MAX-LINEAR MODELS ON INFINITE GRAPHS
GENERATED BY BERNOULLI BOND PERCOLATION

C. KLÜPPELBERG AND E. SÖNMEZ

Abstract. We extend previous work of max-linear models on finite directed acyclic
graphs to infinite graphs, and investigate their relations to classical percolation theory.
We formulate results for the oriented square lattice graph $\mathbb{Z}^2$ and nearest neighbor bond
percolation. Focus is on the dependence introduced by this graph into the max-linear
model. As a natural application we consider communication networks, in particular, the
distribution of extreme opinions in social networks.

1. Introduction

Extreme value theory is concerned with max-stable random elements which occur as
limits of normalized maxima. The theory has progressed in recent years from classical finite
models to infinite-dimensional models (see, for example, [7, 15, 16]). A monograph relevant
in the infinite-dimensional context is [5]. Prominent models are stochastic processes in
space and/or time having finite dimensional max-stable marginal distributions (cf. [4,
8, 12]). Such processes model extreme dependence between process values at different
locations and/or time points.

Max-linear models are natural analogues of linear models in an extreme value frame-
work. Within the class of multivariate extreme value distributions, whose dependence
structures are characterized by a measure on the sphere, they are characterized by the
fact that this measure is discrete (cf. [17]).

We extend previous work of max-linear models on finite directed acyclic graphs (cf. [9,
10, 13]) to infinite graphs. The model allows for finite subgraphs with different dependence
structures, and we envision applications where this may play a role as, for instance, a
hierarchy of communities with different communication structures.

We investigate the relation of the infinite max-linear model to classical percolation
theory, more precisely to nearest neighbor bond percolation (cf. [3, 11]). We focus on
the square lattice $\mathbb{Z}^2$ with edges to the nearest neighbors, where we orient all edges in a
natural way (north-east) resulting in a directed acyclic graph (DAG) on this lattice. On
this infinite DAG a random sub-DAG may be constructed by choosing nodes and edges
between them at random. In a Bernoulli bond percolation DAG edges are independently
declared open with probability $p \in (0, 1)$ and closed otherwise. The random graph consists
then of the nodes and the open edges. The percolation probability is the probability
$P_p(|C(i)| = \infty)$ that a given node $i$ belongs to an infinite open cluster $C(i)$, which is 0 if

Date: April 17, 2018.
2010 Mathematics Subject Classification. Primary 60G70; Secondary 60K35, 62–09.
Key words and phrases. Bernoulli bond percolation, extreme value theory, graphical model, infinite
graph, percolation, recursive max-linear model.
$p \leq 1/2$ and positive for $p > 1/2$. Kolmogorov’s zero-one law entails that an infinite open cluster exists for $p > 1/2$ with probability 1, and otherwise with probability 0.

We combine percolation theory with an infinite max-linear model by assigning to each node a max-stable random variable. Sampling a random graph by Bernoulli bond percolation, we use this subgraph for modelling the dependence in the max-stable process on the oriented square lattice. The max-linear models we envision are recursively constructed from independent standard $\alpha$-Fréchet random variables $(Z_j)_{j \in \mathbb{Z}^2}$ for $\alpha > 0$, which are one-dimensional max-stable. More precisely, each random variable $X_i$ on a node $i \in \mathbb{Z}^2$ with ancestral set $\text{an}(i)$ exhibits the property

$$X_i = \bigvee_{j \in \{i\} \cup \text{an}(i)} b_{ji} Z_j,$$

in distribution on every finite DAG, where $b_{ji}$ are positive coefficients. As this model is defined on a random graph it is a max-linear model in random environment. To the best of our knowledge, it is the first such model considered in extreme value theory.

One prerequisite for this work is the fact that max-stable random variables on different nodes of a DAG are independent if and only if they have no common ancestors ([10, Theorem 2.3]). As a consequence of this and percolation theory we find for the subcritical case where $p < 1/2$ that two random variables become independent with probability 1, whenever their distance tends to infinity. In contrast, for the supercritical case where $p > 1/2$ two random variables are dependent with positive probability, even when their node distance tends to infinity.

Finally, we consider changes in the dependence properties of random variables on a sub-DAG $H$ of a finite or infinite graph on the oriented square lattice $\mathbb{Z}^2$, when enlarging this subgraph. The method of enlargement consists of adding nodes and edges of Bernoulli bond percolation clusters. Here we start with $X_i$ and $X_j$ independent in $H$, and answer the question, whether they can become dependent in the enlarged graph. We evaluate critical probabilities such that $X_i$ and $X_j$ become dependent in the enlarged graph with positive probability or with probability 1. We find in particular that for every DAG $H$ with finite number of nodes, in the enlarged graph $X_i$ and $X_j$ remain independent with positive probability. On the other hand, if $H$ has nodes $\mathbb{Z}^2$ and percolates everywhere; i.e. every connected component of $H$ is infinite, then $X_i$ and $X_j$ become dependent with probability 1 in the enlarged graph.

Our paper is organised as follows. In Section 2 we introduce recursive max-linear models on DAGs in $\mathbb{Z}^2$. In particular, we show that by Kolmogorov’s extension theorem, finite-dimensional max-linear models can be extended to infinite-dimensional models. Section 3 uses the fact that the max-linear coefficients $b_{ji}$ originate from an algebraic path analysis by multiplying edge weights along a path between nodes $j$ and $i$ with $j$ being an ancestor of $i$. This concept, known from finite recursive max-linear models, extends to infinite DAGs. Example 3.1 shows that the important class of max-weighted models can be extended from finite to infinite graphs such that the max-weighted property remains. Recursive max-linear processes on a DAG have the nice property that independence between random variables on two different nodes is characterized by their ancestral sets. This is the starting point of our investigation. Section 4 contains the dependence results. Here we investigate the Bernoulli bond percolation DAGs. In Section 4.1 we prove that nearest neighbor
bond percolation on $\mathbb{Z}^2$ yields independence of $X_i$ and $X_j$ with probability 1 provided $|i - j| \to \infty$ for $p < 1/2$ and dependence with positive probability for $p > 1/2$. In Section 4.2 we investigate for $X_i$ and $X_j$, which are independent in some subgraph $H$, whether enlargement of $H$ can result in dependence between $X_i$ and $X_j$.

2. Max-linear processes on directed acyclic lattice graphs

This section presents a description of infinite max-linear models on directed acyclic lattice graphs. We first explain the structure of the directed graph on a lattice before we define and show the existence of a random field with finite-dimensional distributions entailing a dependence structure of max-linear type encoded in such graphs.

2.1. Graph notation and terminology. Let $\mathbb{Z}^2$ be the oriented square lattice as follows (cf. [1, 3, 6, 11]). We write $i = (i_1, i_2)$ for elements in $\mathbb{Z}^2$ and refer to them as nodes. The distance from $i$ to $j$ is defined as

$$\delta(i, j) = |i_1 - j_1| + |i_2 - j_2|$$

for $i, j \in \mathbb{Z}^2$. We regard $\mathbb{Z}^2$ as a graph by adding edges between all nodes $i, j$ with $\delta(i, j) = 1$. In addition, we assume the edges to be oriented in the following manner. Denote by $\text{pa}(i)$ and $\text{ch}(i)$ the parents and children of node $i = (i_1, i_2)$, respectively. Then $j = (j_1, j_2) \in \text{pa}(i)$ if and only if either $(j_1, j_2) = (i_1 - 1, i_2)$ or $(j_1, j_2) = (i_1, i_2 - 1)$ and, consequently, $j = (j_1, j_2) \in \text{ch}(i)$ if and only if either $(j_1, j_2) = (i_1 + 1, i_2)$ or $(j_1, j_2) = (i_1, i_2 + 1)$. We may write $i \rightarrow j$ if there is a directed edge from $i$ to $j$, that is if $i$ is a parent of $j$. The set of edges in this oriented lattice $\mathbb{Z}^2$ is $E(\mathbb{Z}^2)$, which is a subset of $\mathbb{Z}^2 \times \mathbb{Z}^2$. In this paper we work with graphs $G = (V(G), E(G))$ with $V(G) \subset \mathbb{Z}^2$ and $E(G) \subset E(\mathbb{Z}^2)$, which are directed acyclic lattice graphs. We refer to them simply as DAGs. When there is no ambiguity, we often abbreviate $V = V(G)$ and $E = E(G)$. Thus, every node $i \in V$ has at most two children and two parents, but possibly infinitely many descendants and ancestors, denoted by $\text{de}(i)$ and $\text{an}(i)$, respectively. Moreover, we define $\text{De}(i) = \{i\} \cup \text{de}(i)$ and $\text{An}(i) = \{i\} \cup \text{an}(i)$. Note that such a DAG may have no roots, which proves relevant for the questions we want to answer.

2.2. Infinite recursive max-linear models. We now introduce recursive max-linear processes. Let $G = (V(G), E(G))$ be a DAG with some possibly infinite set of nodes $V(G) \subset \mathbb{Z}^2$ and let $H = (V(H), E(H)) \subset G$ be a finite sub-DAG, that is to say $|V(H)| < \infty$.

**Definition 2.1.** (a) A family of random variables $X := \{X_i : i \in V(G)\}$ is called a recursive max-linear process if for every finite sub-DAG $H$ there exists a matrix $B = B(H)$ with non-negative entries such that the random vector $(X_i : i \in V(H))$ is a recursive max-linear model on $H$ as defined in (b).

(b) Let $H$ be a finite DAG, then a recursive max-linear model on $H$ is defined as

$$X_i = \bigvee_{j \in \text{An}(i)} b_{ij} Z_j, \quad i \in V(H),$$

where $(Z_j)_{j \in V(H)}$ are independent non-negative noise variables with infinite support on $(0, \infty)$ and a max-linear coefficient matrix $B = (b_{ij})_{i, j \in V(H)}$ with non-negative entries.
Note that Definition 2.1(b) coincides with the definition in [9] and thus Definition 2.1(a) can be seen as an extension of the latter to models on infinite DAGs.

We now prove the existence of a stochastic process with the dependence structure described by infinite recursive max-linear processes as in Definition 2.1(a).

**Lemma 2.2.** There exists a stochastic process \( X = \{X_i : i \in V(G)\} \) with the finite-dimensional distributions of a recursive max-linear process as in Definition 2.1.

**Proof.** For a given finite sub-DAG \( H \) with nodes \( V(H) = \{i^1, \ldots, i^d\} \subset \mathbb{Z}^2 \) let \( X = (X_{i^1}, \ldots, X_{i^d}) \) be a recursive max-linear model on \( H \) and let

\[
B_{i^1, \ldots, i^d} = \begin{pmatrix}
  b_{i^1, i^1} & \cdots & b_{i^1, i^d} \\
  \vdots & \ddots & \vdots \\
  b_{i^d, i^1} & \cdots & b_{i^d, i^d}
\end{pmatrix}
\]

be its max-linear coefficient matrix according to [9, Theorem 2.2], that is for every \( i^k \) the random variable \( X_{i^k} \) admits the representation

\[
X_{i^k} = \bigvee_{j \in \text{An}(i^k)} b_{j, i^k} Z_j
\]

with noise variables \( Z_{i^1}, \ldots, Z_{i^d} \), where the corresponding set of ancestors is taken with respect to \( H \). In addition, assume that \( Z_{i^1}, \ldots, Z_{i^d} \) are standard \( \alpha \)-Fréchet distributed noise variables. Then [10, Proposition A.2] identifies the distribution function of \( (X_{i^1}, \ldots, X_{i^d}) \) as

\[
G_{i^1, \ldots, i^d}(x_1, \ldots, x_d) = \exp \left( -\sum_{j \in \text{An}(i^k) \cup \ldots \cup \text{An}(i^d)} \left( \frac{b_{j, i^k}}{x_1} \right)^\alpha \cdots \left( \frac{b_{j, i^d}}{x_d} \right)^\alpha \right).
\]

Note that \( G_{i^1, \ldots, i^d}(x_1, \ldots, x_{d-1}, \infty) = G_{i^1, \ldots, i^{d-1}}(x_1, \ldots, x_{d-1}) \) for all \( x_j > 0 \) and \( i^j \in \mathbb{Z}^2 \) and the latter relation is invariant to permutations, i.e.,

\[
G_{i^{(1)}, \ldots, i^{(d)}}(x_{\pi(1)}, \ldots, x_{\pi(d)}) = G_{i^1, \ldots, i^d}(x_1, \ldots, x_d)
\]

for all permutations \( \pi \) of \( \{1, \ldots, d\} \). Thus, by Kolmogorov’s extension theorem there exists a stochastic process \( \{X_i : i \in V(G)\} \) with finite-dimensional distributions as in equation (2.2). This finishes the proof. \( \square \)

Different blocks of the matrix \( B \) may correspond to distinct communities with different communication structure. The random variables \( X_i \) may correspond to extreme events like extreme opinions expressed at node \( i \) in a social network like Twitter. As there may be different paths leading to \( X_i \) with different coefficients \( (b_{ji})_{j \in \text{An}(i)} \), different opinions may arrive at node \( i \).

For the sake of completeness, we state the following limit result, which can be found in [17, Lemma 2.1(iv)].

**Remark 2.3.** If \( (Z_j)_{j \in \mathbb{Z}^2} \) are independent standard \( \alpha \)-Fréchet random variables and \( (V_n, E_n)_{n \in \mathbb{N}} \) is a sequence of finite sub-DAGs of the oriented square lattice \( \mathbb{Z}^2 \) then from Lemma 2.2 we know that

\[
X_i^{(n)} = \bigvee_{j \in V_n} b_{ji} Z_j, \quad i \in V_n,
\]
Remark 2.4. The finite max-linear coefficient matrix $B$ in Definition 2.1(b) can be calculated from the positive weights $c_{ki} = c_{ki}(H)$, $i \in V(H), k \in \{i\} \cup \text{pa}(i)$, assigned to the edges of $H$, for every given finite DAG $H$ by a simple path analysis. The max-linear coefficient $b_{ji}$ is positive if and only if $j \in \text{An}(i)$ and its value is the maximum of products along all directed paths between $j$ and $i$; cf. [9, Theorem 2.2].

For the infinite DAG $G$, applying this path analysis we immediately see that the max-linear coefficient matrix heavily depends on the chosen finite sub-DAG $H$. Thus for distinct sub-DAGs we obtain different coefficient matrices. In general we cannot identify a recursive max-linear process with a unique max-linear coefficient matrix. In particular, this is not possible if the chosen DAG $G$ has no roots. Therefore, in the next section we first treat the case that $V(G) \subset \mathbb{N}^2_0$, so that every node has at most finitely many ancestors.

3. Infinite coefficient matrices

As motivated in Remark 2.4 we first consider infinite DAGs on $\mathbb{N}^2_0$, which we view as a prototypical sub-DAG with infinitely many nodes of the oriented square lattice $\mathbb{Z}^2$, such that each node has at most finitely many ancestors.

3.1. Infinite max-linear coefficient matrix. Let $G = (V, E)$ be a DAG with $V \subset \mathbb{N}^2_0$ and corresponding edges $E$. Assume a recursive max-linear process $X = \{X_i : i \in V\}$ on $G$. In the following the aim is to give a canonical choice of a possible max-linear coefficient matrix $B$ representing the dependence structure of $X$ and to provide characterizations.

Assume that the edges of $G$ are equipped with positive weights $c_{ki}$ for every $i \in V$ and $k \in \{i\} \cup \text{pa}(i)$. For $n \in \mathbb{N}$ let $G_n = (V_n, E_n)$ be the DAG with nodes $V_n = \{i = (i_1, i_2) \in V : i_1 + i_2 \leq n\}$ and corresponding edges taken from $E$, so that $\lim_{n \to \infty} G_n = G$. By Definition 2.1(b) there are independent non-negative noise variables $(Z_i)_{i \in V_n}$ with infinite support on $(0, \infty)$ and a max-linear coefficient matrix $B = (b_{ij})_{i,j \in V_n}$ with non-negative entries such that $X_i^{(n)}$ as in (2.3). Indeed the entries $b_{ji}$ may be derived from the path analysis mentioned in Remark 2.4. This in particular shows that for $i \in V$ the $b_{ji}$ do not depend on the descendants $\text{de}(i)$. Thus, an infinite max-linear coefficient matrix $B$ is built up from increasing finite blocks representing $V_n$ for increasing $n \in \mathbb{N}$. 

has a $\alpha$-Fréchet distribution with scale parameter $(\sum_{j \in V_i} b_{ji}^\alpha)^{1/\alpha}$. Suppose that the sequence of DAGs $(V_n, E_n)_{n \in \mathbb{N}}$ tends to a DAG $(V, E)$ with infinitely many nodes as $n \to \infty$. Then

$$X_i^{(n)} \xrightarrow{a.s.} X_i, \quad n \to \infty,$$

where $X_i$ has a $\alpha$-Fréchet distribution with scale parameter $(\sum_{j \in V_i} b_{ji}^\alpha)^{1/\alpha} < \infty$. If this series diverges then $X_i^{(n)} \xrightarrow{a.s.} \infty$ as $n \to \infty$.
For a communication network on \( \mathbb{N}_0^2 \) the representation (2.3) reduces to a maximum over finitely many random variables, for instance, the opinion of the root 0 influences all opinions in the network. Hence, if the root node happens to hold the maximum of all \( Z_j \) for \( j \in \mathbb{N}_0^2 \) it may dominate the opinion of the whole network, although by the max-linear coefficient matrix \( B \) the opinions of all other nodes may have different realisations.

We consider in Section 4 percolation (dependence) properties between two fixed nodes \( i \) and \( j \) on \( \mathbb{Z}^2 \). Hence, although the underlying graph is infinite, we can always find a finite graph \( (V_n, E_n) \) as above which contains \( i \) and \( j \). As a consequence, although the matrix \( B \) may be an infinite matrix, we only need finite submatrices of it.

As there may be several paths between nodes with different path-weights, so-called max-weighted models with same paths-weights along all possible directed paths between two nodes play an important role. We now give an example of such a max-linear process relying on the definition of max-weighted models presented in [9, Definition 3.1] and discussed in [10, Section 3]. Resulting as a limit of max-weighted paths, we may call such a process max-weighted. In such a model, the same opinion reaches node \( i \) regardless of the path it takes within the network. This means that it suffices to consider one path, for instance, that from the root 0 to every other node.

**Example 3.1** (Max-weighted process). Let \( V = \mathbb{N}_0^2 \) be the set of nodes and assume oriented edges between all nodes \( i, j \) with \( \delta(i, j) = 1 \). Start with a subgraph in which the set of nodes is bounded and of the form \( V_n = \{(i_1, i_2) \in \mathbb{N}_0^2 : i_1 + i_2 \leq n \} \) for some \( n \in \mathbb{N}_0 \) and the corresponding set of edges is denoted by \( E_n \). Assume that the corresponding model is max-weighted so that every entry of the max-linear coefficient matrix is given by \( b_{ji} = d_p((j_1, j_2), (i_1, i_2)) \), where \( d_p((j_1, j_2), (i_1, i_2)) \) is calculated by a path analysis along the edge-weights as in equation (1.5) in [9]. Since the model is max-weighted, \( d_p((j_1, j_2), (i_1, i_2)) \) is the same value for every path \( p \) from \( i \) to \( j \) and thus we can write \( d_p((j_1, j_2), (i_1, i_2)) = d((j_1, j_2), (i_1, i_2)) \), since the latter value is independent of the chosen path \( p \). We now show that the DAG can be enlarged in such a way that the enlarged new subgraph is again max-weighted. Moreover, this procedure can be executed infinitely often. Let \( n \geq 1 \) and assume that we add a node, say \( (\ell_1, \ell_2) \) which we connect with the nodes \( (\ell_1 - 1, \ell_2) \) and \( (\ell_1, \ell_2 - 1) \) in \( V \) by two edges with corresponding weights \( c((\ell_1 - 1, \ell_2), (\ell_1, \ell_2)) \) and \( c((\ell_1, \ell_2 - 1), (\ell_1, \ell_2)) \). By choosing these appropriately we can ensure that the new model is again max-weighted. More precisely, we choose the weights satisfying

\[
c((\ell_1 - 1, \ell_2), (\ell_1, \ell_2)) = \frac{c((\ell_1, \ell_2 - 1), (\ell_1, \ell_2)) \cdot d((0, 0), (\ell_1 - 1, \ell_2))}{d((0, 0), (\ell_1, \ell_2 - 1))}.
\]

We now show that the enlarged DAG again leads to a max-weighted model. Let \( p_1 \) be a path from the root to \( (\ell_1, \ell_2) \) containing \( (\ell_1 - 1, \ell_2) \) and let \( p_2 \) be such a path containing the node \( (\ell_1, \ell_2 - 1) \). Then we have by definition

\[
d_{p_1}((1, 1), (\ell_1, \ell_2)) = d((1, 1), (\ell_1, \ell_2 - 1)) \cdot c((\ell_1 - 1, \ell_2), (\ell_1, \ell_2))
\]

\[
= c((\ell_1, \ell_2 - 1), (\ell_1, \ell_2)) \cdot d((1, 1), (\ell_1 - 1, \ell_2))
\]

\[
= d_{p_2}((1, 1), (\ell_1, \ell_2)).
\]
Thus every path from the root to \((\ell_1, \ell_2)\) is max-weighted and this shows that the new model is max-weighted.

Next we return to DAGs on \(\mathbb{Z}^2\), which allow for infinitely many ancestors.

### 3.2. Infinite tail dependence coefficient matrix

In this section we let \(G = (V, E)\) be an arbitrary, possibly infinite DAG with nodes \(V \subset \mathbb{Z}^2\) and oriented edges \(E\). Furthermore, let \(X\) be a recursive max-linear process on \(G\) as in Definition 2.1(a).

As in [10] for \(i \in V\) we denote the distribution function of \(X_i\) by \(F_i\) and its generalized inverse by \(F_i^{-}\), so that the tail dependence coefficient of \(X_i\) and \(X_j\) is given by

\[
\chi(i, j) = \chi(j, i) = \lim_{u \uparrow 1} P\left( X_i > F_i^{-}(u) \mid X_j > F_j^{-}(u) \right), \quad i, j \in V.
\]

A tail dependence coefficient is positive if the positive random variables \(X_i, X_j\) are heavy-tailed; more precisely, if the bivariate distribution function is regularly varying; cf. [15, 16]; otherwise, it is zero. We summarize these coefficients in a matrix \((\chi(i, j))_{i,j \in V}\), which we call the infinite tail dependence coefficient matrix. We assume from now on that the max-linear model \(X\) is recursively constructed from independent standard \(\alpha\)-Fréchet noise variables \((Z_j)_{j \in \mathbb{Z}^2}\) for \(\alpha > 0\), which are one-dimensional max-stable. Under this condition the following result is an analogue to [10, Theorem 2.3] and its proof justifies the extension of the arguments to infinite dimension.

**Proposition 3.2.** Let \((Z_j)_{j \in \mathbb{Z}^2}\) be a sequence of independent standard \(\alpha\)-Fréchet random variables for \(\alpha > 0\). The following statements are equivalent.

(i) \(X_i\) and \(X_j\) are independent.

(ii) \(\text{An}(i) \cap \text{An}(j) = \emptyset\).

(iii) \(\chi(i, j) = 0\).

**Proof.** We first show the equivalence of (i) and (ii). Note that \(X_i\) and \(X_j\) are independent if and only if they are independent on every finite sub-DAG \(H\) of \(G\). But by [10, Theorem 2.3] this is the case if and only if \(\text{An}(i) \cap \text{An}(j) \cap V(H) = \emptyset\). Since \(H\) is an arbitrary finite DAG, this is equivalent to \(\text{An}(i) \cap \text{An}(j) = \emptyset\).

Now we prove that (ii) \(\Rightarrow\) (iii). The implication (iii) \(\Rightarrow\) (ii) is proven in a similar manner. So let

\[
B(n) = [-n, n]^2
\]

be the box with centre at 0 and side-length \(2n\), where \(n \in \mathbb{N}\) is large enough such that \(i, j \in B(n)\). Moreover, let \(H(n)\) be the DAG with nodes \(V(H) = B(n) \cap V\) and corresponding edges \(E(H(n))\). We denote by \(\chi_n(i, j)\) the tail dependence coefficient between the components \(Y_i\) and \(Y_j\) of the finite recursive max-linear model \((Y_k : k \in H(n))\) on the DAG \(H(n)\). Assume that \(\text{An}(i) \cap \text{An}(j) = \emptyset\). This in particular implies \(\text{An}(i) \cap \text{An}(j) \cap B(n) = \emptyset\).

By [10, Theorem 2.3] it follows that \(\chi_n(i, j) = 0\) for all sufficiently large \(n \in \mathbb{N}\), entailing that \(\chi(i, j) = 0\). \(\square\)

Having introduced the infinite tail dependence coefficient matrix \(B\) we are now interested in the following. We use Bernoulli bond percolation to generate random DAGs on the oriented square lattice \(\mathbb{Z}^2\) and, thus, random tail dependence coefficient matrices.
As for the communication example mentioned above, we want to answer the following question: given an extreme opinion (fake news) in a community, observed at two nodes $i$ and $j$, is there a common cause in the network (a common ancestor) or not.

4. Bernoulli bond percolation DAGs

The main purpose of this section is to construct max-linear models on randomly obtained DAGs with a possibly infinite number of nodes in order to investigate a randomized dependence structure.

In view of Proposition 3.2 the probability that random variables $X_i$ and $X_j$ on the random graph are dependent is nothing else than the probability that $i$ and $j$ have common ancestors inside the random open cluster containing nodes $i$ and $j$.

4.1. Max-linear models on random open clusters. Recall that we consider the oriented square lattice $\mathbb{Z}^2$. For this oriented model, the open cluster at 0 is usually defined as the set of all points we can reach from the origin by travelling along open edges in the direction of the orientation; see [1, 6], or [11, Section 12.8]. As this open cluster always has root 0, all nodes $i$ and $j$ would have at least common ancestor 0, and would make the problem discussed below trivial. Consequently, we consider unoriented, but not undirected, paths in (4.2) as we will make precise below.

Let us first recall the framework of Bernoulli bond percolation from any book on percolation as e.g. [3, 11]. Given the oriented square lattice $\mathbb{Z}^2$ with edge set $E \subset \mathbb{Z}^2 \times \mathbb{Z}^2$, a (bond) configuration is a function $\omega : E \to \{0, 1\}$, $e \mapsto \omega_e$. An edge $e$ is open in the configuration $\omega$, if and only if $\omega_e = 1$, so configurations correspond to open subgraphs. Recall from Section 2 that in our setting open edges are directed, hence a configuration is a DAG denoted by $(V, E)$ with $V \subset \mathbb{Z}^2$ and directed edges $E$. Each edge is declared open with probability $p$ and closed otherwise, different edges having independent designations. This gives the Bernoulli measure $P_p$, $p \in (0, 1]$ on the space $\Omega = \{0, 1\}^E$ of configurations. The $\sigma$-field $\mathcal{F}$ is generated by the finite-dimensional cylinders of $\Omega$. In summary, the probability space is $(\Omega, \mathcal{F}, P_p)$.

Let $C(k)$ be the open cluster containing the node $k \in V$. The distribution of $C(k)$ is, by the translation-invariance of the measure $P_p$, well-known to be independent of $k \in V$, so that we assume in the following $k = 0 \in V$ without loss of generality. If $|C(0)|$ denotes the (random) number of nodes of $C(0)$, then $P_p(|C(0)| = \infty)$ is called the percolation probability. This probability depends on $p \in [0, 1]$, and Hammersley’s critical percolation probability is defined as

$$p_c^1(V) = \inf\{p \in (0, 1) : P_p(|C(0)| = \infty) > 0\}. \tag{4.1}$$

Thus, for $p > p_c^1(V)$ it is possible to generate infinite open clusters with positive probability. By Kolmogorov’s zero-one law (cf. [11, Theorem 1.11]) there exists an infinite open cluster with probability 1 for $p > p_c^1(V)$, and otherwise with probability 0. Similarly, for two different given nodes $i, j \in V$ we can define $C(i, j)$ as the open cluster containing $i$ and $j$, which may be empty. Again, by translation-invariance its distribution only depends on the edge distance $|i - j|$ and in the following we consider $C(\ell, 0)$ with $|\ell| = |i - j|$. The following definition is related to the radius of a finite open cluster as investigated in [11, Sections 6.1 and 8.4].
As in (4.1) we define the critical probability
\begin{equation}
\tag{4.2}
p_c^2(V) = \inf\{p \in (0, 1) : P_p(|C(\ell, 0)| = \infty) > 0\},
\end{equation}
where we use the convention that $|C(\ell, 0)| > 0$ if and only if there exists a possibly undirected path,
\begin{equation}
\tag{4.3}
[0 \leftrightarrow \ell] := [0 = k_0 \leftrightarrow k_1 \leftrightarrow \cdots \leftrightarrow k_n = \ell]
\end{equation}
of open edges from 0 to $\ell$, called an open path.

Clearly, $p_c^1(V) \leq p_c^2(V)$. Indeed, both probabilities are identical, which can be seen from the following standard argument. Let $A = \{0 \leftrightarrow \ell\}$ be the event that there exists an open path from the origin to node $\ell$. Note that this event has strictly positive probability $P_p(0 \leftrightarrow \ell)$, also called the two-point connectivity function in \cite[Section 8.5]{11}. Thus,
\[ P_p(|C(\ell, 0)| = \infty \mid A) = P_p(|C(0)| = \infty), \]
which is equivalent to
\[ P_p(|C(\ell, 0)| = \infty) = P_p(|C(\ell, 0)| = \infty \cap A) = P_p(|C(0)| = \infty) P_p(A). \]
Since $P_p(A) > 0$ we obtain
\[ P_p(|C(\ell, 0)| = \infty) > 0 \iff P_p(|C(0)| = \infty) > 0 \]
and thus $p_c^1(V) = p_c^2(V)$. Recall that the critical percolation probability $p_c^1(\mathbb{Z}^2)$ on the whole unoriented square lattice $\mathbb{Z}^2$ equals $\frac{1}{2}$ and moreover satisfies $P_0^1(|C(\ell, 0)| = \infty) = 0$ (\cite[Chapter 11]{11}).

Given such an infinite open cluster, we are interested in the probability that the components $X_i$ and $X_j$ on the random DAG are independent, which by Proposition 3.2 implies a zero entry in the random tail dependence coefficient matrix $\chi$. Thus, in the following we investigate the probability $P_p(\chi(\ell, 0) = 0)$. That is to say, we consider the max-linear process $\{X_i : i \in C(\ell, 0)\}$ on the random DAG with nodes $V(C(\ell, 0))$ and edges $E(C(\ell, 0))$.

Let $\omega = (\omega_e)_{e \in E}, \omega' = (\omega'_e)_{e \in E} \in \Omega$ with $\omega_e \leq \omega'_e$ for every $e \in E$. We recall that an event $A \subset \Omega$ is increasing if $\omega \in A$ implies that $\omega' \in A$. We observe that the events
\[ \{\chi(\ell, 0) \neq 0\} = \{\text{An}(\ell) \cap \text{An}(0) \neq \emptyset\} \]
and
\[ \{\text{De}(\ell) \cap \text{De}(0) \neq \emptyset\}. \]
are increasing.

From arguments given below, it is not difficult to see that the probability for a path between 0 and $\ell$ satisfies
\[ \frac{1}{2} P_p(0 \leftrightarrow \ell) \leq P_p(\{\text{An}(\ell) \cap \text{An}(0) \neq \emptyset\}). \]
The following lemma gives a refinement of this bound, which may be of interest in its own right.

**Lemma 4.1.** For $0 \leq p \leq 1$ we have
\[ P_p(\{\text{An}(\ell) \cap \text{An}(0) \neq \emptyset\}) \geq 1 - (1 - P_p(0 \leftrightarrow \ell))^{\frac{1}{2}}. \]
Proof. By translation invariance we find
\[ P\{\{\text{De}(\ell) \cap \text{De}(0) \neq \emptyset\}\} = P\{\{\text{An}(-\ell) \cap \text{An}(0) \neq \emptyset\}\} = P\{\{\text{An}(\ell) \cap \text{An}(0) \neq \emptyset\}\}, \]
more precisely, \{\text{De}(\ell) \cap \text{De}(0) \neq \emptyset\} and \{\text{An}(\ell) \cap \text{An}(0) \neq \emptyset\} are two increasing sets of equal probability. Inequality (11.14) in [11, p. 289] yields
\[ P\{\{\text{An}(\ell) \cap \text{An}(0) \neq \emptyset\}\} \geq 1 - (1 - P\{\{\text{An}(\ell) \cap \text{An}(0) \neq \emptyset\} \cup \{\text{De}(\ell) \cap \text{De}(0) \neq \emptyset\}\})^{\frac{1}{2}} \]
\[ = 1 - (1 - P_p(0 \leftrightarrow \ell))^{\frac{1}{2}}, \]
where the last equality follows from (4.6) below.

\[ \square \]

**Theorem 4.2.** For \( p < \frac{1}{2} \) we have
\[ \lim_{|i-j| \to \infty} P_p(X_i \text{ and } X_j \text{ are independent}) = 1. \]

For \( p > \frac{1}{2} \) there exists a constant \( 0 < C < 1 \) not depending on \( |i-j| \) such that
\[ 0 < P_p(X_i \text{ and } X_j \text{ are independent}) \leq C. \]

**Proof.** By translation-invariance the distribution of the above event only depends on the edge distance \( |\ell| = |i-j| \). Recall from basic results on percolation ([11, Theorem 6.10]) that the probability of an open path from the origin to \( \ell \)
\[ P_p(A) = P_p(\{0 \leftrightarrow \ell\}) \leq \sigma |\ell| \exp \left( -\varphi(p) |\ell| \right) \]
where \( \sigma > 0 \) and \( \varphi(p) \) is a rate function satisfying \( \varphi(p) > 0 \) if \( 0 < p < \frac{1}{2} \) and \( \varphi(p) = 0 \) for \( p \geq \frac{1}{2} \) (cf. [11, Theorem 6.14]). Hence, making the crucial observation that
\[ \{0 \leftrightarrow \ell\} = \{\text{An}(\ell) \cap \text{An}(0) \neq \emptyset\} \cup \{\text{De}(\ell) \cap \text{De}(0) \neq \emptyset\} \]
we get from Proposition 3.2 for \( p < \frac{1}{2} \)
\[ P_p(X_i \text{ and } X_j \text{ are dependent}) = P(\{\text{An}(\ell) \cap \text{An}(0) \neq \emptyset\}) \leq P_p(\{0 \leftrightarrow \ell\}) \leq \sigma |i-j| \exp \left( -|i-j|\varphi(p) \right) \to 0 \]
as \( |i-j| \to \infty \), giving (4.4).

For \( p > \frac{1}{2} \) note that
\[ 1 > P_p(0 \leftrightarrow \ell) \geq P_p(|C(\ell, 0)| = \infty) > 0 \]
for every \( \ell \in \mathbb{Z}^2 \). Applying Lemma 4.1 we obtain
\[ 1 > P_p(0 \leftrightarrow \ell) \geq P_p(\{\text{An}(\ell) \cap \text{An}(0) \neq \emptyset\}) \]
\[ \geq 1 - (1 - P_p(0 \leftrightarrow \ell))^{\frac{1}{2}} \]
\[ \geq 1 - (1 - P_p(|C(0)| = \infty)^2)^{\frac{1}{2}} > 0 \]
for every \( |\ell| = |i-j| \), where the second last inequality follows from the first display of [11, Section 8.5]. Thus, in the supercritical phase, with positive probability one can generate dependence between random variables \( X_i \) and \( X_j \), which proves (4.5).
Theorem 4.2 links the subcritical and supercritical case to probabilities for dependence and independence of $X_i$ and $X_j$.

For the communication in a Bernoulli bond percolation network, we conclude that for edges being open (communication channels) with small probability, extreme opinions at two different nodes become a.s. independent, when nodes are far apart. However, if edges are open with high probability then there is a positive probability that two extreme opinions are expressed dependently; i.e., there may be a common source.

Also further properties of $X_i$ and $X_j$ within the oriented square lattice $\mathbb{Z}^2$ can be derived similarly using percolation properties. The following remark gives an example.

Remark 4.3 (Number of common ancestors per pair of nodes:). Let $0 \leq p \leq 1$ and $A(i, j, n) := |An(i) \cap An(j) \cap B(n)|$ the number of common ancestors of $i$ and $j$ inside the box $B(n)$ as defined in (3.1). Then by an ergodic theorem (cf. [11, Theorem 4.2]) $P_p$-a.s. and in $L^1(P_p)$,

$$\frac{1}{|B(n)|} \sum_{k, \ell \in B(n), |k-\ell|=|i-j|} |A(k, \ell, n)|^{-1} \to E_p(|An(i) \cap An(j)|^{-1}), \quad n \to \infty.$$

4.2. Enlargement of DAGs using Bernoulli percolation. Throughout this section fix two nodes $i, j \in \mathbb{Z}^2$. We are again interested in dependence properties of the random variables $X_i$ and $X_j$. We write $\mathcal{P}$ for the property that $X_i$ and $X_j$ are dependent, and for a DAG $G$ we write $G \in \mathcal{P}$ if a max-linear model $X$ on $G$ has the property that the components $X_i$ and $X_j$ are dependent.

Suppose that $H = (V(H), E(H))$, $V(H) \subset \mathbb{Z}^2$, is a sub-DAG of the oriented square lattice $\mathbb{Z}^2$ containing $i, j$ such that $X_i$ and $X_j$ are independent on $H$, equivalently $An(i) \cap An(j) \cap V(H) = \emptyset$ by Proposition 3.2; i.e., $H \notin \mathcal{P}$. We utilize a method introduced in [14] in order to enlarge the sub-DAG $H$ by adding possibly infinitely many nodes and edges of open clusters and investigate the probability that $X_i$ and $X_j$ become dependent on the randomly enlarged DAG.

In the framework of communication in a network, if two extreme opinions are expressed seemingly independent, we investigate if a possible dependence could arise by a different network (family or friends) of a network member $i$, which are not present in the original network. The following results answer this question.

Recall that for $k \in \mathbb{Z}^2$ the open cluster containing $k$ is denoted by $C(k)$. The following definition goes back to [14, Definition 1.1].

Definition 4.4. For $0 \leq p \leq 1$ let $U(H) = U(\omega, p, H)$ be the random subgraph of the oriented square lattice $\mathbb{Z}^2$ with node set

$$V(U(H)) = \bigcup_{k \in V(H)} V(C(k))$$

and edge set

$$E(U(H)) = E(H) \cup \bigcup_{k \in V(H)} E(C(k)).$$

Note that by definition $U(H)$ is a DAG containing the nodes $i$ and $j$. Furthermore, we add finitely many or possibly infinitely many nodes, according as $p \leq \frac{1}{2}$ or $p > \frac{1}{2}$.
Moreover, Definition 4.4 corresponds to percolation with underlying probability measure $P_p^H$ on $\{0, 1\}^{E(\mathbb{Z}^2)}$ given by

$$(4.7) \quad P_p^H(\omega_\epsilon = 1) = 1 \text{ if } e \in E(H) \quad \text{and} \quad P_p^H(\omega_\epsilon = 1) = p \text{ else.}$$

In addition, we have by definition that

$$(4.8) \quad P_p(U(H) \in \mathcal{P}) = P_p^H\left(\text{An}(i) \cap \text{An}(j) \neq \emptyset\right).$$

One prerequisite is the measurability of the event (4.8), and we verify this by observing that $\{U(H) \in \mathcal{P}\}$ is equivalent to the existence of some $n \in \mathbb{N}$ such that $\text{An}(i) \cap \text{An}(j) \neq \emptyset$ holds on the ball $B(i, n) = \{ y \in \mathbb{Z}^2 : \delta(y, i) \leq n\}$ and, thus, $\{U(H) \in \mathcal{P}\}$ is determined by configurations of edges in a finite ball, and hence measurable.

In analogy to [14, Definition 1.3] we regard certain kinds of critical probabilities

$$(4.9) \quad p_{c,1,p,H} := \inf\{p \in [0,1] : P_p(U(H) \in \mathcal{P}) > 0\}$$

and

$$(4.10) \quad p_{c,2,p,H} := \inf\{p \in [0,1] : P_p(U(H) \in \mathcal{P}) = 1\}.$$ 

We first remark that $\{U(H) \in \mathcal{P}\}$ has positive probability for all $p > 0$, such that $p_{c,1,p,H} = 0$ always holds, and the interesting question is for which choice of sub-DAGs $H$ we have $p_{c,1,p,H} = p_{c,2,p,H}$. As an easy example we might first consider the non-connected DAG $H$ with node set $V(H) = \{i,j\}$ and $E(H) = \emptyset$. It is straightforward to see that $P_p(U(H) \notin \mathcal{P}) > 0$ for every $p \in (0,1)$ and this implies $p_{c,2,p,H} = 1 \neq p_{c,1,p,H}$. On the other hand, the following Lemma gives an example of a DAG, where the latter assertion is not true.

**Lemma 4.5.** Let $H$ be an infinite DAG with nodes $V(H) = \mathbb{Z}^2$ and let $i,j,k \in \mathbb{Z}^2$ such that $i_1 \leq k_1 \leq j_1$. Assume edges $E(H)$ only inside the set

$$\left(\mathbb{Z}^2 \setminus \{(k_1 \pm 1, i_2 - n) : n \in \mathbb{N}_0\}\right) \times \left(\mathbb{Z}^2 \setminus \{(k_1 \pm 1, i_2 - n) : n \in \mathbb{N}_0\}\right).$$

Then $p_{c,2,p,H} = 0$.

**Proof.** Fix $p \in (0,1)$. We show that $p_{c,2,p,H} \leq p$ by calculating $P_p(U(H) \notin \mathcal{P})$. By choice of $H$ the event $\{U(H) \notin \mathcal{P}\}$ does not depend on finitely many edges, see also Figure 1. Hence, by Kolmogorov’s zero-one law,

$$P_p(U(H) \notin \mathcal{P}) \in \{0,1\}.$$ 

From $p \in (0,1)$ we further get $P_p(U(H) \notin \mathcal{P}) < 1$ and therefore $P_p(U(H) \notin \mathcal{P}) = 0$. This yields $P_p(U(H) \in \mathcal{P}) = 1$ for every $p \in (0,1)$ and concludes the proof.

If we inspect the examples presented so far we recognize that the number of nodes and edges of the chosen DAG $H$ has a strong impact on whether we have $p_{c,1,p,H} = p_{c,2,p,H}$ or not. The following result substantiates this observation.

**Theorem 4.6.** Let $H$ be a DAG and $j \in V(H)$ such that the connected component containing $j$ is finite. Then we have $p_{c,2,p,H} = 1$.

**Proof.** Let $p < 1$ and recall that

$$P_p(U(H) \in \mathcal{P}) = P_p^H\left(\text{An}(i) \cap \text{An}(j) \neq \emptyset\right).$$
We prove the assertion by making use of planar duality arguments discussed in [11, Section 1.4]. Let $\mathbb{L}_d$ be the dual graph of $\mathbb{Z}^2$ with nodes given by the set $\{x + (\frac{1}{2}, \frac{1}{2}) : x \in \mathbb{Z}^2\}$ and edges joining two neighboring nodes so that each edge of $\mathbb{L}_d$ is crossed by a unique edge of its dual $\mathbb{Z}^2$. As introduced in [11, Section 1.4, p. 16] an edge of the dual is declared to be open if it crosses an open edge of $\mathbb{Z}^2$ and closed otherwise. Recall that a circuit of $\mathbb{L}_d$ is an alternating sequence $k_0, e_0, k_1, e_1, \ldots, k_n, e_n, k_0$ of nodes $k_0, \ldots, k_n$ and edges $e_0, \ldots, e_n$ forming a cyclic path from $k_0$ to $k_0$.

Let $A$ be the event that there is a sub-path of closed edges of a circuit containing $j$ in its interior and $i$ in its exterior. Since the connected component containing node $j$ is finite, we have
\[
0 < P_H^p(A) \leq P_H^p(\text{An}(i) \cap \text{An}(j) = \emptyset)
\]
which yields
\[
P_p(U(H) \in \mathcal{P}) < 1
\]
for every $p \in [0, 1)$. Thus, by definition we get $p_{c,2,\mathcal{P},H} = 1$ as claimed. \(\square\)

**Corollary 4.7.** Let $H$ be a finite DAG. Then we have $p_{c,2,\mathcal{P},H} = 1$.

**Remark 4.8.** Corollary 4.7 enlightens the fact that the events $\{\text{An}(i) \cap \text{An}(j) \neq \emptyset\}$ and $\{i \leftrightarrow j\}$ are essentially different. Indeed, if we choose a DAG $H \notin \mathcal{P}$ with $\{\text{De}(i) \cap \text{De}(j) \neq \emptyset\}$ we have for every $0 \leq p < 1$,
\[
P_p^H(i \leftrightarrow j) = 1
\]
and
\[
P_p^H(\text{An}(i) \cap \text{An}(j) \neq \emptyset) < 1.
\]

Now we want to examine DAGs with the property that $p_{c,1,\mathcal{P},H} = p_{c,2,\mathcal{P},H} = 0$. In Lemma 4.5 we gave an example of a sub-DAG $H$ satisfying this equality. We can prove the same identity for the class of percolating everywhere subgraphs, which is an analogous result to [14, Theorem 1.13 (i)]. According to [2], a sub-DAG $H$ is called *percolating everywhere* if $V(H) = \mathbb{Z}^2$ and every connected component of $H$ is infinite.
Theorem 4.9. Let $H$ be a percolating everywhere sub-DAG of the oriented square lattice $\mathbb{Z}^2$. Then we have $p_{c,2,\mathcal{P},H} = 0$.

Proof. The proof partially relies on the proof of [14, Theorem 1.13]. As there we work with the probability measure $P_H^p$ on $\{0,1\}^{E(\mathbb{Z}^2)}$ given in (4.7). Let $J$ be the graph with node set

$$V(J) = \{(k_1, k_2) : k_1 \leq i_1, k_2 \leq i_2\} \cup \{(k_1, k_2) : k_1 \leq j_1, k_2 \leq j_2\}$$

Note that if $J$ is connected then $\text{An}(i) \cap \text{An}(j) \neq \emptyset$. Define the equivalence relation $k \sim \ell$ on $\mathbb{Z}^2$ if and only if $P_H^p(k \leftrightarrow \ell) = 1$. Denote by $[k]$ the equivalence class containing $k$ and $\mathbb{Z}^2/\sim = Z' = Z'(\omega)$ the (Bernoulli) quotient graph with node set given by

$$V(Z') = \{[k] : k \in \mathbb{Z}^2\}.$$ 

If $|V(Z')| = 1$ then

$$P_H^p(U(H) \text{ is connected and } U(H) = \mathbb{Z}^2) = 1.$$ 

Thus, with probability one there exists $k \in \text{An}(i) \cap J$ with $k \leftrightarrow j$ so that $\text{An}(i) \cap \text{An}(j) \neq \emptyset$. Now assume that $|V(Z')| \geq 2$. For sets $A, B \subset \mathbb{Z}^2$ let

$$E(A, B) = \{(a, b) \in E(\mathbb{Z}^2) : a \in A, b \in B\}.$$ 

By the same arguments as in the proof of [14, Theorem 1.13] we can choose a partition $V(Z') = A \cup B$, $A \cap B = \emptyset$ with $|E(A, B)| = \infty$. At this point observe that the number of connected components of $H$ is infinite, otherwise we would have $|E(A, B)| < \infty$ for every partition $V(Z') = A \cup B$. Thus, by an application of Kolmogorov’s zero-one law we have

$$P_H^p(\text{An}(i) \cap \text{An}(j) = \emptyset) \in \{0, 1\},$$ 

so that

$$P_p(U(H) \in \mathcal{P}) \in \{0, 1\}.$$ 

This in particular implies that

$$p_{c,2,\mathcal{P},H} = p_{c,1,\mathcal{P},H} = 0$$

by definition and concludes the proof. \qed

Acknowledgement. We take pleasure in thanking the Mathematical Research Institute Oberwolfach for an invitation to the Research in Pairs Programme. This visit was extremely stimulating and fruitful for our research.

References


Claudia Klüppelberg, Center for Mathematical Sciences, Technical University of Munich, Boltzmannstr. 3, D–85748 Garching, Germany  
E-mail address: cklu@tum.de

Ercan Sönmez, Mathematical Institute, Heinrich-Heine-University Düsseldorf, Universitätsstr. 1, D-40225 Düsseldorf, Germany  
E-mail address: ercan.soenmez@hhu.de