Abstract. We study a notion of robustness of a Markov kernel that describes a system of several input random variables and one output random variable. Robustness requires that the behaviour of the system does not change if one or several of the input variables are knocked out. If the system is required to be robust against too many knockouts, then the output variable cannot distinguish reliably between input states and must be independent of the input. We study how many input states the output variable can distinguish as a function of the required level of robustness.

Gibbs potentials allow a mechanistic description of the behaviour of the system after knockouts. Robustness imposes structural constraints on these potentials. We show that interaction families of Gibbs potentials allow to describe robust systems.

Given a distribution of the input random variables and the Markov kernel describing the system, we obtain a joint probability distribution. Robustness implies a number of conditional independence statements for this joint distribution. The set of all probability distributions corresponding to robust systems can be decomposed into a finite union of components, and we find parametrizations of the components. The decomposition corresponds to a primary decomposition of the conditional independence ideal and can be derived from more general results about generalized binomial edge ideals.

1. Introduction

Consider a stochastic system of \( n \) input nodes and one output node:

\[
\begin{array}{cccc}
\text{input:} & X_1 & X_2 & X_3 & \cdots & X_n \\
\text{system} & \downarrow & \downarrow & \downarrow & \cdots & \downarrow \\
\text{output:} & Y
\end{array}
\]

As shown in [1], there are two ingredients to robustness:

1. If one or several of the input nodes are removed, the system behaviour should not change too much (“small exclusion dependence”).
2. A causal contribution of the input nodes on the output nodes.

The second point is strictly necessary: If the behaviour of the output does not depend on the inputs at all, then it is usually not affected by a knockout of a subset of the inputs, but this exclusion independence is trivial.

In this paper we do not use the information theoretic measures proposed in [1]. Instead, we start with a simple model of exclusion independence: We study systems in which the behaviour of the output node does not change when one or more of the...
input nodes are knocked out. We formalize our robustness requirements in terms of a *robustness specification* $\mathcal{R}$, which consists of pairs $(R, x_R)$, where $R$ is a subset of the inputs and $x_R$ is a joint state of the inputs in $R$. Let $\mathcal{S}$ be a set of possible states of the input nodes. The system is $\mathcal{R}$-robust in $\mathcal{S}$, if the behaviour of the system does not change if the inputs not in $R$ are knocked out, provided that the inputs in $R$ are in the state $x_R$ and the current state of all inputs belongs to $\mathcal{S}$.

If the robustness specification $\mathcal{R}$ is too large, or if the set $\mathcal{S}$ is too large, then in any $\mathcal{R}$-robust system the output does not depend on the input at all. In general, the behaviour of the system is restricted by robustness requirements. Therefore, to study the causal contribution of the input nodes on the output nodes, we investigate how varied the behaviour of a system can be, given both robustness specifications imply that the system cannot distinguish all input states, and we may ask how many states the system can discern. This question is related to the topic of error detecting codes, see Remark 6.

This paper is organized as follows: Section 2 contains our basic setting and definitions. We find several equivalent formulations of our notion of robustness. Moreover, we study the question how many states an $\mathcal{R}$-robust system can distinguish. Section 3 shows that our definitions generalize the notions of canalyzing and nested canalyzing functions, which have been studied before in the context of robustness. Section 4 proposes to model the different behaviours of a system under various knockouts using a family of Gibbs potentials. Robustness implies various constraints on these potentials. Section 4 discusses the probabilistic behaviour of the whole system, including its inputs, when the input variables are distributed to some fixed input distribution. The set of all joint probability distributions is found to be a subset of the finite set $\mathcal{J}$.

Some of our results in Section 5 can also be derived from recent algebraic results in [13] about generalized binomial ideals. These ideals generalize the binomial edge ideals of [6] and [12]. Similar ideals have recently been studied in the context of nested canalyzing functions. In this paper we give self-contained proofs that are also accessible to readers not acquainted to the language of commutative algebra. We comment on the relation to the algebraic results in Remark 25.

2. Robustness and canalyzing functions

We consider $n$ input nodes, denoted by $1, 2, \ldots, n$, and one output node, denoted by 0. For each $i = 0, 1, \ldots, n$ the state of node $i$ is a discrete random variable $X_i$ taking values in the finite set $X_i$ of cardinality $d_i$. The input state space is the set $X_{\text{in}} = X_1 \times \cdots \times X_n$, and the joint state space is $X = X_0 \times X_{\text{in}}$. For any subset $S \subseteq \{0, \ldots, n\}$ write $X_S$ for the random vector $(X_i)_{i \in S}$; then $X_i$ is a random variable with values in $X_S = X_{\{0\} \cup S \cup \{i\}}$. For $S = \{n\} := \{1, \ldots, n\}$ we also write $X_{\text{in}}$ instead of $X_{\{n\}}$. For any $x \in X$, the restriction of $x$ to a subset $S \subseteq \{0, \ldots, n\}$ is the vector $x_S \in X_S$ with $(x_S)_i = x_i$ for all $i \in S$. In contrast, the notation $x_S$ will refer to an arbitrary element of $X_S$.

As a model for the computation of the output from the input, we use a stochastic map (Markov kernel) $\kappa$ from $X_{\text{in}}$ to $X_0$, that is, $\kappa$ is a function that assigns to each $x \in X_{\text{in}}$ a probability distribution $\kappa(x)$ for the output $X_0$. Such a stochastic map $\kappa$ can be represented by a matrix, with matrix elements $\kappa(x; x_0)$, $x \in X_{\text{in}}, x_0 \in X_0$, satisfying $\sum_{x_0 \in X_0} \kappa(x; x_0) = 1$ for all $x \in X_{\text{in}}$. For each $x \in X_{\text{in}}$ the probability
distribution $\kappa(x)$ models the behaviour of $X_0$ when the input variables are in the state $x$. When the input is distributed according to some input distribution $p_{\text{in}}$, then the joint distribution $p$ of input and output variables satisfies

$$p(X_0 = x_0, X_{\text{in}} = x) = p_{\text{in}}(X_{\text{in}} = x)\kappa(x; x_0).$$

If $p_{\text{in}}(X_{\text{in}} = x) > 0$, then $\kappa(x)$ can be computed from the joint probability distribution $p$ and equals the conditional distribution of $X_0$, given that $X_{\text{in}} = x$.

When a subset $S$ of the input nodes is knocked out and only the nodes in $R = [n] \setminus S$ remain, then the behaviour of the system changes. Without further assumptions, the post-knockout function is not determined by $\kappa$ and has to be specified separately. We model the post-knockout function by a further stochastic map $\kappa_R : X_R \times X_0 \rightarrow [0, 1]$. A complete specification of the system is given by the family $(\kappa_A)_{A \subseteq [n]}$ of all possible post-knockout functions, which we refer to as functional modalities. As a shorthand notation we denote functional modalities by $(\kappa_A)$. The stochastic map $\kappa$ itself, which describes the normal behaviour of the system without knockouts, can be identified with $\kappa_{[n]}$.

What does it mean for functional modalities to be robust? Assume that the input is in state $x$, and that we knock out a set $S$ of inputs. Denoting the remaining set of inputs by $R$, we say that $(\kappa_A)$ is robust in $x$ against knockout of $S$, if $\kappa(x) = \kappa_R(x|_R)$, that is, if

$$\kappa(x; x_0) = \kappa_R(x|_R; x_0) \quad \text{for all } x_0 \in X_0.$$  

Let $R$ be a collection of pairs $(R, x_R)$, where $R \subseteq [n]$ and $x_R \in X_R$. We call such a collection a robustness specification in the following. We say that $(\kappa_A)$ is $R$-robust in a set $S \subseteq X_{\text{in}}$ if

$$\kappa(x) = \kappa_R(x|_R), \quad \text{whenever } x \in S \text{ and } (R, x|_R) \in R.$$  

The main example in this section will be the robustness structures

$$\mathcal{R}_k := \{(R, x_R) : R \subseteq [n], |R| \geq k, x_R \in X_R\}.$$  

Equation (1) only compares the functional modality $\kappa_R$ after knockout with the stochastic map $\kappa$ that describes the regular behaviour of the unperturbed system. In particular, for $R \subseteq R' \subseteq [n]$, the functional modality $\kappa_{R'}$ is in no way restricted by (1). Therefore, it may happen that a system that is not robust against a knockout of a set $S' = [n] \setminus R'$ recovers its regular behaviour if we knockout even more nodes. However, this is not the typical situation. Therefore, it is natural to assume that the following holds: If $(R, x_R) \in \mathcal{R}$ and if $R \subseteq R' \subseteq [n]$, then $(R', x_R|_{R'}) \in \mathcal{R}$ for all $x_{R'} \in X_{R'}$ with $x_{R'}|_{R} = x_{R}$. In this case we call the robustness specification $\mathcal{R}$ coherent. For example the robustness structures $\mathcal{R}_k$ are coherent. The notion of coherence will not play an important role in the following, but it is interesting from a conceptual point of view. It is related to the notion of coherency as used e.g. in [3].

By definition, for robust functional modalities $(\kappa_A)$ the largest functional modality $\kappa_{[n]}$ determines the smaller ones in the relevant points via (2). This motivates the following definition: A stochastic map $\kappa$ is called $\mathcal{R}$-robust in $S$, if there exist functional modalities $(\kappa_A)$ with $\kappa = \kappa_{[n]}$ that are $\mathcal{R}$-robust in $S$. More directly, $\kappa$ is $\mathcal{R}$-robust in $S$ if and only if

$$\kappa(x) = \kappa(y), \quad \text{whenever } x, y \in S, x|_R = y|_R \text{ and } (R, x|_R) \in \mathcal{R}.$$
When studying robustness of a stochastic map $\kappa$ we may always assume that $\mathcal{R}$ is coherent; for if $x_R = y_R$ implies $\kappa(x) = \kappa(y)$, then $x_{R'} = y_{R'}$ also implies $\kappa(x) = \kappa(y)$, whenever $R \subseteq R' \subseteq [n]$.

For any subset $R \subseteq [n]$ and $x_R \in X_R$ let

$$C(R, x_R) := \{ x \in X_n : x_R = x_R \}.$$

be the corresponding cylinder set. Then $\kappa$ is $\mathcal{R}$-robust in $\mathcal{S}$ if and only if $\kappa(x) = \kappa(y)$ for all $x, y \in \mathcal{S} \cap C(R, x_R)$ and $(R, x_R) \in \mathcal{R}$. In other words, the stochastic map $\kappa$ is constant on $\mathcal{S} \cap C(R, x_R)$ for all $(R, x_R) \in \mathcal{R}$.

The following construction is useful to study robust functional modalities: Given a robustness specification $\mathcal{R}$, define a graph $G_\mathcal{R}$ on $X_n$ by connecting two elements $x, y \in X_n$ by an edge if there is $(R, x_R) \in \mathcal{R}$ such that $x_R = y_R = x_R$. Denote by $G_{\mathcal{R}, S}$ the subgraph of $G_\mathcal{R}$ induced by $S$.

**Example 1.** Assume that $X_i = \{0, 1\}$ for $i = 1, \ldots, n$. Then the input state space $X_n = \{0, 1\}^n$ can be identified with the vertices of an $n$-dimensional hypercube. The graph $G_{\mathcal{R}_{n-1}}$ is the edge graph of this hypercube (Fig. 1b)). Cylinder sets correspond to faces of this hypercube. If $R \subseteq [n]$ has cardinality $n - 1$, then the cylinder set $C(R, x_R)$ is an edge, if $R$ has cardinality $n - 2$, then $C(R, x_R)$ is a two-dimensional face. Fig. 1b) shows an induced subgraph of $G_3$ for $n = 4$. By comparison, the graph $G_{\mathcal{R}_{n-2}}$ has additional edges corresponding to diagonals in the quadrangles of $G_{\mathcal{R}_{n-1}}$. For example, the set of vertices marked black in Figure 1b) is connected in $G_{\mathcal{R}_{n-2}}$ but not in $G_{\mathcal{R}_{n-1}}$ (Fig. 1b)).

**Proposition 2.** The following statements are equivalent for a stochastic map $\kappa$:

1. $\kappa$ is $\mathcal{R}$-robust in $\mathcal{S}$.
2. $\kappa$ is constant on $\mathcal{S} \cap C(R, x_R)$ for all $(R, x_R) \in \mathcal{R}$.
3. $\kappa$ is constant on the connected components of $G_{\mathcal{R}, S}$.
4. For any probability distribution $p_m$ of $X_n$ with $p_m(S) = 1$ and for all $(R, x_R) \in \mathcal{R}$, the output $X_0$ is stochastically independent of $X_{[n]\setminus R}$ given $X_R = x_R$.

**Proof.** The equivalence (1) $\iff$ (2) was already shown.

(2) $\iff$ (3): Condition (2) says that $\kappa$ is constant along each edge of $G_{\mathcal{R}, S}$. By iteration this implies (3). In the other direction, the subgraph of $G_{\mathcal{R}, S}$ induced by $\mathcal{S} \cap C(R, x_R)$ is connected for all $(R, x_R) \in \mathcal{R}$, and therefore (3) implies (2).

(2) $\implies$ (4): For any $x \in X_n$ with $p_m(x) > 0$, the conditional distribution of the output given the input satisfies $p(Y_0 = x_0 | X_n = x) = \kappa(x; x_0)$. By (2), $\kappa(x; x_0)$ is constant on $C(R, x_R) \cap \mathcal{S}$. Hence the conditional distribution does not depend on $X_{[n]\setminus R}$, and so $p(Y_0 = x_0 | X_n = x) = p(X_0 = x_0 | X_R = x_R)$.

![Figure 1. An illustration of Example 1 with $n = 4$.](image-url)
associate a random variable independent of as many connected components as possible in \( \bigcup \mathcal{C}(\mathcal{R}, x_R) \). Therefore, \( \kappa(x) \) is constant on \( S \cap \mathcal{C}(\mathcal{R}, x_R) \).

The choice of the set \( S \) is important: On one hand \( S \) should be large, because otherwise the notion of robustness is very weak. However, if \( S \) is too large, then the equations (1) imply that the output \( X_0 \) is (unconditionally) independent of all inputs. Proposition 2 gives a hint how to choose the set \( S \): The goal is to have as many connected components as possible in \( G_{R,S} \). This motivates the following definition:

**Definition 3.** For any subset \( S \subseteq X_{\text{in}} \), the set of connected components of \( G_{R,S} \) is called an \( \mathcal{R} \)-robustness structure.

Let \( B \) be an \( \mathcal{R} \)-robustness structure, and let \( S = \bigcup \mathcal{B} \). Let \( f_B : S \to B \) be the map that maps each \( x \in S \) to the corresponding block of \( B \) containing \( x \). Any stochastic map \( \kappa \) that is \( \mathcal{R} \)-robust on \( S \) factorizes through \( f_B \), in the sense that there is a stochastic map \( \kappa' \) that maps each block in \( B \) to a probability distribution on \( X_0 \) that satisfies \( \kappa = \kappa' \circ f_B \). Conversely, any stochastic map \( \kappa \) that factorizes through \( f_B \) is \( \mathcal{R} \)-robust.

To any joint probability distribution \( p_{\text{in}} \) on \( X_{\text{in}} \) with \( p(X_{\text{in}} \in S) = 1 \) we can associate a random variable \( B = f_B(X_1, \ldots, X_n) \). If \( \kappa \) is \( \mathcal{R} \)-robust on \( S \), then \( X_0 \) is independent of \( X_1, \ldots, X_n \) given \( B \). Note that the random variable \( B \) is only defined on \( \bigcup \mathcal{B} \), which is a set of measure one with respect to \( p_{\text{in}} \). The situation is illustrated by the following graph:

\[
\begin{align*}
X_1 & \quad X_2 & \quad X_3 & \quad \cdots & \quad X_n \\
\downarrow & \quad \downarrow & \quad \downarrow & \quad \downarrow & \quad \downarrow \\
Y & \quad f_B(X_1, X_2, \ldots, X_n) & \quad & \quad & \quad
\end{align*}
\]

When the robustness specification \( \mathcal{R} \) is fixed, how much freedom is left to choose a robust stochastic map \( \kappa \)? More precisely, how many components can an \( \mathcal{R} \)-robustness structure \( B \) have?

**Lemma 4.** Let \( B \) be a robustness structure of the robustness specification \( \mathcal{R} \). Let \( R \subseteq [n], S = [n] \setminus R \) and \( \mathcal{Y}_R = \{ x_R \in X_R : (R, x_R) \in \mathcal{R} \} \). Then

\[
|B| \leq |\mathcal{Y}_R| + |X_R \setminus \mathcal{Y}_R| \cdot |X_S|.
\]

**Proof.** The set \( S \) is the disjoint union of the \( |\mathcal{Y}_R| \) sets \( C(R, x_R) \cap S \) for \( x_R \in \mathcal{Y}_R \) and the \( |X_R \setminus \mathcal{Y}_R| \cdot |X_S| \) singletons \( \{x\} \subseteq S \) with \( x_R \notin \mathcal{Y}_R \). Each of these sets induces a connected subgraph of \( G_R \). The statement now follows from Proposition 2.

**Example 5.** Suppose that \( S = X_{\text{in}} \). This means that any \( \mathcal{R} \)-robustness structure \( B \) satisfies \( \bigcup \mathcal{B} = X_{\text{in}} \). If \( G_R \) is connected, then \( B \) has just a single block. In this case the bound of Lemma 4 is usually not tight. On the other hand, the bound is tight if \( \mathcal{R} = \{(R, x_R) : x_R \in X_R \} \).
Remark 6 (Relation to coding theory). Assume that all $d_i$ are equal. We can interpret $X_n$ as the set of words of length $n$ over the alphabet $[d_i]$. Consider the uniform case $\mathcal{R} = \mathcal{R}_k$. Then the task is to find a collection of subsets such that any two different subsets have Hamming distance at least $n - k + 1$. A related problem appears in coding theory: A code is a subset $Y$ of $X_n$ and corresponds to the case that each element of $B$ is a singleton. If distinct elements of the code have Hamming distance at least $n - k + 1$, then a message can be reliably decoded even if only $k$ letters are transmitted. If all letters are transmitted, but up to $k$ letters may contain an error, then this error may at least be detected; hence such codes are called error detecting codes. In this setting, the function $f_B$ can be interpreted as the decoding operation. The problem of finding a largest possible code such that all code words have a fixed minimum distance is also known as the sphere packing problem. The maximal size $A_d(n, n - k + 1)$ of such a code is unknown in general.

3. Canalyzing functions

Our notion of $\mathcal{R}$-robust functional modalities naturally generalizes and is motivated by canalyzing [9] and nested canalyzing functions [10]. Let $f : X_n \to X_0$ be a function, also called (deterministic) map. Such a map can be considered as a special case of a stochastic map by identifying $f$ with $\kappa f(x; x_0) := \begin{cases} 1, & \text{if } f(x) = x_0 \\ 0, & \text{otherwise} \end{cases}$.

We say that $f$ is $(R, x_R)$-canalyzing, if the value of $f$ does not depend on the input variables $X_{[n]\setminus R}$ given that the input variables $X_R$ are in state $x_R$. In other words, an $(R, x_R)$-canalyzing function is assumed to be constant on the corresponding cylinder set:

$$x, x' \in C(R, x_R) \Rightarrow f(x) = f(x').$$

Given a robustness specification $\mathcal{R}$, we say that a function $f$ is $\mathcal{R}$-canalyzing if it is $(R, x_R)$-canalyzing for all $(R, x_R) \in \mathcal{R}$. Clearly, the set of $\mathcal{R}$-canalyzing functions strongly depends on $\mathcal{R}$. On one hand, any function is $\mathcal{R}$-canalyzing with respect to

$$\mathcal{R} = \{(i, n), x \in X_n\}.$$

On the other hand, for two different elements $i, j \in [n]$, and

$$\mathcal{R} = ((i) \times X_i) \cup ((j) \times X_j),$$

any $\mathcal{R}$-canalyzing function is constant. Note that constant functions are $\mathcal{R}$-canalyzing for any $\mathcal{R}$.

The following statement directly follows from Proposition 2:

**Proposition 7.** A function $f : X_n \to X_0$ is $\mathcal{R}$-canalyzing if and only if $\kappa^f$ is $\mathcal{R}$-robust in $S = X_n$.

Particular cases of canalyzing functions have been studied in the context of robustness:

**Example 8.** (1) Canalyzing functions. A function $f$ with domain $X_n$ is canalyzing in the sense of [9], if there exist an input node $k \in [n]$, an input value $a \in X_k$, and an output value $b \in X_0$ such that the value of $f$ is independent of $x_{[n]\setminus k}$, given that
where the

The structure of this representation of the stochastic map

where

each

φ

decomposition of the sum that appears in (3).

κ

κ\( (4)\)

(5)

For an arbitrary output

κ

κ\( (2)\)

Nested canalyzing functions

have been studied in [10]. A function f is nested canalyzing in the variable order \(X_1, \ldots, X_n\) with canalyzing input values \(a_1 \in X_1, \ldots, a_n \in X_n\) and canalyzed output values \(b_1, \ldots, b_n\) if f satisfies \(f(x) = b_k\) for all \(x \in X\) satisfying \(x|_k = a_k\) and \(x|_i \neq a_i\) for all \(i < k\). Let \(\mathcal{R} := \bigcup_{k=1}^n \mathcal{R}^{(k)}\), where

\[
\mathcal{R}^{(k)} := \{(R, x_R) : [k] \subseteq R, x_R|_k \neq a_1, \ldots, x_R|_{k-1} \neq a_{k-1}, x_R|_k = a_k\}.
\]

It is easy to see that f is a nested canalyzing function if and only if it is \(\mathcal{R}\)-canalyzing.

The set of Boolean nested canalyzing functions has been described algebraically in [7] as a variety over the finite field \(\mathbb{F}_2\). Here, we use a different viewpoint, which allows to study not only deterministic functions, but also stochastic functions.

4. Robustness and Gibbs representation

Let \((\kappa_A)\) be a collection of functional modalities, as defined in Section 2. Instead of providing a list of all functional modes \(\kappa_A\), one can describe them in more mechanistic terms. To illustrate this, we first consider an example from the field of neural networks: Assume that the output node receives an input \(x = (x_1, \ldots, x_n) \in \{-1, +1\}^n\) and generates the output +1 with probability

\[
\kappa(x_1, \ldots, x_n; +1) := \frac{1}{1 + e^{-\left(\sum_{i=1}^n w_i x_i - \eta\right)x_0}}.
\]

For an arbitrary output \(x_0\) this implies

\[
\kappa(x_1, \ldots, x_n; x_0) := \frac{e^{\frac{1}{2} \left(\sum_{i=1}^n w_i x_i - \eta\right)x_0}}{e^{\frac{1}{2} \left(\sum_{i=1}^n w_i x_i - \eta\right)(-1)} + e^{\frac{1}{2} \left(\sum_{i=1}^n w_i x_i - \eta\right)(+1)}}.
\]

The structure of this representation of the stochastic map \(\kappa\) already suggests what the function should be after a knockout of a set \(S\) of input nodes: Simply remove the contribution of all the nodes in \(S\). The post-knockout function is then given by

\[
\kappa_R(x_R; +1) := \frac{e^{\frac{1}{2} \left(\sum_{i \in R} w_i x_i - \eta\right)x_0}}{e^{\frac{1}{2} \left(\sum_{i \in R} w_i x_i - \eta\right)\left(-1\right)} + e^{\frac{1}{2} \left(\sum_{i \in R} w_i x_i - \eta\right)\left(+1\right)}}, \quad \text{where } R = [n] \setminus S.
\]

These post-knockout functional modalities are based on the decomposition of the sum that appears in (3).

More generally, we consider the following model of \((\kappa_A)\):

\[
\kappa_A(x_A; x_0) = e^{\sum_{B \subseteq A} \phi_B(x_B, x_0)} \sum_{x'_0} e^{\sum_{B \subseteq A} \phi_B(x'_B, x'_0)},
\]

where the \(\phi_B\) are functions on \(X_B \times X_0\). Such a sum decomposition of \(\kappa\) is referred to as a Gibbs representation of \(\kappa\) and contains more information than \(\kappa\) itself. Clearly, each \(\kappa_A\) is strictly positive. Using the Möbius inversion, it is easy to see that each strictly positive family \((\kappa_A)\) has a representation of the form (5) with

\[
\phi_A(x_A, x_0) := \sum_{C \subseteq A} (-1)^{|A| - |C|} \ln \kappa_C(x_A|_C; x_0).
\]
The statement follows from the fact that assuming that (for fixed $x$) for all $x$ belonging to a set $S$ the system can choose its policy within such a restricted set of functional modalities such that ($\kappa_A$) for all $x$ belonging to a set $S$ means

$$\kappa_{(1)}(x_1; x_0) = \kappa_{(1,2)}(x_1, x_2; x_0) = \kappa_{(2)}(x_2; x_0)$$

for all $x_0$ whenever $x_1 = x_2$. By Proposition 9 this translates into the conditions

$$\phi_{(1,2)}(x_1, x_2; x_0) + \phi_{(1)}(x_1; x_0) = 0 = \phi_{(1,2)}(x_1, x_2; x_0) + \phi_{(2)}(x_2; x_0)$$

for all $x_0$ whenever $x_1 = x_2$ for the potentials ($\phi_A$) defined via (6). This means: Assuming that ($\kappa_A$) is $R$-robust, it suffices to specify the four functions

$$\phi_0(x_0), \phi_{(1)}(x_1; x_0), \phi_{(1,2)}(0, 1; x_0), \phi_{(1,2)}(1, 0; x_0).$$

The remaining potentials can be deduced from (7). If only the values of ($\kappa_A$) for $x \in S$ are needed, then it suffices to specify $\phi_0(x_0)$ and $\phi_{(1)}(x_1; x_0)$.

Does $R$-robustness in $x$ imply any structural constraints on ($\kappa_A$)? If ($\kappa_A$) is $R$-robust in $x$ for all $x$ belonging to a set $S$, then the corresponding conditions imposed by Proposition 9 depend on $S$. In this section, we are interested in conditions that are independent of $S$. Such conditions allow to define sets of functional modalities that contain all $R$-robust functional modalities for all possible sets $S$. If $S$ (which will be the support of the input distribution in Section 5) is unknown from the beginning, then the system can choose its policy within such a restricted set of functional modalities. To find results that are independent of $S$, our trick is to find a set $M_R$ of functional modalities such that ($\kappa_A$) can be approximated on $S$ by functional modalities in $M_R$. The approximation will be independent of $S$.

We first consider the special case $R = K_k := \{ (R, x_R) : R \subseteq [n], |R| \geq k, x_R \in X_R \}$. For simplicity, we replace any prefix or subscript $R_k$ by $k$. Denote by $M_{k+1}$ the set of all functional modalities ($\kappa_A$) such that there exist potentials $\phi_A$ of the form

$$\phi_A(x_A; x_0) = \sum_{B \subseteq A} \alpha_{A,B} \Psi_B(x_A|B; x_0),$$

where $\alpha_{A,B} \in \mathbb{R}$ and $\Psi_B$ is an arbitrary function $\mathbb{R}^{X_B \times X_0} \to \mathbb{R}$. The set $M_{k+1}$ is called the family of $(k + 1)$-interaction functional modalities. Note that the functions $\Psi_B$ do not depend on $A$. This ensures a certain interdependence among the functional
modalities $\kappa_A$. The name “(k + 1)-interaction” comes from the fact that each potential $\Psi_B$ depends on the $k$ (or less) variables in $B$ plus the output variable $X_0$. Since $M_{k+1}$ only contains strictly positive functional modalities, we are also interested in the closure of $M_{k+1}$ with respect to the usual topology on the space of matrices, considered as elements of a finite-dimensional real vector space.

Example 11. The functional modalities (4), derived from the classical model (3) of a neural network, belong to $M_2$. To illustrate the difference between $M_2$ and its closure, consider the functional modalities ($\kappa_A$) with

$$\kappa_A(x_1, \ldots, x_n; x_0) := \frac{e^{\sum_{\alpha \in A} w_{\alpha}(x_i - \eta)}}{e^{\sum_{\alpha \in A} w_{\alpha}(x_i - \eta)} + e^{\sum_{\alpha \in A} w_{\alpha}(x_i - \eta)}}.$$

If $w_1, \ldots, w_n$ and $\eta$ are fixed and $\beta \to \infty$, then

$$\kappa_A(x_1, \ldots, x_n; +1) \to \theta(\sum_{i \in A} w_i x_i - \eta),$$

where

$$\theta(x) = \begin{cases} +1, & \text{if } x > 0, \\ \frac{1}{2}, & \text{if } x = 0, \\ 0, & \text{if } x < 0. \end{cases}$$

The functional modalities (8) are deterministic limits of the probabilistic model (3), called linear threshold functions. They lie in the closure of $M_2$, but not in $M_2$ itself.

Linear threshold functions are widely used as elementary building blocks in network dynamics, for example to build simple models of neural networks, metabolic networks or gene-regulation networks. Robustness against knockouts of such networks has been studied in [2], exploring the example of the yeast cell cycle.

Let $\tilde{M}_{k+1}$ be the set of strictly positive functional modalities ($\kappa_A$) such that

$$\kappa_C(x_C; x_0) = \frac{1}{Z_{C,x}} \exp \left( \sum_{B \subseteq C} \frac{1}{|C|} \ln(k_B(x_C|B; x_0)) \right)$$

(9)

$$= \frac{1}{Z_{C,x}} \left( \prod_{B \subseteq C, |B| = k} k_B(x_C|B; x_0) \right)^{1/(|C|)}$$

for all $C \subseteq [n]$ with $|C| > k$, where $Z_{C,x}$ is a normalization constant that ensures that $\kappa_C(x_C)$ is a probability distribution. Note that equations (9) can be used to parametrize the set $\tilde{M}_{k+1}$: The stochastic maps $\kappa_A$ with $|A| \leq k$ can be chosen arbitrarily, while all other stochastic maps $\kappa_C$ with $|C| > k$ can be computed by normalizing the geometric mean of the stochastic maps $\kappa_B$ for $B \subseteq C$ and $|B| = k$.

Lemma 12. $\tilde{M}_{k+1}$ is a subset of $M_{k+1}$. It consists of those functional modalities ($\kappa_A$) where the coefficients $\alpha_{A,B}$ additionally satisfy

$$(-1)^{|A|} \alpha_{A,B} = (-1)^{|A'|} \alpha_{A',B}, \quad \text{whenever } B \subseteq A \cap A' \text{ and } |B| < k,$$

and

$$(-1)^{|A'|} \alpha_{A,B} = (-1)^{|A|} \alpha_{A',B}, \quad \text{whenever } B \subseteq A \cap A' \text{ and } |B| = k.$$

for all $x_B \in X_B$ and $x_0 \in X_0$. 

ROBUSTNESS, ANALYZING FUNCTIONS AND SYSTEMS DESIGN 9
Proof. Assume that the coefficients \( a_{A,B} \) of \( (\kappa_A) \) \( \in \mathcal{M}_{k+1} \) satisfy the conditions stated in the lemma. We may multiply all functions \( \Psi_B \) by scalars and assume

\[
\alpha_{A,B} = (-1)^{|A|-|B|}, \quad \text{if } |B| < k, \quad \alpha_{A,B} = (-1)^{|A|-k} \frac{k}{|A|}, \quad \text{if } |B| = k.
\]

Then \( \ln(\kappa_C(x_C; x_0)) \) equals the logarithm of the normalization constant plus

\[
\sum_{A \subseteq C} \left( \sum_{B \subseteq A \setminus \{C\}} (-1)^{|A|-|B|} \Psi_B(x_C|B; x_0) + \sum_{B \subseteq A \setminus \{C\}} (-1)^{|A|-k} \frac{k}{|A|} \Psi_B(x_C|B; x_0) \right)
\]

\[
= \sum_{B \subseteq C \setminus \{C\}} \left( \sum_{|B| < k} (-1)^{|B|} \right) \Psi_B(x_C|B; x_0)
\]

\[
+ \sum_{B \subseteq C \setminus \{C\}} \left( \sum_{|B| = k} (-1)^{|B|} \frac{k}{|B| + k} \right) \Psi_B(x_C|B; x_0)
\]

\[
= \sum_{B \subseteq C \setminus \{C\}} \left( \sum_{|B| < k} (-1)^{|B|} \right) \Psi_B(x_C|B; x_0)
\]

\[
+ \sum_{B \subseteq C \setminus \{C\}} \left( \sum_{|B| = k} (-1)^{|B|} \frac{k}{|B| + k} \right) \Psi_B(x_C|B; x_0)
\]

(10) \[
= \sum_{B \subseteq C \setminus \{C\}} \delta_{|C|,|B|} \Psi_B(x_C|B; x_0) + \sum_{B \subseteq C \setminus \{C\}} \left( \frac{1}{|C|} \right) \Psi_B(x_C|B; x_0),
\]

where the identity \( \sum_{j=0}^m \binom{m}{j} \frac{(-1)^j}{m+j} = 1/\binom{m+r}{m+1} \) was used and \( \delta_{a,b} \) denotes Kronecker’s delta. For \( |C| > k \) the first sum is empty, and it follows that \( \kappa_C \) satisfies the defining equality of \( \mathcal{M}_{k+1} \).

Conversely, if \( (\kappa_A) \) \( \in \mathcal{M}_{k+1} \), then let \( \alpha_{A,B} \) be as in (10), and let

\[
\Psi_B(x_B; x_0) = \log(\kappa_B(x_B; x_0)), \quad \text{for all } x_0 \in X_0, x_B \in X_B, |B| \leq k.
\]

These coefficients \( \alpha_{A,B} \) and functions \( \Psi_B \) together define an element \( (\tilde{\kappa}_A) \) \( \in \mathcal{M}_{k+1} \). The calculation (11) shows that

\[
\tilde{\kappa}_A(x_A; x_0) = \begin{cases} \frac{1}{|A|} \exp(\Psi_A(x_A; x_0) = \kappa_A(x_A; x_0), & \text{if } |A| \leq k, \\ \frac{1}{|A|} \exp \left( \sum_{B \subseteq A : |B| = k} \ln(\kappa_B(x_B|B; x_0)) \right), & \text{if } |A| > k, \end{cases}
\]

and so \( (\kappa_A) = (\tilde{\kappa}_A) \) belongs to \( \mathcal{M}_{k+1} \) and is of the desired form. \( \square \)

Theorem 13. Let \((\kappa_A)\) be functional modalities. Then there exist functional modalities \((\tilde{\kappa}_A)\) in the closure of \(\mathcal{M}_{k+1}\) such that the following holds: If \((\kappa_A)\) is \(k\)-robust on a set \(S \subseteq X_m\), then \(\kappa_A(x_A|A) = \tilde{\kappa}_A(x_A|A)\) for all \(A \subseteq [n]\) and all \(x \in S\). In particular, \((\tilde{\kappa}_A)\) belongs to the closure of the family of \((k+1)\)-interactions.
Therefore, if

where \( \tilde{A} \) is a normalization constant. By definition, \((\tilde{\kappa})\) lies in the closure of \( \hat{M}_{k+1} \). Let \( x \in S \) and \( C \subseteq [n] \). If \( |C| \leq k \), then \( \kappa_C(x|C) = \kappa_C(x|C) \) by definition of \( \kappa_A \). So assume that \( |C| > k \). By definition of \( k \)-robustness, if \( x \in S \), then \( \kappa_C(x|C) = \kappa_B(x|B) \) for all \( B \subseteq C \) with \( |B| = k \). Therefore, if \( x \in S \) and \( |C| > k \), then

\[
\kappa_C(x|C; x_0) = \left( \prod_{B \subseteq C, |B| = k} \kappa_B(x|B; x_0) \right)^{1/(|C|)}.
\]

Therefore, if \( x \in S \) and \( |C| > k \), then \( Z_{C|x|C} = 1 \) and \( \kappa_C(x|C) = \tilde{\kappa}_C(x|C) \). \( \square \)

Since \( M_{k+1} \) and \( \hat{M}_{k+1} \) are independent of \( S \), Theorem 13 shows that these two families can be used to construct robust systems, when the set \( S \) is not known a priori but must be learnt by the system, or when \( S \) changes with time and the system must adapt.

If we are not interested in all functional modalities but just the stochastic map \( \kappa \) describing the unperturbed system, we can describe \( \kappa \) in terms of low interaction order. The family of \((k + 1)\)-interaction stochastic maps, denoted by \( K_{k+1} \), consists of all strictly positive maps \( \kappa \) such that

\[
\ln \kappa(x; x_0) = \sum_{A \subseteq [n], |A| \leq k} \Psi_A(x|A; x_0)
\]

for some real functions \( \Psi_A : X_A \rightarrow \mathbb{R} \).

**Corollary 14.** Let \( \kappa \) be a stochastic map. For given \( k \) there exists a stochastic map \( \tilde{\kappa} \) in the closure of \( K_{k+1} \) such that the following holds: If \( \kappa \) is \( k \)-robust on a set \( S \), then \( \kappa(x) = \tilde{\kappa}(x) \) for all \( x \in S \).

**Proof.** If \( \kappa \) is \( k \)-robust on \( S \), there exist functional modalities \( \kappa_A \) with \( \kappa = \kappa_{[n]} \). Choose \( \tilde{\kappa}_A \) as in Theorem 13. If \( x \in S \), then \( \kappa(x) = \kappa_{[n]}(x) = \tilde{\kappa}_{[n]}(x) \). Hence the Corollary holds true with \( \tilde{\kappa} = \tilde{\kappa}_{[n]} \). \( \square \)

**Example 15.** The functional modalities \( \psi \) do not lie in \( M_2 \). This does not mean that neural networks are not robust: In fact, it is possible to naturally redefine the functional modalities \( \psi \) such that the new functional modalities lie in \( \hat{M}_2 \).

The construction \( \psi \) identifies the summand \( w_i x_i x_0 \) with \( \phi_{[i]} \). Now we will make another identification: For each \( i \in [n] \) let

\[
k_{[i]}(x_i; x_0) = \frac{1}{Z_{A, x_i}} \exp(n w_i x_i x_0 - \eta) .
\]

The unique extension of these stochastic maps to functional modalities \( \kappa_A \) in \( M_2 \) is given by

\[
\kappa_A(x|A; x_0) = \frac{1}{Z_{A, x_i}} \left( \prod_{i \in A} k_{[i]}(x_i; x_0) \right)^{1/|A|} = \frac{1}{Z_{A, x_i}} \exp \left( \frac{n}{|A|} \sum_{i \in A} w_i x_i x_0 - \eta \right) .
\]
where $Z_{A,x}$ and $Z'_{A,x}$ are constants determined by normalization. The functional modalities defined in this way lie in $\mathcal{M}_2$, and the stochastic map $\kappa_{[n]}$ agrees with (3).

Note that, by tuning the parameters $w_1, \ldots, w_n$, any combination of stochastic maps is possible for $\kappa_1, \ldots, \kappa_n$. This shows that any element of $\mathcal{M}_2$ has a representation of the form (12).

As in example (11) we can scale the weights $w_i$ and the threshold $\eta$ by a factor of $\beta$ and send $\beta \to +\infty$. This leads to the rule

$$\kappa_A(x_A; +1) \to \theta\left(\frac{n}{|A|} \sum_{i \in A} w_i x_i - \eta\right),$$

which is a normalized variant of (8).

The rule (12) implements a renormalization of the effect of the remaining inputs under knockout. Similar renormalization procedures are sometimes used when training neural networks using Hebb’s rule. Usually the total sum of the weights $\sum_i w_i$ is normalized to not grow to infinity. The rule (12) suggests that under knockout all remaining weights are amplified by a common factor.

The ideas leading to Theorem 13 can be applied to more general robustness structures $\mathcal{R}$ as follows: For any $x \in X$ let

$$\mathcal{R}_x := \left\{ \begin{array}{ll} |R \subseteq [n]: (R, x_R) \in \mathcal{R} \}, & \text{if there exists } R \subseteq [n] \text{ with } (R, x_R) \in \mathcal{R}, \\ \{[n]\}, & \text{else}, \end{array} \right.$$ 

and let $\mathcal{R}_x^{\text{min}}$ be the subset of inclusion-minimal elements of $\mathcal{R}_x$. If $(\kappa_A)$ is $\mathcal{R}$-robust in $S$, then

$$\kappa(x; x_0) = \kappa_R(x_R; x_0) \quad \text{for any } R \in \mathcal{R}_x^{\text{min}}, \; x \in S$$

and hence

$$\kappa(x; x_0) = \left( \prod_{R \in \mathcal{R}_x^{\text{min}}} \kappa_R(x_R; x_0) \right)^{1/|\mathcal{R}_x^{\text{min}}|}.$$ 

For any $C \subseteq [n]$ let $\mathcal{R}_x^{\text{min}}(C) = \{ R \in \mathcal{R}_x^{\text{min}} : C \subseteq R \}$. If $\mathcal{R}$ is coherent, then we can deduce

$$\kappa_C(x_C; x_0) = \left( \prod_{R \in \mathcal{R}_x^{\text{min}}(C)} \kappa_R(x_R; x_0) \right)^{1/|\mathcal{R}_x^{\text{min}}(C)|}$$

for all $x \in S$ with $\mathcal{R}_x^{\text{min}}(C) \neq \emptyset$. This motivates the following definition: Denote by $\bar{\mathcal{M}}_{\mathcal{R}}$ the set of all strictly positive functional modalities that satisfy

$$\kappa_C(x_C; x_0) = \frac{1}{Z_{C,x|C}} \left( \prod_{R \in \mathcal{R}_x^{\text{min}}(C)} \kappa_R(x_R; x_0) \right)^{1/|\mathcal{R}_x^{\text{min}}(C)|}$$

for all $x \in X$ and all $C \subseteq [n]$ with $\mathcal{R}_x^{\text{min}}(C) \neq \emptyset$, where $Z_{C,x|C}$ is a suitable normalization constant. The same proof as for Theorem 13 implies:

**Theorem 16.** Let $(\kappa_A)$ be functional modalities, and assume that $\mathcal{R}$ is coherent. Then there exist functional modalities $(\bar{\kappa}_A)$ in the closure of $\bar{\mathcal{M}}_{\mathcal{R}}$ such that the following holds: If $(\kappa_A)$ is $\mathcal{R}$-robust on a set $S \subseteq X$, then

$$\kappa_A(x_A) = \bar{\kappa}_A(x_A) \quad \text{for all } x \in S.$$
As a generalization of Lemma[14] we can also describe \( \tilde{\mathcal{M}}_R \) as a set of functional modalities with limited interaction order. To simplify the presentation, we assume that \( \mathcal{R} \) is saturated, by which we mean the following: If \( (R, x_R) \in \mathcal{R} \) for some \( x_R \in \mathcal{X}_R \), then \( (R, x'_R) \in \mathcal{R} \) for all \( x_R \in \mathcal{X}_R \). In other words, a saturated robustness specification is given by enumerating a family of subsets of \([n] \). For example, the robustness structures \( \mathcal{R}_k \) are saturated, while the robustness structures defining canalyzing and nested canalyzing functions (see Section 3) are not saturated. If \( \mathcal{R} \) is saturated, then \( \mathcal{R}_x \) and \( \mathcal{R}_x^{\text{min}} \) are independent of \( x \in \mathcal{X} \).

Consider the family

\[
\Delta = \left\{ C \subseteq [n] : C \subseteq R \text{ for some } R \in \mathcal{R}_x^{\text{min}} \text{ and } x \in \mathcal{X} \right\},
\]

and let \( \Delta(C) = \{ R \in \Delta : R \subseteq C \} \). Let \( M_\Delta \) be the set of all functional modalities \( (\kappa_C) \) such that there exist potentials \( \phi_A \) of the form

\[
\phi_A(x_A; x_0) = \sum_{B \in \Delta(A)} \alpha_{A,B} \Psi_B(x_A|B; x_0),
\]

where \( \alpha_{A,B} \in \mathbb{R} \) and \( \Psi_B \) is an arbitrary function \( \mathbb{R}^{A \times X_0} \to \mathbb{R} \). We call \( M_\Delta \) the family of \( \Delta \)-interaction functional modalities. Note that the functions \( \Psi_B \) do not depend on \( A \). This ensures a certain interdependence among the functional modalities \( \kappa_C \).

**Lemma 17.** Assume that \( \mathcal{R} \) is coherent and saturated. \( \tilde{\mathcal{M}}_R \) is a subset of \( M_\Delta \).

**Proof.** If \( \mathcal{R}_x = \emptyset \), then \( \Delta \) contains all sets. The Möbius inversion formula shows that \( M_\Delta \) contains all strictly positive functional modalities. Therefore, we may assume that \( \mathcal{R}_x \neq \emptyset \).

Define Gibbs potentials using the Möbius inversion [6]. If \( x \in \mathcal{S} \) and \( A \) is large enough such that \( \mathcal{R}_x^{\text{min}}(A) \neq \emptyset \), then

\[
\sum_{C \subseteq A} (-1)^{|A|} \ln \kappa_C(x_{|C}; x_0) = \sum_{C \subseteq A} (-1)^{|A|} \frac{1}{|\mathcal{R}_x^{\text{min}}(C)|} \sum_{B \in \mathcal{R}_x^{\text{min}}(C)} \ln \kappa_C(x_{|B}; x_0)
\]

\[
= \sum_{C \subseteq A} (-1)^{|A|} \frac{1}{|\mathcal{R}_x^{\text{min}}(C)|} \sum_{B \in \mathcal{R}_x^{\text{min}}(C)} \ln \kappa_B(x_{|B}; x_0)
\]

\[
= \sum_{B \in \mathcal{R}_x^{\text{min}}(C)} \left\{ \sum_{R \subseteq A \setminus B} (-1)^{|A| - |R| - k} \frac{1}{|\mathcal{R}_x^{\text{min}}(B \cup R)|} \right\} \ln \kappa_B(x_{|B}; x_0).
\]

Together with (6) this gives

\[
\phi_A(x_A; x_0) = \sum_{C \subseteq A} \alpha_{A,C} \ln \kappa_C(x_{|C}; x_0) + \sum_{C \subseteq A} \alpha_{A,C} \ln \kappa_C(x_{|C}; x_0),
\]

where

\[
\alpha_{A,C} = \begin{cases} (-1)^{|A| - |C|}, & \text{if } C \notin \mathcal{R}_x, \\ \sum_{R \subseteq A \setminus B} (-1)^{|A| - |R| - k} \frac{1}{|\mathcal{R}_x^{\text{min}}(B \cup R)|} & \text{if } C \in \mathcal{R}_x^{\text{min}}. \end{cases}
\]

This is clearly of the form (15).

In the case \( \mathcal{R} = \mathcal{R}_k \) the sum \( \sum_{R \subseteq A \setminus B} (-1)^{|A| - |R| - k} \frac{1}{|\mathcal{R}_x^{\text{min}}(B \cup R)|} \) that appears in the proof of Lemma[17] can be solved explicitly, resulting in the statement of Lemma[12]. In the general case this is not possible.
Corollary 14 also generalizes. Let $\Lambda$ be as above. The set $K_\Lambda$ of $\Delta$-interactions stochastic maps consists of all strictly positive stochastic maps $\kappa$ such that

$$\ln \kappa(x; x_0) = \sum_{A \in \Lambda} \Psi_A(x|A; x_0)$$

for some real functions $\Psi_A : \mathcal{X}_A \to \mathbb{R}$.

**Corollary 18.** Let $\kappa$ be a stochastic map, and let $\mathcal{R}$ be a coherent and saturated robustness specification. There exists a stochastic map $\tilde{\kappa}$ in the closure of $K_\Lambda$ such that the following holds: If $\kappa$ is $\mathcal{R}$-robust on a set $S$, then $\kappa(x) = \tilde{\kappa}(x)$ for all $x \in S$.

The proof is the same as the proof of Corollary 14.

**Remark 19.** Instead of representing functional modalities as a family $(\kappa_A)$ of stochastic maps, it is possible to use a single stochastic map $\hat{\kappa}$, operating on a larger space, that integrates the information from the family $(\kappa_A)$. The stochastic map $\hat{\kappa}$ can be constructed as follows: For each $i = 1, \ldots, n$ let $\tilde{\mathcal{X}}_i$ be the disjoint union of $\mathcal{X}_i$ and one additional element, denoted by 0. This additional state represents the knockout of $\mathcal{X}_i$. Let $\tilde{\mathcal{X}}_0 = \tilde{\mathcal{X}}_1 \times \cdots \times \tilde{\mathcal{X}}_n$. For each $y \in \tilde{\mathcal{X}}_0$ let $\text{supp}(y) = \{i : y_i \neq 0\}$. We define the stochastic map $\hat{\kappa} : \mathcal{X}_0 \times \tilde{\mathcal{X}}_m$ via

$$\hat{\kappa}(x; x_0) = \kappa_{\text{supp}(x)}(x|\text{supp}(x); x_0).$$

This construction gives a one-to-one correspondence between functional modalities and stochastic maps from $\tilde{\mathcal{X}}_0$ to $\mathcal{X}_0$.

As an example, consider the functional modalities defined in (4). In this example, the construction of $\hat{\kappa}$ is particularly easy: It just amounts to extending the input space to $\{-1, 0, +1\}^n$. Equation (3) remains valid for $\hat{\kappa}$. The construction is more complicated for the functional modalities (12).

More generally, any Gibbs representation for functional modalities $(\kappa_A)$ as in (5) extends to a Gibbs representation of $\hat{\kappa}$: For any $B \subseteq [n]$, $x_0 \in \mathcal{X}_0$ and $x \in \tilde{\mathcal{X}}_m$ let

$$\hat{\phi}_B(x; x_0) = \begin{cases} \phi_B(x|B; x_0), & \text{if sup}(x) \subseteq B, \\ 0, & \text{else.} \end{cases}$$

Then

$$\hat{\kappa}(x; x_0) = \frac{e^{\sum_{B \subseteq [n]} \phi_B(x; x_0)}}{\sum_{y_0 \in \mathcal{X}_0} e^{\sum_{B \subseteq [n]} \phi_B(y; x_0)}}.$$

5. Robustness and Conditional Independence

Given the probability distribution $p_{in}$ of the input variables and a stochastic map $\kappa$ describing the system, the joint probability distribution of the complete system can be computed from

$$p(x_0, x) = \kappa(x; x_0)p_{in}(x), \quad \text{for all } (x_0, x) \in \mathcal{X},$$

As shown in Proposition 2, robustness of stochastic maps is related to conditional independence constraints on the joint distribution. In this section we study the set of all joint distributions that arise from robust systems in this way. Let $\mathcal{R}$ be a robustness specification. By Proposition 2 the stochastic map $\kappa$ is $\mathcal{R}$-robust on supp($p_{in}$) if and only if for all $(R, x_R) \in \mathcal{R}$ the output $X_0$ is (stochastically) independent of $X_{\{n\} \setminus R}$, given that $X_R = x_R$. In the following, this conditional
This observation allows to reformulate the equivalence (1) as follows:

\[ p(x_0 = x_0 \mid X_{in} = x) = p(x_0 = x_0 \mid X_R = x_R), \quad \text{for all } x \in X_{in} \text{ with } p(x) > 0 \]

and \( x_{in} = x_R \). It is often convenient to use another definition that avoids the need to work with conditional distributions: The statement

\[ p(x_0 = x_0 \mid X_{in} = x) = p(x_0 = x_0 \mid X_R = x_R), \quad \text{for all } x \in X_{in} \text{ with } p(x) > 0 \]

and \( x_{in} = x_R \). Each equation (16) imposes conditions on this matrix saying that certain submatrices have rank one. To be precise, for any edge \((x, x')\) in the graph \(G_R\) (defined in Section 2) equations (16) for all \(x_0, x'_0 \in X_0\) require that the submatrix \((p_{kz})_{k \in X_0, z \in \{x, x'\}}\) has rank one. For any \(x \in X_{in}\) denote by \(\tilde{p}_x\) the vector with components \(\tilde{p}_x(z) = p(x_0 = x_0, X_{in} = z)\) for \(x_0 \in X_0\). Then a distribution \(p\) lies in \(\mathcal{P}_R\) if and only if \(\tilde{p}_x\) and \(\tilde{p}_y\) are proportional for all edges \((x, y)\) of \(G_R\). Observe that \(\tilde{p}_x\) and \(\tilde{p}_y\) are proportional if and only if either (i) one of \(\tilde{p}_x\) and \(\tilde{p}_y\) vanishes or (ii) \(\kappa(x) = \kappa(y)\). This observation allows to reformulate the equivalence (1) \(\Leftrightarrow\) (3) of Proposition 2 as follows:

**Lemma 20.** Let \(S = \{x \in X_{in} : \tilde{p}_x \neq 0\}\). A distribution \(p\) lies in \(\mathcal{P}_R\) if and only if \(\tilde{p}_x\) and \(\tilde{p}_y\) are proportional whenever \(x, y \in S\) lie in the same connected component of \(G_{R,S}\).

For any family \(B\) of subsets of \(X_{in}\) let \(\mathcal{P}_B\) be the set of probability distributions \(p\) on \(X_i\) that satisfy the following two conditions:

1. \(\bigcup B = \{x \in X_{in} : \tilde{p}_x \neq 0\}\).
2. \(\tilde{p}_x\) and \(\tilde{p}_y\) are proportional, whenever there exists \(Z \in B\) such that \(x, y \in Z\).

Then \(\mathcal{P}_R = \bigcup_B \mathcal{P}_B\), where the union is over all \(\mathcal{R}\)-robustness structures \(B\). The disadvantage of this decomposition is that there are \(\mathcal{R}\)-robustness structures \(B, B'\) such that \(\mathcal{P}_B\) is a subset of the topological closure \(\overline{\mathcal{P}_B}\) of \(\mathcal{P}_B\). In other words, each \(p \in \mathcal{P}_B\) can be approximated arbitrarily well by elements of \(\mathcal{P}_{B'}\), and therefore in many cases it suffices to only consider \(\mathcal{P}_{B'}\). The following definition is needed:

**Definition 21.** An \(\mathcal{R}\)-robustness structure \(B\) is **maximal** if and only if \(\bigcup B = \bigcup_{Z \in B} Z\) satisfies any of the following equivalent conditions:

1. For any \(x \in X_{in} \setminus \bigcup B\) there are edges \((x, y), (x, z)\) in \(G_R\) such that \(y, z \in \bigcup B\) do not lie in the same connected component of \(G_{\mathcal{R},B}\).
2. For any \(x \in X_{in} \setminus \bigcup B\) the induced subgraph \(G_{\mathcal{R} \cup \bigcup B \cup \{x\}}\) has fewer connected components than \(G_{\mathcal{R} \cup \bigcup B}\).
Lemma 22. \( \mathcal{P}_R \) equals the disjoint union \( \bigcup_B \mathcal{P}_B \), where the union is over all \( R \)-robustness structures. Alternatively, \( \mathcal{P}_R \) equals the (non-disjoint) union \( \bigcup_B \overline{\mathcal{P}}_B \), where the union is over all maximal \( R \)-robustness structures.

Proof. The first statement follows directly from the above considerations. To see that it suffices to take maximal \( R \)-robustness structures in the second decomposition, consider an \( R \)-robustness structure \( B \) that is not maximal. By definition there exists \( x \in X_{\text{in}} \setminus \cup B \) such that the induced subgraph \( G_{R \cup B \cup \{x\}} \) has at least as many connected components as \( G_{R \cup B} \). Let \( B' \) be the family of connected components of \( G_{R \cup B \cup \{x\}} \). If \( G_{R \cup B \cup \{x\}} \) has the same number of connected components as \( G_{R \cup B} \), then there is \( y \in B \) such that \( y \cup \{x\} \in B' \), otherwise let \( y \in B \) be arbitrary. Let \( y \in \mathcal{Y} \). For any \( p \in \mathcal{P}_B \) and \( \epsilon > 0 \) define a probability distribution \( p_\epsilon \) via

\[
p_\epsilon(x_0, z) = \begin{cases} p(x_0, z), & \text{if } z \notin \{x, y\}, \\ (1 - \epsilon)p(x_0, x), & \text{if } z = y, \\ \epsilon p(x_0, x), & \text{if } z = x. \end{cases}
\]

Then \( p_\epsilon \in \mathcal{P}_{B'} \), and hence \( \mathcal{P}_B \subseteq \overline{\mathcal{P}}_{B'} \). If \( B' \) is not maximal, we may iterate the process. \( \square \)

The following lemma sheds light on the structure of \( \overline{\mathcal{P}}_B \):

Lemma 23. Fix an \( R \)-robustness structure \( B \). Then \( \overline{\mathcal{P}}_B \) consists of all probability measures of the form

\[
(17) \quad p(x_0 = x_0, X_{\text{in}} = x) = \begin{cases} \mu(\mathcal{Z})\lambda_\mathcal{Z}(x)p_{\mathcal{Z}}(x_0), & \text{if } x \in \mathcal{Z} \in B, \\ 0, & \text{if } x \in X_{\text{in}} \setminus \cup B, \end{cases}
\]

where \( \mu \) is a probability distribution on \( B \) and \( \lambda_\mathcal{Z} \) is a probability distribution on \( \mathcal{Z} \) for each \( \mathcal{Z} \in B \) and \( (p_{\mathcal{Z}})_{\mathcal{Z} \in B} \) is a family of probability distributions on \( X_0 \).

Proof. It is easy to see that (17) defines indeed a probability distribution. By Lemma 20 it belongs to \( \mathcal{P}_B \). In the other direction, any probability measure can be written as a product

\[
p(x_0, x_1, \ldots, x_n) = p(\mathcal{Z})p(x_1, \ldots, x_n|X_1, \ldots, X_n) \in \mathcal{Z}) p(x_0|x_1, \ldots, x_n),
\]

if \( (x_1, \ldots, x_n) \in \mathcal{Z} \in B \), and if \( p \) is an \( R \)-robust probability distribution, then \( p_{\mathcal{Z}}(x_0) : = p(x_0|x_1, \ldots, x_n) \) depends only on the block \( \mathcal{Z} \) in which \( (x_1, \ldots, x_n) \) lies. \( \square \)

Lemma 22 decomposes the set \( \mathcal{P}_R \) of robust probability distributions into the closures of the smooth manifolds \( \mathcal{P}_B \), where \( B \) runs over the maximal \( R \)-robustness structures. Lemma 23 gives natural parametrizations of these manifolds.

By comparison, Theorem 16 and Lemma 17 describe robustness from a different point of view. The result can be translated to the setting of this section as follows:

Corollary 24. Suppose that \( R \) is a coherent and saturated robustness structure, and define \( \Delta \) as in Section 2. If \( p \in \mathcal{P}_B \), then there exists a stochastic map \( k \in K_\Delta \) such that \( p(x_0|x) = k(x; x_0) \) for all \( x \in \cup B \).

In the statement of the corollary note that \( p(X_{\text{in}} = x) > 0 \) for all \( x \in \cup B \), and hence the conditional distribution \( p(x_0|x) \) is well-defined in this case.
Corollary 24 can also be viewed from the perspective of hierarchical models: Let $\hat{\Delta} = \{1, \ldots, n\} \cup \{S \cup \{0\} : S \in \Delta\}$. The hierarchical loglinear model $E_{\hat{\Delta}}$ consists of all probability distributions $p$ on $X$ of the form

$$\log(p(x)) = \sum_{A \subseteq \hat{\Delta}} \phi_A(x|A),$$

where $\phi_A$ is a real function with domain $X_A$. By the results of this section, $E_{\hat{\Delta}}$ is a smooth manifold containing $\mathcal{P}_R$ in its closure. See [11, 4] for more on hierarchical loglinear models.

Remark 25. It is also possible to derive the decomposition in Lemma 22 from results from commutative algebra. Since the equations (16) that describe conditional independence are algebraic, they generate a polynomial ideal, called conditional independence ideal. In this case the ideal is a generalized binomial edge ideal, as defined in [13]. For such ideals, the primary decomposition is known and corresponds precisely to the decomposition of the set of robust distributions as presented in Lemma 22. The parametrization of Lemma 23 can be considered as a surjective polynomial map and shows that all components of the decomposition are rational.

6. $k$-robustness

In this section we consider the symmetric case $R = R_k$. As above, we replace any prefix or subscript $R$ by $k$.

If $k = 0$, then any pair $(x, y)$ is an edge in $G_0$. This means that any 0-robustness structure $B$ contains only one set. There is only one maximal 0-robustness structure, namely $B = \{X_{\text{in}}\}$. The set $R_0$ is irreducible. This corresponds to the fact that $\mathcal{P}_0$ is defined by $X_0 \perp X_{\text{in}}$.

If $k = 1$, then any pair $(x, y)$ is an edge in $G_0$. This means that any 0-robustness structure $B$ contains only one set. There is only one maximal 0-robustness structure, namely $B = \{X_{\text{in}}\}$. The set $R_0$ is irreducible. This corresponds to the fact that $\mathcal{P}_0$ is defined by $X_0 \perp X_{\text{in}}$.

In [5] a different description is given: The block $S_1 \times S_2$ can be identified with the complete bipartite graph on $S_1$ and $S_2$. In this way, every maximal 1-robustness structure corresponds to a collection of complete bipartite subgraphs with vertices in $X_1 \cup X_2$ such that every vertex in $X_1$ and $X_2$, respectively, is part of one such subgraph. Figure 2 shows an example.

This result generalizes in the following way:

Lemma 26. A 1-robustness structure $B$ is maximal if and only if the following statements hold:

- Each $B \in B$ is of the form $B = S_1 \times \cdots \times S_n$, where $S_1 \subseteq X_1, S_2 \subseteq X_2$.
- For every $x_1 \in X_1$ there exists $B \in B$ and $x_2 \in X_2$ such that $(x_1, x_2) \in B$, and conversely.

In [5] a different description is given: The block $S_1 \times S_2$ can be identified with the complete bipartite graph on $S_1$ and $S_2$. In this way, every maximal 1-robustness structure corresponds to a collection of complete bipartite subgraphs with vertices in $X_1 \cup X_2$ such that every vertex in $X_1$ and $X_2$, respectively, is part of one such subgraph. Figure 2 shows an example.

This result generalizes in the following way:

Lemma 26. A 1-robustness structure $B$ is maximal if and only if the following statements hold:

- Each $B \in B$ is of the form $B = S_1 \times \cdots \times S_n$, where $S_i \subseteq X_i$.
- $\bigcup_{S_1 \times \cdots \times S_n \in B} S_i = X_i$ for all $i \in [n]$. 

This shows that each condition follows from the fact that we assume to the contrary that there are \( i \in X \) such that maximality we can conclude \( y = \emptyset \). By maximality we need to show that \( G_{R,Y} \) is connected and that \( G_{R,Z,Y} \) is not connected for all \( Z \in B \setminus \{Y\} \). The first condition follows from the fact that \( G_{R,Y} \) is connected. For the second condition assume to the contrary that there are \( x \in Y \) and \( y \in Z \) such that \( x = (x_1, \ldots, x_n) \) and \( y = (y_1, \ldots, y_n) \) disagree in at most \( n - 1 \) components. Then there exists a common component \( x_i = y_i \). By construction there exists \( z = (z_1, \ldots, z_n) \in Y \) such that \( z_i = y_i = x_i \), hence \( Y \cup Z \) is connected, in contradiction to the assumptions. This shows that each \( Y \) has a product structure.

Write \( Y = S^Y_1 \times \cdots \times S^Y_n \) for each \( Y \in B \). Obviously \( S^Y_i \cap S^Z_i = \emptyset \) for all \( i \in [n] \) and all \( Y, Z \in B \) if \( Y \neq Z \). For the second assertion, assume to the contrary that \( I \in X \) is contained in no \( S^Y_i \). Take any \( Y \in B \) and define \( Y' := S^Y_\emptyset \times \cdots \times (S^Y_I \cup \{I\}) \times \cdots \times S^Y_n \). Then \( (B \setminus \{Y\}) \cup (Y') \) is another 1-robustness structure with the same number of components as \( B \), contradicting the assumptions.

Conversely, assume that \( B \) is a 1-robustness structure satisfying the two assertions of the theorem. For any \( x \in X \setminus \cup B \) there exist \( y_1, \ldots, y_n \in \cup B \) such that \( x_i = y_{i_1}, \ldots, x_n = y_{i_n} \). Since \( x \notin \cup B \) the points \( y_1, \ldots, y_n \) cannot all belong to the same block of \( B \). If \( y_i \) and \( y_j \) belong to different blocks of \( B \), then the two edges \( (x, y_i) \) and \( (x, y_j) \) of \( G_1 \) show that \( B \) is maximal.

The last result can be reformulated in terms of \( n \)-partite graphs generalizing [5]: Namely, the 1-robustness structures are in one-to-one relation with the \( n \)-partite subgraphs of \( M_{d_1, \ldots, d_n} \) such that every connected component is itself a complete \( n \)-partite subgraph \( M_{e_1, \ldots, e_n} \) with \( e_i > 0 \) for all \( i \in [n] \). Here, an \( n \)-partite graph is a graph which can be coloured by \( n \) colours such that no two vertices with the same colour are adjacent.

Unfortunately the nice product form of the maximal 1-robustness structures does not generalize to \( k > 1 \):

\[ \text{Example 27 (Binary inputs). If } n = 3 \text{ and } d_1 = d_2 = d_3 = 2, \text{ then the graph } G_2 \text{ is the graph of the cube. For a maximal 2-robustness structure } B \text{ the set } X_{\text{in}} \setminus \cup B \text{ can be any one of the following (see Fig. [5]):} \]

- The empty set
- A set of cardinality 4 corresponding to a plane leaving two connected components of size 2
- A set of cardinality 4 containing all vertices with the same parity.

\[ \text{Figure 2. A 1-robustness structures for two variables. a) The graph } G_{1,S}. \text{ b) The representation in terms of bipartite graphs.} \]
Figure 3. The four symmetry classes of maximal 2-robustness structures of three binary inputs, see Example 27.

Figure 4. A maximal 3-robustness structure for four binary inputs.

Figure 5. The 2-robustness structure from Example 28. The graph $G_2$ is the graph of a hypercube of dimension four, where diagonals have been added to the two-dimensional faces. Only the edges of $G_2$ that connect vertices of Hamming distance one are shown, and the edges of $G_{2,\cup B}$. The two blocks are marked in green and red.

- A set of cardinality 3 cutting off a vertex.

In the last case only the isolated vertex has a product structure (Fig. 4d).

If $n = 4$ and $d_1 = d_2 = d_3 = d_4 = 2$, then the graph $G_3$ is the graph of a hyper-cube. Figure 4 shows how a maximal 3-robustness structure can look like.

$k$-robustness implies $(k + 1)$-robustness, and therefore $P_k \subseteq P_{k+1}$. This does not mean that all $k$-robustness structures are also $(k + 1)$-robustness structures, for the following reason: If $B$ is a $k$-robustness structure and $S = \cup B$, then $G_{k+1,S}$ may have more connected components than $G_{k,S}$.

**Example 28.** Consider $n = 4$ binary random variables $X_1, \ldots, X_4$. Then

$$B := \{(1, 1, 1, 1), (2, 2, 1, 1), (1, 2, 2, 2), (2, 1, 2, 2)\}$$

is a maximal 2-robustness structure. Both elements of $B$ are connected in $G_2$, but not in $G_3$, see Fig. 5.

Nevertheless, the notions of $l$-robustness and $k$-robustness for $l > k$ are related as follows:

**Lemma 29.** Assume that $d_1 = \cdots = d_n = 2$, and let $B$ be a maximal $k$-robustness structure of binary random variables. Then each $B \in B$ is connected as a subset of $G_s$, for all $s \leq n - 2k + 1$.

**Proof.** We can identify elements of $X_n$ with binary strings of length $n$. Denote by $I_l$ the string $1 \ldots 10 \ldots 0$ of $r$ ones and $n - r$ zeroes in this order. Without loss of generality assume that $I_0, I_l$ are two elements of $B \in B$, where $k \leq n - l < s \leq$
Assume that

\[ D-04103 \text{ L} \]

S. L. Lauritzen, I. Swanson and A. Taylor, “Minimal primes of ideals arising from conditional independence

J. Rauh, “Generalized binomial edge ideals,”

M. Ohtani, “Graphs and ideals generated by some 2-minors.”

S. Kaufmann, C. Peterson, B. Samuelsson, and C. Troein, “Genetic networks with analyzing Boolean rules are always stable,” PNAS, vol. 101, no. 49, pp. 17 102–17 107, 2004.

S. L. Lauritzen, Graphical Models, 1st ed., ser. Oxford Statistical Science Series. Oxford University Press, 1996.

M. Ohtani, “Graphs and ideals generated by some 2-minors.” Commun. Algebra, vol. 39, no. 3, pp. 905–917, 2011.

J. Rauh, “Generalized binomial edge ideals,” Advances in Applied Mathematics, accepted, 2012.

I. Swanson and A. Taylor, “Minimal primes of ideals arising from conditional independence statements,” Preprint: arXiv:1107.5604v3, 2011.

Acknowledgement. This work has been supported by the Volkswagen Foundation and the Santa Fe Institute. Nihat Ay thanks David Krakauer and Jessica Flack for many stimulating discussions on robustness.

References

[1] N. Ay and D. C. Krakauer, “Geometric robustness theory and biological networks,” Theory in Biosciences, vol. 125, no. 2, pp. 93 – 121, 2007.

[2] G. Boldhaus, N. Bertschinger, J. Rauh, E. Olbrich, and K. Klemm, “Robustness of boolean dynamics under knockouts,” Phys. Rev. E, vol. 82, no. 2, p. 021916, 2010.

[3] E. S. de Cabezón and H. Wynn, “Algebraic reliability based on monomial ideals: A review,” in Harmony of Gröbner Bases and the Modern Industrial Society, T. Hibi, Ed. World Scientific, 2012, pp. 314–335.

[4] M. Drton, B. Sturmfels, and S. Sullivant, Lectures on Algebraic Statistics, 1st ed., ser. Oberwolfach Seminars. Birkhäuser, Basel, 2009, vol. 39.

[5] A. Fink, “The binomial ideal of the intersection axiom for conditional probabilities,” Journal of Algebraic Combinatorics, vol. 33, no. 3, pp. 455–463, 2011.

[6] J. Herzog, T. Hibi, F. Hreinsdóttir, T. Kahle, and J. Rauh, “Binomial edge ideals and conditional independence statements,” Advances in Applied Mathematics, vol. 45, no. 3, pp. 317 – 333, 2010.

[7] A. S. Jarrah and R. C. Laubenbacher, “Discrete models of biochemical networks: The toric variety of nested canalyzing functions,” in Algebraic Biology, 2007, pp. 15–22.

[8] A. S. Jarrah, B. Raposa, and R. Laubenbacher, “Nested canalyzing, unate cascade, and polynomial functions,” Physica D, vol. 233, no. 2, pp. 167 – 174, 2007.

[9] S. Kaufman, The Origins of Order: Self-Organization and Selection in Evolution. Oxford University Press, 1993.

[10] S. Kaufman, C. Peterson, B. Samuelsson, and C. Troein, “Genetic networks with analyzing Boolean rules are always stable,” PNAS, vol. 101, no. 49, pp. 17 102–17 107, 2004.

[11] S. L. Lauritzen, Graphical Models, 1st ed., ser. Oxford Statistical Science Series. Oxford University Press, 1996.

[12] M. Ohtani, “Graphs and ideals generated by some 2-minors.” Commun. Algebra, vol. 39, no. 3, pp. 905–917, 2011.

[13] J. Rauh, “Generalized binomial edge ideals,” Advances in Applied Mathematics, accepted, 2012.

[14] I. Swanson and A. Taylor, “Minimal primes of ideals arising from conditional independence statements,” Preprint: arXiv:1107.5604v3, 2011.