witsenhausen rate

The relevant properties of H are described by a directed graph $\vec{G}_H$ having the (common input and output) alphabet as its vertex set and the following edge set. An ordered pair (a, b) of two letters forms a directed edge of $\vec{G}_H$ if and only if $b \neq a$ and the output of H can be b when it is fed by a at the input.

Remark 2: As usual we will use $V(\vec{F})$ to denote the vertex set and $E(\vec{F})$ to denote the edge set of a directed graph $\vec{F}$. We will use similar notation for undirected graphs that we always consider to be the same as a *symmetrically directed graph*. In such a graph an ordered pair (u, v) of two vertices forms a directed edge if and only if the reversely ordered pair (v, u) is also present in the digraph as a directed edge. We will use the term *oriented graph* for directed graphs that do not contain any edge together with its reversed version. That is $\vec{F}$ is an oriented graph if $(u, v) \in E(\vec{F})$ implies $(v, u) \notin E(\vec{F})$. As it is also customary, the term *digraph* will be used as a synonym for "directed graph".
When we say underlying undirected graph of a directed graph it means the undirected graph we obtain on the same vertex set if we connect by an undirected edge exactly those vertices that are adjacent in the directed graph (in either or both directions). □

To express R_{D} as a graph parameter we need the following notion.

Definition 1: The AND product $\vec{F} \wedge \vec{G}$ of two directed graphs $\vec{F}$ and $\vec{G}$ is defined as follows. The vertex set of $\vec{F} \wedge \vec{G}$ is the direct product $V(\vec{F}) \times V(\vec{G})$ and vertex (f, g) sends a directed edge to (f', g') iff either $(f, f') \in E(\vec{F})$ and $(g, g') \in E(\vec{G})$ or $(f, f') \in E(\vec{F})$ and $g = g'$ or $f = f'$ and $(g, g') \in E(\vec{G})$. The t-th AND power of a digraph $\vec{G}$, denoted by $\vec{G}^{\wedge t}$ is the t-wise AND product of digraph $\vec{G}$ with itself.

Observe that this graph exponentiation extends to sequences of letters the relation between individual letters f and f' expressing that feeding f to the noisy channel H may result in observing letter f' at the output. A sequence of letters at the input of H can result in another such sequence at the output of H if at each coordinate the character in the first sequence can result in the corresponding character of the second sequence. (This includes the possibility that the character does not change when sent through H.)

Remark 3: The terminology of graph products is not completely standardized. The AND product we just defined is also called normal product [4], strong direct product [26], or strong product [18]. We follow the paper of Alon and Orlitsky [3] when use the name AND product, because we find this name informative. A similar remark applies to the OR product that we will introduce later in Definition 3. □

Recall that the chromatic number $\chi(F)$ of a graph F is the minimal number of colors that suffice to color the vertices of F so that adjacent vertices get different colors. If $\vec{F}$ is a digraph, its chromatic number $\chi(\vec{F})$ is understood to be the chromatic number of the underlying undirected graph.

Proposition 1:

$$R_{\mathrm{D}}(H) = \lim_{t\to\infty} \frac{1}{t} \log \chi(\vec{G}_H^{\wedge t}).$$

Remark 4: It is easy to see that the above limit always exists. (The reason is the submultiplicative behaviour of the chromatic number under the AND product). □

Proof: Alice and Bob can agree in advance in a proper coloring of $\vec{G}_H^{\wedge t}$ with $\chi(\vec{G}_H^{\wedge t})$ colors. Alice can send Bob the color of the vertex belonging to the original source output using $\lceil \log \chi(\vec{G}_H^{\wedge t}) \rceil$ bits. Bob compares this to the color of the vertex representing the sequence he obtained. If the latter color is identical to the one Alice has sent him, then he can be sure that his decoding was error-free. This is because any other sequence that could result in his decoded sequence is adjacent (in $\vec{G}_H^{\wedge t}$) to this decoded sequence, so its color is different.

On the other hand, if Alice sent a shorter message through the noiseless channel, then she could not have $\chi(\vec{G}_H^{\wedge t})$ distinct messages and thus there must exist two adjacent vertices in

$\vec{G}_H^{\wedge t}$ that are encoded to the same codeword by Alice (for the noiseless channel). Then one of the two sequences represented by these two adjacent vertices could result in the other one, while this other one could also result in itself. Thus Bob cannot make the difference between these two sequences, one of which is the correct source output sequence while the other one differs from it. So receiveing this message Bob could not be sure whether his decoding was error-free or not. ∎

The right hand side expression in Proposition 1 can be considered as a digraph parameter that we will call the Dilworth rate of the digraph $\vec{G}_H$.

Definition 2: For a directed graph $\vec{G}$ we define its (logarithmic) *Dilworth rate* to be

$$R_{\mathrm{D}}(\vec{G}) := \lim_{t\to\infty} \frac{1}{t} \log \chi(\vec{G}^{\wedge t}).$$

The non-logarithmic Dilworth rate is

$$r_{\mathrm{D}}(\vec{G}) := \lim_{t\to\infty} \sqrt[t]{\chi(\vec{G}^{\wedge t})}.$$

Obviously, $R_{\mathrm{D}}(\vec{G}) = \log r_{\mathrm{D}}(\vec{G})$.

Remark 5: Let $\vec{L}$ be the directed graph on 2 vertices with a single directed edge. If we consider the vertices of $\vec{L}^{\wedge t}$ as characteristic vectors of subsets of a t-element set then $R_{\mathrm{D}}(\vec{L})$ can be interpreted as the asymptotic exponent of the minimum number of antichains (sets of pairwise incomparable elements) in the Boolean lattice of these subsets that can cover all the subsets. (This correspondence becomes clear by realizing that antichains of the Boolean lattice belong to sets of pairwise non-adjacent points of $\vec{L}^{\wedge t}$ and thus the minimum number of antichains in said covering is just the chromatic number of $\vec{L}^{\wedge t}$.) The exact value of this minimum number (which is $t+1$), is given (easily) by a special case of what is called the "dual of Dilworth's theorem" [13]. (This theorem is also called Mirsky's theorem, see [28], and it states that the minumum number of antichains covering a partially ordered set is equal to the length of its longest chain. Dilworth's theorem is the analogous result with the role of chains and antichains exchanged.) This connection to Dilworth's celebrated result is the reason for calling our new parameter Dilworth rate. Note that the name Sperner capacity was picked by the authors of [16] for analoguous reasons: the Sperner capacity of the digraph $\vec{L}$ has a similar relationship with Sperner's theorem [37]. □

The AND product is also defined for undirected graphs. Considering undirected graphs as symmetrically directed graphs the definition is straightforward.

Witsenhausen considered the "zero-error side-information problem" that led him to introduce the quantity

$$R_{\mathrm{W}}(G) = \lim_{t\to\infty} \frac{1}{t} \log \chi(G^{\wedge t})$$

that is called the Witsenhausen rate of (the undirected) graph G.

It is straightforward from the definitions that if $\vec{G}$ is a symmetrically directed graph and G is the underlying undirected graph (that we consider equivalent), then $R_{\mathrm{W}}(G) = R_{\mathrm{D}}(\vec{G})$. Thus Dilworth rate is indeed a generalization of Witsenhausen rate to directed graphs.

Remark 6: We note that Nayak and Rose [30] define what they call "the Witsenhausen rate of a set of directed graphs". Though formally this gives the Dilworth rate of a directed graph, the focus of [30] is elsewhere. When its motivating setup results in a family consisting of a single digraph, then this digraph is symmetrically directed. (See also Theorem 20 in Section IV.) □

III. Bounds on the Dilworth rate

A. Relation to Sperner capacity and a lower bound

Sperner capacity was introduced by Gargano, Körner and Vaccaro [16]. Traditionally this parameter is defined by using the OR product.

Definition 3: The OR product $\vec{F} \vee \vec{G}$ of directed graphs $\vec{F}$ and $\vec{G}$ has vertex set $V(\vec{F}) \times V(\vec{G})$ and (f,g) sends a directed edge to (f',g') iff either $(f,f') \in E(F)$ or $(g,g') \in E(\vec{G})$. The t-th OR power $\vec{G}^{\vee t}$ is the t-wise OR product of digraph $\vec{G}$ with itself.

Let $\vec{K}_n$ denote the complete directed graph on n vertices, that is the one we obtain from a(n undirected) complete graph K_n when substituting each of its edges $\{a,b\}$ by the two oriented edges (a,b) and (b,a). The (directed) complement of a digraph $\vec{G}$ is the directed graph $\vec{G}^c$ on vertex set $V(\vec{G})$ having edge set $E(\vec{G}^c) = E(\vec{K}_n) \setminus E(\vec{G})$.

Now we note the straightforward relation of the AND and OR powers that $(\vec{G}^{\vee t})^c = (\vec{G}^c)^{\wedge t}$.

The (logarithmic) Sperner capacity of digraph $\vec{G}$ is defined (see [16], [17]) as

$$\Sigma(\vec{G}) := \lim_{t\to\infty} \frac{1}{t} \log \omega_s(\vec{G}^{\vee t}),$$

where $\omega_s(\vec{F})$ denotes the *symmetric clique number*, that is the cardinality of the largest symmetric clique in digraph $\vec{F}$: the size of the largest set $U \subseteq V(\vec{F})$ where for each $f, f' \in U$ both (f,f') and (f',f) are edges of $\vec{F}$.

Using the above relation of the AND and OR products, Sperner capacity (of the complementary graph $\vec{G}^c$) can also be defined as

$$\Gamma(\vec{G}) := \Sigma(\vec{G}^c) = \lim_{t\to\infty} \frac{1}{t} \log \alpha(\vec{G}^{\wedge t}),$$

where $\alpha(\vec{F})$ stands for the independence number (size of the largest edgeless subset of the vertex set) of graph $\vec{F}$. This is the definition given in [7]. (The authors of [7] call this value the Sperner capacity of $\vec{G}$.)

When G is an undirected (or symmetrically directed) graph, then $\Gamma(G) = C(G)$, the *Shannon capacity* of graph G (see [33]). (Recall, that Shannon capacity is the zero-error capacity of a channel where the pairwise confusability or distinguishability of the input letters is described by the graph the Shannon capacity of which we are talking about. Its formal definition is just the above formula whith orientations ignored.) As already mentioned in the Introduction, the main motivation for introducing Sperner capacity came from extremal set theory, cf. [17], it turned out that it has its own relevance (that is not only that via Shannon capacity) in information theory, cf. [12], [30], [7].

We will need a sort of probabilistic refinement of our capacity-like parameters called their "within-a-type" versions, see [10]. First we need the concept of (P,ε)-typical sequences, cf. [11].

Definition 4: Let V be a finite set. The *type* of a sequence $\boldsymbol{x}$ in V^t is the probability distribution $P_{\boldsymbol{x}}$ on V defined by $P_{\boldsymbol{x}}(a) = \frac{1}{t}N(a|\boldsymbol{x})$ for every $a \in V$, where $N(a|\boldsymbol{x}) = |\{i : x_i = a\}|$. Given a probability distribution P on V, and $\varepsilon > 0$, a sequence $\boldsymbol{x}$ in V^t is said to be (P,ε)-typical if for every $a \in V$ we have $|P_{\boldsymbol{x}}(a) - P(a)| \leq \varepsilon$. We denote the set of (P,ε)-typical sequences in V^t by $\mathcal{T}^t(P,\varepsilon)$. When $\varepsilon = 0$ we also write $\mathcal{T}_P^t$ for $\mathcal{T}^t(P,0)$.

Let $\vec{G}^{\odot t}$ stand for either $\vec{G}^{\wedge t}$ or $\vec{G}^{\vee t}$. For a directed (or undirected) graph $\vec{F}$ and $U \subseteq V(\vec{F})$ we denote by $\vec{F}[U]$ the digraph induced by $\vec{F}$ on the subset U of the vertex set. We also use the shorthand notation $\vec{F}^{\odot t}_{P,\varepsilon} = \vec{F}^{\odot t}[\mathcal{T}^t(P,\varepsilon)]$.

Let $\beta(\vec{G})$ be either of the following graph parameters of the directed graph $\vec{G}$: independence number, clique number, chromatic number, clique cover number (which is the chromatic number of the complementary graph), symmetric clique number, or transitive clique number. (The latter is the size of the largest subset U of $V(\vec{G})$ the elements of which can be linearly ordered so that if u precedes v then the oriented edge (u,v) is present in $E(\vec{G})$.)

Let the asymptotic parameter $Z(\vec{G})$ be defined as

$$Z(\vec{G}) := \limsup_{t\to\infty} \frac{1}{t} \log \beta(\vec{G}^{\odot t}).$$

Definition 5: The parameter $Z(\vec{G},P)$ of a digraph $\vec{G}$ within a given type P is the value

$$Z(\vec{G},P) = \lim_{\varepsilon\to 0} \limsup_{t\to\infty} \frac{1}{t} \log \beta(\vec{G}^{\odot t}_{P,\varepsilon}).$$

We note that for several of the allowed choices of $\beta(\vec{G})$ and $\vec{G}^{\odot t}$ we obtain a graph parameter that already exists in the literature. For example, when $\beta(\vec{G}) = \omega_s(\vec{G})$ and the power we look at is the OR power, we get Sperner capacity within a given type, that has an important role in the main results of the papers [16], [17].

If we choose $\beta(\vec{G}) = \chi(\vec{G})$ and the OR power, we obtain the functional called graph entropy, which is defined in [20] and has several nice properties, see [34], [35], as well as important applications, see e.g. [19]. When $\beta(\vec{G}) = \chi(\vec{G})$ but the exponentiation is the AND power, then we arrive to the within a type version of Dilworth rate $R_{\mathrm{D}}(\vec{G},P)$. The special case of this for an undirected graph G was already known under name "complementary graph entropy" that could justifiably be called "Witsenhausen rate within a given type". This parameter was introduced by Körner and Longo [22] and further investigated by Marton [27]. Although this within-a-type version of Witsenhausen's invariant was introduced earlier than the non-probabilistic version (cf. [22], [38]), for the sake of consistancy we denote it by $R_W(G,P)$.

Note that Marton [27] proved the important identity

$$R_W(G,P) + C(G^c,P) = H(P),$$

where $H(P) = -\Sigma_{i=1}^n p_i \log p_i$ is the entropy of the probability distribution $P = (p_1,\ldots,p_n)$. This holds for any probability distribution P on $V(G)$. Along the same lines one can also prove the following theorem, the proof of which we will give for the sake of completeness.

Theorem 2: Let $\vec{G}$ be a directed graph and P an arbitrary fixed probability distribution on $V(\vec{G})$. Then

$$R_{\mathrm{D}}(\vec{G},P) + \Gamma(\vec{G},P) = H(P).$$

We will use the notion of fractional chromatic number $\chi_f(G)$ in the proof. Let $S(G)$ denote the set of independent sets in G. A function (a "weighting" of the independent sets) $g : S(G) \to R_+ \cup \{0\}$ is a *fractional coloring* of G if for every vertex $v \in V(G)$ we have $\sum_{v\in A\in S(G)} g(A) \geq 1$, that is the sum of the weights g puts on independent sets containing v is at least 1. (A proper coloring is also a fractional coloring: the color classes get weight 1, the other independent sets get weight 0.) The fractional chromatic number is $\chi_f(G) = \min_g \sum_{A\in S(G)} g(A)$, that is the minimum (taken over all fractional colorings) of the total weight put on independent sets by a fractional coloring g. (Formally we should write infimum but it is known that the minimum is always attained. See the book [32] for a detailed account on fractional graph parameters.)

We will need the following properties of the fractional chromatic number.

Definition 6: A directed graph $\vec{G}$ is vertex-transitive if for any two vertices $u,v \in V(\vec{G})$ it admits an automorphism that maps u to v.

If F is a vertex-transitive graph, then $\chi_f(F) = \frac{|V(F)|}{\alpha(F)}$. (For a proof see [32], Proposition 3.1.1 on page 41.)

For every graph F we have

$$\lim_{t\to\infty} \sqrt[t]{\chi(F^{\odot t})} = \lim_{t\to\infty} \sqrt[t]{\chi_f(F^{\odot t})}.$$

The latter follows from Lovász's result [24] stating that $\chi(F) \leq \chi_f(F)(1 + \ln \alpha(F))$ and the obvious inequality $\chi_f(F) \leq \chi(F)$ that holds for all (finite simple) graphs.

Proof of Theorem 2: Note that by the well-known (and more or less trivial) inequality $\chi(F) \geq \frac{|V(F)|}{\alpha(F)}$ for every graph F, we have $\chi(F^{\wedge t}_{P,\varepsilon}) \geq \frac{|V(F^{\wedge t}_{P,\varepsilon})|}{\alpha(F^{\wedge t}_{P,\varepsilon})}$. Clearly, this relation also holds if we have a directed graph $\vec{F}$ in place of the undirected graph F. This is straightforward since $\chi(\vec{F})$ and $\alpha(\vec{F})$ are defined to be identical to the corresponding parameter of the underlying undirected graph F. It is also well-known (cf. e.g. [11]) that $\lim_{\varepsilon\to 0}\lim_{t\to\infty} \frac{1}{t}\log(|\mathcal{T}^t(P,\varepsilon)|) = H(P)$. The last two relations immediately give $R_{\mathrm{D}}(\vec{G},P) + \Gamma(\vec{G},P) \geq H(P)$. For the reverse inequality let us fix a sequence of probability distributions P_t on the vertex set of our graph so that

$$\lim_{t\to\infty} \max_{a\in V(G)} |P(a) - P_t(a)| = 0$$

and

$$\lim_{t\to\infty}\frac{1}{t}\log\chi(\vec{G}^{\wedge t},P_t)=R_{\mathrm{D}}(\vec{G},P).$$

Notice that $\vec{G}^{\wedge t}(P_t,0)$ is a vertex-transitive graph, since every sequence forming an element of $\mathcal{T}^t(P_t,0)$ can be transformed into any other such sequence by simply permuting the coordinates.

Thus

$$\begin{aligned}R_{\mathrm{D}}(\vec{G},P)= &\ \lim_{t\to\infty}\tfrac{1}{t}\log\chi(\vec{G}^{\wedge t},P_t)\\ = &\ \lim_{t\to\infty}\tfrac{1}{t}\log\chi_f(\vec{G}^{\wedge t},P_t)\\ = &\ \lim_{t\to\infty}\tfrac{1}{t}\log\tfrac{|V(\vec{G}^{\wedge t},P_t)|}{\alpha(\vec{G}^{\wedge t},P_t)}\\ = &\ \lim_{t\to\infty}\tfrac{1}{t}\log|V(\vec{G}^{\wedge t},P_t)|-\\ - &\ \lim_{t\to\infty}\tfrac{1}{t}\log\alpha(\vec{G}^{\wedge t},P_t)\\ = &\ H(P)-\Gamma(\vec{G},P).\end{aligned}$$

■

Using standard techniques of the method of types, cf. [9], [11] we can already state our lower bound on $R_{\mathrm{D}}(\vec{G})$. We need the fact that the number of distinct types that a t-length sequence over some fixed alphabet can have is only a polynomial funcion of t (cf. the Type Counting Lemma 2.2 in [11]), while the parameters we investigate are asymptotic exponents of some graph parameters that grow exponentially as t tends to infinity. With this in mind it follows that

$$R_{\mathrm{D}}(\vec{G})=\sup_P R_{\mathrm{D}}(\vec{G},P),\qquad \Gamma(\vec{G})=\sup_P\Gamma(\vec{G},P).$$

Theorem 3:

$$R_{\mathrm{D}}(\vec{G})\geq\log|V(\vec{G})|-\Gamma(\vec{G}).$$

Proof: Using the above equalities, we obtain $R_{\mathrm{D}}(\vec{G})+\Gamma(\vec{G}) = \sup_P R_{\mathrm{D}}(\vec{G},P)+\sup_P\Gamma(\vec{G},P)\geq \sup_P(R_{\mathrm{D}}(\vec{G},P)+\Gamma(\vec{G},P)) = \sup_P H(P) = \log|V(\vec{G})|$. This gives the lower bound in the statement. ■

Note that Sperner capacity is unkown for many graphs, so the lower bound above usually does not give a known numerical value. Still, there are some examples of graphs where Sperner capacity is known and is non-trivial. A basic example is the cyclically oriented triangle, or more generally, any cyclically oriented cycle.

First we formulate a consequence of the above formula.

Corollary 4: If $\vec{G}$ is a vertex-transitive digraph then

$$R_{\mathrm{D}}(\vec{G})=\log|V(\vec{G})|-\Gamma(\vec{G}).$$

Proof: Let P_U denote the uniform distribution on the vertex set of $\vec{G}$. If $\vec{G}$ is vertex-transitive then by symmetry (see more details below) $R_{\mathrm{D}}(\vec{G})=R_{\mathrm{D}}(\vec{G},P_U)$ and $\Gamma(\vec{G})=\Gamma(\vec{G},P_U)$. Combining these equalities with Theorem 2 we obtain $R_{\mathrm{D}}(\vec{G})+\Gamma(\vec{G})=H(P_U)=\log|V(\vec{G})|$ and thus the statement.

The use of symmetry can be justified as follows. If the maximum of $R_D(\vec{G},P)$ over P would be achieved by a non-uniform distribution Q, then we could take an optimal construction and use time sharing for a realization of this construction according to all possible automorphisms of the digraph $\vec{G}$. There are a finite number of automorphisms, so this gives a construction of finite length where no vertex can appear a different number of times than any other, so we achieve the same optimal rate with uniform distribution. ■

Now we will use this Corollary to determine the Dilworth rate of the cyclically oriented k-length cycle $\vec{C}_k$ for every k. Note that the complement of a cyclically oriented cycle is a cyclically oriented cycle of the same length together with all diagonals as bidirected (or equivalently, undirected) edges. For $k=3$ there are no diagonals, so the cyclic triangle is isomorphic to its complement.

The Sperner capacity of the cyclic triangle was determined in [8], cf. also [5], and its value is $\log 2$. The upper bound part of this result was generalized by Alon [1], who proved the following.

Theorem 5: ([1]) The Sperner capacity of a directed graph $\vec{G}$ always satisfies

$$\Sigma(\vec{G})\leq\log\min\{\Delta_+(\vec{G}),\Delta_-(\vec{G})\}+1,$$

where $\Delta_+(\vec{F})$ and $\Delta_-(\vec{F})$ stand for the maximum outdegree and maximum indegree of $\vec{F}$, respectively, i.e., the number of edges at v that are oriented outwards or towards v, respectively.

Remark 7: See [23] for a generalization of Theorem 5, where the maximum in- and outdegrees are substituted with what is called the *directed local chromatic number* of digraph $\vec{G}$. □

On the other hand, Sperner capacity is bounded from below by (the logarithm of) the transitive clique number, the number of vertices in a largest transitively directed complete subgraph, denoted by $\omega_{\mathrm{tr}}(\vec{G})$. (This is an easy observation which implies that substituting $\omega_{\mathrm{s}}(\vec{G}^{\vee t})$ by $\omega_{\mathrm{tr}}(\vec{G}^{\vee t})$ in the definition of Sperner capacity gives the same value, i.e. it gives an alternative definition of Sperner capacity, see [21], [14], [31] and also Proposition 4 and the Remark following it in [15].) Note that a transitively directed complete subgraph meant here is not necessarily induced. It is allowed that some reverse edges are also present on the same subset of vertices.

Corollary 6: The Dilworth rate of the cyclically oriented k-cycle is

$$R_{\mathrm{D}}(\vec{C}_k)=\log\frac{k}{k-1}.$$

Proof: Let the directed complement of $\vec{C}_k$ be denoted by $\vec{S}_k$. Since $\Delta_+(\vec{S}_k)=\Delta_-(\vec{S}_k)=k-2$, Theorem 5 implies that the Sperner capacity of $\vec{S}_k$ is at most $\log(k-1)$.

It is easy to see that $\omega_{\mathrm{tr}}(\vec{S}_k) = k-1$, so the lower bound mentioned above is also $\log(k-1)$. Since the above two bounds coincide, the Sperner capacity of $\vec{S}_k$ is equal to $\log(k-1)$.

Using that $\vec{C}_k$ is vertex transitive Corollary 4 implies the statement. ■

Note that Corollary 6 shows that the Dilworth rate is a strict generalization of Witsenhausen rate since $\log\frac{k}{k-1}<\log 2\leq R_{\mathrm{W}}(C_k)$ if $k\geq 3$.

Definition 7: Call a subset of the vertex set of a directed graph $\vec{G}$ acyclic if it induces an acyclic subgraph. The latter means that there is no oriented cycle on these vertices. The acyclicity number $a(\vec{G})$ of a directed graph $\vec{G}$ is the number of vertices in a largest acyclic subset of $V(\vec{G})$.

Note that unlike for a transitive clique we do not allow reverse edges in an acyclic subgraph. The following lemma is from [7] (see also the discussion before Corollary 6 for an equivalent statement concerning the complementary digraph).

Lemma 7: ([7]) For all directed graphs $\vec{G}$ we have

$$\Gamma(\vec{G}) \geq \log a(\vec{G}).$$

Let $m \geq 1$ be an odd number. The following tournaments (oriented complete graphs) are also generalizations of the cyclic triangle. (They are, in fact, called *cyclic tournaments*.) Let $V(\vec{T}_m) = \{0, 1, \ldots, m-1\}$ and (i, j) is an edge iff $j - i \equiv r \pmod{m}$ for some $1 \leq r \leq \frac{m-1}{2}$. (Figure 1 shows the tournament $\vec{T}_5$.) Note that it holds for every directed graph that reversing all of its edges does not change the value of either its Sperner capacity or of its Dilworth rate. This implies that if $\vec{T}$ is a tournament then we have $\Sigma(\vec{T}^c) = \Sigma(\vec{T})$ and $\Gamma(\vec{T}^c) = \Gamma(\vec{T})$. By $\Gamma(\vec{T}) = \Sigma(\vec{T}^c)$ all the four values are equal.

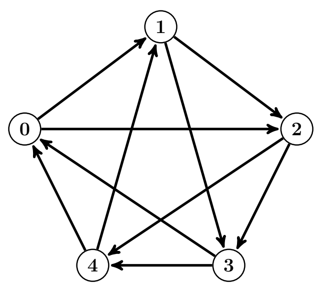

Fig. 1. The tournament $\vec{T}_5$.

Corollary 8: For all odd integers $m > 0$ we have

$$R_{\mathrm{D}}(\vec{T}_m) = \log \frac{2m}{m+1}.$$

Proof: Lemma 7 gives $\Gamma(\vec{T}_m) \geq \log \frac{m+1}{2}$.

By $\Gamma(\vec{T}_m) = \Sigma(\vec{T}_m)$ (see the note before stating the Corollary) Alon's Theorem 5 can be applied implying that our lower bound is sharp. Since $\vec{T}_m$ is vertex-transitive we can apply Corollary 4 to complete the proof. ∎

Observe that Corollary 8 shows not only that the value of the Dilworth rate of an oriented graph may differ from the Witsenhausen rate of the underlying undirected graph, but also that they can differ arbitrarily. The latter is meant in the strong sense that the Witsenhausen rate cannot be bounded from above by any function of the Dilworth rate. Indeed, denoting the complete graph on m vertices by K_m we have $\log \frac{2m}{m+1} < \log m = R_{\mathrm{W}}(K_m)$ for every $m \geq 2$. The left hand side of the inequality is bounded above by $\log 2$, while the right hand side goes to infinity with m.

B. Dichromatic number and upper bounds

Now we show that the (logarithm of the) *dichromatic number* defined in [29] is an upper bound on the Dilworth rate.

Definition 8: The *dichromatic number* $\chi_{\mathrm{dir}}(\vec{G})$ of a directed graph $\vec{G}$ is the minimum number of acyclic subsets that cover $V(\vec{G})$. A partition of $V(\vec{G})$ into acyclic subsets will be called a *directed coloring* or *dicoloring*.

We note that an undirected edge (meaning a bidirected edge) is considered to be a 2-length cycle, therefore its two endpoints cannot be both contained in an acyclic set. This shows that for undirected (equivalently, symmetrically directed) graphs the dichromatic number is equal to the chromatic number.

Remark 8: We do not use the term "acyclic coloring", because it is already used for a completely different concept, see [2]. □

Theorem 9: For any directed graph $\vec{G}$

$$r_{\mathrm{D}}(\vec{G}) \leq \chi_{\mathrm{dir}}(\vec{G}).$$

Proof: Let us fix a directed coloring of digraph $\vec{G}$ consisting of $k := \chi_{\mathrm{dir}}(\vec{G})$ acyclic subsets ("color classes"). For each $v \in V(\vec{G})$ let $g(v)$ denote the color class that contains v.

Now consider $\vec{G}^{\wedge t}$. Its vertex set is $[V(\vec{G})]^t$. For each sequence $(a_1, \ldots, a_t) \in V(\vec{G}^{\wedge t})$ we attach the sequence of colors $(g(a_1), \ldots, g(a_t))$. There are k^t such color sequences, so this gives a partition of $V(\vec{G}^{\wedge t})$ into k^t partition classes. We also give another partition of $V(\vec{G}^{\wedge t})$ according to types. Two vertices are in the same partition class if their type is the same. The Type Counting Lemma 2.2 in [11] states that the number of possible types a sequence of length t over a k-element alphabet can have is not more than $(t+1)^k$. Thus we know that the latter partition has at most $(t+1)^{|V(\vec{G})|}$, that is a polynomial number (in t) of classes. Now let $\mathcal{Q} = (Q_1, \ldots, Q_s)$ be the common refinement of these two partitions. We have $s \leq (t+1)^{|V(\vec{G})|} k^t$ by the foregoing. Now we show that each partition class Q_i induces an independent set in $\vec{G}^{\wedge t}$. Let two sequences $\boldsymbol{a} = (a_1, \ldots, a_t)$ and $\boldsymbol{b} = (b_1, \ldots, b_t)$ belong to the same Q_i, that is, they have the same type and $\forall i : g(a_i) = g(b_i)$. Let j be an index for which $a_j \neq b_j$. Since a_j and b_j are in the same color class of a valid dicoloring, we cannot have both (a_j, b_j) and (b_j, a_j) present in the graph as a directed edge. If neither is present then there is no edge between $\boldsymbol{a}$ and $\boldsymbol{b}$. If $(a_j, b_j) \in E(\vec{G})$, then we know that $(b_j, a_j) \notin E(\vec{G})$, so the oriented edge $(\boldsymbol{b}, \boldsymbol{a})$ cannot be an edge of $\vec{G}^{\wedge t}$. We need to show that neither the opposite oriented edge $(\boldsymbol{a}, \boldsymbol{b})$ can be present in $\vec{G}^{\wedge t}$. If $(a_j, b_j) \in E(\vec{G})$, then we claim that there should be another position ℓ for which $(b_\ell, a_\ell) \in E(\vec{G})$. If this is true, then $(a_\ell, b_\ell) \notin E(\vec{G})$ and so $(\boldsymbol{a}, \boldsymbol{b}) \notin E(\vec{G}^{\wedge t})$. Now we prove the above claim. Consider all coordinates h, for which $g(a_h) = g(a_j) = g(b_j) = g(b_h)$. Denote the set of these h's (including j itself) L. Since vertices $v \in V(\vec{G})$ with the same "color" $g(v)$ induce an acyclic subdigraph, we can put these vertices into a linear order so, that $(v, v') \in E(\vec{G})$ implies that v precedes v' in this linear order. So if $(a_j, b_j) \in E(\vec{G})$, then for our edge (a_j, b_j) a_j precedes b_j. However, since $\boldsymbol{a}$ and $\boldsymbol{b}$ have the same type it cannot happen that for each $h \in L$ a_h precedes b_h in this linear order. So there must be a coordinate ℓ where b_ℓ precedes a_ℓ and this implies the claimed properties. Thus each partition class Q_i is independent indeed, so the undirected graph underlying $\vec{G}^{\wedge t}$ can be properly colored with $s \leq (t+1)^{|V(\vec{G})|}k^t$ colors. This implies that $r_{\mathrm{D}}(\vec{G}) = \lim_{t\to\infty} \sqrt[t]{\chi([\vec{G}_H]^{\wedge t})} \leq \liminf_{t\to\infty} \sqrt[t]{(t+1)^{|V(\vec{G})|}k^t} = k = \chi_{\mathrm{dir}}(\vec{G})$. ■

As a strengthening of the previous theorem we will show that we can also write a natural fractional relaxation of the dichromatic number on the right hand side above. (We could not prove this right away, as the weaker statement will be used in the proof.) To prove this stronger statement we need some preparation, in particular we will use the following observations.

First note, that $\chi_{\mathrm{dir}}(\vec{F}) \leq \chi(\vec{F})$ holds for any digraph $\vec{F}$. This is simply because independent sets in $\vec{F}$ are special acyclic sets, so any proper coloring of $\vec{F}$ is also a directed coloring of $\vec{F}$.

Proposition 10: The dichromatic number is submultiplicative with respect to the AND product, i.e.

$$\chi_{\mathrm{dir}}(\vec{F} \wedge \vec{G}) \leq \chi_{\mathrm{dir}}(\vec{F})\chi_{\mathrm{dir}}(\vec{G}).$$

In particular,

$$\chi_{\mathrm{dir}}(\vec{F}^{\wedge t}) \leq [\chi_{\mathrm{dir}}(\vec{F})]^t.$$

A straightforward consequence of Proposition 10 is that the limit $\lim_{t\to\infty} \sqrt[t]{\chi_{\mathrm{dir}}(F^{\wedge t})}$ exists.

In the proof we will use the following simple fact.

Claim 11: The AND product of two acyclic subdigraphs, A of $\vec{F}$ and B of $\vec{G}$ results in an acyclic subdigraph of $\vec{F} \wedge \vec{G}$.

Proof: Assume for contradiction that $(\vec{F} \wedge \vec{G})[A \times B]$ contains a directed cycle. (Recall that $\vec{Y}[U]$ denotes the digraph $\vec{Y}$ induces on $U \subseteq V(\vec{Y})$.) Let its vertices be $(a_1, b_1), \ldots, (a_k, b_k)$ in the (cyclic) order the cycle defines, i.e. $((a_i, b_i), (a_{i+1}, b_{i+1}))$ is an edge of $\vec{F} \wedge \vec{G}$ for all $i \in \{1, \ldots, k\}$ where addition is intended modulo k. We may assume without loss of generality that not all a_i's are equal. Then in the sequence $a_1, a_2, \ldots, a_k$ we have for all $i \in \{1, \ldots, k\}$ either $a_i = a_{i+1}$ or $(a_i, a_{i+1}) \in E(\vec{F})$ (addition is again modulo k) and for some i the second case occurs. But then there must be a directed cycle in $\vec{F}[\{a_1, \ldots, a_k\}]$ contradicting the assumption that A is acyclic. ■

Proof of Proposition 10: Let $c_{\vec{F}} : V(\vec{F}) \to \{1, \ldots, \chi_{\mathrm{dir}}(\vec{F})\}$ and $c_{\vec{G}} : V(\vec{G}) \to \{1, \ldots, \chi_{\mathrm{dir}}(\vec{G})\}$ be optimal directed colorings of the digraphs $\vec{F}$ and $\vec{G}$, respectively. Using these colorings we define the function $\hat{c} : V(\vec{F}) \times V(\vec{G}) \to \{1, \ldots, \chi_{\mathrm{dir}}(\vec{F})\chi_{\mathrm{dir}}(\vec{G})\}$ as follows. For $(u, v) \in V(\vec{F}) \times V(\vec{G})$ let $\hat{c} : (u, v) \mapsto (c_{\vec{F}}(u), c_{\vec{G}}(v))$. Claim 11 implies that $\hat{c}$ is a directed coloring of $\vec{F} \wedge \vec{G}$. As it uses $\chi_{\mathrm{dir}}(\vec{F})\chi_{\mathrm{dir}}(\vec{G})$ colors the statement is proved. ■

Lemma 12: For any digraph $\vec{F}$ and positive integer k we have

$$r_{\mathrm{D}}(\vec{F}^{\wedge k}) = [r_{\mathrm{D}}(\vec{F})]^k.$$

Proof: Fix an arbitrary positive integer k. We can write

$$\begin{aligned} r_{\mathrm{D}}(\vec{F}) = \;& \lim_{m\to\infty} \sqrt[mk]{\chi(\vec{F}^{\wedge mk})} \\ = \;& \lim_{m\to\infty} \sqrt[k]{\sqrt[m]{\chi([\vec{F}^{\wedge k}]^{\wedge m})}} \\ = \;& \sqrt[k]{\lim_{m\to\infty} \sqrt[m]{\chi([\vec{F}^{\wedge k}]^{\wedge m})}} \\ = \;& \sqrt[k]{r_{\mathrm{D}}(\vec{F}^{\wedge k})}, \end{aligned}$$

that implies the statement. ■

Proposition 13: For any digraph $\vec{F}$ we have

$$\lim_{t\to\infty} \sqrt[t]{\chi_{\mathrm{dir}}(\vec{F}^{\wedge t})} = r_{\mathrm{D}}(\vec{F}).$$

Proof: By $\chi_{\mathrm{dir}}(\vec{F}) \leq \chi(\vec{F})$ we have

$$\lim_{t\to\infty} \sqrt[t]{\chi_{\mathrm{dir}}(\vec{F}^{\wedge t})} \leq \lim_{t\to\infty} \sqrt[t]{\chi(\vec{F}^{\wedge t})} = r_{\mathrm{D}}(\vec{F}).$$

For the reverse inequality we can write

$$r_{\mathrm{D}}(\vec{F}) = \lim_{t\to\infty} \sqrt[t]{r_{\mathrm{D}}(\vec{F}^{\wedge t})} \leq \lim_{t\to\infty} \sqrt[t]{\chi_{\mathrm{dir}}(\vec{F}^{\wedge t})},$$

where the equality follows by Lemma 12 and the inequality is a consequence of $r_{\mathrm{D}}(\vec{G}) \leq \chi_{\mathrm{dir}}(\vec{G})$ (see Theorem 9) applied for $\vec{G} = \vec{F}^{\wedge t}$. ■

Definition 9: Let the set of subsets of the vertex set inducing an acyclic subgraph in a digraph $\vec{G}$ be $\mathcal{A}(\vec{G})$. A function $g : \mathcal{A}(\vec{G}) \to R_+ \cup \{0\}$ is called a *fractional directed coloring* (or fractional dicoloring) if $\forall v \in V(\vec{G})$ we have $\Sigma_{v\in U\in\mathcal{A}(\vec{G})} g(U) \geq 1$. The *fractional dichromatic number* of $\vec{G}$ is

$$\chi_{\mathrm{dir},f}(\vec{G}) = \min_g \Sigma_{U\in\mathcal{A}(\vec{G})} g(U),$$

where the minimum is taken over all fractional directed colorings g of $\vec{G}$.

Note the obvious inequality $\chi_{\mathrm{dir},f}(\vec{G}) \leq \chi_{\mathrm{dir}}(\vec{G})$ for any digraph $\vec{G}$.

We will need the following lemma.

Lemma 14: For any digraphs $\vec{F}$ and $\vec{G}$ we have

$$\chi_{\mathrm{dir},f}(\vec{F} \wedge \vec{G}) \leq \chi_{\mathrm{dir},f}(\vec{F})\chi_{\mathrm{dir},f}(\vec{G}).$$

Proof: Let f and g be optimal fractional directed colorings of $\vec{F}$ and $\vec{G}$, respectively.

Recall that by Claim 11 if $A \in \mathcal{A}(\vec{F})$ and $B \in \mathcal{A}(\vec{G})$ then the direct product $A \times B$ is in $\mathcal{A}(\vec{F} \wedge \vec{G})$, i.e. $A \times B$ induces an acyclic subgraph in $\vec{F} \wedge \vec{G}$.

Now give the following weights w to the acyclic sets of $\vec{F} \wedge \vec{G}$. If $H \in \mathcal{A}(\vec{F} \wedge \vec{G})$ has a product structure, i.e. $H = A \times B$ for some $A \in \mathcal{A}(\vec{F})$ and $B \in \mathcal{A}(\vec{G})$, then let $w(H) = f(A)g(B)$. If H is not of this form, then let $w(H) = 0$. For any $(a, b) \in V(\vec{F} \wedge \vec{G})$ we have $\sum_{H\ni(a,b), H\in\mathcal{A}(\vec{F}\wedge\vec{G})} w(H) = (\sum_{A\ni a} f(A))(\sum_{B\ni b} f(B)) \geq$

1, thus w is a fractional dicoloring of $\vec{F} \wedge \vec{G}$. Now we have $\chi_{\mathrm{dir},f}(\vec{F} \wedge \vec{G}) \leq (\sum_{A\in\mathcal{A}(\vec{F})} f(A))(\sum_{B\in\mathcal{A}(\vec{G})} g(B)) = \chi_{\mathrm{dir},f}(\vec{F})\chi_{\mathrm{dir},f}(\vec{G})$. This completes the proof. ■

Corollary 15: For any digraph $\vec{G}$ and any positive integer t we have

$$\chi_{\mathrm{dir},f}(\vec{G}^{\wedge t}) \leq [\chi_{\mathrm{dir},f}(\vec{G})]^t.$$

■

We also need the following result.

Proposition 16: For any digraph $\vec{F}$ we have

$$\lim_{t\to\infty} \sqrt[t]{\chi_{\mathrm{dir},f}(\vec{F}^{\wedge t})} = r_{\mathrm{D}}(\vec{F}).$$

For the proof we need some preparation.

A hypergraph $\mathcal{H} = (V, \mathcal{E})$ consists of a vertex set $V = V(\mathcal{H})$ and an edge set $\mathcal{E}$, where the elements of $\mathcal{E}$ are subsets of V. A covering of hypergraph $\mathcal{H}$ is a set of edges the union of which contains all elements of $V(\mathcal{H})$. Let $k(\mathcal{H})$ denote the minimum number of edges in a covering of $\mathcal{H}$. A fractional covering of a hypergraph $\mathcal{H} = (V, \mathcal{E})$ is a function $g : \mathcal{E} \to R_+ \cup \{0\}$ satisfying for every $v \in V$ that $\sum_{v\in E\in\mathcal{E}} g(E) \geq 1$. The fractional covering number is $k_f(\mathcal{H}) := \min_g \sum_{E\in\mathcal{E}} g(E)$ where the minimization is over all fractional covers g. Clearly, $k_f(\mathcal{H}) \leq k(\mathcal{H})$. Lovász proved in [24] (cf. also [32]) that

$$k(\mathcal{H}) \leq k_f(\mathcal{H})(1 + \ln \mu(\mathcal{H})),$$

where $\mu(\mathcal{H}) = \max\{|E| : E \in \mathcal{E}(\mathcal{H})\}$, that is the cardinality of a largest edge in $\mathcal{H}$.

For a directed graph $\vec{G}$ let $\mathcal{H}_{\vec{G}} = (V(\vec{G}), \mathcal{E}_{\vec{G}})$ where $\mathcal{E}_{\vec{G}} = \mathcal{A}(\vec{G})$, i.e. it consists of the acyclic subsets of vertices in $\vec{G}$. It is straightforward that $k(\mathcal{H}_{\vec{G}}) = \chi_{\mathrm{dir}}(\vec{G})$ and $k_f(\mathcal{H}_{\vec{G}}) = \chi_{\mathrm{dir},f}(\vec{G})$ while $\mu(\mathcal{H}) = a(\vec{G})$. Thus the above result implies that

$$\chi_{\mathrm{dir}}(\vec{G}) \leq \chi_{\mathrm{dir},f}(\vec{G})(1 + \ln a(\vec{G})).$$

Proof of Proposition 16: We have $\lim_{t\to\infty} \sqrt[t]{\chi_{\mathrm{dir},f}(\vec{F}^{\wedge t})} \leq \lim_{t\to\infty} \sqrt[t]{\chi_{\mathrm{dir}}(\vec{F}^{\wedge t})} = r_{\mathrm{D}}(\vec{F})$ by Proposition 13 and the obvious inequality $\chi_{\mathrm{dir},f}(\vec{G}) \leq \chi_{\mathrm{dir}}(\vec{G})$ applied to $\vec{F}^{\wedge t}$.

For the reverse inequality we write

$$\begin{aligned} r_{\mathrm{D}}(\vec{F}) = &\ \lim_{t\to\infty} \sqrt[t]{\chi_{\mathrm{dir}}(\vec{F}^{\wedge t})} \\ \leq &\ \lim_{t\to\infty} \sqrt[t]{\chi_{\mathrm{dir},f}(\vec{F}^{\wedge t})(1 + \ln a(\vec{F}^{\wedge t}))} \\ = &\ \left(\lim_{t\to\infty} \sqrt[t]{\chi_{\mathrm{dir},f}(\vec{F}^{\wedge t})}\right) \times \\ &\ \times \left(\lim_{t\to\infty} \sqrt[t]{(1 + \ln a(\vec{F}^{\wedge t}))}\right) \\ = &\ \lim_{t\to\infty} \sqrt[t]{\chi_{\mathrm{dir},f}(\vec{F}^{\wedge t})}, \end{aligned}$$

where the last equality follows from the fact, that the second term of the multiplication before it is 1, since $\ln a(\vec{F}^{\wedge t})$ is only linear in t. ■

Theorem 17: For any directed graph $\vec{G}$

$$r_{\mathrm{D}}(\vec{G}) \leq \chi_{\mathrm{dir},f}(\vec{G}).$$

Proof: Using Corollary 15 we obtain

$$\begin{aligned} r_{\mathrm{D}}(\vec{G}) &= \lim_{t\to\infty} \sqrt[t]{\chi_{\mathrm{dir},f}(\vec{G}^{\wedge t})} \\ &\leq \lim_{t\to\infty} \sqrt[t]{[\chi_{\mathrm{dir},f}(\vec{G})]^t} = \chi_{\mathrm{dir},f}(\vec{G}). \end{aligned}$$

■

There are several directed graphs $\vec{G}$ for which the above upper bound is sharp. In particular, Corollaries 6 and 8 can be proved using Theorem 17 instead of vertex-transitivity. (An optimal fractional dicoloring of $\vec{T}_5$ is shown on Figure 2. Here we put weight $\frac{1}{3}$ on all the five acyclic sets formed by three consecutive vertices.)

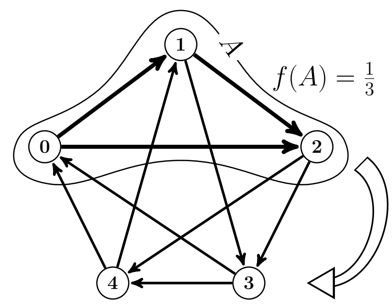

Fig. 2. An optimal fractional dicoloring of $\vec{T}_5$.

Note however, that Theorem 17 is not always tight: for the graph (symmetrically directed digraph) C_5, we have $R_{\mathrm{W}}(C_5) = \log\sqrt{5}$ by results in [38] and [25], while $\chi_{\mathrm{dir},f}(C_5) = \frac{5}{2}$.

We present another such example which is not symmetrically directed. Let the 5-length cycle be oriented in an (as much as possible) alternating manner, that is so that only one of its vertices will have outdegree 1 (implying that two of the 4 others will have outdegree 2, and the remaining 2 have outdegree 0). Denote this oriented graph by $\vec{A}_5$. (See Figure 3.)

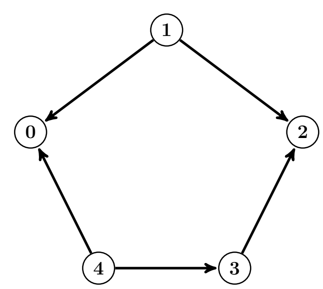

Fig. 3. The directed graph $\vec{A}_5$.

We know that $\Sigma(\vec{A}_5) = \log\sqrt{5}$. (This is proven as Proposition 4 in [15], see [31] for more details on the Sperner capacity of oriented self-complementary graphs. All other orientations of the 5-cycle have Sperner capacity $\log 2$, see [15] and [23]. By Theorems 3 and 17 this implies that the Dilworth rate of their complements is $\log\frac{5}{2}$.) Thus by Theorem 3 for the complement of $\vec{A}_5$ we have $R_{\mathrm{D}}(\vec{A}_5^{\mathrm{c}}) \geq \log 5 - \log\sqrt{5}$ or equivalently $r_{\mathrm{D}}(\vec{A}_5^{\mathrm{c}}) \geq \sqrt{5}$. (The digraph $\vec{A}_5^{\mathrm{c}}$ is shown on Figure 4.)

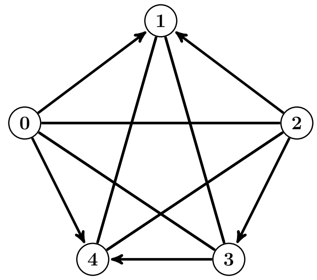

Fig. 4. The directed graph $\vec{A}_5^c$ (complement of $\vec{A}_5$). Bidirected edges are shown as undirected ones.

Proposition 18:

$$\sqrt{5} \leq r_D(\vec{A}_5^c) \leq \sqrt{6} < \frac{5}{2} = \chi_{\mathrm{dir},f}(\vec{A}_5^c).$$

Proof: The first inequality was already given above. To prove the second inequality we give 6 acyclic sets of vertices in the second power $[\vec{A}_5^c]^{\wedge 2}$ of our graph, that cover all vertices in $V([\vec{A}_5^c]^{\wedge 2})$. The existence of this covering implies that $r_{\mathrm{D}}([\vec{A}_5^c]^{\wedge 2}) \leq \chi_{\mathrm{dir}}([\vec{A}_5^c]^{\wedge 2}) \leq 6$, thus by Lemma 12 we get $r_{\mathrm{D}}(\vec{A}_5^c) \leq \sqrt{6}$.

Let us denote the vertices of $V(\vec{A}_5^c) = V(\vec{A}_5)$ by $0,1,2,3,4$ in their cyclic order, so that in $\vec{A}_5$ we have $d_+(3) = 1$ and the unique outneighbor of 3 is 2. (That is, the outdegree 1 vertex is 3, the outdegree 2 vertices are 4 and 1, while 2 and 0 have outdegree 0.) The following six subsets of $V(\vec{A}_5) \times V(\vec{A}_5)$ induce acyclic subgraphs of $V([\vec{A}_5^c]^{\wedge 2})$ that entirely cover its vertex set:

$$44, 31, 10, 23, 02;$$

$$14, 43, 01, 30, 22$$

$$41, 33, 32, 20$$

$$13, 21, 42, 00$$

$$11, 12, 04, 03$$

$$34, 24, 40.$$

One can check that within all these five sets if xy is to the left of zw in the same line above (x,y,z,w may not all be different), then either (x,z) or (y,w) (or both) form an edge of $\vec{A}_5$, thus this is a missing edge in $\vec{A}_5^c$. This implies that as a vertex of $[\vec{A}_5^c]^{\wedge 2}$ the pair(x,y) does not send an edge to vertex (z,w), therefore the corresponding set of vertices induces an acyclic subgraph in $[\vec{A}_5^c]^{\wedge 2}$. This completes the proof of the second inequality.

To see that $\chi_{\mathrm{dir},f}(\vec{A}_5^c) \geq \frac{5}{2}$ it is enough to realize that any 3 vertices of $\vec{A}_5^c$ contains a bidirected edge, thus any acyclic induced subgraph has at most two vertices. To see equality we can put weight $\frac{1}{2}$ on all the five 2-element acyclic subsets. ■

Remark 9: Getting the same upper bound for the values determined in Corrollaries 6 and 8 in two different ways above is not pure coincidence. It follows from the fact that if $\vec{G}$ is vertex-transitive then

$$\chi_{\mathrm{dir},f}(\vec{G}) = \frac{|V(\vec{G})|}{a(\vec{G})}.$$

Note that this is a generalization of the relation, that for every vertex-transitive (undirected) graph G

$$\chi_f(G) = \frac{|V(G)|}{\alpha(G)}$$

that we already referred to right after Definition 6 in Subsection III-A. This latter equality is presented in [32] (Proposition 3.1.1 on page 41) as a consequence of a more general equality concerning vertex-transitive hypergraphs (see Proposition 1.3.4 on page 7 of [32]).

A vertex-transitive hypergraph is a hypergraph that attains for every pair u,v of its vertices an automorphism that maps u to v. Proposition 1.3.4 in [32] states that if $\mathcal{H} = (V, \mathcal{E})$ is a vertex-transitive hypergraph then

$$k_f(\mathcal{H}) = \frac{|V|}{\mu(\mathcal{H})},$$

where (as before; cf. the discussion after stating Proposition 16) $\mu(\mathcal{H}) = \max_{E \in \mathcal{E}} |E|$. For a directed graph $\vec{G}$ we attach again the hypergraph $\mathcal{H}_{\vec{G}} = (V(\vec{G}), \mathcal{E}_{\vec{G}})$ where $\mathcal{E}_{\vec{G}} = \mathcal{A}(\vec{G})$. It is straightforward that if $\vec{G}$ is vertex-transitive then so is $\mathcal{H}_{\vec{G}}$. The equality quoted for k_f from [32] gives the stated equality $\chi_{\mathrm{dir},f}(\vec{G}) = \frac{|V(\vec{G})|}{a(\vec{G})}$ for vertex-transitive digraphs $\vec{G}$. □

IV. Compound systems

Imagine that the handwritten message is left to Bob by one of his three secretaries but it is not known in advance which one. Their handwriting is rather different and this has two consequences that are important for us. One is that the possible mistakes Bob can make when decoding the message are different depending on which secretary wrote him the message. (For example, in the first secretary's handwriting a 7 can be thought to be a 1, while the second secretary "crosses" the leg of 7, so it can never look like a 1, however it can be confused with a 4, etc.) This means that in place of the noisy channel H we had so far, now there are three distinct channels H_1, H_2, and H_3 and we do not know in advance which one will be used. The other important consequence of the secretaries' handwriting being different is that Bob recognizes who wrote the message, i.e., he will know which one of the three noisy channels models the actual situation. The relevant characteristics (the graphs G_{H_i}) of each of these channels are known by Bob and also by his bank. Now it is the bank that will send the second, error-free but expensive message to Bob. Although the bank knows the characteristics it does not know which secretary left the first message. So the second message should make Bob able to decide whether his decoding (of the first message) was correct irrespective of which secretary wrote it. As before, we are interested (asymptotically) in the shortest possible message the bank can send to satisfy the requirements.

Notice that this scenario is basically that of having a compound channel for the first communication. See [17], [30] for more on compound channels from a zero-error point of view.

Here is the abstract setting for the above situation. We have k distinct noisy channels described by the family $\mathcal{H} = \{H_1, \ldots, H_k\}$. The relevant properties of this set are characterised by the family of directed graphs $\vec{\mathcal{G}}_{\mathcal{H}} = \{\vec{G}_{H_1}, \ldots, \vec{G}_{H_k}\}$.

Definition 10: (cf. [30] and [36]) The Dilworth rate of a family of directed graphs $\vec{\mathcal{G}} = \{\vec{G}_1, \ldots, \vec{G}_k\}$ all having the same vertex set V, is

$$R_{\mathrm{D}}(\vec{\mathcal{G}}) = \lim_{t\to\infty} \frac{1}{t} \log \chi(\cup_i \vec{G}_i^{\wedge t}),$$

where $\cup_i \vec{G}_i^{\wedge t}$ denotes the graph on the common vertex set V^t of the graphs $\vec{G}_i^{\wedge t}$ with edges set $\cup_i E(\vec{G}_i^{\wedge t})$.

Proposition 19: If $m_{\mathcal{H}}(t)$ is the shortest possible message the bank should send to inform Bob about the correctness of his decoding of the handwritten message for t consecutive rounds, then

$$\lim_{t\to\infty} \frac{m_{\mathcal{H}}(t)}{t} = R_{\mathrm{D}}(\vec{\mathcal{G}}_{\mathcal{H}}).$$

Proof: It is enough to prove $m_{\mathcal{H}}(t) = \lceil \chi(\cup_i G_{H_i}^{\wedge t})\rceil$.

Let a proper coloring of the graph $\cup_i \vec{G}_{H_i}^{\wedge t}$ be fixed and agreed on by Bob and the bank in advance. Bob knows that he received the first message via, say, H_j. Since the fixed coloring is a proper coloring of $\vec{G}_{H_j}^{\wedge t}$, Proposition 1 implies that the right hand side is an upper bound. If $m_{\mathcal{H}}(t)$ would be smaller, then there is some j for which $\vec{G}_{H_j}^{\wedge t}$ has two adjacent vertices for which the bank sends the same message. If the channel in use is just H_j then Proposition 1 implies that the right hand side above is also a lower bound. ∎

The interesting fact about the above quantity is that it is not more then its obvious lower bound.

Theorem 20: ([30], cf. also [36]) For every finite family of directed graphs $\vec{\mathcal{G}} = \{\vec{G}_1, \ldots, \vec{G}_k\}$ we have

$$R_{\mathrm{D}}(\vec{\mathcal{G}}) = \max_{\vec{G}_i \in \vec{\mathcal{G}}} R_{\mathrm{D}}(\vec{G}_i).$$

The analogous result for Witsenhausen rate (that is the special case of the above when all graphs are undirected) is proven in [36]. The above general form is already stated by Nayak and Rose in [30] (cf. Remark 6 of the present paper). They write that the proof uses essentially the same argument as in [36] and they omit it for the sake of brevity. We do the same.

V. Connections to extremal set theory

As is the case with Sperner capacity, Dilworth rate also has relevance in extremal set theory. (Recall that both notions got their name from this relationship, cf. Remark 5.) These connections become clear when we consider the t-length sequences of vertices of a (di)graph $\vec{G}$ as characteristic vectors of partitions of a t-element set. We already mentioned that if $\vec{G}$ is the digraph consisting of two vertices and a single oriented edge between them, then the Dilworth rate is just the asymptotic exponent of the minimum number of Sperner systems (antichains in the Boolean lattice) that cover all subsets of a t-element set (the elements of the Boolean lattice). This is known (and easy to prove) to be $t+1$, that is the asymptotic exponent is 0. (The situation with Sperner capacity is similar: its value for the above mentioned single edge graph is the asymptotic exponent of the size of a largest Sperner system on a t-element set which is easy to see to be 1.)

Here we present another example that we believe to be interesting. Let us call a family of pairs of disjoint subsets (A_i, B_i) of a t-element set cross-intersecting if for every two pairs (A_i, B_i) and (A_j, B_j) both of the intersections $A_i \cap B_j$ and $A_j \cap B_i$ are nonempty. (In other words, $A_k \cap B_\ell = \emptyset$ iff $k = \ell$.) Bollobás [6] proved that for such a family $\sum_i \frac{1}{\binom{|A_i|+|B_i|}{|A_i|}} \leq 1$. Now we ask, what is the minimum number of cross-intersecting families that can cover all possible pairs of disjoint subsets of a t-element set. If we are satisfied with determining the asymptotic exponent (i.e. not the exact value) of this number, then this question is equivalent to asking the Dilworth rate of an appropriate graph.

Proposition 21: Let $B(t)$ denote the minimum number of cross-intersecting families that cover all pairs of disjoint subsets of a t-element set. Then

$$\lim_{t\to\infty} \frac{1}{t} \log B(t) = 1.$$

Proof: Let $\vec{F}$ be the following directed graph. The vertex set of $\vec{F}$ is $\{0,1,2\}$ and the edge set is $E(\vec{F}) = \{(0,1),(1,0),(0,2),(2,0),(1,2)\}$. That is $\vec{F}$ has two undirected (bidirected) edges connecting 0 to the other two vertices and one oriented edge from 1 to 2. If we encode pairs of disjoint sets of a t-element set by ternary sequences (the positions of 1's are the elements of A_i and the positions of 2's are the elements of B_i in the ternary sequence encoding the pair (A_i, B_i)), then it is immediate to see that $B(t)$ is just the chromatic number of $\vec{F}^{\wedge t}$. Thus $R_{\mathrm{D}}(\vec{F})$ can indeed be interpreted as the limit in the statement.

Now we have to show that $R_{\mathrm{D}}(\vec{F}) = 1$. We have $\chi_{\mathrm{dir}}(\vec{F}) = 2$, so we have $R_{\mathrm{D}}(\vec{F}) \leq 1$ by Theorem 9. Since $\vec{F}$ contains an undirected edge, $\vec{F}^{\wedge t}$ contains a symmetric clique of size 2^t. This implies $\chi(\vec{F}^{\wedge t}) \geq 2^t$ and thus $R_{\mathrm{D}}(\vec{F}) \geq 1$. The two inequalities prove $R_{\mathrm{D}}(\vec{F}) = 1$. ∎

VI. Complete zero-error decoding

Here we consider the more ambitious setup, where Bob, otherwise in the same situation as described in the Introduction, should decode the actual message with zero-error. (Not only getting to know whether his earlier decoding was correct or not.)

A. The closure graph

It remains true that all the relevant information to solve this problem is contained in the directed graph $\vec{G}_H$ defined at the beginning of Subsection II-B. We will need the following operation on directed graphs.

Definition 11: Let $\vec{F}$ be a directed graph on vertex set V. Let the *closure graph* $\mathrm{cl}(\vec{F})$ of $\vec{F}$ be the following undirected graph.

$$V(\mathrm{cl}(\vec{F})) := V(\vec{F}) = V$$

and

$$E(\mathrm{cl}(\vec{F})) := \{\{a,b\} : (a,b) \in E(\vec{F})\} \cup$$
$$\cup \{\{a,b\} : \exists v \in V \text{ s.t. } (a,v),(b,v) \in E(\vec{F})\}.$$

Note that if $\vec{F} = \vec{G}_H$ then $\mathrm{cl}(\vec{F}) = \mathrm{cl}(\vec{G}_H)$ is the graph where two vertices a and b are connected if and only if the input letters they represent can result in the same output letter. This output letter can be one of a and b but also a third element v of the alphabet. (Recall that the input and output alphabets of the noisy channel H are identical.) The last possibility means that $\mathrm{cl}(\vec{G}_H)$ may have edges the two endpoints of which are not adjacent in $\vec{G}_H$ in either direction.

For example, if $\vec{G}_H$ has three vertices, a, b, c and only two (directed) edges (a,c) and (b,c), then $\vec{G}_H$ is a bipartite graph, while $\mathrm{cl}(\vec{G}_H)$ is the complete (undirected) graph on 3 vertices.

B. Relevance of the Witsenhausen rate in this case

Now we are ready to state the graph theoretic solution of the problem considered here.

Theorem 22: Let $h_c(t)$ denote the minimum number of bits Alice should send to Bob via the noiseless channel for making Bob able to decode a t-length sequence of the source output with zero-error. (The subscript c stands for "complete".) Then

$$\lim_{t\to\infty} \frac{h_c(t)}{t} = R_{\mathrm{W}}(\mathrm{cl}(\vec{G}_H)),$$

the Witsenhausen rate of the closure graph $\mathrm{cl}(\vec{G}_H)$.

Proof: Assume that a t-length source output is sent through channel H, and the second message sent by Alice is shorter than $\log \chi([\mathrm{cl}(\vec{G}_H)]^{\wedge t})$. Then there are two t-length source outputs, that is two sequences $\boldsymbol{x}, \boldsymbol{y}$ in $V([\mathrm{cl}(\vec{G}_H)]^{\wedge t})$ that are adjacent in $[\mathrm{cl}(\vec{G}_H)]^{\wedge t}$ and for which Alice sends the same message when encoding either of them for the noiseless channel. The adjacency of $\boldsymbol{x}$ and $\boldsymbol{y}$ in $[\mathrm{cl}(\vec{G}_H)]^{\wedge t}$ means that for every i there is a $v_i \in V(\mathrm{cl}(\vec{G}_H))$ such that both x_i and y_i can result in v_i when sent through the noisy channel H. (The reason of this can be that $x_i = y_i = v_i$ or that (x_i, y_i) is an edge of $\vec{G}_H$, in which case $v_i = y_i$ or (y_i, x_i) is an edge of $\vec{G}_H$ and $v_i = x_i$ or we have $(x_i, v_i), (y_i, v_i) \in E(\vec{G}_H)$, where v_i differs from both x_i and y_i.) Thus if Bob's original decoding of Alice's (first) message was $\boldsymbol{v} = (v_1, \dots, v_t)$ then he knows that the message sent could be either of $\boldsymbol{x}$ or $\boldsymbol{y}$. Since Alice's second message for $\boldsymbol{x}$ is identical to that for $\boldsymbol{y}$, Bob will not know even after receiving the second message whether the original message was $\boldsymbol{x}$ or $\boldsymbol{y}$.

On the other hand, if the length of Alice's second message is at least $\log \chi([\mathrm{cl}(\vec{G}_H)]^{\wedge t})$ then Alice can make Bob able to decide for sure what the original message was. Indeed, fix a proper coloring of $[\mathrm{cl}(\vec{G}_H)]^{\wedge t}$ with $\chi([\mathrm{cl}(\vec{G}_H)]^{\wedge t})$ colors in advance that is known by both parties. Encode each color by a (distinct) sequence of $\lceil \log \chi([\mathrm{cl}(\vec{G}_H)]^{\wedge t}) \rceil$ bits. If the original message was $\boldsymbol{z} = (z_1, \dots, z_t)$ then send Bob the (codeword for the) color of $\boldsymbol{z}$. Since as a vertex of $[\mathrm{cl}(\vec{G}_H)]^{\wedge t}$ $\boldsymbol{z}$ is connected to all those sequences that could result in the same sequence when sent through H what $\boldsymbol{z}$ can result in, all these sequences have a different color than $\boldsymbol{z}$ in our coloring of $[\mathrm{cl}(\vec{G}_H)]^{\wedge t}$. Thus when Bob gets to know the color of $\boldsymbol{z}$ from Alice's second message he will know that whatever he saw at the output of H could only arise from $\boldsymbol{z}$ as the input. So he will decode $\boldsymbol{z}$ with zero-error.

Thus we proved that

$$h_c(t) = \lceil \log \chi([\mathrm{cl}(\vec{G}_H)]^{\wedge t}) \rceil.$$

So $\lim_{t\to\infty} \frac{h_c(t)}{t} = \lim_{t\to\infty} \frac{1}{t} \log \chi([\mathrm{cl}(\vec{G}_H)]^{\wedge t}) = R_{\mathrm{W}}(\mathrm{cl}(\vec{G}_H))$ as stated. ■

C. What graphs can be closure graphs?

Not every graph can appear as the closure graph $\mathrm{cl}(\vec{G})$ of some directed graph $\vec{G}$.

Proposition 23: Let G be a(n undirected) bipartite graph with $|E(G)| \geq |V(G)| + 1$. Then G cannot be the closure graph of any directed graph.

Proof: Let $\mathrm{cl}(\vec{F})$ be the closure graph of a directed graph $\vec{F}$. Observe that if $\mathrm{cl}(\vec{F})$ has an edge e connecting two vertices that were not adjacent (in either direction) in $\vec{F}$, then e is contained in a triangle in $\mathrm{cl}(\vec{F})$. Let G be a bipartite graph with more edges than vertices. By bipartiteness G contains no triangle, so if it is a closure graph of some graph $\vec{G}$, then $\vec{G}$ is just a directed version of G. If any vertex has indegree at least 2 in $\vec{G}$ that would generate a triangle in $\mathrm{cl}(\vec{G})$, so the closure graph could not be G itself. Since the sum of indegrees equals the number of edges, we cannot avoid having a vertex with indeegree at least two if $|E(G)| > |V(G)|$. This proves the statement. ■

To give a complete characterization of those graphs that can arise as a closure graph seems tedious and complicated. It is certainly not a family of graphs possessing the nice property that it would be closed under taking induced subgraphs. In fact, the following statement is true.

Proposition 24: For any finite simple undirected graph G, there exists a directed graph $\vec{F}$ such that $\mathrm{cl}(\vec{F})$ contains G as an induced subgraph.

Proof: Let G be an arbitrary finite simple undirected graph. For every edge $e = \{a, b\} \in E(G)$ consider a new vertex v_e. We add the oriented edges (a, v_e) and (b, v_e) to our graph G. Now delete the edges of G thus obtaining a graph $\vec{F}$ on vertex set $V(G) \cup \{v_e : e \in E(G)\}$ containing only the $2|E(G)|$ oriented edges leading to some vertex v_e. It is straightforward to see, that $\mathrm{cl}(\vec{F})$ contains graph G as an induced subgraph. ■

VII. Open problems

The general problem concerning the Dilworth rate is to determine it for specific directed graphs. Since this is a difficult and mostly open problem to the related notions of Shannon and Sperner capacities as well as for the Witsenhausen rate, we cannot expect that this problem is easy. Nevertheless, we have seen some digraphs for which it was solvable (at least when using some non-trivial results already established for Sperner capacity). Still, there are some directed graphs for which determining the Dilworth rate seems particularly interesting.

Problem 1: What is the Dilworth rate of the graph $\vec{A}_5^c$ we presented in Subsection III-B? Recall that we know $\sqrt{5} \leq r_D(\vec{A}_5^c) \leq \sqrt{6}$.

Tournaments play a special role in our setting, because they are exactly those oriented graphs the complement of which is also an oriented graph (that is one without bidirected edges). So it may have some particular interest how their Dilworth rate behave.

Problem 2: Is there a tournament $\vec{T}$ for which $r_D(\vec{T})$ is strictly smaller than $\chi_{\mathrm{dir},f}(\vec{T})$?

ACKNOWLEDGEMENT

Useful discussions with Imre Csiszár are gratefully acknowledged. We also thank two anonymous referees for their thorough reading of our manuscript and their detailed comments.

REFERENCES

[1] N. Alon, On the capacity of digraphs, *European J. Combin.*, 19 (1998), 1–5.
[2] N. Alon, C. McDiarmid, B. Reed, Acyclic coloring of graphs, *Random Structures and Algorithms*, 2 (1991), 277–288.
[3] N. Alon and A. Orlitsky, Source coding and graph entropies, *IEEE Trans. Inform. Theory*, 42 (1996), 1329-1339.
[4] C. Berge, *Graphs and Hypergraphs*, North-Holland, Amsterdam, 1973.
[5] A. Blokhuis, On the Sperner capacity of the cyclic triangle, *J. Algebraic Combin.*, 2 (1993), 123–124.
[6] B. Bollobás, On generalized graphs, *Acta Math. Acad. Sci. Hungar.*, 16 (1965), 447–452.
[7] C. Bunte, A. Lapidoth, A. Samorodnitsky, The zero-undetected-error capacity approaches the Sperner capacity, *IEEE Trans. Inform. Theory*, 60 (2014), 3825–3833; arXiv:1309.4930 [cs.IT].
[8] R. Calderbank, P. Frankl, R. L. Graham, W. Li, L. Shepp, The Sperner capacity of the cyclic triangle for linear and non-linear codes, *J. Algebraic Combin.*, 2 (1993), 31–48.
[9] I. Csiszár, The method of types, *IEEE Trans. Inform. Theory*, Vol. 44, No. 6 (October 1998, commemorative issue), 2505–2523.
[10] I. Csiszár and J. Körner, On the capacity of the arbitrarily varying channel for maximum probability of error, *Z. Warsch. Verw. Gebiete*, 57 (1981), 87–101.
[11] I. Csiszár and J. Körner, *Information Theory. Coding Theorems for Discrete Memoryless Systems*, Second Edition, Cambridge University Press, Cambridge, 2011.
[12] I. Csiszár and P. Narayan, Channel capacity for a given decoding metric, *IEEE Trans. Inform. Theory*, 41 (1995), 35–43.
[13] R. P. Dilworth, A decomposition theorem for partially ordered sets, *Annals of Mathematics*, 51 (1950), 161–166.
[14] E. Fachini and J. Körner, Colour number, capacity, and perfectness of directed graphs, *Graphs Combin.*, 16 (2000), 389–398.
[15] A. Galluccio, L. Gargano, J. Körner, G. Simonyi, Different capacities of a digraph, *Graphs Combin.*, 10 (1994), 105-121.
[16] L. Gargano, J. Körner, U. Vaccaro, Sperner theorems on directed graphs and qualitative independence, *J. Combin. Theory Ser. A*, 61 (1992), 173–192.
[17] L. Gargano, J. Körner, U. Vaccaro, Capacities: from information theory to extremal set theory, *J. Combin. Theory Ser. A*, 68 (1994), 296–316.
[18] R. Hammack, W. Imrich, S. Klavžar, *Handbook of Product Graphs. Second Edition*, CRC Press, Boca Raton, 2011.
[19] J. Kahn and J. H. Kim, Entropy and sorting, *J. Comput. Syst. Sci.*, 51 (1995), 390–399.
[20] J. Körner, Coding of an information source having ambiguous alphabet and the entropy of graphs, in: *Transactions of the 6th Prague Conference on Information Theory, etc.*, 1971, Academia, Prague, (1973), 411–425.
[21] J. Körner, On clique growth in products of directed graphs, *Graphs Combin.*, 14 (1998), 25–36.
[22] J. Körner and G. Longo, Two-step encoding of finite memoryless sources, *IEEE Trans. Inform. Theory*, 19 (1973), 778-782.
[23] J. Körner, C. Pilotto, G. Simonyi, Local chromatic number and Sperner capacity, *J. Combin. Theory Ser B*, 95 (2005), 101–117.
[24] L. Lovász, On the ratio of optimal integral and fractional covers, *Discrete Math.*, 13 (1975), 383–390.
[25] L. Lovász, On the Shannon capacity of graphs, *IEEE Trans. Inform. Theory*, 25 (1979), 1–7.
[26] L. Lovász, *Combinatorial Problems and Exercises. Second Edition*, North-Holland, Amsterdam, 1993.
[27] K. Marton, On the Shannon capacity of probabilistic graphs, *J. Combin. Theory Ser. B*, 57 (1993), 183–195.
[28] L. Mirsky, A dual of Dilworth's decomposition theorem, *American Mathematical Monthly*, 78 (1971), 876–877.
[29] V. Neumann-Lara, The dichromatic number of a digraph, *J. Combin. Theory Ser. B.*, 33 (1982), 265–270.
[30] J. Nayak and K. Rose, Graph capacities and zero-error transmission over compound channels, *IEEE Trans. Inform. Theory*, 51 (2005), 4374–4378.
[31] A. Sali and G. Simonyi, Orientations of self-complementary graphs and the relation of Sperner and Shannon capacities, *European J. Combin.*, 20 (1999), 93–99.
[32] E. R. Scheinerman and D. H. Ullman, *Fractional Graph Theory*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley and Sons, Chichester, 1997.
[33] C. E. Shannon, The zero error capacity of a noisy channel, *IRE Trans. Inform. Theory*, 2 (1956), 8–19.
[34] G. Simonyi, Graph entropy: a survey, in: *Combinatorial Optimization*, (W. Cook, L. Lovász, P. Seymour eds.), DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Volume 20, AMS, Providence, RI, 1995, 399-441.
[35] G. Simonyi, Perfect Graphs and Graph Entropy. An Updated Survey, Chapter 13 in: *Perfect Graphs* (J. Ramirez-Afonsín, B. Reed, eds.), John Wiley and Sons, 2001, 295–328.
[36] G. Simonyi, On Witsenhausen's zero-error rate for multiple sources, *IEEE Trans. Inform. Theory*, 49 (2003), 3258-3261.
[37] E. Sperner, Ein Satz über Untermengen einer endlichen Menge, *Mathematische Zeitschrift*, 27 (1928), 544–548, (in German).
[38] H. S. Witsenhausen, The zero-error side-information problem and chromatic numbers, *IEEE Trans. Inform. Theory*, 22 (1976), 592–593.

Gábor Simonyi received his PhD in Mathematics from the Hungarian Academy of Sciences (HAS) in 1991 and the Doctor of HAS degree in 2009. He has been with the Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences since 1989. He is also professor at the Budapest University of Technology and Economics. In 1986 and 1987–88 he spent altogether eight months at the École Nationale Supérieure des Télécommunications in Paris. In the academic year 1992–93, he was a DIMACS postdoc at Rutgers University, New Jersey. He has held visiting appointments at Simon Fraser University in Vancouver, Canada and at Sapienza University in Rome. His research interest includes zero-error information theory, graph theory, and extremal combinatorics.

Ágnes Tóth received her PhD degree in Mathematics from the Budapest University of Technology and Economics. At the time of this research she was a junior researcher at the Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences.