Robustness to Dependency in Influence Maximization

Louis Chen[1], Chee Chin Lim[2], Divya Padmanabhan[3], and Karthik Natarajan[4]

[1]Operations Research Department, Naval Postgraduate School, Monterey, California, 93943

[2]Institute of Operations Research and Analytics, National University of Singapore, Singapore 117602, Singapore

[3]School of Mathematics and Computer Science at Indian Institute of Technology (IIT), Goa, India

[4]Engineering Systems and Design,Singapore University of Technology and Design, Singapore 487372, Singapore

Abstract

In this paper, we pursue a correlation robust study of the influence maximization problem. Departing from the classic independent cascade model, we study a diffusion process adversarially-adapted to the choice of seed set. More precisely, rather than the independent coupling of known individual edge probabilities, we now evaluate a seed set's expected influence under all possible correlations - specifically, the one that presents the worst-case. We find that the worst-case expected influence can be efficiently computed, its NP-hard optimization done approximately $(1 - 1/e)$ with greedy construction, and we provide complete, efficient characterizations of the adversarial coupling, the random graph, and the random number of influenced nodes. But most importantly, upon mixing the independent cascade with the worst-case, we attain a tunable and more comprehensive model better suited for real-world diffusion phenomena than the independent cascade alone, and without increased computational complexity. Extensions to the correlation robust study of risk follow, along with numerical experiments on network data sets, with demonstration of how our models can be tuned.

1 Introduction

Today's world is an increasingly connected and complex network, presenting both exciting possibilities as well as concerning challenges. Online platforms, in particular, now present a powerful medium for the flow of information, and motivates study of propagation in a network under influences both natural and adversarial. We find this challenge in many settings such as the combat of the spread of epidemics [Hoffmann and Caramanis, 2018], but perhaps most notably so in the use of social networks [Hunter and Zaman, 2022], [Chen et al., 2011] to influence or spread opinions, to test the effectiveness of policies, to promote adoption of medical innovations and to conduct viral marketing campaigns [Mallipeddi et al., 2022].

There are now many models for diffusion and the spread of influence in graphs [Watts, 2002], [Kempe et al., 2003], [Chen et al., 2009], [Li et al., 2018]. Central to such models is a directed graph $G = (V, E)$ where V is the set of nodes (denoting members or users) and E is the set of edges (denoting influence relationships). While we focus on directed edges, it is straightforward to extend all results to undirected graphs. In progressive diffusion models, a subset of the edges are randomly deemed "live" via a binary valued random vector $\tilde{\boldsymbol{c}}$ of size $|E|$, in which $\tilde{c}_{ij} = 1$ with probability (w.p.) p_{ij} if and only if edge (i, j) is live. When $\tilde{c}_{ij} = 1$, node i can influence node j and when $\tilde{c}_{ij} = 0$, node i cannot influence node j. Given a seed set $S \subseteq V$, all nodes reachable along live-edge paths from S (including S itself) are activated, or "influenced". We note that this type of modeling does not take into account the role of time, unlike models in [Bass, 1969], [Akbarpour and Jackson, 2018].

The formalization and study of influence maximization began with the seminal work of [Kempe et al., 2003], [Kempe et al., 2015], which focused on, among other diffusion models, the well-known *Independent Cascade (IC)* model, in which all edges are independently live; equivalently, the components of $\tilde{\boldsymbol{c}}$ are mutually independent Bernoulli random variables, and we denote this by $\tilde{\boldsymbol{c}} \sim \theta_{ic}$. The problem is to find a k-sized seed set $S \subseteq V$ that maximizes the number of influenced nodes (denoted by the random variable $R(\tilde{\boldsymbol{c}}, S)$) in expectation under θ_{ic}:

$$\max_{S:|S|\leq k} E_{\theta_{ic}}\left[R(\tilde{\boldsymbol{c}}, S)\right]. \qquad \text{(IC Influence Maximization)}$$

The IC influence maximization is known to be NP-hard [Kempe et al., 2003]; in fact, even evaluating the objective function $f^{ic}(S) := \mathbb{E}_{\theta_{ic}}[R(\tilde{\boldsymbol{c}}, S)]$ for a fixed seed set S is #P-hard [Chen et al., 2010]. Importantly, since the objective function was shown to be a nondecreasing submodular function by [Kempe et al., 2003], the greedy algorithm of [Nemhauser et al., 1978] provides provable approximation guarantees for this problem. Currently the best known implementation of the greedy algorithm in conjunction with sampling based methods provides a $(1 - 1/e - \epsilon)$ approximation for any $\epsilon > 0$ in time $\mathcal{O}((|V| + |E|)\log(|V|)\frac{1}{\epsilon^3})$ or $\mathcal{O}((|V| + |E|)k\log(|V|)\frac{1}{\epsilon^2})$ for IC Influence Maximization (see [Borgs et al., 2014]). It is also known that it is NP-hard to approximate the optimal value to within a factor better than $1 - 1/e$. The polynomial dependence of the running time on $1/\epsilon$ and not $\log(1/\epsilon)$ is a consequence of the #P-hardness of computing expectations [Klein Haneveld, 1986,Meilijson and Nádas, 1979] as in $f^{ic}(\mathcal{S})$ [Chen et al., 2010,Kempe et al., 2003].

Despite the attention devoted to the independent cascade model, the assumption that edges independently realize can certainly be inappropriate, for reasons natural and/or adversarial. For example, activity in social networks is often "bursty" - see [Akbarpour and Jackson, 2018] - so that times of activity may rarely overlap, especially if, say, two members live in different time zones so that their presences online are negatively correlated. As well, influence relationships often depend on the particular idea/product/news to be propagated; for instance, political news might be shared among members with similar views only. These examples serve to illustrate that while the set E may outline connections in a social network, hidden latent variables can cause these connections to display correlated capacities to propagate influence. Beyond these natural phenomena are also the possibilities of adversarial intervention as well; for example, the controversial practice of "shadow banning" in some online platforms moderates the visibility of shared content in subtle, discreet ways.

The aim of this work is to develop robust methodologies for decision-making when the fit of any particular model θ, e.g., the independent cascade θ_{ic}, is uncertain. Indeed, it is often the case that the true dependency structure cannot be reliably learned. Motivated by this, we propose a new set of robust models for study in which we assume knowledge of the probabilities p_{ij} but make no assumptions on the dependence of the $\tilde{c}_{ij}$ random variables, admitting all joint distributions θ consistent with the marginal probabilities - written $\theta \in \Theta$.

Under this agenda, we first propose and study the *correlation robust* influence, in which the influence of a seed set is now taken to be the expected influence evaluated with respect to the most antagonistic selection of (consistent) joint distribution. Secondly, as correlation robust influence possesses many tractable properties, we next propose the *mixture of the independent cascade and correlation robust* influence, a new robust model of influence that provides a tunable level of conservatism without elevation of computational complexity beyond that of the independent cascade. Thirdly, we propose a robust risk measure with our study of the *correlation robust CVaR*, in which the CVaR for the random influence of a seed set is minimized over all joint distributions. For each of these models, we provide explicit characterizations of the adversarially-selected distribution, among other properties. Then we study their optimization problems, which all reduce to finding a k-sized seed set S in anticipation of some kind of adversarially- selected dependence structure. Though they are NP-Hard, we show they admit greedy approximation guarantees of $(1 - 1/e)$.

The techniques used in this work fall under the realm of distributionally robust optimization (DRO), a research area concerned with optimizing an expected objective function value under the most adverse distribution from a family of distributions. By introducing such an adversarial selection, decision-makers now seek decisions that would be "robust" to assuming an incorrect distribution. The family of distributions we consider in this work is commonly referred to as the Fréchet class (all joint distributions consistent with the given marginals) in the literature [Meilijson and Nádas, 1979], [Weiss, 1986], [Klein Haneveld, 1986], [Natarajan et al., 2009], [Rũschendorf, 2013], [Chen et al., 2022].

An accompanying aim of this study is to characterize and analyze the cost to a decision-maker that may be incurred when making the independence assumption despite there being uncertainty around the fit of the independent cascade. In other words, we study the extent of sub-optimality that may result when the diffusion process in fact departs from a zero correlation situation - as is assumed in the classical independent cascade model ([Kempe et al., 2003]). This line of questioning was similarly posed in [Agrawal et al., 2012], where the effect of departure from independence in stochastic programming was investigated. We show that the independent cascade model can lead to arbitrarily sub-optimal decisions. And this motivates consideration of moderating the assumption of independence, which we propose with the mixture model $f^{Mix(\lambda)}$.

Experiments on various datasets conclude the paper, in which we provide comparisons of computation and solution performance for all our models and the independent cascade. We also illustrate how we may select various parameters in practice through case studies on real-world datasets.

1.1 Contributions

The main contributions of this work and the structure of the paper are listed below:

1. We formulate a correlation robust model for influence maximization in Section 3 after reviewing related work in Section 2. This model has a distinct computational advantage over classical approaches because it is in fact efficient to compute the influence. More precisely, we show that by solving a polynomial time solvable linear program, we can compute the worst-case expected influence function for any seed set S. The linear program contains $O(|V|)$ variables and $O(|E| + |V|)$ constraints and provides interpretable information on the structure of the correlation robust distribution. In fact, we provide a complete analytical characterization of the *correlation robust distribution*, down to the recipe of how to generate random graphs under it and even its support of graphs. Furthermore it is possible to compute the worst-case expected influence function using shortest path algorithms which makes it practically feasible to implement even for very large networks and online platforms.

2. The influence maximization problem attached to this model is studied and compared to that of classical models. In Section 4, we show that while finding an optimal seed set S in the correlation robust model is NP-hard, the worst-case expected influence function is still submodular in S. By using the greedy algorithm, we obtain an improved approximation guarantee of $(1 - 1/e)$. This contrasts with the IC model where the approximation guarantee is $(1 - 1/e - \epsilon)$ for $\epsilon > 0$ with running time that has polynomial dependence on $1/\epsilon$. This improvement arises from the fact that we do not need to resort to simulation methods, and we use an exact optimization-based method in the evaluation of correlation robust influence. Finally, our model also admits a mixed integer linear program formulation, providing practitioners alternative methods (e.g., branch and bound) for optimization that is exact, if it is so desired and/or within computational means - something not available in classical approaches like the Independent Cascade.

3. In Section 5, we quantify using examples and theory -via the *price of correlations (POC)*- the extent to which operating under the assumption of independence can hurt influence maximization when the correlation is in fact unknown. We hence study the mixing of the independent cascade and correlation robust influence, which provides a tunable departure from the independent cascade model in the direction of the worst-case so as to attain a model more suited for real-world diffusion. This λ- weighted convex combination offers control of conservatism without elevation of computational complexity, and with the greedy performance guarantees of its optimization (for each λ) inherited. Further, operating under a carefully selected λ', we can generate seed sets with performance guarantees across a collection of $\lambda \neq \lambda'$ (Theorem 5).

4. In Section 6, we discuss how the approach can be extended to optimize the worst-case conditional value at risk (CVaR) objective in influence maximization. We show that the computation of the worst-case CVaR, like the expected influence, is also possible in polynomial time and is submodularity-preserving. Interestingly, this is in contrast to

the IC model's conditional value at risk, where it is known that submodularity is not preserved [Maehara, 2015], meaning the greedy performance guarantee is inherited in the worst-case but not in the independent cascade. A comparison of the key differences between the independent cascade and correlation robust model is provided in Table 1. Finally, perhaps surprisingly, we show that the *correlation robust distribution* is in fact an extremal distribution achieving the worst-case CVaR.

5. An experimental study of using the robust influence models on real world datasets is provided in Section 7. Insights into tuning and model performance are discussed.

	Independent Cascade Influence	Correlation Robust Influence
Objective	$f^{ic}(S) := \mathbb{E}_{\theta_{ic}}[R(\tilde{\mathbf{c}}, S)]$	$f^{corr}(S) := \min_{\theta\in\Theta} \mathbb{E}_{\theta}[R(\tilde{\mathbf{c}}, S)]$
Evaluation	#P-hard	P
Optimization	NP-hard	NP-hard
Submodularity	Yes	Yes
Approximation	Greedy $(1 - 1/e - \epsilon)$	Greedy $(1 - 1/e)$
	Independent Cascade CVaR	Correlation Robust CVaR
Objective	$CVaR_{\alpha}\Big(R(\tilde{\mathbf{c}}, S)\Big), \ \ \tilde{c} \sim \theta_{ic}$	$CVaR^{corr}_{\alpha}(S) := \min_{\theta\in\Theta} CVaR_{\alpha}\Big(R(\tilde{\mathbf{c}}, S)\Big)$
Evaluation	#P-hard	P
Optimization	NP-hard	NP-hard
Submodularity	No	Yes
Approximation	$1 - 1/e - \epsilon$ using randomization	Greedy $(1 - 1/e)$

Table 1: Comparison of independent cascade and correlation robust model.

2 Related Work and Preliminaries

There is an extensive literature on influence maximization models. Robustness in influence maximization first received attention through the parametric interval uncertainty model [He and Kempe, 2014] where for each edge $(i, j) \in E$, the probability p_{ij} is not known exactly, but rather known to lie in an interval $[l_{ij}, r_{ij}] \subseteq [0, 1]$. The objective therein is to obtain the best seed set under the IC model while accounting for all possible values the marginal probability vector $\mathbf{p}$ can take. Here $\mathbf{p} = (p_e)_{e\in E}$ is a vector of size $|E|$ where $p_e = p_{ij}$ denotes the activation probability of edge $e = (i, j) \in E$ with $0 \leq p_e \leq 1$. More general treatments in which $\mathbf{p}$ lies in a set $\mathcal{P}$ can be found in [Chen et al., 2016], [He and Kempe, 2016], [Kalimeris et al., 2019]. Specifically, these works focus on solving either

$$\max_{S:|S|\leq k} \min_{\mathbf{p}\in\mathcal{P}} f^{ic}_{\mathbf{p}}(S), \qquad \text{(Robust Influence Maximization)}$$

where $f^{ic}_{\mathbf{p}}(S)$ is the expected influence under the independent distribution with marginal probability vector $\mathbf{p}$ or alternatively

$$\max_{S:|S|\leq k} \min_{\mathbf{p}\in\mathcal{P}} \frac{f^{ic}_{\mathbf{p}}(S)}{f^{ic}_{\mathbf{p}}(S^*)}, \qquad \text{(Robust Ratio Influence Maximization)}$$

where $S^* = \arg\max_{S:|S|\leq k} f^{ic}_{\mathbf{p}}(S)$. In the second formulation, one seeks to find a set of nodes that minimizes the suboptimality gap across the collection of models in a simultaneous way [Chen et al., 2016], [He and Kempe, 2016]. However these problems have well-established hardness results - their forms are large, discrete, non-convex, and sometimes no longer involve monotone, nor submodular objectives. Consequently, such studies commonly resort to bi-criteria approximation guarantees [Krause et al., 2008]. Another study of robustness has been performed with the linear threshold model in [Nannicini et al., 2019] where the parameters are assumed to be uncertain.

While the majority of studies that focus on robustness in influence maximization have considered parameter uncertainty by way of the marginal probabilities (edge likelihoods), and hence still assume a fixed correlation structure - namely, independent edge propagation- we study the "reverse"' problem by assuming the edge likelihoods are fixed and the uncertainty lies in how they are correlated. Indeed, edge likelihoods are amenable to estimation individually while estimation of multivariate joint distributions is generally intractable. There has been prior interest in modeling the role that correlations play [Vaswani et al., 2017], [Aral and Dhillon, 2018]. In particular, [Vaswani et al., 2017] replace the influence function with a surrogate function that provides the most optimistic expected spread of influence. This, too, is a reversal of this work's focus - a consideration of the pessimistic expected spread of influence. Another related model is studied in [Staib et al., 2019] who develop a distributionally robust optimization approach to tackle influence maximization. The main difference of their work from the current paper is that they assume the availability of samples of the distribution (past scenarios where edges are live or not). Using these samples, an empirical distribution is constructed and a set of candidate distributions within a fixed χ^2 divergence of the empirical distribution is considered. Their objective is to then maximize the worst expected influence in this set and propose a suitably adapted Frank-Wolfe algorithm for this. The size of their formulation grows with the sample size. In contrast, the size of the optimization formulations developed in this paper do not depend on sample sizes.

We refer the reader to the Appendix for the proofs of all lemmas, corollaries, and theorems stated in this paper that warrant detail.

2.1 Reachable Nodes and Maximum Flow

While influence maximization is described in a stochastic setting, it will be useful to start by considering a deterministic instance using the maximum flow problem. Given the graph $G = (V, E)$, a set of seed nodes $S \subseteq V$, and a binary vector $\boldsymbol{c}$ with $|E|$ components, start by adding two distinct nodes s and t. Form the auxiliary graph $G'(\boldsymbol{c}) = (V', E'(\boldsymbol{c}))$ as follows: set $V' = V \cup \{s, t\}$ and $E'(\boldsymbol{c}) := \{(i, j) \in E : c_{ij} = 1\} \cup \{(s, i) : i \in S\} \cup \{(j, t) : j \in V \backslash S\}$. Then it is straightforward to see that the s-t maximum flow problem on $G'(\boldsymbol{c})$ provides the number of nodes that the seed set S influences (apart from itself), i.e., $R(\mathbf{c}, S) - |S|$ under realization $\mathbf{c}$. We state this clear fact formally (without proof):

Lemma 1 *Let $Z(\mathbf{c}, S)$ denote the optimal value of the following s-t max flow problem on the* *graph* $G'(\boldsymbol{c}) = (V', E'(\boldsymbol{c}))$*:*

$$Z(\mathbf{c}, S) = \max_{\boldsymbol{x} \in \mathbb{R}^{E'(\boldsymbol{c})}, v \in \mathbb{R}} \quad v$$

$$\text{s.t.} \quad \sum_{j:(i,j)\in E'(\boldsymbol{c})} x_{ij} - \sum_{j:(j,i)\in E'(\boldsymbol{c})} x_{ji} = \begin{cases} v, & i = s, \\ 0, & i \in V, \\ -v, & i = t, \end{cases} \tag{1}$$

$$x_{jt} \leq 1 \ \ \forall j \in V \backslash S$$

$$x_{ij} \geq 0 \ \ \forall (i,j) \in E'(\boldsymbol{c}).$$

The number of nodes reachable from S *along the live edges under realization* $\mathbf{c}$ *is* $R(\mathbf{c}, S) = |S| + Z(\mathbf{c}, S)$*.*

The max flow problem provides a natural way to describe $R(\mathbf{c}, S)$. Given any optimal flow $\boldsymbol{x}^* \in \mathbb{R}^{E'(\boldsymbol{c})}$, then $\{j \in V \backslash S : x^*_{jt} = 1\}$ is precisely the set of nodes that are reachable from s in the graph $G'(\boldsymbol{c})$. Equivalently, these are the nodes reachable from S in the graph G along those edges in E that are live, equivalently, $\{(i,j) \in E : c_{ij} = 1\}$. An example of the transformation and the max flow problem in Lemma 1 is illustrated in Figure 1.

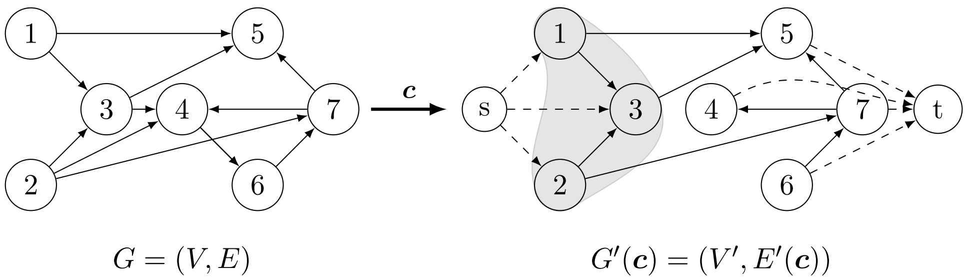


Figure 1: Construction of the auxiliary graph G'. The bold edges in $G'(\boldsymbol{c})$ denote the live edges $(i,j) \in E$ such that $c_{ij} = 1$. For the seed set $S = \{1, 2, 3\}$ under the realization $\boldsymbol{c}$, the maximum flow value is $Z(\mathbf{c}, S) = 3$ and $R(\mathbf{c}, S) = 6$. Node 6 is not reachable under realization $\boldsymbol{c}$ while it is reachable in the original graph.

3 The Correlation Robust Influence Function

Let $\mathcal{C}$ denote the set of all binary vectors of length $|E|$ (here $|\mathcal{C}| = 2^{|E|}$). Denote the set of all joint probability distributions over $\mathcal{C}$ consistent with the marginal probability vector $\mathbf{p}$ as:

$$\Theta := \left\{ \theta \in \mathbb{R}^{\mathcal{C}}_{+} : \sum_{\mathbf{c} \in \mathcal{C}: c_{ij} = 1} \theta(\mathbf{c}) = p_{ij} \ \ \forall (i,j) \in E, \sum_{\mathbf{c} \in \mathcal{C}} \theta(\mathbf{c}) = 1 \right\}. \tag{2}$$

Any θ in Θ provides a distribution over the set of $2^{|E|}$ graph realizations where $\theta(\mathbf{c})$ denotes the probability assigned to the realization $\mathbf{c}$. The set Θ is non-empty as $\theta_{ic} \in \Theta$ and is uncountable in general. Define for any seed set $S \subseteq V$:

$$f^{corr}(S) := \min_{\theta \in \Theta} \mathbb{E}_{\theta}[R(\tilde{\mathbf{c}}, S)]. \qquad \text{(Correlation Robust Influence Function)}$$

The correlation robust influence function is the smallest (worst-case) expected number of influenced nodes among all consistent joint distributions (and in turn over all possible correlation structures). According to Lemma 1, $f^{corr}(S) = \min_{\theta\in\Theta} \mathbb{E}_\theta[R(\tilde{\mathbf{c}}, S)] = |S| + \min_{\theta\in\Theta} \mathbb{E}_\theta[Z(\tilde{\mathbf{c}}, S)]$ amounts to evaluating a worst-case, expected max-flow. More precisely, the value $\min_{\theta\in\Theta} \mathbb{E}_\theta[Z(\tilde{\mathbf{c}}, S)]$ is exactly the optimal value of an instance of the distributionally robust max flow problem studied in [Chen et al., 2020]. Leveraging this connection, we obtain an efficient linear program to compute it, as stated in the following theorem. This implies that, for any seed set S, $f^{corr}(S)$ is computable in polynomial time with linear programming. This is in contrast to $f^{ic}(S) = \mathbb{E}_{\theta_{ic}}[R(\tilde{\mathbf{c}}, S)]$, which is #P-hard to compute. Distributionally robust bounds for the maximum flow problem have also been studied in other contexts such as in process flexibility in bipartite graphs [Wang and Zhang, 2015].

Theorem 1 *Given $G = (V, E)$, a seed set $S \subseteq V$ and marginal probability vector $\mathbf{p} \in [0, 1]^E$, the correlation robust influence function is $f^{corr}(S) = |S| + \min_{\theta\in\Theta} \mathbb{E}_\theta[Z(\tilde{\mathbf{c}}, S)]$ where the second term is given by the optimal value of the following polynomial sized linear program:*

$$\begin{aligned} \min_{\theta\in\Theta} \mathbb{E}_\theta[Z(\tilde{\mathbf{c}}, S)] \quad = \quad \min_{\boldsymbol{\pi}\in\mathbb{R}^V} \quad & \sum_{i\in V\setminus S} \pi_i \\ s.t \quad & \pi_i = 1 \ \ \forall i \in S, \\ & \pi_i - \pi_j \le 1 - p_{ij} \ \ \forall (i,j) \in E, \\ & 0 \le \pi_i \le 1 \quad \forall i \in V. \end{aligned} \tag{3}$$

As will be seen, an optimal solution π^* to the above linear program is essential for the characterizations and constructions to come.

3.1 Correlation-Robust Distributions: Properties and Characterization

Given a seed set S and a collection of marginals $\{p_{ij}\}_{(i,j)\in E}$, we say a distribution $\theta \in \Theta$ is *correlation-robust* (against seed set S) if $\theta \in \arg\min_{\theta\in\Theta} \mathbf{E}_{\tilde{\mathbf{c}}\sim\theta}\left[Z(\tilde{\mathbf{c}}, S)\right]$ or $\arg\min_{\theta\in\Theta} \mathbb{E}_\theta[R(\tilde{\mathbf{c}}, S)]$. The next results will describe properties to the family of correlation-robust distributions, and then we will provide an explicit characterization to obtain a specific correlation-robust distribution.

We define the *correlation robust likelihood* as the probability that a node i is influenced by S, under a(ny) correlation-robust distribution. This is well-defined because, as we will see shortly, all correlation-robust distributions will produce the same collection of likelihoods. We will arrive at this through the use of the linear program (3). Consider any distribution $\theta \in \Theta$ with $\tilde{\mathbf{c}} \sim \theta$. Let $G(\tilde{\mathbf{c}}) = (V, E(\tilde{\mathbf{c}}))$ be the random graph induced by $\tilde{\mathbf{c}}$ in which $E(\tilde{\mathbf{c}}) := \{(i,j) \in E : \tilde{c}_{ij} = 1\}$. In the graph $G(\tilde{\mathbf{c}})$, a node i is influenced by S if and only if there exists a directed path $\gamma = (i_0 \to i_1 \to i_2 \to \ldots \to i_\lambda = i)$ with positive length λ_γ from some node $i_0 \in S$ to i. This implies that any such directed path γ presents the way to a lower bound on the likelihood of node i's influence. More precisely, for every $\theta \in \Theta$,

$$\begin{aligned} \mathbb{P}_\theta(\text{Node } i \text{ is reachable from } S \text{ in } G(\tilde{\mathbf{c}})) &\geq \mathbb{P}_\theta(G(\tilde{\mathbf{c}}) \text{ contains path } \gamma) \\ &= 1 - \mathbb{P}_\theta\left(\cup_{l=0}^{\lambda_\gamma - 1} \{(i_l, i_{l+1}) \notin E(\tilde{\mathbf{c}})\}\right) \\ &\geq 1 - \sum_{l=0}^{\lambda_\gamma - 1} (1 - p_{i_l, i_{l+1}}), \end{aligned} \tag{4}$$

which follows from the union bound. This inequality can be understood as expressing the intuition that in order to minimize the likelihood of γ's existence, the adversary needs only to make the disruptions of the individual arcs "as disjoint as possible" (if not possible, then the lower bound is necessarily 0). It is natural then to consider the greatest of these lower bounds; hence, we denote the collection of all such directed paths from S to i in the original graph G as $\Gamma(S, i)$, and define for any $\gamma \in \Gamma(S, i)$:

$$L(\gamma) := 1 - \sum_{l=0}^{\lambda_\gamma - 1} (1 - p_{i_l, i_{l+1}}). \tag{5}$$

Then we obtain the following lower bound on the probability that node i is influenced by the set of nodes S:

$$\begin{aligned} \mathbb{P}_\theta(\text{Node } i \text{ is reachable from } S \text{ in } G(\tilde{\mathbf{c}})) &= \mathbb{P}_\theta\left(\cup_{\gamma \in \Gamma(S,i)} \{G(\tilde{\mathbf{c}}) \text{ contains path } \gamma\}\right) \\ &\geq \left[\max_{\gamma \in \Gamma(S,i)} L(\gamma)\right]^+, \end{aligned} \tag{6}$$

so that

$$\min_{\theta \in \Theta} \mathbb{P}_\theta(\text{Node } i \text{ is reachable from } S \text{ in } G(\tilde{\mathbf{c}})) \geq \left[\max_{\gamma \in \Gamma(S,i)} L(\gamma)\right]^+, \tag{6'}$$

where $x^+ = \max(0, x)$. This gives the following lower bound on the expected number of nodes influenced outside the seed set S:

$$\begin{aligned} \min_{\theta \in \Theta} \mathbb{E}_\theta[Z(\tilde{\mathbf{c}}, S)] &= \min_{\theta \in \Theta} \sum_{i \in V \backslash S} \mathbb{P}_\theta(\text{Node } i \text{ is reachable from } S \text{ in } G(\tilde{\mathbf{c}})) \\ &\geq \sum_{i \in V \backslash S} \min_{\theta \in \Theta} \mathbb{P}_\theta(\text{Node } i \text{ is reachable from } S \text{ in } G(\tilde{\mathbf{c}})) \\ &\overset{(6')}{\geq} \sum_{i \in V \backslash S} \left[\max_{\gamma \in \Gamma(S,i)} L(\gamma)\right]^+. \end{aligned} \tag{7}$$

The following corollary shows that all the inequalities in (7) and (6) turn out to be tight when θ is correlation-robust.

Corollary 1 (Correlation Robust Influence Likelihood) *For an arbitrary seed set S and marginal probability vector $\mathbf{p} \in [0,1]^E$, there is a unique solution π^* to the linear program (3), and for each $i \in V \backslash S$, the optimal decision variable π_i^* satisfies :*

$$\begin{aligned} \pi_i^* &= \left[\max_{\gamma \in \Gamma(S,i)} L(\gamma)\right]^+, \\ &= \mathbb{P}_{\theta^*}\left(\textit{Node } i \textit{ is reachable from } S \textit{ in } G(\tilde{\mathbf{c}})\right), \end{aligned} \tag{8}$$

where θ^* *is any correlation-robust distribution, equivalently, optimal solution to* $\min_{\theta\in\Theta} \mathbb{E}_\theta \left[Z(\tilde{\mathbf{c}}, S)\right]$. *In particular, it is upper bounded by the probability under the independent cascade model:*

$$\pi_i^* \leq \mathbb{P}_{\theta_{ic}}(\text{Node } i \text{ is reachable from } S).$$

The next result establishes a kind of "dual" object to the influence likelihoods π_i^*. It will highlight how correlation robust influence reduces to shortest path computations (used in Section 7 experiments). As well, the supermodularity property that is identified will factor crucially in the forthcoming Theorem 3.

Corollary 2 (Correlation Robust Non-influence Likelihood) *For an arbitrary seed set* S *and marginal probability vector* $\mathbf{p} \in [0,1]^E$, *let*

$$\phi_i^S := \begin{cases} \min_{\gamma\in\Gamma(S,i)} \sum_{l=0}^{\lambda_\gamma - 1} 1 - p_{i_l, i_{l+1}} & i \in V \backslash S, \\ 0 & i \in S. \end{cases} \tag{9}$$

If $\boldsymbol{\pi}^*$ *solves the linear program (3), then* $\pi_i^* = \left[1 - \phi_i^S\right]^+$ *for all* $i \in V$. *Thus,*

$$\min(\phi_i^S, 1) = \mathbb{P}_{\theta^*}(\text{Node } i \text{ is not reachable from } S \text{ in } G(\tilde{\mathbf{c}})),$$

where θ^* *is any correlation-robust distribution, equivalently, optimal solution to* $\min_{\theta\in\Theta} \mathbb{E}_\theta \left[Z(\tilde{\mathbf{c}}, S)\right]$. *Finally, the function* ϕ_i^S *is a non-increasing, supermodular set function on sets* $S \subseteq V$, *for any* $i \in V$.

At this point we may observe some structural comparisons between the correlation-robust and independent cascade distribution θ_{ic}. Let $\eta_i^S := \mathbb{P}_{\theta_{ic}}\Big(\text{Node is NOT reachable from S in } G(\tilde{\mathbf{c}})\Big)$. For fixed $i \in V$, η_i^S is - like ϕ_i^S- supermodular. To see this, let $S \subseteq T \subseteq V$ and $v \in V \setminus T$. Then

$$\begin{aligned} \eta_i^S - \eta_i^{S\cup\{v\}} &= \mathbb{P}_{\theta_{ic}}\Big(\text{node i is not reachable from S, but is reachable from node v}\Big) \\ &\geq \mathbb{P}_{\theta_{ic}}\Big(\text{node i is not reachable from T, but is reachable from node v}\Big) \\ &= \eta_i^T - \eta_i^{T\cup\{v\}}. \end{aligned}$$

Further, observe that in the case of ϕ_i^S only those paths that produce the minimum will determine the likelihood, whereas in the case of η_i^S all paths are factored into the likelihood. This feature is explored more in the next result, which concerns the structure of the random graph $G(\tilde{\mathbf{c}})$ under any correlation-robust distribution. In contrast to the IC model, not all paths in $\Gamma(S,i)$ contribute to the likelihood that a node i is influenced; in fact, as the corollary shows, only a subset of the paths ever manifest, and when they do, they always appear together with positive probability.

Corollary 3 (Path existence under correlation robustness) *Let* S *be an arbitrary seed set,* $\boldsymbol{\pi}^*$ *be the optimal solution to (3) and let* θ^* *be any optimal solution to* $\min_{\theta\in\Theta} \mathbb{E}_\theta \left[Z(\tilde{\mathbf{c}}, S)\right]$. *For every node* $i \notin S$, *let* $\bar{\Gamma}(S,i) := \arg\max_{\gamma\in\Gamma(S,i)} L(\gamma)$ *denote the set of paths which attain the maximum value. If* $i \notin S$ *and* $\pi_i^* = \max_{\gamma\in\Gamma(S,i)} L(\gamma) > 0$, *then:*

$$\pi_i^* = \mathbb{P}_{\theta^*}(\cup_{\gamma\in\bar{\Gamma}(S,i)} \{G(\tilde{\mathbf{c}}) \text{ contains path } \gamma\}) = \mathbb{P}_{\theta^*}(\cap_{\gamma\in\bar{\Gamma}(S,i)} \{G(\tilde{\mathbf{c}}) \text{ contains path } \gamma\}),$$

Furthermore, for any path $\gamma \in \bar{\Gamma}(S,i)$, at most one of the arcs in the path is missing in the random graph $G(\tilde{\mathbf{c}})$, θ^* - almost surely.

That either all the paths in $\bar{\Gamma}(S,i)$ are live or dead (under θ^*) highlights the fact that it only takes one path between S and i for i to be influenced. Consequently, we see that redundancy is being exploited by any $\theta^* \in \arg\min_{\theta\in\Theta} \mathbb{E}_\theta\left[R(\tilde{\mathbf{c}}, S)\right]$.

In the following corollary, we provide a complete characterization of one particular member θ^* of the class of correlation robust distributions, i.e., $\theta^* \in \arg\min_{\theta\in\Theta} \mathbb{E}_\theta[R(\tilde{\mathbf{c}}, S)]$. From this point forward, the reader may understand all mentions of *the* correlation robust distribution as references to this one described here.

Corollary 4 (Correlation Robust Distribution) *Given an arbitrary seed set S and marginal probability vector $\mathbf{p} \in [0,1]^E$, let $\boldsymbol{\pi}^*$ be the optimal solution to (3). Define the following random sets for $\tilde{q} \sim$ Uniform$[0,1]$:*

$$V(\tilde{q}) := \{i \in V : \tilde{q} < \pi_i^*\}$$

and $E(\tilde{q}) := \{(k,j) \in E : \pi_k^* > \pi_j^*, \tilde{q} \notin [\pi_k^* - 1 + p_{kj}, \pi_k^*]\} \cup \{(k,j) \in E : \pi_k^* \le \pi_j^*, \tilde{q} \in (0, p_{kj}]\}$.

As well, let $\mathbf{c}(\tilde{q}) \in \{0,1\}^E$ be a random vector in which $c_{ij}(\tilde{q}) = 1$ if and only if $(i,j) \in E(\tilde{q})$. Then

1. *The joint distribution of $c(\tilde{q})$, call it θ^*, is worst-case, i.e., $\theta^* \in \arg\min_{\theta\in\Theta} \mathbb{E}_\theta\left[Z(\tilde{\mathbf{c}}, S)\right]$.*
2. *$V(\tilde{q})$, the set of all nodes reachable from S (including all members of S) in the graph $G(\tilde{q}) = (V, E(\tilde{q}))$ satisfies $V(\tilde{q}) = S \cup \{i \in V \setminus S : 1 - \tilde{q} > \phi_i^S\}$.*
3. *$R(\tilde{\mathbf{c}}(\tilde{q}), S) = |V(\tilde{q})| = |S| + l$ if and only if $1 - \tilde{q} \in (\phi_{i(l)}^S, \phi_{i(l+1)}^S]$, where $i(\cdot)$ enumerates $V \setminus S$ such that $0 \le \phi_{i(1)}^S \le \phi_{i(2)} \le \dots \le \phi_{i(|V\setminus S|)}^S$.*

   In other words, when $\tilde{c} \sim \theta^*$, the cumulative distribution function is given by

$$F_{R(\tilde{\mathbf{c}},S)}(\tau) = \lim_{z\downarrow\tau} \min(\phi_{i(\lceil z-|S|\rceil)}, 1) = \begin{cases} \min(\phi_{i(\lceil \tau-|S|\rceil)}, 1) & |S| < \tau \notin \mathbb{Z} \\ \min(\phi_{i(\tau-|S|+1)}, 1) & |S| \le \tau \in \mathbb{Z} \\ 0 & o.t.w. \end{cases}$$

The characterization presented in Corollary 4 can in fact be used to construct the support of the correlation robust distribution. To do so, we note that the breakpoints used in the definition of $E(\tilde{q})$ partition the [0,1] interval into sub-intervals for which $E(\tilde{q})$ is constant; in other words, any given sub-interval corresponds to a graph in the support, and its length yields the probability that θ^* assigns to it. We refer the reader to Section E.3's Figure 15 for an example application of this result. As well, it is worth mentioning that it easily follows that the support of the correlation robust distribution is linear in $|E|$.

Corollary 4 also provides a form that stands in interesting contrast to the linear threshold model (LTM) proposed in [Kempe et al., 2003]. In particular, while Corollary 4 uses a *single* random number $\tilde{q} \in [0,1]$ to characterize both the random set of influenced nodes $V(\tilde{q})$ and the random live edges $E(\tilde{q})$ that allow S to reach them, in the LTM, independent draws from Uniform$[0,1]$ for *every* node are performed. Further, in LTM at most one live edge enters any node i, while under correlation robustness either all the paths in $\bar{\Gamma}(S,i)$ are live simultaneously or i is not reached at all (from Corollary 3).

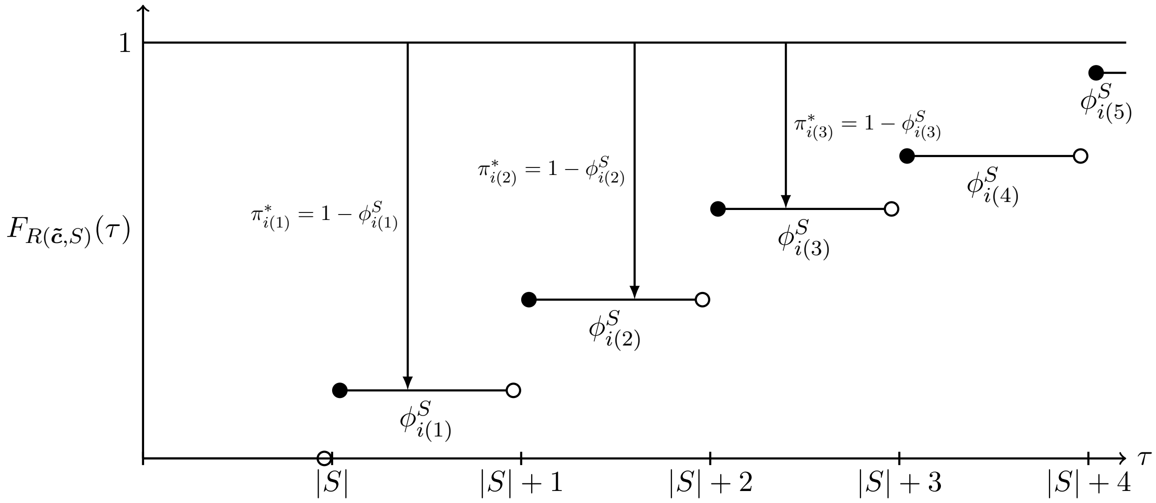


Figure 2: Schematic illustration of the CDF $F_{R(\tilde{\mathbf{c}},S)}(\tau)$ of $R(\tilde{\mathbf{c}}, S)$ where $\tilde{\mathbf{c}}$ is distributed according to the correlation-robust distribution of Corollary 4.

4 Optimization

4.1 The Correlation Robust Influence Maximization Problem

We formulate the corresponding optimization problem as:

$$\max_{S:|S|\leq k} f^{corr}(S). \qquad \text{(Correlation Robust Influence Maximization)}$$

Correlation robust influence maximization models the problem of finding the set of nodes that are influential regardless of the correlation structure, by maximizing the worst-case expected influence.

As we will see, in terms of complexity and approximation, this problem shares similarities with the problem of maximizing the independent cascade's influence $f^{ic}(S)$. However, to be sure, the different influence dynamics result in different seed selections.

Example 1 *Consider finding the optimal $k = 1$ seed set for the graph in Figure 3 which is composed of two connected components: one that is a star graph of 10 nodes centered at node a; one that is a collection of 4 line graphs with common origin at node b. In words, the decision is between node a's high local but low global reach versus node b's high global but low local reach. The marginal probabilities are homogeneous at $p = 1 - 1/z$, for $z \in \mathbb{Z}_{>0}$.*

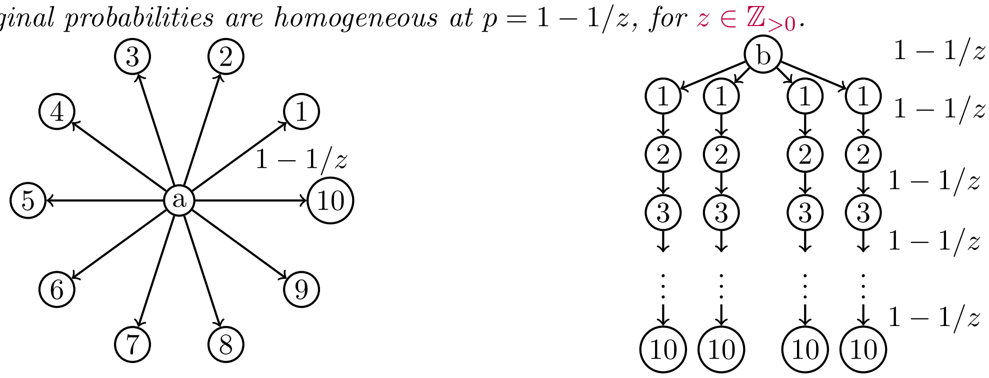


Figure 3: A graph with two components: a star graph with origin a, and a collection of 4 line graphs with common origin b.

In the case of independent cascade, $f^{ic}(\{a\}) = 1 + 10 \cdot (1 - 1/z)$, *while* $f^{ic}(\{b\}) = 1 + 4\left[\sum_{i=1}^{10-1} i\left[(1-1/z)^i \cdot (1/z)\right] + 10 \cdot (1-1/z)^{10}\right]$.

It thus follows that $\arg\max_{S:|S|=1} f^{ic}(S) = \{a\}\mathbb{1}_{z\leq 2} + \{b\}\mathbb{1}_{z\geq 3}$. *On the other hand, in the case of robust influence,* $f^{corr}(\{a\}) = 1+10\cdot(1-1/z)$, *while* $f^{corr}(\{b\}) = 1+4\cdot\left[min(z,10) - 1/z \cdot \sum_{l=1}^{min(z,10)} l\right]$. *It thus follows that* $\arg\max_{S:|S|=1} f^{corr}(S) = \{a\}\mathbb{1}_{z<5} + \{\{a\},\{b\}\}\mathbb{1}_{z=5} + \{b\}\mathbb{1}_{z>5}$. *As we can see, the decision to seed node* b *requires a higher marginal probability* p *in the case of robust influence than in the case of independent cascade.*

Though Example 1 illustrates a qualitative difference, the optimization for the robust case is NP-hard just as with independent cascade. However for networks that are small-sized, the optimal seed set can be obtained by solving a mixed integer linear program, as stated by the theorem below.

Theorem 2 *Given the graph* $G = (V, E)$ *with marginal probability vector* $\mathbf{p} \in [0,1]^E$ *and integer* k, *the problem* $\max_{S:|S|\leq k} f^{corr}(S)$ *is NP-Hard. In particular, we can solve the exact formulation using the following mixed-integer linear program (MILP):*

$$\begin{aligned}
\max_{S:|S|\leq k} f^{corr}(S) \quad = \quad \max_{\mathbf{x},\mathbf{y},\mathbf{w},\mathbf{q}\in\mathbb{R}^V,\mathbf{z}\in\mathbb{R}^E} \quad & \sum_{(i,j)\in E} z_{ij}(p_{ij}-1) + \sum_{i\in V} w_i - q_i \\
s.t. \quad & 1 - y_i - \sum_{j:(j,i)\in E} z_{ji} + \sum_{j:(i,j)\in E} z_{ij} + q_i \geq 0 \;\forall i \in V \\
& |V|x_i + y_i - |V| \leq w_i \;\; \forall i \in V \\
& w_i \leq |V|x_i \;\; \forall i \in V \\
& w_i \leq y_i \;\; \forall i \in V \\
& \sum_{i\in V} x_i \leq k, x_i \in \{0,1\} \;\forall i \in V \\
& y_i \geq 0, q_i \geq 0 \;\forall i \in V, z_{ij} \geq 0 \;\forall (i,j) \in E
\end{aligned}$$

Let $\mathbf{x}^*, \mathbf{y}^*, \mathbf{w}^*, \mathbf{q}^*, \mathbf{z}^*$ *be the optimal solution to the MILP. The optimal seed set is* $S_{corr} = \{i : x_i^* = 1\}$.

Next we show that the function $f^{corr}(\cdot)$ is a submodular function. In general, if functions $g(\cdot)$ and $h(\cdot)$ are submodular set functions, the pointwise minimum $\min(g(\cdot), h(\cdot))$ need not be submodular. But interestingly, we show that $f^{corr}(S) = \min_{\theta\in\Theta} \mathbb{E}_\theta[R(\tilde{\mathbf{c}}, S)]$ which is a pointwise minimum of submodular functions $\mathbb{E}_\theta[R(\tilde{\mathbf{c}}, S)]$ over $\theta \in \Theta$ is in fact submodular for the specific choice of Θ considered in this paper.

Theorem 3 *The correlation robust influence function* $f^{corr}(\cdot)$ *is a nondecreasing submodular function.*

Although f^{ic} can be shown to be nondecreasing, submodular with basic graph arguments as in [Kempe et al., 2003], we may alternatively obtain the conclusion by finding $f^{ic}(S) = \sum_{i\in V}[1 - \eta_i^S]$, further illustrating parallels between the correlation robust and independent cascade settings. But as we will see in Section 6, the similarities end when generalizing to the conditional value at risk.

A major consequence of the above two theorems is that though the correlation robust maximization problem is NP-hard, a greedy algorithm (see Algorithm 1) provides a useful approximation guarantee (as in the IC setting).

Algorithm 1 Greedy Algorithm for Maximization of a set function $f : V \to \mathbb{R}$

$S^{(0)} := \emptyset$
for $i = 1 : k$ **do**
  $S^{(i)} = S^{(i-1)} \cup \arg\max_{v \in V \setminus S^{(i-1)}} f(S^{(i-1)} \cup \{v\})$
end for
Return $S^{(k)}$

Corollary 5 *Let S^g_{corr} denote the seed set generated upon termination of the greedy algorithm for maximization of $f^{corr}(\cdot)$. Then, $f^{corr}(S^g_{corr}) \geq (1 - 1/e) \max_{|S| \leq k} f^{corr}(S)$.*

Note that a key difference from the IC model is that there is no ϵ loss in the approximation guarantee of the greedy algorithm. In the IC model this occurs due to sampling errors in the estimation of $f^{ic}(S)$ [Kempe et al., 2003]. Under correlation robustness, the sampling error does not appear, as $f^{corr}(\cdot)$ is polynomial time computable with linear programming.

5 Price of Correlations and Mixing Independence with Dependence

In this section, we examine the extent to which the assumption of independence can cost the decision maker when there is dependence among the influence relationships. Indeed, in the previous section's example, we saw how the decision of which nodes to seed may differ according to the metric f^{ic} versus f^{corr}. And now we see to what degree of suboptimality those differences can incur. Suppose the decision maker computes the seed set $S_{ic} \in \arg\max_{S:|S| \leq k} f^{ic}(S)$, as opposed to $S_{corr} \in \arg\max_{S:|S| \leq k} f^{corr}(S)$. Then the price of correlations (POC) ratio (see [Agrawal et al., 2012]) characterizes the suboptimality of S_{ic} in the optimization of the $f^{corr}(\cdot)$:

$$\text{POC} = \frac{f^{corr}(S_{ic})}{f^{corr}(S_{corr})}. \qquad \text{(Price of Correlations)}$$

Intuitively, POC describes the cost of using a seed set optimal to an 'incorrect' diffusion model. A related concept is the correlation gap which for any seed set S is denoted by $\kappa(S)$ (see [Agrawal et al., 2012]). In the influence maximization setting:

$$\kappa(S) = \frac{f^{corr}(S)}{f^{ic}(S)}.$$

Since $\max_{S:|S| \leq k} f^{corr}(S) = f^{corr}(S_{corr}) \leq f^{ic}(S_{corr}) \leq f^{ic}(S_{ic})$, we have,

$$1 \geq \text{POC} = \frac{f^{corr}(S_{ic})}{f^{corr}(S_{corr})} \geq \frac{f^{corr}(S_{ic})}{f^{ic}(S_{ic})} = \kappa(S_{ic}) \geq 0. \tag{10}$$

A POC value closer to 1 indicates that we do not suffer much by resorting to S_{ic} when the underlying diffusion process corresponds to an adversarial model. But a POC value close to zero indicates major loss. We demonstrate that both of these scenarios are possible for appropriate graphs.

Example 2 *We first consider the series graph in Figure 4 with $n \geq 2$. Let $p_{ij} = 1 - 1/n$ for all the edges and let $k = 1$. It is easy to verify by inspection that $S_{corr} = S_{ic} = \{1\}$. Then, the correlation gap $\kappa(S_{ic})$ can be computed as:*

$$\kappa(S_{ic}) = \frac{f^{corr}(S_{ic})}{f^{ic}(S_{ic})} = \frac{1 + \sum_{i=2}^{n} 1 - \frac{i-1}{n}}{1 + \sum_{i=1}^{n-1}(1 - 1/n)^i} = \frac{1 + (n-1)/2}{n\,[1 - (1 - 1/n)^n]}$$

Therefore $\lim_{n\to\infty} \kappa(S_{ic}) = 1/2 \cdot e/(e-1) \approx 0.791$.

Example 3 ([Ma, 2020]) *We next consider the tree in Figure 5 where the root node contains l children. There are a total of l paths from the root to all the leaf nodes. Each path contains $m+2$ nodes (apart from the root). The labels on the nodes depict the "type" of each node. Edges between nodes of type 0 and 1 as well as between type 1 and type 2 nodes have 0.5 probability of being live. For all other edges, the probability is 1. The total number of nodes in the graph is $n = l(m+2)+1$, where we let $\frac{4m}{m+3} \leq l \leq 2m$. Then if $k = 1$, S_{corr} is any one of the type 2 nodes while $S_{ic} = \{0\}$ (details in Appendix C.1). Thus $POC = ((l/2)+1)/(m+1)$. If $l = 4m/(m+3)$, $POC = (2m+3)/((m+1)(m+3))$ which tends to zero as $m \to \infty$.*

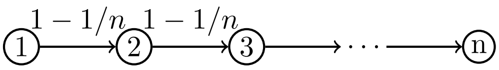


Figure 4: Example 1

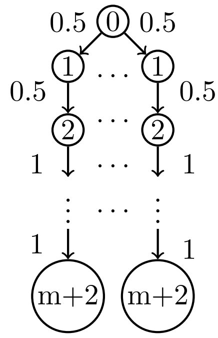


Figure 5: Example 2 with $l = 2$

This example leads to the following theorem.

Theorem 4 *There exists a graph on n nodes with price of correlations $O(1/n)$.*

The theorem reveals that in general, the POC may be arbitrarily close to zero. In other words, S_{ic} can be arbitrarily sub-optimal, if used under an adversarial diffusion process.

5.1 Mixture of Independent and Correlation Robust Models

It may be the case that a decision maker finds f^{corr} to be too pessimistic but at the same time wishes to not operate under the assumption of independence. In this section, we propose a mixture of the independent cascade and correlation robust models. For a decision-maker less than confident in making an independence assumption, they can turn to this mixture model to scalably depart from the independent cascade model, or conversely, temper the pessimism of the worst-case by effectively constraining the departure from the independent cascade.

Consider the problem:

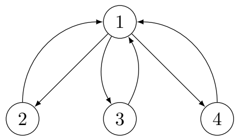
$$\max_{S:|S|\leq k} \left(f^{Mix(\lambda)} := \lambda E_{\theta_{ic}} \left[R(\tilde{\mathbf{c}}, S) \right] + (1-\lambda) \min_{\theta \in \Theta} \mathbb{E}_{\theta}[R(\tilde{\mathbf{c}}, S)] \right), \tag{Mixture Influence Maximization}$$

where $\lambda \in [0,1]$ is a parameter in which $\lambda = 1$ recovers the independent cascade model and $\lambda = 0$ recovers the correlation robust model.

For any value of $\lambda \in [0,1]$, the objective function for a given seed set S is exactly $\lambda f^{ic}(S) + (1-\lambda) f^{corr}(S)$. This is exactly the expected influence for the mixture of two distributions where the mixture weights are $(\lambda, 1-\lambda)$ and the distributions are $(\theta_{ic}, \arg\min_{\theta\in\Theta} \mathbb{E}_{\theta}[R(\tilde{\mathbf{c}}, S)])$. It easily follows (see the proof of Theorem 2) that the optimization $\max_{S:|S|\leq k} f^{Mix(\lambda)}$ is NP-Hard for any λ, and the optimal value is nondecreasing in λ; however, since the objective function is a nonnegative sum of submodular functions, the greedy algorithm and its guarantee apply just as in the correlation-robust or the independent cascade settings alone.

We conclude this subsection with an illustrative example. For a larger computational example, we refer the reader to Appendix C.2.

Example 4 ([Maehara, 2015]) *Consider the network in Figure 6 with all* $p_{ij} = 1/2$ *and* $k = 2$*. Then* S_{ic} *is any one of* $\{2,3\}, \{2,4\}$ *or* $\{3,4\}$*, and* S_{corr} *is any one of* $\{1,2\}, \{1,3\}$ *or* $\{1,4\}$*.* S_λ *on the other hand would be any of the* S_{ic} *sets when* $\lambda \in (0.8, 1]$*, otherwise it would be any of the* S_{corr} *sets when* $\lambda \in [0, 0.8)$*. In other words, the optimal solution structure changes once at least 20% of the worst-case model is injected into the mixture.*



S	$f^{ic}(S)$	$f^{corr}(S)$	$f^{Mix(.85)}$	$f^{Mix(.5)}$
{1,2}	3	3	3	3
{1,3}	3	3	3	3
{1,4}	3	3	3	3
{2,3}	3.125	2.5	3.031	2.813
{2,4}	3.125	2.5	3.031	2.813
{3,4}	3.125	2.5	3.031	2.813

Figure 6: Example 4

5.1.1 Tuning the Mixture Parameter λ

Let S^g_λ denote a seed set generated upon termination of the greedy algorithm for maximization of $f^{Mix(\lambda)}$, for $\lambda \in [0,1]$. It follows from Corollary 5 and [Kempe et al., 2003] that if there exists $\lambda', C \in [0,1]$ and subset $\Lambda \subseteq [0,1]$ such that

$$f^{Mix(\lambda)}(S^g_{\lambda'}) \geq C \cdot f^{Mix(\lambda)}(S^g_\lambda) \quad \forall \lambda \in \Lambda,$$

then we may say that the mixture λ' affords a performance guarantee that is uniform over Λ in the following sense:

$$f^{Mix(\lambda)}(S^g_{\lambda'}) \geq C \cdot (1 - 1/e) \max_{|S|\leq k} f^{Mix(\lambda)}(S) \;\; \forall \lambda \in \Lambda.$$

The following theorem indicates that it is possible to efficiently identify a greedy set with a performance guarantee that is uniform over a collection of mixture levels.

Theorem 5 *Let $\Lambda \subseteq [0,1]$ be a finite set. If λ' solves the maximization problem*

$$C := \max_{\lambda' \in \Lambda} \min_{\bar{\lambda} \in \Lambda} \left(\frac{f^{Mix(\bar{\lambda})}(S^g_{\lambda'})}{f^{Mix(\bar{\lambda})}(S^g_{\bar{\lambda}})} \right), \tag{11}$$

then $\qquad f^{Mix(\lambda)}(S^g_{\lambda'}) \geq C \cdot (1 - 1/e) \max_{|S| \leq k} f^{Mix(\lambda)}(S) \qquad \forall \lambda \in \Lambda.$

Towards operationalizing this insight, we consider the following procedure: (1) we outline a finite set Λ consisting of possible mixtures we are concerned about or wish to address; (2) we run the greedy algorithm for each member of Λ to obtain a collection of seed sets $\{S^g_\lambda\}_{\lambda \in \Lambda}$; (3) we solve the maximization in (11) to arrive at a λ'. We refer the reader to Section 7.2 for an example application of this procedure for the selection of λ'.

6 Correlation Robust CVaR

In this section we will examine the computation and optimization of the influence function's conditional value-at-risk (CVaR), but in a worst case form. This new objective will subsume f^{corr}. But before we formally define this worst case form of CVaR, we briefly recall for the reader the original concept of CVaR and the VaR metric it is derived from, which will require establishing some notational decisions since the random variable of interest $R(\tilde{\boldsymbol{c}}, S)$ will always be a discrete random variable, in which case multiple variants to the CVaR and VaR concepts arise. For example, we will proceed with a definition of VaR_α that is alternatively referred to as the "upper VaR_α" elsewhere in the literature, e.g., [Rockafellar and Uryasev, 2002].

CVaR is a well-studied risk measure amenable to optimization arising in the finance and insurance domains [Rockafellar and Uryasev, 2000], [Acerbi and Tasche, 2002]. For any real-valued, random variable $\tilde{x}$ with CDF $F(\cdot)$, let the inverse CDF be given by $F^{-1}(q) = \inf\{t : F(t) > q\}$. Then, given a level $\alpha \in (0,1]$, the (left-tail) conditional value at risk and value at risk at α denoted, respectively, $CVaR_\alpha(\tilde{x})$ and $VaR_\alpha(\tilde{x})$, are defined as:

$$VaR_\alpha(\tilde{x}) \;\; := \;\; F^{-1}(\alpha), \quad CVaR_\alpha(\tilde{x}) \;\; := \;\; \frac{1}{\alpha} \int_0^\alpha F^{-1}(q) dq, \tag{12}$$

in which $CVaR_\alpha(\tilde{x}) \leq VaR_\alpha(\tilde{x})$ follows, and in the special case of $\alpha = 1$, $VaR_\alpha(\tilde{x}) = +\infty$ and $CVaR_\alpha(\tilde{x}) = \mathbb{E}\tilde{x}$.

We remark that in the traditional definition for the CVaR and VaR risk measure, the random variable of interest is a loss function (smaller the better) and so the risk measure is traditionally defined on the right tail which is then minimized. In contrast, the random variable of interest in this work is the number of influenced nodes $R(\tilde{\boldsymbol{c}}, S)$ (larger the better) and so the risk measure is defined on the left tail which is then maximized. In words, $VaR_\alpha(\tilde{x})$ is the smallest right endpoint t of any half-interval $(-\infty, t]$ for which the probability that $\tilde{x} \in (-\infty, t]$ exceeds α, while $CVaR_\alpha(\tilde{x})$ is, loosely speaking, the average of the random variable $\tilde{x}$ conditional on it not exceeding $VaR_\alpha(\tilde{x})$. Intuitively, for any $\tau > VaR_\alpha(\tilde{x})$, the forward derivative of $\mathbb{E}\left[\tau - X\right]^+$ is greater than α, while for any $\tau < VaR_\alpha(\tilde{x})$ the forward derivative is less than α. The following lemma exploits this, providing a known, alternative characterization of $CVaR_\alpha(\tilde{x})$ that is useful in the optimization context [Rockafellar and Uryasev, 2000].

Lemma 2 *For $\alpha \in (0,1)$ and real-valued random variable $\tilde{x}$, $CVaR_\alpha(\tilde{x})$ can be computed as:*

$$CVaR_\alpha(\tilde{x}) = \max_{\tau \geq 0} \left(\tau - \frac{1}{\alpha} \mathbb{E}\left[\tau - \tilde{x}\right]^+ \right),$$

The optimal solution $\tau^* = VaR_\alpha(\tilde{x})$.

Here our interest is on $\tilde{x} = R(\tilde{\mathbf{c}}, S)$ which is a random, integer number of reachable nodes from (and including) seed set S when $\tilde{\mathbf{c}}$ denotes the underlying random vector. Let Θ denote the set of all joint distribution over $\tilde{\mathbf{c}}$ consistent with the known marginals. Then $CVaR_\alpha(R(\tilde{\mathbf{c}}, S))$ captures the average influence of S, conditional on the worst α fraction of outcomes. When $\alpha = 1$, this simply reduces to the expected influence function $E[R(\tilde{\mathbf{c}}, S)]$. We will look at finding the best seed set S that maximizes this objective function in the presence of an adversarial intervention. Prior studies have also considered using the CVaR objective in influence maximization under the IC model (see [Maehara, 2015], [Ohsaka and Yoshida, 2017], [Wilder, 2018]). However these studies have adopted a strategy of choosing a portfolio of seed sets (randomized solution) to provide approximation guarantees for the optimization of the conditional value at risk, under independence assumptions. This stems from the challenge that in the IC model, submodularity breaks down for the CVaR objective. On the other hand, we will show that there is a deterministic seed set with provable approximation guarantees for the optimization of the worst (lowest) value of $CVaR_\alpha\Big(R(\tilde{\mathbf{c}}, S)\Big)$ over all distributions θ in Θ.

Given $\alpha \in (0,1]$ and the marginal probability vector $\mathbf{p} = (p_{ij})_{(i,j)\in E}$, we define the correlation-robust CVaR function $CVaR_\alpha^{corr} : 2^V \to \mathbb{R}$ as:

$$CVaR_\alpha^{corr}(S) := \min_{\theta \in \Theta} CVaR_\alpha\Big(R(\tilde{\mathbf{c}}, S)\Big). \tag{13}$$

We will show that the computation of $CVaR_\alpha^{corr}(S)$ can be done efficiently given any seed set S (Theorem 6), that a minimizing distribution $\theta \in \Theta$ solving (13) can be found in the correlation robust distribution θ^* of Corollary 4 (Theorem 7), that the influence function's corresponding value at risk VaR_α under this distribution may be also be found efficiently. Finally, we will conclude by establishing that $CVaR_\alpha^{corr}(\cdot)$ is a nondecreasing, submodular function (Theorem 8), making the greedy algorithm effective for approximate optimization.

6.1 Efficient Computation

Using Lemma 2 and a straightforward application of linear programming duality (alternatively, an appeal to Sion's Minimax Theorem - [Sion, 1958]) gives the equivalence:

$$\begin{aligned} CVaR_\alpha^{corr}(S) &= \min_{\theta \in \Theta} \max_{\tau \geq 0} \left(\tau - \frac{1}{\alpha} \mathbb{E}_\theta \left[\tau - R(\tilde{\mathbf{c}}, S)\right]^+ \right) && \text{(Lemma 2)} \\ &= \max_{\tau \geq 0} \left(\tau + \frac{1}{\alpha} \min_{\theta \in \Theta} \mathbb{E}_\theta \left[\min(R(\tilde{\mathbf{c}}, S) - \tau, 0)\right] \right). && (14) \end{aligned}$$

We now focus on the inner minimization term in (14). Given any threshold $\tau \geq 0$, the random variable $\min(R(\tilde{\mathbf{c}}, S) - \tau, 0)$ expresses the amount of nodes influenced by S (including S) that falls short of the threshold τ; for shorthand, let $f_{\leq \tau}^{corr}(S) := \min_{\theta \in \Theta} \mathbb{E}_\theta \left[\min(R(\tilde{\mathbf{c}}, S) - \tau, 0)\right]$ denote the *correlation-robust τ- shortfall influence function.*

One may intuit that for any $\tau \geq 0$, $f^{corr}_{\leq\tau}(S)$ should be similar to $f^{corr}(S)$ in admitting a linear program form (recall Theorem 1), if only because $\min(R(\tilde{\mathbf{c}}, S) - \tau, 0)$ is a piecewise-linear concave transformation of $R(\tilde{\mathbf{c}}, S)$. This intuition is in fact correct (Corollary 6), the consequence of which is that the computation of $CVaR^{corr}_{\alpha}(\cdot)$ reduces to a tractable linear program (Theorem 6) that is not unlike that of Theorem 1.

Corollary 6 *Given $G = (V, E)$, a seed set $S \subseteq V$, $\tau \geq 0$, and marginal probability vector $\mathbf{p} \in [0,1]^E$, $f^{corr}_{\leq\tau}(S)$ is equal to the optimal value of the following polynomial sized linear program:*

$$\begin{aligned} \min\nolimits_{\theta\in\Theta} \mathbb{E}_{\theta}\left[\min(R(\tilde{\mathbf{c}}, S) - \tau, 0)\right] = \min_{\boldsymbol{\pi}\in\mathbb{R}^V, d\in\mathbb{R}} \quad & d(|S| - \tau) + \textstyle\sum_{i\in V\setminus S} \pi_i \\ s.t \quad & \pi_i = d \quad \forall i \in S \\ & \pi_i - \pi_j \leq 1 - p_{ij} \ \forall (i,j) \in E \\ & 0 \leq d \leq 1, \quad 0 \leq \pi_i \leq d \ \ \forall i \in V. \end{aligned}$$

Given an optimal solution d^* and π^*, it is the case that $\pi^*_i = \left[d^* - \phi^S_i\right]^+$.

Corollary 6 may be compared and contrasted with the findings of Theorem 1 and Corollary 2; in particular, d^* now replaces 1 in the closed-form to π^*_i. We will explore how this optimal variable d^* behaves with respect to τ in the larger maximization effort of (14) in the next subsection. But in the meantime, the immediate consequence of Corollary 6 is the efficiency of $CVaR^{corr}_{\alpha}(S)$.

Theorem 6 *Given $G = (V, E)$, a seed set $S \subseteq V$, marginal probability vector $\mathbf{p} \in [0,1]^E$ and $\alpha \in (0,1)$, $CVaR^{corr}_{\alpha}(S)$ admits a polynomial time solvable linear programming formulation:*

$$\begin{aligned} CVaR^{corr}_{\alpha}(S) \quad = \quad \max_{\tau, z, \boldsymbol{q}, \boldsymbol{g}, \boldsymbol{w}} \quad & \tau - \frac{1}{\alpha}\left(\sum_{(i,j)\in E} (1 - p_{ij}) w_{ij} + z\right) \\ s.t \quad & \sum_{i\in V} q_i - \sum_{i\in S} g_i - z + \tau \leq |S| \\ & \sum_{j:(j,i)\in E} w_{ji} - \sum_{j:(i,j)\in E} w_{ij} + g_i \leq q_i \, \forall i \in S, \\ & \sum_{j:(j,i)\in E} w_{ji} - \sum_{j:(i,j)\in E} w_{ij} - q_i \leq 1 \, \forall i \in V\setminus S, \\ & \tau, z \geq 0, \ q_i \geq 0 \, \forall i \in V, \ w_{ij} \geq 0 \, \forall (i,j) \in E. \end{aligned}$$

6.2 The Worst Case Correlation for CVaR

We proceed now with solving for the extremal distribution in (12), which will admit an efficient construction. The analysis will hold for arbitrary $\alpha \in (0,1)$, and hence complements the insights we found for $f^{corr}(S) = CVaR^{corr}_{\alpha=1}(S)$ in previous sections. Furthermore, this characterization will pay dividends in revealing the monotonicity and submodularity of $CVaR^{corr}_{\alpha}(\cdot)$.

The first step will be to return to equation (14) and note that Corollary 6 provides the reduction to

$$CVaR^{corr}_{\alpha}(S) = \max_{\tau\geq 0}\left\{\tau - \frac{1}{\alpha}\max_{d\in[0,1]}\left\{d\cdot(\tau - |S|) - \sum_{i\in V\setminus S}\left[d - \phi^S_i\right]^+\right\}\right\}. \tag{15}$$

We now characterize the optimal $d_S(\tau)$ variable to the inner maximization above. As it turns out, it is connected to the CDF of $R(\tilde{\mathbf{c}}, S)$ when $\tilde{\mathbf{c}}$ is distributed according to the correlation-robust distribution θ^* of Corollary 4.

Lemma 3 *For any $\tau \geq 0$, the optimal solution $d_S(\tau)$ of the inner maximization in* (15) *is given by*

$$d_S(\tau) := \begin{cases} \min(\phi^S_{i(\lceil \tau - |S| \rceil)}, 1) & \text{if } |S| < \tau \leq |V| \\ 0 & \text{if } \tau \leq |S| \\ 1 & \text{if } \tau > |V|, \end{cases} \tag{16}$$

where $\phi_i^S = \min_{\gamma \in \Gamma(S,i)} \sum_{l=0}^{\lambda_\gamma - 1} 1 - p_{i_l, i_{l+1}}$, and $i(\cdot)$ is an enumeration of $V \setminus S$ such that

$$0 \leq \phi^S_{i(1)} \leq \phi^S_{i(2)} \leq \ldots \leq \phi^S_{i(|V \setminus S|)}.$$

In Figure 7, we provide the graph of d_S. We highlight its similarity to Figure 2 and recall the insights of Corollary 4 to observe that the upper-semicontinuous closure of d_S yields the CDF $F_{R(\tilde{\mathbf{c}},S)}$ of $R(\tilde{\mathbf{c}}, S)$ when $\tilde{\mathbf{c}}$ is distributed according to the correlation-robust distribution θ^* of Corollary 4. This is a critical insight that yields several important conclusions that are summarized in the following main theorem of this section.

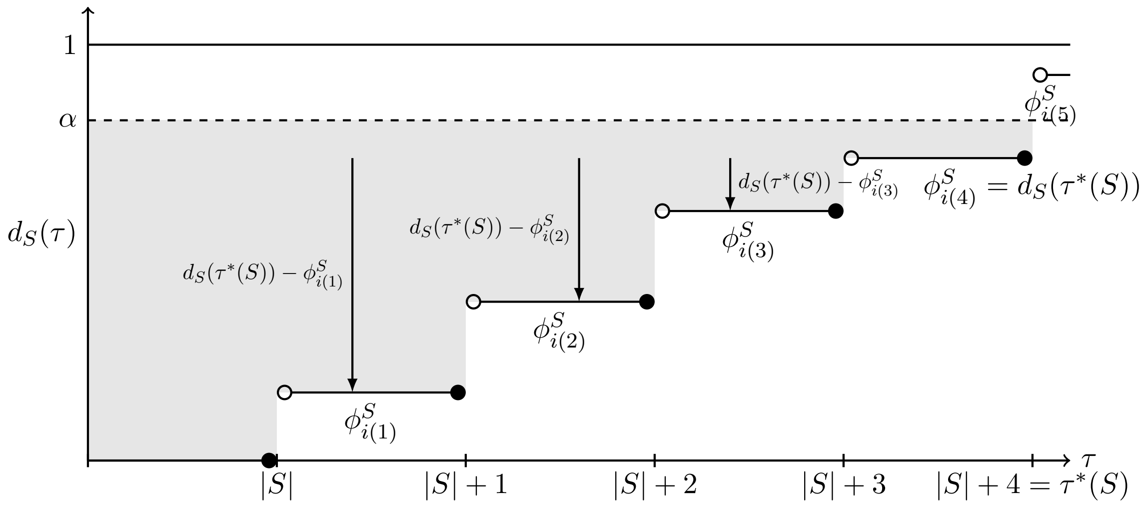


Figure 7: Schematic illustration of $d_S(\tau)$ as a function of τ. The graph coincides (up to discontinuities) with the CDF of $R(\tilde{\mathbf{c}}, S)$ where $\tilde{\mathbf{c}} \sim \theta^*$, the correlation-robust distribution of Corollary 4. The area of the shaded region is $\alpha \cdot CVaR^{corr}_\alpha(S)$ - see Eq. (17)

Theorem 7 *Consider a graph $G = (V, E)$, seed set $S \subseteq V$, marginal probability vector $\mathbf{p} \in [0,1]^E$. Let θ^* denote the correlation robust distribution of Corollary 4. Then*

(a) For $\alpha \in (0, 1]$, θ^* solves (13), i.e.,

$$CVaR^{corr}_\alpha(S) = \min_{\theta \in \Theta} \Big[CVaR_\alpha\Big(R(\tilde{\mathbf{c}}, S)\Big)\Big] = CVaR_\alpha\Big(R(\tilde{\mathbf{c}}, S)\Big), \quad \tilde{\mathbf{c}} \sim \theta^*$$

(b) $\tau^*(S) := \sup\{z \in \mathbb{Z}_+ : d_S(z) \leq \alpha\}$ solves (15). Further,

- $\tau^*(S) = VaR_\alpha(R(\tilde{\mathbf{c}}, S))$, *when* $\tilde{\mathbf{c}} \sim \theta^*$
- $d_S(\tau^*(S)) = \mathbb{P}_{\tilde{c}\sim\theta^*}\Big(R(\tilde{\mathbf{c}}, S) < \tau^*(S)\Big)$,

and both these quantities can be computed in polynomial time in the size of the graph ($|V|$ and $|E|$).

(c) For $\alpha \in (0, 1)$,

$$CVaR_\alpha^{corr}(S) = |S| + \frac{1}{\alpha}\sum_{l=1}^{\tau^*(S)-|S|}(\alpha - \phi^S_{i(l)}) = \frac{1}{\alpha}\sum_{i\in V}\left[\alpha - \phi_i^S\right]^+. \tag{17}$$

For $\alpha = 1$,

$$CVaR_\alpha^{corr}(S) = \sum_{i\in V}\left[1 - \phi_i^S\right]^+ = f^{corr}(S).$$

Remark 1 *Theorem 7 comments on the meaning or interpretation behind the optimal choice of τ and d variables that appear in (15), and they can be computed using the linear programs of Corollary 6 and Theorem 6. Specifically, $\tau^*(S) = VaR_\alpha(R(\tilde{\mathbf{c}}, S))$, but is crucially distinct from* $\min_{\theta\in\Theta} VaR_\alpha(R(\tilde{\mathbf{c}}, S))$*; in other words, θ^* is guaranteed to provide the worst $CVaR_\alpha$ at all levels of α, but not necessarily the worst $VaR_\alpha(R(\tilde{\mathbf{c}}, S))$. We refer the reader to the Appendix D.4 for an example of this.*

We note that both quantities in Theorem 7(b) can also be computed efficiently using equivalent shortest path computations. This can be exploited in an efficient implementation of greedy optimization, explored in detail in Appendix D.5. Further illustrative numerical experiments can be found in Appendix D.6. In the next section we exploit the fact that Theorem 7's equation (17) clearly exposes the monotonicity and submodularity of $CVaR_\alpha^{corr}$.

6.3 On Submodularity

In this section, we comment on the monotonicity and submodularity of $CVaR_\alpha^{corr}(\cdot)$. As previously seen these properties are desirable because this in turn implies that the problem of maximizing $CVaR_\alpha^{corr}(\cdot)$ admits a greedy algorithm approximation guarantee of $(1 - 1/e)$ (see [Nemhauser et al., 1978]). In particular, we have already seen that the expected influences f^{corr} and f^{ic} share in this guarantee by both being submodular (Theorem 3).

However, when it comes to the conditional value at risk of the influence, only one of these models continues to present submodularity. Indeed, under the independent cascade setting, the conditional value at risk is no longer necessarily submodular (see the counterexample on page 528 in [Maehara, 2015] and Example 4). We briefly comment on this breakdown by the independent cascade model.

6.3.1 Independent Cascade's Loss of Submodularity

Since $\mathbb{E}_{\tilde{\mathbf{c}}\sim\theta}\left[R(\tilde{\mathbf{c}}, S)\right] = \sum_{i\in V}\mathbb{P}_{\tilde{c}}\Big(\text{Node } i \text{ is influenced by } S\Big)$ for any $\tilde{\mathbf{c}} \sim \theta$, it follows that the expectations $f^{corr}(S) = \sum_{i\in V}[1 - \phi_i^S]^+$ and $f^{ic}(S) = \sum_{i\in V}(1 - \eta_i^S)$ display an (additively) separable form that ultimately reveals both are sums of submodular functions. Theorem 7c and equation (17) indicates this separable form is not guaranteed when transitioning to the CVaR metric however. Thus in Example 4, we find that when $\tilde{\mathbf{c}} \sim \theta_{ic}$, and $1/8 < \alpha \le 1/4$, then

$$CVaR_\alpha\Big(R(\tilde{\mathbf{c}}, \{2,3\})\Big) + CVaR_\alpha\Big(R(\tilde{\mathbf{c}}, \{3,4\})\Big) = 1/\alpha\Big(2\alpha + [\alpha - 1/4]^+\Big) + 1/\alpha\Big(2\alpha + [\alpha - 1/4]^+\Big)$$
$$< 1/\alpha\Big(\alpha + [\alpha - 1/2]^+\Big) + 1/\alpha\Big(3\alpha + [\alpha - 1/8]^+\Big) = CVaR_\alpha\Big(R(\tilde{\mathbf{c}}, \{3\})\Big) + CVaR_\alpha\Big(R(\tilde{\mathbf{c}}, \{2,3,4\})\Big),$$

indicating violation of submodularity by the independent cascade's CVaR.

6.3.2 $CVaR_\alpha^{corr}$'s Preservation of Submodularity

We now formally establish the submodularity property of $CVaR_\alpha^{corr}$ as well as the $1 - 1/e$ guarantee the greedy algorithm yields. We refer the reader to the Appendix D.5 for a run-time analysis.

Theorem 8 *For any given value $\alpha \in (0,1)$, the function $CVaR_\alpha^{corr}(S)$ is nondecreasing and submodular as a function of S. In particular, if S_α^g is any set that is produced upon termination of the greedy procedure for the optimization of $CVaR_\alpha^{corr}(S)$, then $CVaR_\alpha^{corr}(S_\alpha^g) \ge (1 - 1/e) \cdot \max_{|S| \le k} CVaR_\alpha^{corr}(S)$.*

Analogously to Theorem 5, we may also select a greedy set with a performance guarantee that is uniform over a collection of risk levels.

Theorem 9 *Let $\Omega \subseteq [0,1]$ be a finite set. If β solves the maximization problem*

$$C := \max_{\beta \in \Omega} \min_{\alpha \in \Omega} \left(\frac{CVaR_\alpha^{corr}(S_\beta^g)}{CVaR_\alpha^{corr}(S_\alpha^g)} \right), \tag{18}$$

then $\quad CVaR_\alpha^{corr}(S_\beta^g) \ge C \cdot (1 - 1/e) \max_{S \le k} CVaR_\alpha^{corr}(S) \qquad \forall \alpha \in \Omega.$

7 Data Experiments

We now discuss our findings from numerical experiments. In the discussions to follow, S_{ic} and S_{corr} denote optimal seed sets for the optimization of f^{ic} and f^{corr}, respectively. Similarly, S_{ic}^g and S_{corr}^g denote greedy-optimal seed sets. In Section 7.2, S_λ and S_λ^g denote, respectively, optimal and greedy-optimal seed sets for the optimization of $f^{Mix(\lambda)}$, for any $\lambda \in (0,1)$. Although both the mixture parameter λ and the risk level α are numbers in $(0,1)$, the context will always be clear as to whether we are referring to mixture or risk. For this reason, in Section 7.3, we also let S_α and S_α^g denote, respectively, optimal and greedy-optimal seed sets for the optimization of $CVaR_\alpha^{corr}$, for any $\alpha \in (0,1)$. For a discussion on the homophily phenomenon as it relates to our robust models, we refer the interested reader to Appendix E.3.

7.1 f^{ic} versus f^{corr}

Datasets: We now discuss numerical experiments to compare and contrast the IC (f^{ic}) and correlation robust influence (f^{corr}) models [1] on two publicly available datasets: `wikivote`

[1]Code available at `https://github.com/justanothergithubber/corr-im/`

and `polblogs`. We briefly describe these datasets below, and Table 2 summarizes the graphs contained therein.

(1) `wikivote`: Here each node denotes a user and each edge (i, j) denotes the action of user i voting for user j to be an admin [Leskovec and Krevl, 2014]. As in [Yan et al., 2011,Zhang et al., 2014] we reverse the edges so that edge (i, j) in the original graph becomes (j, i). Indeed, this reverse direction more aptly captures a notion of influence, as user i's vote for j establishes that user j has influence over user i.

(2) `polblogs`: Each node denotes a blog and each edge (i, j) denotes that blog i references blog j via a hyperlink [Adamic and Glance, 2005]. Since a highly referenced blog is "influential". Just as in `wikivote`, we reversed the direction of all edges.

Dataset	$\|V\|$	$\|E\|$	Min Deg	Average Deg	Max Deg
`wikivote`	7115	103689	1	29.146	1167
`polblogs`	1490	19022	0	25.532	467

Table 2: Numerical Summaries of the Graphs from the Datasets

Computational Method: For the computation of $f^{corr}(S)$, following Corollary 2, this was done deterministically with shortest path calculations, after generation of the p_{ij} values (experimental details below). In particular, the edge weights for the equivalent shortest path calculations are $1 - p_{ij}$ (see (9)). We favored this style of implementation over Theorem 1's efficient linear program for the fact that it can be efficiently updated between iterations of greedy maximization. We used the `igraph` Python library [Csardi and Nepusz, 2006] to represent the graphs and for the shortest path calculations. For (theoretical) details on implementation as well as the run-time of computing $f^{corr}(S)$, including its greedy maximization, we refer the reader to Section D.5.

As for the computation of $f^{ic}(S) = \mathbb{E}_{\theta_{ic}}[R(\tilde{\mathbf{c}}, S)]$, we implemented 10 batches of the pruned Monte Carlo algorithm of [Ohsaka et al., 2014], each batch having algorithm parameter setting of $R = 10^4$ for the number of DAGs generated. We used an ASUS laptop with i7-7500U processor for all experiments- with reported computational times in Appendix E.1.

Experiment and Results - Seed Performance and Graph Statistics: We examine properties to the seed sets S^g_{ic} and S^g_{corr} under three different schemes of generating the edge probabilities p_{ij}:

(1) Unif$(0, 1)$: p_{ij} drawn i.i.d. from Unif$(0, 1)$;

(2) Trivalency: p_{ij} drawn i.i.d. from Unif$\{0.1, 0.01, 0.001\}$;

(3) Weighted cascade: $p_{ij} = 1/\text{deg}(i)$, deg(i) denotes the number of edges directed into i.

In Table 3 we report various metrics for each of the three seed sets S^g_{corr}, S^g_{ic}, and S^g_{deg}. We summarize some takeaways.

- $\mathbf{S^g_{corr}}$ **versus** $\mathbf{S^g_{ic}}$: Nearly always $\frac{f^{ic}(S^g_{corr})}{f^{ic}(S^g_{ic})} > \frac{f^{corr}(S^g_{ic})}{f^{corr}(S^g_{corr})}$, indicating the performance of S^g_{corr} is more robust than that of S^g_{ic}.
- $\mathbf{S^g_{corr}}$ **versus** $\mathbf{S_{deg}}$: Nearly always $\frac{f^{ic}(S^g_{corr})}{f^{ic}(S^g_{ic})} > \frac{f^{ic}(S_{deg})}{f^{ic}(S^g_{ic})}$, highlighting the utility of the probability information p_{ij} taken into account by S^g_{corr} as opposed to S_{deg}. The only instance in which this inequality did not hold was in the Trivalency case on the `wikivote`

graph, which might be expected, in light of the fact that the p_{ij} values were very small. As well, it was always the case that $\frac{f^{corr}(S_{deg})}{f^{corr}(S^g_{corr})} < 1$, and this ratio was significantly smaller in the case of W.C.-generated p_{ij} values.

For a comparison of the degree distributions to S^g_{corr} and S^g_{ic}, we refer the reader to Appendix E.1's histograms of Figures 11a and 11b, which detail the experiment on `polblogs`.

Seed	Dataset	p	$\frac{f^{ic}(S^g_{corr})}{f^{ic}(S^g_{ic})}$	Min Deg	Average Deg	Max Deg	Diam
S^g_{corr}	`wikivote`	Unif(0,1)	0.988	41	159.475	472	3
		Trivalency	0.928	104	288.75	1167	2
		W.C.	0.948	29	217.375	537	3
	`polblogs`	Unif(0,1)	0.981	1	98.025	383	6
		Trivalency	0.933	15	159.575	467	3
		W.C.	0.964	4	140.0	467	5

Seed	Dataset	p	$\frac{f^{corr}(S^g_{ic})}{f^{corr}(S^g_{corr})}$	Min Deg	Average Deg	Max Deg	Diam
S^g_{ic}	`wikivote`	Unif(0,1)	0.976	12	116.85	331	3
		Trivalency	0.908	93	192.325	472	2
		W.C.	0.949	60	271.975	1167	3
	`polblogs`	Unif(0,1)	0.957	1	29.825	143	7
		Trivalency	0.928	42	130.275	383	3
		W.C.	0.96	15	163.225	467	4

Seed	Dataset	p	$\frac{f^{corr}(S_{deg})}{f^{corr}(S^g_{corr})}$	$\frac{f^{ic}(S_{deg})}{f^{ic}(S^g_{ic})}$	Min Deg	Average Deg	Max Deg	Diam
S_{deg}	`wikivote`	Unif(0,1)	0.952	0.952	178	329.975	1167	2
		Trivalency	0.942	0.995	178	329.975	1167	2
		W.C.	0.746	0.909	178	329.975	1167	2
	`polblogs`	Unif(0,1)	0.937	0.927	103	192.075	467	3
		Trivalency	0.959	0.918	103	192.075	467	3
		W.C.	0.810	0.919	103	192.075	467	3

Table 3: Properties of S^g_{ic}, S^g_{corr}, and S_{deg} for non-identical edge probabilities generated using either: (1) independent Unif(0,1) draws; (2) Trivalency; (3) Weighted Cascade. $k = 40$. Deg measures total degree, including both in- and out-edges.

7.2 $f^{Mix(\lambda)}$: Mixing f^{ic} and f^{corr}

We now discuss experiments conducted on Amazon's Co-Purchasing Network provided in the `amazon-meta` dataset [Leskovec et al., 2007a]. These experiments demonstrate the applicability of robust influence maximization in areas like advertising, and also demonstrate how in practice a decision-maker may tune the λ parameter of $f^{Mix(\lambda)}$ to attain decisions exhibiting desired levels of robustness.

Dataset: The `amazon-meta` dataset consists of product metadata on the e-commerce website Amazon from the year 2003. For each product i, there is a list consisting of products j for which j is "frequently" found to be purchased, given i is purchased. While the exact numerical frequencies of such co-purchasing events are not provided in this historical dataset, such data is precisely the kind of information that tech firms like Amazon could readily supplement in modern-day practice. Thus, we content ourselves with the following devised experiment as proof-of-concept for application.

Computational Method: The computations of $f^{ic}(S)$ and $f^{corr}(S)$ for any set S were performed just as in Section 7.1, after generation of the p_{ij} values (experimental details below). Consequently, the computation of $f^{Mix(\lambda)}(S)$ followed immediately as a convex combination of these two values, once computed. The computations of greedy sets S^g_λ thus also follow similarly.

7.2.1 Data Experiment: Amazon's DVD Co-Purchasing Network

In this experiment, we consider the influence maximization framework as a methodology for deriving a set of product recommendations. More precisely, a seed set S in this context will refer to a set of products recommended to a customer. Upon visiting the page of a product i, there is some likelihood that the customer subsequently visits the page of j, and so on. We adopt the network structure of `amazon-meta` as well as the notion of co-purchasing captured therein as, more generally, *page-traversing* activity. In doing so, the influence maximization problem manifests in this context as devising a seed set S such that the average number of products (including those in S) viewed by the customer via page-traversing is maximized.

From the Amazon dataset, we derived a directed graph $G = (V, E)$ for the influence maximization framework in the following way. First, we let V consist of only those nodes(/products) belonging to the DVD group. Secondly, for any node i, we created a directed arc from i to some j if and only if DVD j was among i's ordered list of co-purchased items. Recall that this ordered list details the nodes j in which i and j are part of so-called "frequent" co-purchasing events. And while `amazon-meta` provides no numerical frequencies associated with entries on this list (as explained above), we assumed the ordering in this list was the ordering of the associated numerical frequencies, and we subsequently used this to construct the p_{ij} probabilities associated with arcs (i, j) exiting node i. More precisely, since for DVD nodes, any i's list of neighbors was no larger than 5, we assigned to each directed arc (i, j) one of five prescribed frequency values $\{v_\ell\}_{\ell=1}^5$ satisfying $1 > v_1 > v_2 > \ldots > v_5 > 0$. That is, (i, j) was prescribed $p_{ij} = v_\ell$ if and only if j was the ℓ-th member on the ordered list of neighbors of i. As for the precise values of $\{v_\ell\}_{\ell=1}^5$, we experimented with two regimes - "large" and "small".

1. "Large" edge probabilities: $v_1 = .95, v_2 = .85, v_3 = .75, v_4 = .65, v_5 = .55$
2. "Small" edge probabilities: $v_1 = .1, v_2 = .075, v_3 = .05, v_4 = .025, v_5 = .01$

The resulting sub-network was comprised of 19,828 nodes (DVDs) and 50,908 arcs. With this setup, we greedily solved $\max_{|S|\leq 40} f^{Mix(\lambda)}(S)$ for $\lambda \in \{0, 0.25, 0.5, 0.75, 1\}$ to observe the effect of varying the mixture parameter.

Experiment Results: Recall that S^g_λ denotes a greedy solution obtained in the greedy maximization of $f^{Mix(\lambda)}$. In Table 4 (for "large" regime) and Table 5 (for "small" regime), we display $f^{Mix(\lambda)}\left(S^g_{\lambda'}\right)$ for $\lambda, \lambda' \in \Lambda := \{0, 0.25, 0.5, 0.75, 1\}$. For example, the second row corresponding to set $S^g_{0.25}$ provides its performance $f^{Mix(\lambda)}\left(S^g_{0.25}\right)$ under varying mixtures λ, and the first column corresponding to function $f^{Mix(0)}$ provides the performance $f^{Mix(0)}\left(S^g_{\lambda'}\right)$ of various seed sets $\{S^g_{\lambda'}\}_{\lambda' \in \Lambda}$ under this choice of function.

We highlight the fact that S_{deg}, the benchmark seed set composed of the nodes with the highest out-degrees consistently performed the worst, as might be expected since no probability information informs its construction. Further, we note that in the "small" probability regime, the performance of all seed sets were virtually the same, as might also be expected. In fact, we see that the computationally efficient $f^{Mix(0)}$ differed very little from the expensive $f^{Mix(1)}$, suggesting an advantage of f^{corr} over f^{ic} in the "small" probability regime.

$S^g_{\lambda'}$ | $f^{Mix(\lambda)}$	$f^{Mix(0)}$	$f^{Mix(0.25)}$	$f^{Mix(0.50)}$	$f^{Mix(0.75)}$	$f^{Mix(1)}$
$S^g_{corr} = S^g_0$	1077.65 (1)	1541.07 (.985)	2004.48 (.95)	2467.9 (.909)	2931.32 (.874)
$S^g_{0.25}$	1035.85 (.961)	1564.31 (1)	2092.78 (.988)	2621.24 (.965)	3149.7 $\boxed{.939}$
$S^g_{0.50}$	964.20 (.89)	1541.11 (.985)	2118.02 (1)	2694.93 (.992)	3271.84 (.976)
$S^g_{0.75}$	858.55 (.797)	1477.55 (.945)	2096.54 (.990)	2715.54 (1)	3334.54 (.995)
$S^g_{ic} = S^g_1$	794.4 (.737)	1433.95 (.917)	2073.51 (.979)	2713.06 (.999)	3352.61 (1)
S_{deg}	502	922.57	1343.14	1763.71	2184.28

Table 4: ("Large" edge probabilities) $f^{Mix(\lambda)}\left(S^g_{\lambda'}\right)$ for $\lambda' \in \Lambda := \{0, 0.25, 0.5, 0.75, 1\}$. The ratio $\frac{f^{Mix(\lambda)}(S^g_{\lambda'})}{f^{Mix(\lambda)}(S^g_{\lambda})}$ is in parentheses. Boxed is $\max_{\lambda'\in\Lambda}\min_{\lambda\in\Lambda}\left(\frac{f^{Mix(\lambda)}(S^g_{\lambda'})}{f^{Mix(\lambda)}(S^g_{\lambda})}\right)$.

$S^g_{\lambda'}$ | $f^{Mix(\lambda)}$	$f^{Mix(0)}$	$f^{Mix(0.25)}$	$f^{Mix(0.50)}$	$f^{Mix(0.75)}$	$f^{Mix(1)}$
$S^g_{corr} = S^g_0$	50.4 (1)	51.0 (.994)	51.6 (.988)	52.19 (.982)	52.79 (.976)
$S^g_{0.25}$	50.4 (1)	51.32 (1)	52.23 (1)	53.15 (1)	54.07 (1)
$S^g_{0.50}$	50.4 (1)	51.32 (1)	52.23 (1)	53.15 (1)	54.07 (1)
$S^g_{0.75}$	50.4 (1)	51.32 (1)	52.23 (1)	53.15 (1)	54.07 (1)
$S^g_{ic} = S^g_1$	50.32 (.998)	51.25 (.999)	52.19 (.999)	53.13 (.999)	54.07 (1)
S_{deg}	50.08	50.75	51.42	52.09	52.76

Table 5: ("Small" edge probabilities) $f^{Mix(\lambda)}\left(S^g_{\lambda'}\right)$ for $\lambda' \in \Lambda := \{0, 0.25, 0.5, 0.75, 1\}$. The ratio $\frac{f^{Mix(\lambda)}(S^g_{\lambda'})}{f^{Mix(\lambda)}(S^g_{\lambda})}$ for $\lambda \in \Lambda$ is in parentheses.

Tuning λ: Finding Mixture-Robust Influence.

In this subsection, we illustrate application of Theorem 5 to choose from a collection of greedy sets $\{S^g_{\lambda'}\}_{\lambda'\in\Lambda}$ so as to obtain the best available performance guarantee. Towards making such a selection, Theorem 5 proposes consideration of the ratios $\frac{f^{Mix(\lambda)}(S^g_{\lambda'})}{f^{Mix(\lambda)}(S^g_{\lambda})}$ for $\lambda \in \Lambda$ displayed in Table 4 (resp. Table 5) in parentheses; namely, we solve (11) by determining the row whose smallest parenthetical entry is largest. In doing so for Tables 4 and 5, we determine that $S^g_{0.25}$ provides the best uniform guarantee both in the "large" probabilities regime,

$$f^{Mix(\lambda)}\left(S^g_{0.25}\right) \geq 0.939 \cdot (1 - 1/e) \max_{S \leq k} f^{Mix(\lambda)}(S), \text{for all } \lambda \in \Lambda,$$

as well as in the "small" probabilities regime,

$$f^{Mix(\lambda)}\left(S^g_{0.25}\right) \geq 1 \cdot (1 - 1/e) \max_{S \leq k} f^{Mix(\lambda)}(S), \text{for all } \lambda \in \Lambda.$$

Tuning λ: Insights into DVD Recommendations

Beyond granting the ability to obtain a uniform performance guarantee, the mixture parameter λ in this DVD recommendation context also offers some operational insights into the genres of movies that are most effective for influence spread. We compared and contrasted the seed sets $S_0^g, S_{0.5}^g$, and S_1^g. The results are presented in two Venn Diagrams (one for the "large" and for the "small" probability regime) in Appendix E.2. In summary, we found that in the "large" probability regime, there were more DVDs found at the various intersections of the three seed sets. In contrast, in the "small" probability regime, $S_0^g \cap S_1^g = S_0^g \cap S_{0.5}^g = \emptyset$. Further, the "large" probability regime had greedy seed sets displaying a far more diverse set of genres being recommended than in the "small" probability regime. One commonality to both regimes, however, was that the genres being represented the most were: Comedy, Drama, and Action/Adventure.

7.3 $CVaR_\alpha^{corr}$: Tuning α.

Dataset: `polblogs`
Computational Method: Theorem 6 indicates that computation may be performed with an efficient linear program. However, (17), Corollary 2 and Theorem 7 indicate an even more efficient procedure that leverages shortest path calculations and is conducive for implementation in a greedy maximization framework. For (theoretical) details on implementation as well as the run-time of computing $CVaR_\alpha^{corr}$, including its greedy maximization, we refer the reader to Section D.5.
Experiment and Results: We performed two experiments, differentiated by how the edge probabilities were randomly generated - Unif(0,1) and Trivalency - done just as in Section 7.1. Upon randomly generating the edge probabilities (either Unif(0,1) or Trivalency), we derived the greedy sets S_β^g for greedy optimization of $CVaR_\beta^{corr}$, over various β. Table 6 (resp. 7) compiles the computations of $CVaR_\alpha^{corr}(S_\beta^g)$ for combinations of $\alpha, \beta \in \Omega_{unif} := \{0.01, 0.25, 0.5, 0.75, 1\}$ (resp. $\Omega_{tri} := \{0.90, 0.92, 0.94, 0.96, 1.0\}$) when the edge probabilities were generated via iid draws from Unif(0,1) (resp. Unif{0.1, 0.01, 0.001}). For comparison, in the last row of each table we also provide $CVaR_\alpha^{corr}(S_{deg})$ for varying $\alpha \in \Omega$.

We remark that in the case of iid drawn edge probabilities from Unif{0.1, 0.01, 0.001}, according to Theorem 7's Eq. (17), for $\alpha \leq 0.90$, the function $CVaR_\alpha^{corr}$ is identically equal to $k = 40$. This is because $\phi_{i(1)}^S \geq 0.9$ in such a setup, and it is the reason for why Table 7 examines a different range for α and β than in Table 6. Other consequences of this difference in edge probabilities can be found in the performance of S_{deg}. In Table 6's case of Unif(0,1)-generated edge probabilities, while S_{deg} was not the worst performing seed set for low $\alpha \in \{0.01, 0.25\}$, it was for all $\alpha \geq 0.5$. On the other hand, in Table 7's case of Trivalency-generated edge probabilities, S_{deg} performed closer to the best performing set for all α. Highlighted in parentheses are the ratios $\frac{CVaR_\alpha^{corr}(S_\beta^g)}{CVaR_\alpha^{corr}(S_\alpha^g)}$, and boxed is the max-min value found by searching for the row of values whose minimum is largest. Following Theorem 9 and the computation of (18), we conclude that the max-min search in Table 6 indicates $S_{0.01}^g$ grants a $0.92 \cdot (1 - 1/e)$ guarantee uniformly over $\alpha \in \Omega_{unif}$, that is, $CVaR_\alpha^{corr}(S_{0.01}^g) \geq 0.92 \cdot (1 - 1/e) \max_{S \leq k} CVaR_\alpha^{corr}(S)$, for all $\alpha \in \Omega_{unif}$. Interestingly, in the case of Table 7, we see that S^g_{corr} was consistently the best performer, endowing a uniform guarantee of $1 \cdot (1 - 1/e)$ by way of Theorem 9, making a case for consideration of the correlation robust model in cases of small edge probabilities.

$S^g_\beta \vert \alpha$	$\alpha = 0.01$	$\alpha = 0.25$	$\alpha = 0.50$	$\alpha = 0.75$	$\alpha = 1.0$
$S^g_{0.01}$	94.42 (1)	534.38 $\boxed{(.92)}$	700.54 (.94)	789.15 (.95)	845.55 (.95)
$S^g_{0.25}$	73.52 (.779)	582.73 (1)	741.45 (.991)	824.06 (.987)	875.92 (.984)
$S^g_{0.50}$	70.35 (.745)	580.71 (.997)	747.82 (1)	833.03 (.998)	885.49 (.994)
$S^g_{0.75}$	65.44 (.693)	571.58 (.981)	745.26 (.997)	834.63 (1)	889.49 (.999)
S^g_{corr}	64.62 (.684)	565.74 (.971)	742.59 (.993)	834.05 (.999)	890.43 (1)
S_{deg}	69.99	535.62	694.15	779.49	834.62

Table 6: $CVaR^{corr}_\alpha(S^g_\beta)$ with edge probabilities drawn iid from $Unif(0,1)$. The ratio $\frac{CVaR^{corr}_\alpha(S^g_\beta)}{CVaR^{corr}_\alpha(S^g_\alpha)}$ is in parentheses. Boxed is $\max_{\beta\in\Omega_{unif}} \min_{\alpha\in\Omega_{unif}} \left(\frac{CVaR^{corr}_\alpha(S^g_\beta)}{CVaR^{corr}_\alpha(S^g_\alpha)}\right)$

$S^g_\beta \vert \alpha$	$\alpha = 0.90$	$\alpha = 0.92$	$\alpha = 0.94$	$\alpha = 0.98$	$\alpha = 1.0$
$S^g_{0.90}$	40 (1)	42.33 (.756)	44.55 (.625)	48.73 (.487)	51.26 (.446)
$S^g_{0.92}$	40 (1)	55.96 (1)	71.23 (.999)	99.92 (.999)	114.69 (.999)
$S^g_{0.94}$	40 (1)	55.96 (1)	71.23 (.999)	99.92 (.999)	114.67 (.999)
$S^g_{0.98}$	40 (1)	55.96 (1)	71.23 (.999)	99.92(.999)	114.7 (.999)
S^g_{corr}	40 $\boxed{(1)}$	55.98 $\boxed{(1)}$	71.28 $\boxed{(1)}$	100 $\boxed{(1)}$	114.84 $\boxed{(1)}$
S_{deg}	40	54.93	69.23	96.08	110.13

Table 7: $CVaR^{corr}_\alpha(S^g_\beta)$ with edge probabilities iid from $Unif\{0.1, 0.01, 0.001\}$. The ratio $\frac{CVaR^{corr}_\alpha(S^g_\beta)}{CVaR^{corr}_\alpha(S^g_\alpha)}$ is in parentheses. Boxed is $\max_{\beta\in\Omega_{tri}} \min_{\alpha\in\Omega_{tri}} \left(\frac{CVaR^{corr}_\alpha(S^g_\beta)}{CVaR^{corr}_\alpha(S^g_\alpha)}\right)$

8 Conclusions and Extensions

We have proposed a correlation robust model - f^{corr} - for influence maximization in which the activation probabilities of the edges are known, but the joint distribution of these activations is unknown, and in fact, adversarially chosen upon selection of a seed set. This model thus accounts for the possibility of correlated edge activations (unlike IC), and we find it presents several properties favorable for computation, analytical characterization, and optimization. We used the POC metric to illustrate the extent that the IC model may mislead a decision maker in cases of correlations, and subsequently with the mixture model $f^{Mix(\lambda)}$, we offer a way to tunably introduce pessimism into the IC model to achieve more robust decisions. Our results and insights are extended to the problem of optimizing $f^{Mix(\lambda)}$, as well as the robust risk measure $CVaR^{corr}_\alpha$. Finally, our experiments provide insights on real datasets, and showcases how we may tune parameters like λ and α for robust performance guarantees in practice.

Acknowledgments

We thank Will Ma for useful discussions and suggestions, especially in providing Example 2 in the price of correlations section of our work. We are also grateful to the referees and associate editor for useful suggestions that have led to significant enhancement of the paper.

References

[Acerbi and Tasche, 2002] Acerbi, C. and Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503.

[Adamic and Glance, 2005] Adamic, L. A. and Glance, N. (2005). The Political Blogosphere and the 2004 U.S. Election: Divided They Blog. In *Proceedings of the 3rd International Workshop on Link Discovery*, LinkKDD '05, page 36–43, New York, NY, USA. Association for Computing Machinery.

[Agrawal et al., 2012] Agrawal, S., Ding, Y., Saberi, A., and Ye, Y. (2012). Price of Correlations in Stochastic Optimization. *Operations Research*, 60(1):150–162.

[Akbarpour and Jackson, 2018] Akbarpour, M. and Jackson, M. O. (2018). Diffusion in networks and the virtue of burstiness. *Proceedings of the National Academy of Sciences*, 115(30):E6996–E7004.

[Aral and Dhillon, 2018] Aral, S. and Dhillon, P. S. (2018). Social influence maximization under empirical influence models. *Nature Human Behaviour*, 2(6):375–382.

[Bass, 1969] Bass, F. M. (1969). A new product growth for model consumer durables. *Management science*, 15(5):215–227.

[Borgs et al., 2014] Borgs, C., Brautbar, M., Chayes, J., and Lucier, B. (2014). Maximizing social influence in nearly optimal time. In *Proceedings of the 2014 Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '03, page 946–957, USA. Society for Industrial and Applied Mathematics.

[Chen et al., 2022] Chen, L., Ma, W., Natarajan, K., Simchi-Levi, D., and Yan, Z. (2022). Distributionally robust linear and discrete optimization with marginals. *Operations Research*, 70(3):1822–1834.

[Chen et al., 2020] Chen, L. L., Ma, W., Orlin, J. B., and Simchi-Levi, D. (2020). Distributionally robust max flows. In *Symposium on Simplicity in Algorithms*, pages 81–90. SIAM.

[Chen et al., 2011] Chen, W., Collins, A., Cummings, R., Ke, T., Liu, Z., Rincón, D., Sun, X., Wang, Y., Wei, W., and Yuan, Y. (2011). Influence maximization in social networks when negative opinions may emerge and propagate. In *Proceedings of the 2011 siam international conference on data mining*, pages 379–390. SIAM.

[Chen et al., 2016] Chen, W., Lin, T., Tan, Z., Zhao, M., and Zhou, X. (2016). Robust influence maximization. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 795–804.

[Chen et al., 2010] Chen, W., Wang, C., and Wang, Y. (2010). Scalable influence maximization for prevalent viral marketing in large-scale social networks. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '10, page 1029–1038, New York, NY, USA. Association for Computing Machinery.

[Chen et al., 2009] Chen, W., Wang, Y., and Yang, S. (2009). Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 199–208.

[Csardi and Nepusz, 2006] Csardi, G. and Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, Complex Systems:1695.

[Floyd, 1962] Floyd, R. W. (1962). Algorithm 97: shortest path. *Communications of the ACM*, 5(6):345.

[Fu et al., 2012] Fu, F., Nowak, M. A., Christakis, N. A., and Fowler, J. H. (2012). The evolution of homophily. *Scientific reports*, 2(1):1–6.

[He and Kempe, 2014] He, X. and Kempe, D. (2014). Stability of Influence Maximization. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, page 1256–1265, New York, NY, USA. Association for Computing Machinery.

[He and Kempe, 2016] He, X. and Kempe, D. (2016). Robust influence maximization. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 885–894, New York, NY, USA. Association for Computing Machinery.

[Hoffmann and Caramanis, 2018] Hoffmann, J. and Caramanis, C. (2018). The cost of uncertainty in curing epidemics. *Proceedings of the ACM Measurement and Analysis of Computing Systems*, 2(2).

[Hunter and Zaman, 2022] Hunter, D. S. and Zaman, T. (2022). Optimizing opinions with stubborn agents. *Operations Research*, 70(4):2119–2137.

[Johnson, 1977] Johnson, D. B. (1977). Efficient algorithms for shortest paths in sparse networks. *Journal of the ACM (JACM)*, 24(1):1–13.

[Kalimeris et al., 2019] Kalimeris, D., Kaplun, G., and Singer, Y. (2019). Robust influence maximization for hyperparametric models. In *ICML*, volume 97, pages 3192–3200, Long Beach, California, USA. PMLR.

[Kempe et al., 2003] Kempe, D., Kleinberg, J., and Tardos, É. (2003). Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146.

[Kempe et al., 2015] Kempe, D., Kleinberg, J., and Tardos, É. (2015). Maximizing the spread of influence through a social network. *Theory of Computing*, 11(4):105–147.

[Klein Haneveld, 1986] Klein Haneveld, W. K. (1986). *Robustness against dependence in PERT: An application of duality and distributions with known marginals*, pages 153–182. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Kossinets and Watts, 2009] Kossinets, G. and Watts, D. J. (2009). Origins of homophily in an evolving social network. *American journal of sociology*, 115(2):405–450.

[Krause et al., 2008] Krause, A., McMahan, H. B., Guestrin, C., and Gupta, A. (2008). Robust Submodular Observation Selection. *Journal of Machine Learning Research*, 9(Dec):2761–2801.

[Leskovec et al., 2007a] Leskovec, J., Adamic, L. A., and Huberman, B. A. (2007a). The dynamics of viral marketing. *ACM Transactions on the Web (TWEB)*, 1(1):5–es.

[Leskovec et al., 2007b] Leskovec, J., Krause, A., Guestrin, C., Faloutsos, C., VanBriesen, J., and Glance, N. (2007b). Cost-effective outbreak detection in networks. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '07, page 420–429, New York, NY, USA. Association for Computing Machinery.

[Leskovec and Krevl, 2014] Leskovec, J. and Krevl, A. (2014). SNAP Datasets: Stanford large network dataset collection. `http://snap.stanford.edu/data`.

[Li et al., 2018] Li, Y., Fan, J., Wang, Y., and Tan, K.-L. (2018). Influence maximization on social graphs: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 30(10):1852–1872.

[Ma, 2020] Ma, W. (2020). Personal communication with the authors on an example graph indicating arbitrarily small Price of Correlations, March 2020.

[Maehara, 2015] Maehara, T. (2015). Risk averse submodular utility maximization. *Operations Research Letters*, 43(5):526–529.

[Mallipeddi et al., 2022] Mallipeddi, R. R., Kumar, S., Sriskandarajah, C., and Zhu, Y. (2022). A framework for analyzing influencer marketing in social networks: Selection and scheduling of influencers. *Management Science*, 68(1):75–104.

[McCormick, 1976] McCormick, G. P. (1976). Computability of global solutions to factorable nonconvex programs: Part I — Convex underestimating problems. *Mathematical Programming*, 10(1):147–175.

[McPherson et al., 2001] McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27:415–444.

[Meilijson and Nádas, 1979] Meilijson, I. and Nádas, A. (1979). Convex majorization with an application to the length of critical paths. *Journal of Applied Probability*, 16(3):671–677.

[Minoux, 1978] Minoux, M. (1978). Accelerated greedy algorithms for maximizing submodular set functions. In Stoer, J., editor, *Optimization Techniques*, pages 234–243, Berlin, Heidelberg. Springer Berlin Heidelberg.

[Nannicini et al., 2019] Nannicini, G., Sartor, G., Traversi, E., and Wolfler-Calvo, R. (2019). An exact algorithm for robust influence maximization. In Lodi, A. and Nagarajan, V., editors, *Integer Programming and Combinatorial Optimization*, pages 313–326, Cham. Springer International Publishing.

[Natarajan et al., 2009] Natarajan, K., Song, M., and Teo, C.-P. (2009). Persistency model and its applications in choice modeling. *Management Science*, 55(3):453–469.

[Nemhauser et al., 1978] Nemhauser, G. L., Wolsey, L. A., and Fisher, M. L. (1978). An analysis of approximations for maximizing submodular set functions–I. *Mathematical programming*, 14(1):265–294.

[Ohsaka et al., 2014] Ohsaka, N., Akiba, T., Yoshida, Y., and Kawarabayashi, K.-i. (2014). Fast and accurate influence maximization on large networks with pruned monte-carlo simulations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28.

[Ohsaka and Yoshida, 2017] Ohsaka, N. and Yoshida, Y. (2017). Portfolio optimization for influence spread. In *Proceedings of the 26th International Conference on World Wide Web*, pages 977–985.

[Rũschendorf, 2013] Rũschendorf, L. (2013). *Mathematical Risk Analysis: Dependence, Risk Bounds, Optimal Allocations and Portfolios*. Springer, 2 edition.

[Rockafellar, 1997] Rockafellar, R. T. (1997). *Convex analysis*, volume 11. Princeton university press.

[Rockafellar and Uryasev, 2000] Rockafellar, R. T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2(3):21–41.

[Rockafellar and Uryasev, 2002] Rockafellar, R. T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of banking & finance*, 26(7):1443–1471.

[Sion, 1958] Sion, M. (1958). On general minimax theorems. *Pacific Journal of Mathematics*.

[Staib et al., 2019] Staib, M., Wilder, B., and Jegelka, S. (2019). Distributionally robust submodular maximization. In *Proceedings of the International Workshop on Artificial Intelligence and Statistics*.

[Vaswani et al., 2017] Vaswani, S., Kveton, B., Wen, Z., Ghavamzadeh, M., Lakshmanan, L. V., and Schmidt, M. (2017). Model-independent online learning for influence maximization. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70*, ICML'17, page 3530–3539. JMLR.org.

[Wang and Zhang, 2015] Wang, X. and Zhang, J. (2015). Process flexibility: A distribution-free bound on the performance of k-chain. *Operations Research*, 63(3):555–571.

[Watts, 2002] Watts, D. J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9):5766–5771.

[Weiss, 1986] Weiss, G. (1986). Stochastic bounds on distributions of optimal value functions with applications to PERT, network flows and reliability. *Operations Research*, 34(4):595–605.

[Wilder, 2018] Wilder, B. (2018). Risk-sensitive submodular optimization. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

[Yan et al., 2011] Yan, Q., Guo, S., and Yang, D. (2011). Influence Maximizing and Local Influenced Community Detection Based on Multiple Spread Model. In Tang, J., King, I., Chen, L., and Wang, J., editors, *Advanced Data Mining and Applications*, pages 82–95, Berlin, Heidelberg. Springer Berlin Heidelberg.

[Zhang et al., 2014] Zhang, P., Chen, W., Sun, X., Wang, Y., and Zhang, J. (2014). Minimizing seed set selection with probabilistic coverage guarantee in a social network. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, page 1306–1315, New York, NY, USA. Association for Computing Machinery.