To appear in *International Journal of Disaster Risk Reduction*

Cascading Failures in Hierarchical Networks with Unity of Command: An Info-Gap Analysis

Yakov Ben-Haim
Yitzhak Moda'i Chair in Technology and Economics
Technion — Israel Institute of Technology
Haifa, Israel
yakov@technion.ac.il

Contents

Abstract Cascading failures occur in networks of interacting agents in which failure at one node can cause further failures. We define the 'degree of cascading failure' as the fraction of nodes that *could* fail as a result of one single failure. This refers to the *possibility* of failure, and thus involves the *uncertainty* of failures. We emphasize vulnerability to uncertainty, and employ the concept of

\papers\emergency-readiness\workshop2018\cas-fail015.tex. 10.8.2019

robustness as developed in info-gap theory, to study uncertain cascading failures. We study hierarchical networks with unity of command, which means that each node in the hierarchy receives a message from at most one other node. Our concern is in designing the network to adequately manage cascading failures. We explore a situation where the decision maker must choose between design alternatives that entail a dilemma: choose the putatively better but more uncertain option, or choose the putatively worse but more reliable one? The info-gap robustness analysis offers a resolution of this dilemma. This analysis underlies a critique of conventional optimization in which one uses the best data, knowledge and understanding to prioritize the decision alternatives based on predicted outcomes.

Keywords: cascading failures; uncertainty; info-gap; hierarchical networks; unity of command

1 Introduction: Cascading Failures

Networks of interacting nodes occur in a great diversity of situations. For instance, a supply network has nodes of producers and consumers. A producer may face failure if its product is not purchased, and a consumer may face failure if its demand is not satisfied. This is an “individualistic” network in the sense that failure may occur at individual nodes rather than in the system as a whole, though one could consider overall system failure. In contrast, a national defense network has nodes of commanders and operational units and the system as a whole fails if national security is not achieved, even though individual nodes may survive. This is a “dedicated” network, though one could analyze a defense network from an individualistic perspective. In this paper we focus on individualistic networks.

Cascading failures occur in individualistic networks of interacting agents when failure at one node causes further failures. We define the ‘degree of cascading failure’ as the fraction of nodes that *could* fail as a result of one single failure. This refers to the *possibility* of failure, and thus involves the *uncertainty* of failures at individual nodes.

Probabilistic analysis has been fruitfully employed to identify generic characteristics of cascading failures. Buldyrev *et al.* (2010) consider cascading failures of interconnected networks in which nodes are randomly connected. They consider the probability distribution of the number of nodes connected to each node in a network. As that probability distribution gets broader, the vulnerability to cascading failures in the interconnected networks increases. Hines *et al.* (2017) study cascading failures in electric power networks in which failures propagate non-locally: sequential failures can be distant geographically and topologically because of the nature of power flows. They develop a probabilistic Markov model of how cascading failures propagate in such networks.

Rinaldi *et al.* (2001) have demonstrated that cascading failures depend not only on the topology of the network, but also on many other factors such as the social-political environment, the state of repair, the type of failure, whether the interconnections of nodes are physical or logical, loose or tight, and many other factors. Under ordinary situations, for which extensive historical data are available, one can calibrate empirical probabilistic models to account for these factors.

Probabilistic models are less accessible under highly irregular, unprecedented and severe situations. We consider situations of deep uncertainty in which likelihoods cannot be accurately assessed because historical precedents are lacking and the domain of possible disruptive events is poorly understood. Economic upheavals, severe terror attacks, rare combinations of natural catastrophes, disruptive technological innovations, or other major surprises are sources of deep uncertainty for which probability models may be quite uncertain. Our focus in this paper is on the support of design and policy decisions when probability distributions are uncertain or lacking.

The propensity for cascading failures of large degree may tend to increase with the size of the network. However, this may not be true if the network is highly fragmented, or if nodes cooperate or compensate or adapt in real time. Network size, topology and rules of behavior are important factors in characterizing cascading failures. Nonetheless, our emphasis in this paper is not on identifying generic traits of cascading failures, but rather on managing the deep uncertainty that makes these failures pernicious.

We emphasize the distinction between vulnerability to threats, as distinct from vulnerability to uncertainty. A threat is an event or situation that may cause disruption or damage. A threat acts in the environment of the system in question; it is a substantive situation relating to that system. An uncertainty is a lack of knowledge or understanding on the part of a decision maker. Uncertainty is not a substantive threat, but it entails possible threats that are not known or identified. If there were no threats, then uncertainty would be inconsequential, so threats are important. However, design or planning decisions that are based on one's best assessment of the threats — without also managing the deep uncertainty in this assessment — is irresponsible. Under deep uncertainty one's best assessment may well be quite wrong, implying that prediction of outcomes may be quite unreliable. One's best assessment is only the starting point for managing the potential error in the assessment, as expressed by one's uncertainty.

We employ the concept of robustness, as developed in info-gap theory, to study cascading failures under deep uncertainty that cannot be adequately represented probabilistically (the info-gap conception of uncertainty will be elaborated in section 3.1 and subsequently). This method attempts to satisfy an outcome requirement, and to maximize the robustness against uncertainty. This is a procedural optimization rather than a substantive optimization. The outcome is the substantive 'good' that one seeks, for instance, requiring that the degree of cascading failure is low. In the methodology of info-gap robust-satisficing we only attempt to make the outcome good enough, while the robustness is optimized. Vulnerability to deep uncertainty motivates the maximization of robustness against surprise, rather than maximization of the quality of the outcome. The robustness is an aspect of the procedure of reaching a decision, and is not an attribute of the substantive outcome. What constitutes a good enough outcome can, of course, be as demanding or lax as one wants.

Info-gap theory is a method for prioritizing options and making choices and decisions under severe uncertainty (Ben-Haim, 2006, 2010, 2018). The options might be operational alternatives (implement a policy, choose a budget, decide to intervene or not, etc.) or more abstract decisions (choose a model structure, make a forecast, formulate a policy, etc.). Decisions are based on data, scientific theories, empirical relations, knowledge and contextual understanding, all of which we'll refer to as one's *models,* and these models often recognize and quantify uncertainty.

Info-gap theory has been applied to decision problems in many fields, including various areas of engineering (Kanno and Takewaki, 2006; Chinnappen-Rimer and Hancke, 2011; Harp and Vesselinov, 2013), biological conservation (Burgman, 2005), economics (Knoke, 2008; Ben-Haim, 2010), medicine (Ben-Haim *et al.,* 2012), national security (Moffitt, Stranlund, and Field, 2005), public policy (Hall *et al.,* 2012), and more (info-gap.com). Info-gap robust-satisficing has been discussed non-technically elsewhere (Schwartz, Ben-Haim, and Dacso, 2011; Ben-Haim, 2012 a, b, 2018).

We illustrate three properties of info-gap robustness analysis: zeroing, trade off, and preference reversal.

'Zeroing' asserts that predicted outcomes (e.g. degree of cascading failure) have no robustness against uncertainty. Predicted outcomes, based on one's best knowledge (which may be probabilistic), are often used to design or evaluate a network. However, under deep uncertainty the knowledge may err greatly, so predicted outcomes may not be a good basis for decision making, as stated by the zeroing property.

'Trade off' asserts that greater robustness is obtained only in exchange for more modest outcome requirements (e.g. accepting greater degree of cascading failure). Analysis and decision making under deep uncertainty can be assisted by evaluating this trade off, as we will show. Specifically, the choice that one makes between design or policy alternatives may differ from the putatively optimal choice based on the best available knowledge. Zero robustness of the putative outcome, and the trade off between outcome and robustness, can justify the choice of a putatively sub-optimal alternative.

'Preference reversal' is the selection of a sub-optimal alternative over the putatively optimal choice. Preference reversal between design or policy alternatives may occur in situations where one alternative is putatively better but more uncertain than another option. The decision maker faces a dilemma: choose the putatively better but more uncertain option, or choose the putatively worse but more reliable option? The info-gap robustness analysis offers a resolution of this dilemma. This analysis underlies a critique of conventional optimization in which one uses the best data, knowledge and understanding to prioritize the decision alternatives based on predicted outcomes.

Fraid (2016) has studied the application of info-gap decision theory to cascading failures in dynamic networks involving supply and demand requirements at each node. In this paper we use examples of static networks to explore the implications of deep uncertainty, and to demonstrate the info-gap analysis of robustness to support decision making. The first example, in section 3, explores a general serial network with uncertain probability of successful transmission from one node to the next. Section 4 generalizes this to hierarchical networks with unity of command. In all examples we illustrate the properties of zeroing, trade off, and preference reversal, and we explore how the robustness analysis supports design and policy decisions. On a more general level, our analysis is a critique of conventional optimization, as mentioned above and elaborated in the concluding discussion, section 5.

We begin by defining and illustrating the concept of the degree of cascading failure of a network.

2 Degree of Cascading Failure: Definition and Examples

Consider a network with n nodes. The network has **cascading failures of degree** ϕ if ϕ is the largest fraction of nodes in the network that *could* fail as the result of failure of one node of the network.

Example 1 Consider the game of 'telephone' in which n people sit in a row. The first person has "received" a message (perhaps by inspiration) and whispers the message to their neighbor, who whispers the message to the next neighbor, etc. The last person to receive the message announces it out loud. That announcement often differs greatly from the original version, to everyone's delight. This game involves n transmissions of the message (including the final announcement), any one of which could corrupt the message. We will assume that no corruption could correct a previous

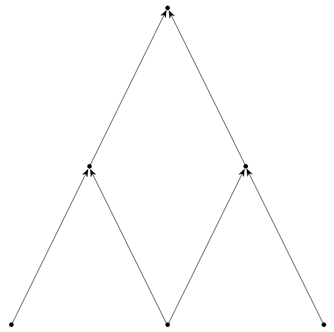

Figure 1: Triangular hierarchical network for example 3.

corruption.[1] A node fails if the message that it transmits is a corruption of the original message. The first transmission could be corrupted, meaning that all n transmissions of nodes $1, \ldots, n$ would be corrupted. Thus this is a network with cascading failure of degree $\phi = 1$. ▮

Example 2 Consider n observers watching the same scene, for instance a volcano. Assume that this scene is in one of two well defined states. For instance, either "eruption" or "no eruption" of the volcano. Observer i fails if it records the scene incorrectly. There is no communication between observers, so failure of any one observer has no effect on any other observer. Hence this network has cascading failure of degree $\phi = 1/n$. ▮

Example 3 *Hierarchical network.* Consider a triangular hierarchical system such as shown in fig. 1. Each node on the bottom row has received a message that it passes to either 1 or 2 nodes in the layer just above. Every receiving node does the same. Each node may corrupt the message it has received. A single node in the bottom row of a 2-row network (e.g. the top two rows in fig. 1) could pass a corrupted message to the top node so the degree of cascading failure of this 3-node network is $\phi = 2/3$. In the 3-row network with 6 nodes, the central node in the bottom row could "infect" three higher nodes, so the degree of cascading failure is $\phi = 4/6 = 2/3$. One can readily show that the generalization of the network in fig. 1, to have $2n$ or $2n + 1$ rows, has degree of cascading failure:

$$\phi_{2n} = \phi_{2n+1} = \frac{n+1}{2n+1}, \quad n = (0,)\, 1,\, 2,\, \ldots \tag{1}$$

where the expression for ϕ_{2n+1} holds also when $n = 0$. ▮

The degree of cascading failure, ϕ, is a deterministic property of the network: the largest possible fraction of nodes that could fail in a cascading failure. However, the fraction of nodes involved in a specific cascading failure can be smaller than ϕ. Let ψ denote the fractional size of a cascading failure in a network with n nodes. ψ can take values from among $0/n$, $1/n$, $2/n$, ..., ϕ.

[1] For example, if the message is a single binary bit, a corruption would change the bit, and a subsequent corruption would change the bit back to its original value, which thus corrects the previous corruption. We exclude this possibility by considering only complex messages.

3 Serial Network with Uncertain Transmissions

3.1 Cascading Failures and Uncertainty

Let us refine our definition of corruption of a message. A message is corrupted if it is significantly different from the original message. For instance, if the original message was "I *really* want to talk to Ted", and the transmission is "I *very much* want to talk to Ted", then we might not consider this to be a corruption; the meaning is really (or very much) the same. However, sequential insignificant alterations of a message may accumulate to a significant alteration.

To formalize this, let μ_0 denote the original message that the first node intended to transmit, and let μ_i denote the message transmitted by node i, for $i = 1, \ldots, n$. Message μ_i is not corrupted if and only if:

$$|\mu_i - \mu_0| \le \delta \tag{2}$$

where $|\cdot|$ is a distance function in a compact metric space of messages and δ is the largest alteration of a message that is not a corruption. For any messages in the space, the distance function $|\cdot|$ is a non-negative real function with 3 properties. Unique zero: $|\zeta - \nu| = 0$ if and only if $\zeta = \nu$. Symmetry: $|\zeta - \nu| = |\nu - \zeta|$. Triangle inequality: $|\zeta - \nu| \le |\zeta - \eta| + |\eta - \nu|$ (see Kolmogorov and Fomin, 1975, p.37).

For clarity, we explain that compactness of a metric space is a generalization of the concept of a space (such as a Euclidean space) which is closed (all its limit points belong to the space) and bounded (all points in the space lie within some finite distance of one another).

Let $\mu = (\mu_1, \ldots, \mu_n)$ denote the vector of messages that are transmitted in the serial network, and let $\psi(\mu)$ denote the fraction of corrupted messages in μ according to the definition of corruption in eq.(2). We can formalize $\psi(\mu)$ as follows. Let $\ell(\mu)$ be the lowest index of a corrupted message in μ, where we define $\ell(\mu) = n + 1$ if no messages in μ are corrupted. Recall that if the ith message is corrupted, then all subsequent messages are also corrupted. Thus, the fraction of failed transmissions for a vector of messages μ is:

$$\psi(\mu) = \frac{n + 1 - \ell(\mu)}{n} \tag{3}$$

The process and extent of corruption of messages at each node is highly uncertain, either a priori or after the fact: we don't know what messages will be, or were, corrupted. The message transmitted by node i, μ_i, may differ by an unknown amount from the message received in that node, μ_{i-1}, for $i = 1, \ldots, n$. We do not know the distance between these messages, $|\mu_i - \mu_{i-1}|$, and it may be small or large. An info-gap model expressing this unbounded uncertainty in the corruption of the messages is:

$$\mathcal{U}(h) = \{\mu : \ |\mu_i - \mu_{i-1}| \le h, \ i = 1, \ldots, n\}, \quad h \ge 0 \tag{4}$$

$\mathcal{U}(h)$ is the set of all message vectors in the space, μ, in which each message differs from its predecessor by no more than h (which is a non-negative real number). (Note that, because the space is compact, it contains its limit points and thus $\mathcal{U}(h)$ contains messages for which the inequality in eq.(4) is an equality.) When h is small, then the set $\mathcal{U}(h)$ contains all messages whose corruption is at most small. As h gets larger, the set contains messages with greater degree of corruption. However, the value of h is unknown, because the degree of corruption is unknown. Hence the info-gap model is not a single set, but rather the unbounded family of nested sets of messages $\mathcal{U}(h)$, for $h \ge 0$. We refer to h as the horizon of uncertainty because the range of uncertain variation of the messages increases as h increases.

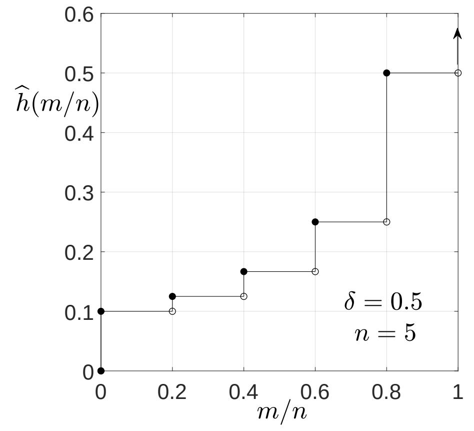

Figure 2: Robustness curve with 5 nodes, eq.(6). Calculated with tele002.m

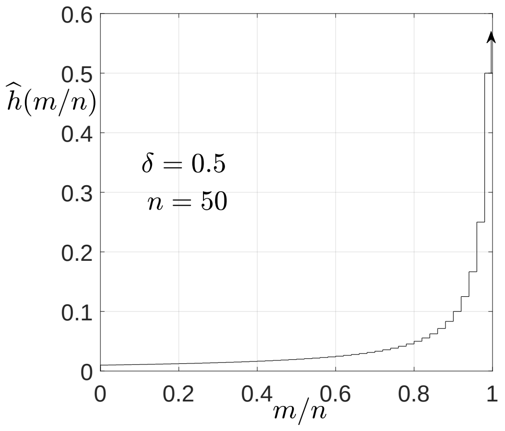

Figure 3: Robustness curve with 50 nodes, eq.(6). Calculated with tele002.m

3.2 Robustness of Degree of Cascading Failures

No nodes will fail in the absence of error because each message is accurately transmitted, in which case $\psi = 0/n$. However, uncertainty is rampant so failure can occur. We would like the fraction of failed nodes to be small, and the largest acceptable fraction is m/n. The "robustness question" is: how much uncertainty can we tolerate and the fraction of failed nodes will be no greater than m/n? More formally, our requirement is that the fraction of failed transmissions, $\psi(\mu)$, not exceed the value m/n:

$$\psi(\mu) \le \frac{m}{n} \tag{5}$$

The robustness for satisfying this requirement is the greatest horizon of uncertainty, h, up to which all message vectors μ in the uncertainty set $\mathcal{U}(h)$ satisfy this requirement:

$$\widehat{h}(\psi \le m/n) = \max\left\{ h : \left(\max_{\mu \in \mathcal{U}(h)} \psi(\mu) \right) \le \frac{m}{n} \right\} \tag{6}$$

It is crucial to understand the relation between the horizon of uncertainty, h, and the robustness, $\widehat{h}$. The horizon of uncertainty is the *unknown* level of variation between successive messages, as stated in the info-gap model of uncertainty in eq.(4); all we know about h is $h \ge 0$. The robustness, $\widehat{h}$, is the greatest horizon of uncertainty, (the greatest value of h), up to which the requirement in eq.(5) is guaranteed to hold. Thus $\widehat{h}$ is a specific value of h. The robustness is *known* (or at least calculable) while the horizon of uncertainty is *not known.* In other words, we *do not know* the level of message corruption (h), but we *do know* how much corruption we can tolerate without violating the performance requirement ($\widehat{h}$). Stated differently, the robustness function expresses the degree of confidence in achieving acceptable outcome: fractional failure no greater than m/n. Alternatively, one can use the robustness function to identify outcomes (fractional failures) that can be confidently expected. The inner maximum in the definition of the robustness function is derived in appendix A, and from that the robustness function is deduced.

Figs. 2 and 3 shows the robustness function of eq.(6) for a telephone game with $n = 5$ and $n = 50$ nodes, respectively, and maximum acceptable alteration of messages of $\delta = 0.5$, as evaluated with eqs.(35) and (36) in appendix A. The positive slope shows the trade off between robustness and the degree of cascading failure: greater robustness is obtained only by allowing greater cascading failure. Comparison of the two figures shows that the robustness decreases as the size of this linear network increases, as stated in eq.(37) and proven at the end of appendix A.

3.3 Uncertain Probability of False Transmission

We have considered uncertainty in the transmission of messages in a serial network, as represented by the non-probabilistic info-gap model in eq.(4), $\mathcal{U}(h)$ for $h \geq 0$. We now consider the probability of accurate transmission, π, but treat the value of this probability as uncertain. We will employ an info-gap model to represent this non-probabilistic uncertainty in the value of the probability, and we will then evaluate the robustness to this uncertainty. That is, we embed an uncertain probabilistic model in a non-probabilistic info-gap model of uncertainty. We now formulate these ideas precisely.

3.3.1 Formulation

Let π denote the probability that any single node accurately transmits the message it received. We will assume that the probability of accurately transmitting the received message is the same for all nodes, and that correct transmission is statistically independent between nodes.

As noted at the end of section 2, the degree of cascading failure, ϕ, is not a random variable; it is a deterministic property of the network: the largest possible fraction of nodes that could fail in a cascading failure. However, the fraction of nodes involved in a specific cascading failure can be smaller than ϕ. Let ψ denote the fractional size of a cascading failure in a serial network with n nodes. In this formulation of the example, ψ is a random variable with values $0/n$, $1/n$, $2/n$, ..., ϕ. As shown in appendix B, the probability of cascading failures larger than i/n is:

$$P\left(\psi > \frac{i}{n}\right) = 1 - \pi^{n-i}, \quad i = 0,\, 1,\, \ldots,\, n \tag{7}$$

At fixed network size, n, and fixed probability π, the probability of cascading failures larger than i/n decreases as i/n increases.

3.3.2 Cascading Failures and Robustness

We suppose that we have an estimate, $\widetilde{\pi}$, of the probability of accurate transmission from one node to the next in the serial network, and an estimated error of this estimate, a positive number s. Roughly speaking, π is estimated as $\widetilde{\pi} \pm s$, where we recognize that the actual error may exceed s. More precisely, the fractional error of π with respect to $\widetilde{\pi}$ is $|\pi - \widetilde{\pi}|/s$, whose magnitude is unknown. Thus the info-gap model of uncertainty is:

$$\mathcal{U}(h) = \left\{ \pi : \ \pi \in [0,1], \ \ \left| \frac{\pi - \widetilde{\pi}}{s} \right| \leq h \right\}, \quad h \geq 0 \tag{8}$$

Like all info-gap models of uncertainty, this is an unbounded family of nested sets.[2] When $h = 0$ the set contracts to the estimated value: $\mathcal{U}(0) = \{\widetilde{\pi}\}$. As h increases the sets become more inclusive. Thus h is called the horizon of uncertainty, and its value is unknown; there is no known worst case.

[2] The family of sets is unbounded in the space of possible probability values, $[0,\ 1]$.

We would like to design or manage the network so that the probability of large cascades is small. More precisely, we require that the probability of cascading failures with degree larger than i/n have probability no greater than a critical value $P_{\rm c}$. That is, we require:

$$P\left(\psi > \frac{i}{n}\right) \le P_{\rm c} \tag{9}$$

For any choice of this critical probability, $P_{\rm c}$, we would like to know if this requirement is feasible in light of the deep uncertainty about the probability of accurate transmission. Alternatively, we would like to know what $P_{\rm c}$ values are feasible. The answers to these questions are provided by the robustness function.

The robustness is the greatest horizon of uncertainty, h, in the info-gap model of eq.(8), up to which the requirement in eq.(9) is guaranteed. Formally, the robustness to uncertainty in π is defined as:

$$\widehat{h}(P_{\rm c}, i) = \max\left\{h :\ \left(\max_{\pi\in\mathcal{U}(h)} P\left(\psi > \frac{i}{n}\right)\right) \le P_{\rm c}\right\} \tag{10}$$

As shown in appendix C, the robustness function is:

$$\widehat{h}(P_{\rm c}, i) = \frac{1}{s}\left(\widetilde{\pi} - (1-P_{\rm c})^{1/(n-i)}\right) \tag{11}$$

or zero if this is negative.

The best-estimated prediction of the probability that the degree of cascading failure exceeds i/n is, from eq.(7), $P(\psi > i/n) = 1 - \widetilde{\pi}^{n-i}$. From eq.(11) we see that the robustness reaches the horizontal axis precisely when the critical probability, $P_{\rm c}$, equals this predicted value. That is, if $1-\widetilde{\pi}^{n-i}$ is adopted as the performance requirement, $P_{\rm c}$ in eq.(9) — and some would consider this sensible because it is the best estimate of the probability that the degree exceeds i/n — then the robustness for satisfying this requirement is zero. This is the **zeroing property:** best-model predicted outcomes have no robustness against uncertainty underlying the predictions.

Eq.(11) also shows the irrevocable **trade off** between robustness and performance: as the performance requirement becomes more demanding (as $P_{\rm c}$ is decreased) the robustness becomes smaller ($\widehat{h}(P_{\rm c}, i)$ goes down). This is a trade **off** because we would like $P_{\rm c}$ to be small and $\widehat{h}$ to be large.

The robustness function of eq.(11) is shown in fig. 4 for $n = 100$, $\widetilde{\pi} = 0.999$, $s = 0.05$, and three different values of i: 10, 20 and 30. As we expect from eq.(11) and the zeroing property, the robustness vanishes when $P_{\rm c} = 1 - \widetilde{\pi}^{n-i}$, which equals 0.087, 0.078 and 0.068 for $i =$ 10, 20 and 30, respectively. Furthermore, the robustness increases as $P_{\rm c}$ increases as seen from the positive slopes of the curves, expressing the trade off property. We also see that the robustness increases as i increases. This is because the probability of cascading failures larger than i/n decreases as i/n increases, as seen from eq.(7).

We now consider two different implementations of the serial network, both with $\widetilde{\pi} = 0.999$. In one we have better knowledge so the error estimate, s, is smaller, but the number of nodes is greater, so n is larger. In both cases $i/n = 0.2$. Specifically, in the first implementation, $n_1 = 100$, $s_1 = 0.05$ and $i = 20$ as before (middle curve of fig. 4), and the estimated probability that the degree of cascading failure exceeds i/n is $P(\psi > i/n) = 0.078$. In the second implementation $n_1 = 60$, $s_1 = 0.15$ and $i = 12$ and the estimated probability is $P(\psi > i/n) = 0.048$. That is, the poorer information in the second case ($s_2 > s_1$) is compensated by the smaller number of players ($n_2 < n_1$).

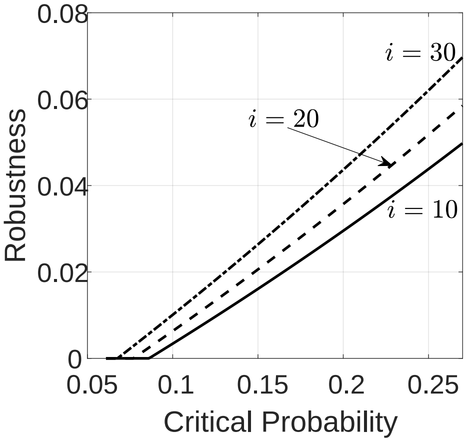

Figure 4: Robustness curves for 3 values of i. $\widetilde{\pi} = 0.999$, $n = 100$, $s = 0.05$. Calculated with cf003.m

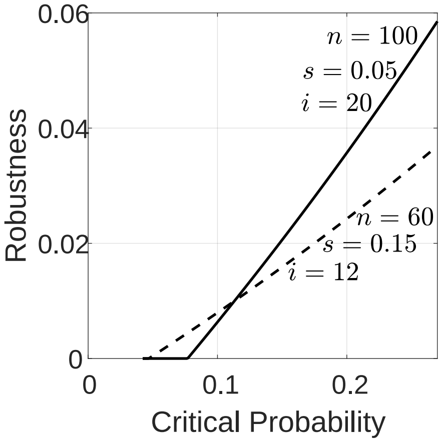

Figure 5: Robustness curves for 2 values of n, s and i but $i/n = 0.2$ in both cases. $\widetilde{\pi} = 0.999$. Calculated with cf004.m

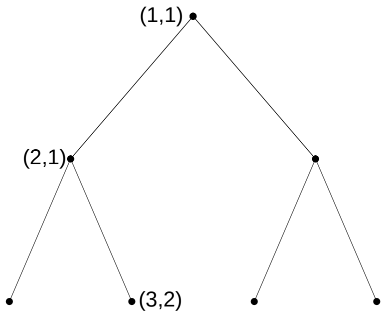

Figure 6: Hierarchical network with 3 rows and 2 branches at each node.

Plots of robustness curves for these two implementations are shown in fig. 5. The robustness curves cross one another, implying the potential for a **reversal of preference** between these two options. The second option (dashed curve) is putatively better: its predicted probability that ψ exceeds i/n is lower because of the smaller size of the network, as we expect from eq.(7). That is, the best assessment is that the second option is less vulnerable to failure. However, from the zeroing property we know that predicted outcomes have no robustness to uncertainty. From the trade off property we know that only poorer outcomes (larger P_c) have greater robustness to uncertainty. However, the trade off is more severe for the putatively better case (dashed curve) because the uncertainty, as expressed by s_2, is greater. That is, the second option is more vulnerable to uncertainty. As a consequence, the robustness curves cross one another at $P_\mathrm{c} = 0.11$. If one needs a critical probability less than 0.11, then option 2 is more robust, though its robustness is small. If greater P_c would be acceptable then option 1 is more robust.

4 Hierarchical Networks with Unity of Command

We now generalize the serial networks studied in section 3 by considering hierarchical networks with unity of command.

4.1 Introduction

We study hierarchical networks as in the simple example of fig. 6, though generalizing to arbitrary number of rows and branching ratio. The single node in the highest row receives a message from an external source and passes it along to all the nodes in the 2nd row. Each node passes the message it receives to nodes in the next lower row until the message reaches the bottom of the network. Throughout the network, each node receives a message from exactly one other node, which expresses the unity of command in the network.

Unity of command is prevalent in militaries, occurring widely throughout the history of war, but also in commercial and industrial organizations (Takahashi, 1986) as elaborated by Henri Fayol (see Godwin *et al.,* 2017). The prominence of unity of command is strongly supported by battlefield experience. For instance, Murray and Hsieh (2016, pp.119–120) discuss situations in the American Civil War in which adverse outcomes resulted from absence of unity of command. Unity of command has also been adopted in disaster management to manage or ameliorate volatile civilian catastrophes (Pal, Ghosh and Ghosh, 2017). Unity of command comprising both civilian and military elements is important for successful nation-building against an insurgency when the insurgents target military and non-military units indiscriminately, as done by the Taliban in Afghanistan. Without civilian-military unity of command the military mission may succeed while the governance and humanitarian missions may fail (Welle, 2010).

The generalization of section 3 is not only in the topology — hierarchical and branching rather than serial. The concept of a message now refers to actions as well as statements, either of which induce a subsequent action or statement. For instance, Diener (2018) discusses promotion in a military hierarchy. When a vacancy occurs at one rank, a person in the next lower echelon is promoted, which creates a new vacancy, and subsequent promotion down to the lowest level. Each vacancy creates a generic message that propagates down the hierarchy. Furthermore, the message is intended to evolve in a pre-planned manner. For instance, a command at one node is intended to elicit a specific action at the next node which in turn should elicit a specific response at the next node, and so on. Actions and statements may deviate from the plan, and central to our analysis is the uncertainty that accompanies this process.

Consider two examples. The top node receives a message from an external source. For a company president this might be the stock holders demanding greater profits. For a military general this might be the civilian commander in chief demanding successful conclusion of the war. The general then gives orders (messages) to the special operations division, to the tank division, and so on. Special Ops then gives orders to intelligence units, activates commando units etc., while the tank commander activates field commanders, logistic units, and so on. Different orders are given at each level and on each branch of every node, but all are intended to conform to an overall plan.

As in section 3, deviation of a message from the plan at any particular node occurs if either that node significantly alters the message (responding differently than intended), or faithfully transmits a corrupted message (responding correctly to a faulty message). As before, we assume that corruption of a corrupted message cannot correct that message. The serial network of section 3 now is one path within the hierarchy, albeit with the generalized conception of a message. We formalize this process in section 4.2.

Our focus on networks with unity of command is motivated in part by their continuing relevance in many situations. But also by the fact that even this simple traditional structure is fraught with challenges in assessing vulnerability to uncertainty.

4.2 Formulation

4.2.1 Topology

The network has R rows, which can be thought of as echelons in an organizational hierarchy such as an army. The first row has only one node, and each node (except in the bottom row) branches into B nodes in the next lower row; we refer to B as the branching ratio. (For simplicity we do not consider varying branching ratio among the nodes, though the analysis can be extended to include this.) Fig. 6 shows an example with $R = 3$ and $B = 2$.

The number of nodes in row i is $N_i = B^{i-1}$ for $i = 1, \ldots, R$. If the branching ratio, B, is greater than 1, then the total number of nodes in the network is:

$$N = \sum_{i=1}^{R} B^{i-1} = \frac{B^R - 1}{B - 1} \tag{12}$$

If $B = 1$ then the number of nodes equals the number of rows: $N = R$.

4.2.2 Messages and Corruption

Let μ_{ij} denote the message — action or statement — transmitted by node j in row i, for $i = 1, \ldots, R$ and $j = 1, \ldots, N_i$. As explained in section 4.1, this message may take different forms in the context of the different branches to which it is transmitted. Each receiving node is intended to understand the message in the context of its location in the network or its role in the overall plan. Each node in the bottom row also issues a message that is an action or statement. The external message received by the highest node is μ_0.

In analogy to eq.(2), message μ_{ij} is not corrupted if and only if:

$$|\mu_{ij} - \mu_0| \leq \delta \tag{13}$$

The distance function, $|\cdot|$, assesses the degree of deviation of the message (action or statement) transmitted by node (i, j) from the intended message according to the overall plan inherent in the initiating message, μ_0. As in section 3, we never actually have to specify this distance function, other than to assert its obeyance of the triangle inequality and other basic properties of a compact metric space.

Because of the unity of command, there is one and only one path — chain of command — from the single node in the top row, to each node in the bottom row. Likewise, each path contains exactly one node in each row because of the unity of command. There are R rows and N_R nodes in the bottom row, so there are exactly N_R paths in the network. Let C_j denote the ordered sequence of node indices along the path to node j in the bottom row, for $j = 1, \ldots, N_R$. For instance, C_2 in fig. 6 is:

$$C_2 = ((1,1),\ (2,1)\ (3,2)) \tag{14}$$

More formally, the elements of C_j are denoted:

$$C_j = (c_{1,j},\, c_{2,j},\, \ldots,\, c_{R,j}), \quad j = 1, \ldots, N_R \tag{15}$$

Note that $c_{1,j} = (1,1)$ and $c_{R,j} = (R, j)$ for all paths. For C_2 in fig. 6 the elements are $c_{1,2} = (1,1)$, $c_{2,2} = (2,1)$ and $c_{3,2} = (3,2)$.

The state of the network is specified by the set S containing the message transmitted by each node. Thus S contains N messages because there are N nodes. For any given set S, Let S_j denote the subset of S containing the messages along path C_j. There is one such subset for each of the N_R nodes in the bottom row. Formally:

$$S_j = \left(\mu_{c_{1j}},\, \mu_{c_{2j}},\, \ldots,\, \mu_{c_{Rj}}\right), \quad j = 1, \ldots, N_R \tag{16}$$

For instance, for C_2 in fig. 6, and referring to eq.(14), we have:

$$S_2 = (\mu_{1,1},\ \mu_{2,1}\ \mu_{3,2}) \tag{17}$$

Let $\ell(S_j)$ denote the highest row with a corrupted message in the set of messages S_j. Define $\ell(S_j) = R+1$ if no message in S_j is corrupted. Let $\psi(S_j)$ denote the fraction of corrupted messages in S_j. Reasoning as in eq.(3), we find:

$$\psi(S_j) = \frac{R + 1 - \ell(S_j)}{R} \tag{18}$$

for each $j = 1, \ldots, N_R$.

4.3 Network Robustness and First Generic Result

In section 4.1 we explained that the concept of a message includes actions and statements, and that the message is intended to change along each path according to an overall plan. Furthermore, we need to account for different difficulties, uncertainties, and propensities for deviation of the messages along the different paths. For instance, the challenges, potential disruptions, and uncertainties in military special operations are different from those of military logistics. Interfering factors include weather, civilians, the enemy, deficient morale, and so on, and these act differently in different paths of the network.

As in section 3, the messages are uncertain, though we must extend the info-gap model of uncertainty in eq.(4) to account for different challenges at different nodes along different paths. We do this as follows.

Consider a message, $\mu_{c_{ij}}$, in path C_j. Its deviation from the prior message, $\mu_{c_{i-1,j}}$, is $|\mu_{c_{ij}} - \mu_{c_{i-1,j}}|$. (When $i = 1$ we define $\mu_{c_{0j}} = \mu_0$ which is the external initiating message.) In line with the explanation of eq.(13), this distance is the degree of deviation of message $\mu_{c_{i,j}}$ from the intended message at that node, according to the overall plan inherent in message $\mu_{c_{i-1,j}}$. This distance will be zero in the absence of uncertainty. In many situations we are able to identify messages that are more vulnerable to uncertainty, and others that are less so. Moreover, we can make judgments for each message about its relative vulnerability, as expressed by a typical degree of deviation at node c_{ij}, denoted by a positive “uncertainty weight” $w_{c_{ij}}$. At any horizon of uncertainty, h, the range of uncertain deviation at node c_{ij} is:

$$\frac{|\mu_{c_{ij}} - \mu_{c_{i-1,j}}|}{w_{c_{ij}}} \le h \tag{19}$$

A large value of $w_{c_{ij}}$ implies that $\mu_{c_{ij}}$ tends to vary greatly even at a small horizon of uncertainty h; a small uncertainty weight implies small variation.

We are considering situations of deep uncertainty, so the horizon of uncertainty, h, is unknown and unbounded. Hence the uncertainty weights, $w_{c_{ij}}$, do not establish maximal deviations, but only relative distances of deviation at any value of h. We now arrive at the info-gap model of uncertainty of the messages in path C_j, in analogy to eq.(4):

$$\mathcal{U}_j(h) = \left\{ S_j : \ \frac{|\mu_{c_{ij}} - \mu_{c_{i-1,j}}|}{w_{c_{ij}}} \le h, \ i = 1, \ldots, R \right\}, \quad h \ge 0 \tag{20}$$

for each $j = 1, \ldots, N_R$.

We first consider the robustness of a single path, C_j. As in eq.(5), we require that no more than a fraction m/R of the messages in path C_j are corrupted, where m is an integer from 0 to R. That is, for each path C_j, we require:

$$\psi(S_j) \le \frac{m}{R} \tag{21}$$

In analogy to eq.(6), the robustness of path C_j is:

$$\widehat{h}_j(\psi \le m/R, \delta) = \max \left\{ h : \ \left(\max_{S_j \in \mathcal{U}_j(h)} \psi(S_j) \right) \le \frac{m}{R} \right\} \tag{22}$$

for each $j = 1, \ldots, N_R$.

The overall network robustness could be defined in different ways. The total number of nodes in the network is N, so one definition of overall robustness would be the greatest horizon of uncertainty up to which the fraction of all failed nodes is no greater than m/N, where m is an integer from 0 to N. Some paths may have low fractional failure and others high fractional failure, but network failure is defined to occur only if the total fractional failure exceeds m/N.

Alternatively, each path can be thought of as a chain of command to a final action node in the bottom row. In this case, one would define the overall network robustness as the greatest horizon of uncertainty up to which no path has fractional failure greater than m/R, where m is an integer from 0 to R. We adopt this concept of overall robustness, whose definition is:

$$\widehat{h}(\psi \le m/R, \delta) = \max \left\{ h : \left(\max_{1 \le j \le N_R} \; \max_{S_j \in \mathcal{U}_j(h)} \psi(S_j) \right) \le \frac{m}{R} \right\} \tag{23}$$

This overall network robustness equals the robustness of the most vulnerable path because all paths must have fractional failure no greater than m/R. Hence:

$$\widehat{h}(\psi \le m/R, \delta) = \min_{1 \le j \le N_R} \widehat{h}_j(\psi \le m/R, \delta) \tag{24}$$

The robustness function in eq.(22) is a step-wise discontinuous function, derived in appendix D, resulting in eq.(56) when $\delta = 0$, and eq.(66) when $\delta > 0$. In the latter case an explicit expression for the robustness is:

$$\widehat{h}_j(\psi \le m/R, \delta) = \frac{\delta}{\displaystyle\sum_{i=1}^{R-m} w_{c_{i,j}}}, \quad m = 0,\, 1,\, \ldots,\, R \tag{25}$$

We define the sum in the denominator to be zero when $m = R$, resulting in infinite robustness if we are willing to accept any number of message distortions along the path.

The following proposition gives us our first generic insight into the info-gap robustness function and its implications for designing the hierarchical topology. The proof appears in appendix E.

Proposition 1 ***Given:***

1. *A hierarchical network with unity of command containing R rows and branching ratio B.*
2. *The threshold for corruption of a message is positive: $\delta > 0$ in eq.(13).*
3. *The info-gap model of eq.(20) where all uncertainty weights are positive: $w_{c_{i,j}} > 0$.*
4. *The definition of the robustness of path C_j, $\widehat{h}_j(\psi \le m/R, \delta)$ in eq.(22).*
5. *The definition of the overall network robustness, $\widehat{h}(\psi \le m/R, \delta)$ in eq.(23).*

Then:

1. *The overall network robustness, $\widehat{h}(\psi \le m/R, \delta)$, equals the robustness of the path whose average uncertainty weight of the first $R - m$ nodes is maximal.*
2. *If $R > 1$ then this is not necessarily the path containing the node with maximal uncertainty weight from among all nodes in the network.*

This proposition provides insight into the perspective provided by the analysis of robustness, and how that perspective may alter one's choice of the network topology.

Cascading failures occur when one node fails to adequately transmit its message, causing subsequent nodes to fail as well. Vulnerable nodes — whose estimated uncertainties are large as expressed by their uncertainty weights w_{ij} — are foremost candidates for initiating a cascading failure. Paths containing these especially vulnerable nodes might reasonably draw particular attention, leading perhaps to design alterations or resource allocations to modify those vulnerable nodes.

However, the first assertion of proposition 1 states that the robustness depends on the path-average uncertainty weight, not the network-maximal uncertainty weight. The second assertion states that the least robust path — which determines the network robustness — need not include the most vulnerable node. Consequently, the robustness analysis may lead to design alterations or resource allocations that differ from choices based on identifying the most vulnerable nodes. Choices based on robustness will enhance the immunity to deep uncertainty, while choices based on attention solely to the most vulnerable nodes may not.

4.4 Example: Identical Nodes

As a simple example we consider the case of identical nodes. The uncertainty weights $w_{c_{ij}}$ in the info-gap model of eq.(20) are all the same, and we can assign each of them the value of 1. All the paths are the same, and identical to the single-path serial example in section 3 when we replace n with R: The info-gap model of eq.(4) is the same as eq.(20), and the robustness in eq.(6) is the same as eq.(22). Because the paths are identical, eq.(24) implies that the network robustness in eq.(23) is the same as the serial robustness, eq.(6).

From eq.(35) in appendix A, and in light of the equivalence of the robustness functions that we just explained, we see that the network robustness for no deviant messages is:

$$\widehat{h}(\psi = 0, \delta) = \frac{\delta}{R} \tag{26}$$

We can combine this relation with the expression for the total number of nodes in the network, eq.(12), to explore the trade off between network topology, R and B, and network robustness, $\widehat{h}(\psi = 0, \delta)$. Table 1 shows results.

R	$\widehat{h}(\psi = 0, \delta)$	$N(R, B = 2)$	$N(R, B = 3)$	$N(R, B = 4)$	$N(R, B = 5)$
2	0.50	3	4	5	6
3	0.34	7	*13*	21	**31**
4	0.25	*15*	**40**	85	*156*
5	0.20	**31**	*121*	341	781
6	0.17	63	364	1365	3906
7	0.15	*127*	1093	5461	19531
8	0.13	255	3280	21845	97656
9	0.12	511	9841	87381	488281
10	0.10	1023	29524	349525	2441406

Table 1: Number of nodes, N, vs. R and B, eq.(12). Network robustness for $\delta = 1$.

Each row of table 1 corresponds to a value of the number of rows in the hierarchical network, R in column 1. The 2nd column is the robustness for no deviant messages, with $\delta = 1$, which depends on R as seen in eq.(26). Columns 3 to 6 show the total number of nodes, N in eq.(12), for the corresponding number of rows, R, and branching ratio, B.

R, B and N are all necessarily integers, yet one can find combinations of R and B for which N is roughly constant. For instance, the 3 bold face numbers in table 1 represent hierarchies with 31, 40, and 31 total nodes, where the topologies are $(R, B) = (5, 2)$, $(4, 3)$ and $(3, 5)$, respectively. Not surprisingly, the total number of nodes is roughly constant as the number of rows decreases and the number of branches increases. The same trend is seen in the 3 N-values in boxes, and the 2 underlined values.

What is particularly significant in table 1 is the relation between topology, R and B, and robustness, $\widehat{h}(\psi = 0, \delta)$. Keeping the total number of nodes constant does *not* by any means keep the robustness constant. The robustness at $(R, B) = (5, 2)$ is $\widehat{h} = 0.20$, while the robustness at $(R, B) = (3, 5)$ is $\widehat{h} = 0.34$, though $N = 31$ in both cases. Similar increase of robustness is seen in the other examples as well.

At a purely technical level this is explained by two observations. First, the robustness is independent of the network branching because it is purely a property of each serial path. Second, the robustness increases as the serial path becomes shorter. Thus, to keep N roughly constant one moves up in the table (which increases the robustness) and moves right (which has no effect on the robustness).

The strong relation between topology and robustness is significant because robustness should be a significant factor in choosing a topology when the network faces deep uncertainty, such as we are considering. Topology itself may be functionally significant, and the number of nodes reflects this functionality in part. But functionality also depends on reliable transmission of messages. Hence the robustness is important in assuring reliable functionality, and may tip the balance between alternative topologies.

For example, Salmerón and Appleget (2014) discuss the choice between "keeping a larger number of BCTs [brigade combat teams] that have predominantly two maneuver battalions, and creating a smaller number of BCTs that all have three maneuver battalions." This choice of hierarchical topology can, and should, be supplemented with the analysis of robustness.

4.5 Second Generic Result

In this section we present and discuss a proposition that generalizes the conclusions discussed in section 4.4 that was limited to the special case of identical nodes. We now consider a hierarchical network with unity of command in which each node has its own uncertainty weight as in the info-gap model of eq.(20).

The main conclusion from the example in section 4.4, seen in table 1, was that the robustness can change substantially when the number of rows, R, and branches per node, B, are varied to keep the total number of nodes at least roughly constant. This conclusion holds also for the general case, as we now explain. This is important because it shows that the choice of hierarchical topology can, and should, be supplemented with the analysis of robustness to uncertainty.

The three assertions in the following proposition underlie our discussion. The proof of this proposition appears in appendix F.

Proposition 2 ***Given:***

1. *A hierarchical network with unity of command containing R rows and branching ratio B.*
2. *The threshold for corruption of a message is positive: $\delta > 0$ in eq.(13).*
3. *The info-gap model of eq.(20) where all uncertainty weights are positive: $w_{c_{i,j}} > 0$.*
4. *The definition of the robustness of path C_j, $\widehat{h}_j(\psi \le m/R, \delta)$ in eq.(22), for $m \le R$.*

Then:

1. *Assertion 1. Each path robustness strictly increases as the number of rows decreases:*

$$\widehat{h}_j(\psi \le m/R, \delta) > \widehat{h}_j(\psi \le m/R', \delta) \quad \textit{for} \quad R < R' \tag{27}$$

2. *Assertion 2. The robustness of any specific path in the given network is independent of the branching ratio.*
3. *Assertion 3. Each path robustness strictly increases as the greatest acceptable number of corrupted messages, m/R, increases:*

$$\widehat{h}_j(\psi \leq m/R, \delta) > \widehat{h}_j(\psi \leq n/R, \delta) \quad \textit{for} \quad n < m \leq R \tag{28}$$

The first assertion in proposition 2 states that short paths are more robust than long paths. Eq.(25) shows that the robustness of a path can also be increased by reducing the uncertainty weights, $w_{c_{i,j}}$, through improving the understanding of the circumstances in which that path operates. However, achieving this enhanced understanding may be costly or infeasible. The first assertion in proposition 2 provides an alternative: robustness of a path is increased by reducing its length. As a military example, special operations units should have short command chains because of great ambient uncertainty, while logistics units can tolerate longer command chains due to lower uncertainty. Beyond this basic intuition, the info-gap analysis provides quantitative assessment of robustness, eq.(25), and thus quantitative evaluation of the utility of specific organizational changes.

The second assertion in proposition 2 states that the path robustness does not depend on the number of branches at each node. This is a distinctive feature of networks with unity of command.

The first two assertions together demonstrate that the robustness will increase as R is decreased and B is increased while keeping the total number of nodes (roughly) constant. The discussion of table 1, showing increasing robustness along diagonals in the table, holds also for general hierarchical networks with unity of command. The total number of nodes often results from resource constraints. The effectiveness of those nodes depends in part on the reliable transmission of the generic message at each node. Hence the robustness should inform the allocation of resources in designing the topology of the network, R and B.

The third assertion in proposition 2 states that, at fixed path length, allowing more corrupted messages is strictly more robust than requiring fewer corruptions. Fewer corrupted messages is better than more corruptions, and robustness can be increased by topological changes or by improving one's information. However, for a given topology and state of knowledge, the third assertion offers the commander or manager the basis for deciding if the requirement must be relaxed (because the robustness is currently quite small), or whether it can be made stricter (because the robustness is currently very large). The robustness function, eq.(25), provides quantitative assessment in support of this judgment.

For example, suppose that the commander initially requires that the degree of cascading failure not exceed a specified value, say 3/10 (maximum of 3 corruptions in a path with 10 nodes). Suppose further that the robustness for achieving this is enormous. The third assertion clearly suggests that the requirement can be made stricter, perhaps 1 or 0 corruptions in the path. Alternatively, suppose the commander initially requires no corruptions, but the robustness for this requirement is tiny or even zero. If topological change or information acquisition are not feasible, then the third assertion suggests exploring the extent to which the requirement must be relaxed in order to attain good robustness.

5 Discussion

We have used the term “models” to refer to one’s data, knowledge and understanding. The conventional optimization of a decision begins by using the best models that one has to predict the outcomes of the decision alternatives. One then chooses the option whose predicted outcome is best. This “best-model optimization” makes sense when one’s models are fairly good. However, when the models are subject to deep uncertainty, then the zeroing property of the info-gap robustness function demonstrates that the predictions have no immunity against this uncertainty: their robustness is precisely zero. This means that the prioritization of the decision alternatives, based on their best-model predicted outcomes, is unreliable.

Furthermore, the trade off property of the robustness function implies that outcomes less desirable than the predicted outcome can have positive robustness against uncertainty. Decision analysis must, therefore, be based on exploring the robustness curve, rather than just its end-point at the predicted outcome. The robustness curves of different decision alternatives may cross one another. When this happens, the analyst may prefer one option over one range of required outcomes, while preferring the other option over a different range of outcome requirements. In other words, the analyst may encounter a reversal of preference between one option (e.g. the putative optimum) and a different option (which is putatively sub-optimal). We have seen this crossing of robustness curves and consequent potential for preference reversal in the examples.

The methodology that is developed in this paper is to satisfy a performance requirement on the outcome — rather than trying to optimize the outcome — and to maximize the robustness against uncertainty. This is a procedural optimization rather than a substantive optimization. The outcome is the substantive “good” that one seeks, e.g. small probability of large-degree cascading failures. In the robust-satisficing approach we only attempt to make the substantive outcome good enough. In contrast to the substantive outcome, the robustness is an aspect of the procedure of reaching a decision, and the robustness is optimized, not the outcome. Once the decision is made and implemented, the robustness is of no substantive consequence. Optimizing the robustness and satisficing the substantive outcome is justified when uncertainty is the main source of vulnerability.

6 References

1. Ben-Haim, Yakov. (2006) *Info-Gap Decision Theory: Decisions Under Severe Uncertainty,* 2nd edition, Academic Press, London.

2. Ben-Haim, Yakov. (2010). *Info-Gap Economics: An Operational Introduction,* Palgrave-Macmillan, London.

3. Ben-Haim, Yakov. (2018). *The Dilemmas of* Wonderland: *Decisions in the Age of Innovation,* Oxford University Press.

4. Ben-Haim, Yakov, Zetola, Nicola M. & Dacso, Clifford. (2012). Info-gap management of public health policy for TB with HIV-prevalence, *BMC Public Health,* 12, 1091. DOI: 10.1186/1471-2458-12-1091.

5. Buldyrev, Sergey V., Roni Parshani, Gerald Paul, H. Eugene Stanley and Shlomo Havlin, 2010, Catastrophic cascade of failures in interdependent networks, *Nature,* 464: 1025–1028.

6. Burgman, Mark. (2005). *Risks and Decisions for Conservation and Environmental Management,* Cambridge University Press, Cambridge.

7. Chinnappen-Rimer, S. & Hancke, G.P. (2011). Actor coordination using info-gap decision theory in wireless sensor and actor networks, *Intl. J. of Sensor Networks,* 10(4), 177–191.

8. Diener, Ross, 2018, A solvable model of hierarchical workforces employed by the Canadian Armed Forces, *Military Operations Research,* vol.23 no.3, pp.47–57.

9. Fraid, Oded (2016). Info-gap analysis of cascading failures in networks, MSc thesis, Technion—Israel Institute of Technology.

10. Godwin, Achinivu, Okwu E. Handsome, Wey A. Ayomide, Akpan E. Enobong, Fasan O. Johnson, 2017, Application of the Henri Fayol principles of management in startup organizations, *IOSR Journal of Business and Management,* vol 19, issue 10, pp.78–85.

11. Hall, Jim W., Lempert, Robert J., Keller, Klaus, Hackbarth, Andrew, Mijere, Christophe & McInerney, David J. (2012). Robust climate policies under uncertainty: A comparison of robust decision making and info-gap methods, *Risk Analysis,* 32 (10), 1657–1672.

12. Harp, Dylan R. & Vesselinov, Velimir V. (2013). Contaminant remediation decision analysis using information gap theory, *Stochastic Environmental Research and Risk Assessment,* 27(1), 159–168.

13. Hines, Paul D. H., Ian Dobson and Pooya Rezaei, 2017, Cascading power outages propagate locally in an influence graph that is not the actual grid topology, *IEEE Transactions on Power Systems,* 32(2): 958–967.

14. Kanno, Y. & Takewaki, I. (2006). Robustness analysis of trusses with separable load and structural uncertainties, *Intl. J. of Solids and Structures,* 43(9), 2646–2669.

15. Knoke, Thomas. (2008). Mixed forests and finance — Methodological approaches, *Ecological Economics,* 65(3), 590–601.

16. Kolmogorov, A.N. and S.V. Fomin, 1975, *Introductory Real Analysis*, Dover Publications, New York.

17. Moffitt, L. Joe, Stranlund, John K. & Field, Barry C. (2005). Inspections to avert terrorism: Robustness under severe uncertainty, *J. Homeland Security and Emergency Management,* vol. 2, no. 3.

    http://www.bepress.com/jhsem/vol2/iss3/3.

18. Murray, Williamson and Wayne Wei-siang Hsieh, 2016, *A Savage War: A Military History of the Civil War*, Princeton University Press.

19. Pal, Indrajit , Tuhin Ghosh and Chandan Ghosh, 2017, Institutional framework and administrative systems for effective disaster risk governance—Perspectives of 2013 Cyclone Phailin in India, *Intl J Disaster Risk Reduction,* 21: 350–359.

20. Rinaldi, Steven M., James P. Peerenboom, and Terrence K. Kelly, 2001, Identifying, understanding, and analyzing critical infrastructure interdependencies, *IEEE Control Systems Magazine,* 21: 11–25.

21. Salmerón, Javier and Jeff Appleget, 2014, Reshaping the US Army: Brigade Combat Team optimization, *Military Operations Research,* vol.19(3): 51–65.

22. Schwartz, Barry, Ben-Haim, Yakov & Dacso, Cliff. (2011). What Makes a Good Decision? Robust Satisficing as a Normative Standard of Rational Behaviour, *The J. for the Theory of Social Behaviour,* 41(2): 209–227.

23. Takahashi, Nobuo, 1986, On the principle of unity of command: Application of a model and empirical research, *Behavioral Science,* vol. 31, pp.42–51.

24. Welle, Joshua W., 2010, Civil-military integration in Afghanistan: Creating unity of command, *Joint Force Quarterly,* U.S. National Defense University, issue 56, 1st quarter, 2010, pp.54–59.

A Derivation of the Inner Maximum of the Robustness Function in Eq.(6), Section 3.2

Let $M(h)$ denote the inner maximum in the definition of the robustness function, eq.(6): $M(h) = \max\limits_{\mu \in \mathcal{U}(h)} \psi(\mu)$. This is the inner maximum in the definition of $\widehat{h}(\psi \le m/n, \delta)$ at fixed n and δ. That is, a plot of h vs. $M(h)$ is identical to a plot of $\widehat{h}(\psi \le m/n, \delta)$ vs. m/n.[3] We now derive an expression for $M(h)$.

The original message, μ_0, is not uncertain (it is known; given).

The uncertainty set $\mathcal{U}(h)$ contains all messages μ_1 such that $|\mu_1 - \mu_0| \le h$.

The uncertainty set $\mathcal{U}(h)$ contains all messages μ_2 such that $|\mu_2 - \mu_1| \le h$.

By the triangle inequality, $\mathcal{U}(h)$ contains all messages μ_2 such that $|\mu_2 - \mu_0| \le 2h$.

By induction, we see that $\mathcal{U}(h)$ contains all messages μ_i such that $|\mu_i - \mu_0| \le ih$ for $i = 1, \ldots, n$.

Furthermore, because the space is compact it contains its limit points, so we can assert that:

$$\max_{\mu \in \mathcal{U}(h)} |\mu_i - \mu_0| = ih, \quad i = 1, \ldots, n \tag{29}$$

Thus $\mathcal{U}(h)$ contains a corrupted message from node i, according to the definition in eq.(2), if and only if $ih > \delta$ or if and only if:

$$i > \delta/h \tag{30}$$

For any real number x, define $\lceil x \rceil$ as the lowest integer strictly greater than x.

Thus from eq.(30) we see that the lowest index of a node that contains a corrupted message in $\mathcal{U}(h)$ is $\lceil \delta/h \rceil$. If $\lceil \delta/h \rceil > n$ then no node is corrupted.

From eq.(3), the inner maximum in the robustness occurs when $\ell(\mu)$ is the lowest index of a corrupted message. Thus the inner maximum is:

$$M(h) = \frac{n + 1 - \lceil \delta/h \rceil}{n} \tag{31}$$

[3] This well-known relation is explained in Ben-Haim, 2006, p.81.

or zero if this is negative. A plot of h vs. $M(h)$ is identical to a plot of $\widehat{h}(\psi \le m/n, \delta)$ vs. m/n at fixed δ and n, as mentioned at the start of the proof.

When $\delta = 0$, then $\lceil \delta/h \rceil = 1$ and eq.(31) shows that:

$$M(h) = 1 \quad \text{for all} \quad h \ge 0 \tag{32}$$

When $\delta > 0$, we proceed as follows to derive an explicit expression for $M(h)$.

From eq.(31), note that:

$$n \le \frac{\delta}{h} \quad \text{implies} \quad \left\lceil \frac{\delta}{h} \right\rceil \ge n+1 \quad \text{implies} \quad M(h) = \frac{0}{n} \tag{33}$$

and

$$n-i \le \frac{\delta}{h} < n-i+1 \quad \text{implies} \quad \left\lceil \frac{\delta}{h} \right\rceil = n-i+1 \quad \text{implies} \quad M(h) = \frac{n+1-(n-i+1)}{n} = \frac{i}{n}, \quad i = 1, 2, \ldots, n \tag{34}$$

Eqs.(33) and (34) can be re-written more conveniently as:

$$h \le \frac{\delta}{n} \quad \text{implies} \quad M(h) = \frac{0}{n} \tag{35}$$

and

$$\frac{\delta}{n-i+1} < h \le \frac{\delta}{n-i} \quad \text{implies} \quad M(h) = \frac{i}{n}, \quad i = 1, 2, \ldots, n \tag{36}$$

These two equations define $M(h)$, which is the inner maximum in the definition of $\widehat{h}(\psi \le m/n, \delta)$ at fixed n and fixed $\delta > 0$.

We now point out a general property of the robustness function in this example. From eq.(31) we can prove:

$$n < n' \text{ and } h \le \delta \quad \Longrightarrow \quad M(h,n) < M(h,n') \tag{37}$$

Proof:

1. $h \le \delta$ implies that $\lceil \delta/h \rceil > 1$ which implies that $1 - \lceil \delta/h \rceil < 0$.

2. Define $\gamma = 1 - \lceil \delta/h \rceil$, which is negative, and note that $M(h,n) = (n+\gamma)/n$. Thus:

$$\gamma n' < \gamma n \quad \Longrightarrow \quad nn' + \gamma n' < nn' + \gamma n \quad \Longrightarrow \quad \frac{n+\gamma}{n} < \frac{n'+\gamma}{n'} \quad \Longrightarrow \quad M(h,n) < M(h,n') \tag{38}$$

3. $M(h,n)$ is the inner maximum in the definition of the robustness function, eq.(6). A plot of h vs. $M(h,n)$ is identical to a plot of $\widehat{h}(m/n, \delta, n)$ vs. m/n. Likewise, a plot of h vs. $M(h,n')$ is identical to a plot of $\widehat{h}(m/n, \delta, n')$ vs. m/n. Hence, from the last relation in eq.(38), we conclude:

$$\widehat{h}(m/n, \delta, n) > \widehat{h}(m/n, \delta, n') \tag{39}$$

B Derivation of the Probability Distribution in Eq.(7), Section 3.3

A cascading failure has fractional size of zero if no nodes in the telephone chain fail. Thus $P(\psi = 0) = \pi^n$.

A cascading failure has fractional size of i/n, for $i = 1, \ldots, n$, if the first $n-i$ nodes do not fail, for which the probability is π^{n-i}, and the next node does fail for which the probability is $1-\pi$. Failure of nodes is statistically independent, so the probability that $\psi = i/n$ is the product of these two terms: $P(\psi = i/n) = \pi^{n-i}(1-\pi)$ for $i = 1, \ldots, n$.

The cumulative probability distribution is the sum of terms of the probability distribution:

$$P\left(\psi \le \frac{i}{n}\right) = \sum_{j=0}^{i} P\left(\psi = \frac{j}{n}\right), \quad i = 0,\, 1,\, \ldots,\, n \tag{40}$$

$$= \pi^n + \sum_{j=1}^{i} \pi^{n-j}(1-\pi) \tag{41}$$

$$= \pi^n + (1-\pi)\pi^n \sum_{j=1}^{i} \left(\frac{1}{\pi}\right)^j \tag{42}$$

The geometric sum in eq.(42) is:

$$\sum_{j=1}^{i} \left(\frac{1}{\pi}\right)^j = \frac{(1/\pi) - (1/\pi)^{i+1}}{1-(1/\pi)} = \frac{1-\pi^i}{\pi^i(1-\pi)} \tag{43}$$

Thus eq.(42) becomes:

$$P\left(\psi \le \frac{i}{n}\right) = \pi^n + \pi^{n-i}(1-\pi^i), \quad i = 0,\, 1,\, \ldots, n \tag{44}$$

$$= \pi^n + \pi^{n-i} - \pi^n \tag{45}$$

$$= \pi^{n-i} \tag{46}$$

Eq.(7) is the complement of eq.(46).

C Derivation of the Robustness Function in Eq.(11), Section 3.3

Let $m(h)$ denote the inner maximum in eq.(10). From eq.(7) we see that this inner maximum occurs when π is as small as possible in the uncertainty set $\mathcal{U}(h)$ of eq.(8). This occurs when $\pi = (\tilde{\pi} - sh)^+$ where we have defined $x^+ = x$ if $x \ge 0$ and $x^+ = 0$ otherwise. Thus:

$$m(h) = 1 - \left[(\tilde{\pi} - sh)^+\right]^{n-i} \tag{47}$$

The robustness is the greatest value of h at which this expression does not exceed P_{c}:

$$1 - \left[(\tilde{\pi} - sh)^+\right]^{n-i} \le P_{\mathrm{c}} \tag{48}$$

For $h \le \tilde{\pi}/s$, solving this relation at equality yields the robustness in eq.(11). This expression is no greater than $\tilde{\pi}/s$ so we needn't consider greater values of h. This completes the derivation of eq.(11).

D Derivation of the Inner Maximum of the Robustness Function in Eq.(22), Section 4.3

Let $M(h)$ denote the inner maximum in the definition of the robustness function, eq.(22). This is the inner maximum in the definition of $\widehat{h}_j(\psi \le m/R, \delta)$ at fixed R and δ. That is, a plot of h vs. $M(h)$ is identical to a plot of $\widehat{h}_j(\psi \le m/R, \delta)$ vs. m/R. We now derive an expression for $M(h)$.

The original message, μ_0, is not uncertain.

The uncertainty set $\mathcal{U}_j(h)$ contains messages $\mu_{c_{1,j}}$ such that $|\mu_{c_{1,j}} - \mu_0| \le w_{c_{1,j}} h$.

The uncertainty set $\mathcal{U}_j(h)$ contains messages $\mu_{c_{2,j}}$ such that $|\mu_{c_{2,j}} - \mu_{c_{1,j}}| \leq w_{c_{2,j}} h$.

By the triangle inequality, $\mathcal{U}_j(h)$ contains messages $\mu_{c_{2,j}}$ such that $|\mu_{c_{2,j}} - \mu_0| \leq (w_{c_{1,j}} + w_{c_{2,j}})h$.

By induction, we see that $\mathcal{U}_j(h)$ contains messages $\mu_{c_{i,j}}$ such that $|\mu_{c_{i,j}} - \mu_0| \leq h \sum_{k=1}^{i} w_{c_{k,j}}$ for $i = 1, \ldots, R$.

Furthermore, the uncertainty set contains messages for which this last relation is an equality: $|\mu_{c_{i,j}} - \mu_0| = h \sum_{k=1}^{i} w_{c_{k,j}}$. More specifically, we can assert that:

$$\max_{S_j \in \mathcal{U}_j(h)} |\mu_{c_{i,j}} - \mu_0| = h \sum_{k=1}^{i} w_{c_{k,j}}, \quad i = 1, \ldots, R \tag{49}$$

Thus $\mathcal{U}_j(h)$ contains a corrupted message from node $c_{i,j}$ in path C_j, according to the definition in eq.(13), if and only if:

$$\sum_{k=1}^{i} w_{c_{k,j}} > \frac{\delta}{h} \tag{50}$$

As in eq.(18), define $\ell(S_j)$ as the highest row with a corrupted message in the set of messages S_j. At horizon of uncertainty h we see that $\ell(S_j)$ is the smallest integer value of i satisfying eq.(50). That is:

$$\sum_{k=1}^{\ell(S_j)} w_{c_{k,j}} > \frac{\delta}{h} \geq \sum_{k=1}^{\ell(S_j)-1} w_{c_{k,j}} \tag{51}$$

Define $\ell(S_j) = R + 1$ if no node is corrupted in path C_j.

Employing $\ell(S_j)$ in eq.(18) yields the inner maximum in the definition of the robustness for path j:

$$M(h) = \frac{R + 1 - \ell(S_j)}{R} \tag{52}$$

A plot of h vs. $M(h)$ is identical to a plot of $\widehat{h}_j(\psi \leq m/R, \delta)$ vs. m/R at fixed δ and R.

$M(h)$ is the inner maximum in the definition of the robustness function for path j. We can reformulate eq.(52) to obtain an explicit expression for the robustness of path j as follows.

Define the quantities:

$$W_{k,j} = \sum_{i=1}^{k} w_{c_{i,j}} \quad \text{for } k = 1, \ldots, R \tag{53}$$

and define $W_{0,j} = 0$.

Thus the definition of $\ell(S_j)$ in eq.(51) becomes:

$$W_{\ell(S_j),j} > \frac{\delta}{h} \geq W_{\ell(S_j)-1,j} \tag{54}$$

and $\ell(S_j) = R + 1$ if no value of $\ell(S_j) \leq R$ satisfies this relation, meaning that no node is corrupted.

When $\delta = 0$, eq.(54) implies that $\ell(S_j) = 1$, and eq.(52) implies that:

$$M(h) = 1 \quad \text{for all } h \geq 0 \tag{55}$$

Thus the robustness function for path j, when $\delta = 0$, is:

$$\widehat{h}_j(\psi, 0) = \begin{cases} 0 & , \quad \psi < \frac{R}{R} \\ \infty & , \quad \psi = \frac{R}{R} \end{cases} \tag{56}$$

When $\delta > 0$ we proceed as follows.

If:

$$\frac{\delta}{h} \geq W_{R,j} \quad \Longrightarrow \quad \ell(S_j) = R+1 \quad \Longrightarrow \quad M(h) = \frac{0}{R} \tag{57}$$

If, for some $k = 1, \ldots, R$:

$$W_{k,j} > \frac{\delta}{h} \geq W_{k-1,j} \quad \Longrightarrow \quad \ell(S_j) = k \ \text{ and } \ M(h) = \frac{R+1-k}{R} \tag{58}$$

Eqs.(57) and (58) can be re-written as:

$$h \leq \frac{\delta}{W_{R,j}} \qquad \Longrightarrow \qquad M(h) = \frac{0}{R} \tag{59}$$

$$\frac{\delta}{W_{k,j}} < h \leq \frac{\delta}{W_{k-1,j}} \qquad \Longrightarrow \qquad M(h) = \frac{R+1-k}{R}, \quad k = 1, \ldots, R \tag{60}$$

We can now invert eqs.(59) and (60) to obtain explicit expressions for the robustness of path j:

$$\widehat{h}_j(\psi, \delta) \leq \frac{\delta}{W_{R,j}} \qquad \text{for} \qquad \psi = \frac{0}{R} \tag{61}$$

$$\frac{\delta}{W_{k,j}} < \widehat{h}_j(\psi, \delta) \leq \frac{\delta}{W_{k-1,j}} \qquad \text{for} \qquad \psi = \frac{R+1-k}{R}, \quad k = 1, \ldots, R \tag{62}$$

As stated in eqs.(61) and (62), the robustness for path j is a step-wise discontinuous function taking all non-negative real values. It is convenient to identify the values of the robustness function at its kinks. From the "$\leq$" relations in these equations we can write:

$$\widehat{h}_j(\psi, \delta) = \frac{\delta}{W_{k-1,j}} \quad \text{for} \quad \psi = \frac{R+1-k}{R}, \ k = 1, \ldots, R+1 \tag{63}$$

Define the following change of indices:

$$m = R+1-k \quad \text{for} \quad k = 1, \ldots, R+1 \tag{64}$$

So:

$$k = R+1-m \quad \text{and} \quad \psi = \frac{R+1-k}{R} = \frac{m}{R} \quad \text{for} \quad m = 0, 1, \ldots, R \tag{65}$$

Thus eq.(63), the robustness at the kinks, becomes:

$$\widehat{h}_j(\psi \leq m/R, \delta) = \frac{\delta}{W_{R-m,j}} \quad \text{for} \quad m = 0, 1, \ldots, R \tag{66}$$

where R and δ are fixed. Note that eq.(66) implies that $\widehat{h}_j(\psi \leq R/R, \delta) = \infty$ because $W_{0,j} = 0$.

E Proof of Proposition 1, section 4.3

Proof of proposition 1.

Assertion 1. This derives directly from eqs.(24) and (25).

If $m = R$ then the robustness is infinite for all paths and the assertion is true because all path robustnesses are the same.

If $m < R$ then eq.(25) shows that the path with lowest robustness is the path whose average uncertainty weight of the first $R - m$ nodes is maximal. Eq.(24) then shows that this is the overall network robustness.

Assertion 2. We define the following two quantities. Let $(i^\star, j^\star)$ denote the coordinates of the node with maximal uncertainty weight from among all nodes in the network:

$$(i^\star, j^\star) = \arg \max_{1 \le j \le N_R} \max_{1 \le i \le R} w_{c_{i,j}} \tag{67}$$

Let $\overline{j}_m$ denote the index of the path whose average uncertainty weight of the first $R-m$ nodes is maximal, recalling that $m < R$:

$$\overline{j}_m = \arg \max_{1 \le j \le N_R} \sum_{i=1}^{R-m} w_{c_{i,j}} \tag{68}$$

It is clear that many choices of positive uncertainty weights of nodes in the network allow $j^\star$ to differ from $\overline{j}_m$. ∎

F Proof of Proposition 2, Section 4.5

Proof of assertion 1. It is sufficient to prove the proposition for the robustness values at the kinks, because $\widehat{h}_j(\psi \le m/R, \delta)$ is step-wise discontinuous at these kinks.

From the expression for the robustness at a kink, eq.(66) in appendix D, we can write:

$$\widehat{h}_j(\psi \le m/R, \delta) = \frac{\delta}{\displaystyle\sum_{i=1}^{R-m} w_{c_{i,j}}}, \quad m = 1, \ldots, R-1 \tag{69}$$

If $m = R$ then the denominator of $\widehat{h}_j(\psi \le R/R, \delta)$ in eq.(69) is zero so the robustness is infinite: $\widehat{h}_j(\psi \le R/R, \delta) = \infty$. But because $R < R'$, $\widehat{h}_j(\psi \le m/R', \delta)$ is finite, which completes the proof for this case.

If $m < R$ the the denominator of eq.(69) contains $R-m$ strictly positive terms. At fixed m, this denominator strictly decreases as R decreases. Thus $\widehat{h}_j(\psi \le m/R, \delta)$ strictly increases as R decreases, which completes the proof.

Proof of assertion 2. It is sufficient to prove the proposition for the robustness values at the kinks, because $\widehat{h}_j(\psi \le m/R, \delta)$ is step-wise discontinuous at these kinks.

The expression for the robustness at a kink, eq.(66) in appendix D, can be written as eq.(69) in the proof of the first assertion in proposition 2. This expression is independent of the number of branches at each node, which completes the proof.

Proof of assertion 3. It is sufficient to prove the proposition for the robustness values at the kinks, because $\widehat{h}_j(\psi \le m/R, \delta)$ is step-wise discontinuous at these kinks.

The expression for the robustness at a kink, eq.(66) in appendix D, can be written as eq.(69) in the proof of the first assertion in proposition 2. The denominator of eq.(69) contains $R-m$ strictly positive terms.

If $m = R$ then the denominator of $\widehat{h}_j(\psi \le R/R, \delta)$ in eq.(69) is zero so the robustness is infinite: $\widehat{h}_j(\psi \le R/R, \delta) = \infty$. But because $n < m$, $\widehat{h}_j(\psi \le n/R, \delta)$ is finite, which completes the proof for this case.

If $m < R$ then, at fixed R, this denominator strictly decreases as m increases. Thus $\widehat{h}_j(\psi = m/R, \delta)$ strictly increases as m increases, which completes the proof. ∎