To appear in *Military Operations Research Journal*

Decapitation Paradox 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: We study decapitation in a hierarchical network with unity of command. Decapitation means removing a node and the sub-network below that node. Unity of command means that each node receives a command from only one other node. We focus on the question of vulnerability to decapitation: given two different networks, which network is more vulnerable to decapitation? Vulnerability to decapitation can be assessed by counting the number (or fraction) of lost nodes resulting from a specific decapitation. Two propositions prove, however, that the result is paradoxical: A different answer is obtained if one selects the decapitated node by counting i rows from the top of each network, or i rows from the bottom. Alternatively, one can assess the average loss based on the probabilities for decapitation of the different nodes. However, the probability distribution may be uncertain, in which case the assessment of vulnerability is also uncertain. In that case one can

\papers\emergency-readiness\workshop2018\dcdox005.tex. Derived from hinuc005.tex, which was derived from cas-fail012.tex. 12.7.2019 assess vulnerability to decapitation by evaluating the robustness to uncertainty in the probability distribution, while satisfying a performance requirement on the average loss. This analysis is based on info-gap decision theory. We discuss several examples, including choosing the number of brigades in a brigade combat team.

1 INTRODUCTION

Steinbruner (1981-1982) emphasized that "U.S. military planners have long understood the vulnerability of the U.S. command structure" to "nuclear decapitation" by the Soviet Union. He stressed, however, that "the importance of command coordination for overall strategic capability is not easily calculated and cannot be included in standard measures of the strategic balance." Furthermore, Steinbruner noted the lack of a "theory of behavior robust enough to yield a confident prediction" of the impact of decapitation. The goal of this paper is to explain part of the difficulty, and to provide a quantitative tool for responding to the challenge.

By "decapitation" we mean the removal of a command node in a hierarchical network. By "unity of command" we mean that each unit in the network receives orders from only one other unit (Leighton, 1952, p.401). Thus decapitation entails the removal of a node in the hierarchy and removal of the entire sub-network under that node.

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). This is based in part on a long tradition and strong social norms in numerous societies throughout history.[1] The prominence of military 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 attack 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).

We focus on the vulnerability to decapitation from the viewpoint of organizational structure of a hierarchical network with unity of command. The central question we address is: given two alternative network structures, which is more vulnerable to decapitation? For example, should one have "a larger number of BCTs [brigade combat teams] that have predominantly two maneuver battalions, [or] . . . a smaller number of BCTs that all have three maneuver battalions." (Salmerón and Appleget, 2014, p.51).

We employ the concept of robustness, as developed in info-gap theory which is a method for prioritizing options and making choices and decisions under deep uncertainty (Ben-Haim, 2006, 2010, 2018). The options might be operational alternatives (implement a policy, choose a budget,

[1] Hierarchical networks with unity of command are ancient social structures. Imperial armies of Sumer, Akkad, and Babylon surely used such command structures. Similarly, Jethro advised Moses to establish a judicial hierarchy. "So Moses hearkened to the voice of his father in law, and did all that he had said. And Moses chose able men out of all Israel, and made them heads over the people, rulers of thousands, rulers of hundreds, rulers of fifties, and rulers of tens. And they judged the people at all seasons: the hard causes they brought unto Moses, but every small matter they judged themselves." (Exodus, 18: 24–26, King James Version.)

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 may be highly uncertain.

Section 2 takes a first look at the decapitation paradox. In section 3 we formulate our representation of hierarchical networks with unity of command. Section 4 presents and discusses two propositions quantifying the decapitation paradox. Sections 5 and 6 explore average loss due to decapitation with various probability distributions. Section 7 explores the robustness to uncertainty in the probability distribution when assessing average vulnerability to decapitation. Section 8 is a concluding discussion.

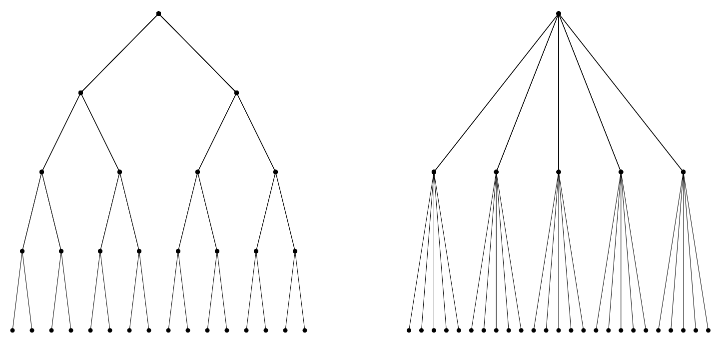

Figure 1: Hierarchical networks with $(R, B) = (5, 2)$ (left) and $(R, B) = (3, 5)$ right. Each network has 31 nodes.

2 THE PARADOX: A FIRST LOOK

The two hierarchical networks in fig. 1 each have 31 nodes and unity of command. Consider decapitation of a node 1 row above the bottom. The left network loses 3 nodes, while in the right network loses 6 nodes. The right network is more vulnerable to decapitation. Likewise, remove a node that is 2 rows above the bottom. The left network loses 7 nodes and the right network loses all 31 nodes. Again the right network is more vulnerable.

Now consider decapitation of a node 1 row below the top. The left network loses 15 nodes while the right network loses only 6 nodes. The left network is more vulnerable. Likewise, remove a node that is 2 rows below the top. The left network loses 7 nodes while the right network loses only 1 node. Again the left network is more vulnerable.

Any node is uniquely identified by its row number and its position in that row; counting rows from the top or from the bottom makes no difference. However, which of the networks in fig. 1 is more vulnerable to decapitation seems to depend on whether one identifies the decapitated node by counting from the bottom or from the top. Assessment of vulnerability to decapitation is important in selecting between hierarchical topologies. Hence resolution of this paradox is central to our study of the reliable functionality of hierarchical networks with unity of command.

We are evaluating the impact of decapitation in terms of the number (or fraction) of nodes loss. While the number of nodes lost does not reflect the full impact on effectiveness of the organization, it is a reasonable basis for comparative evaluation: greater numerical loss will usually correspond to greater loss of effectiveness.

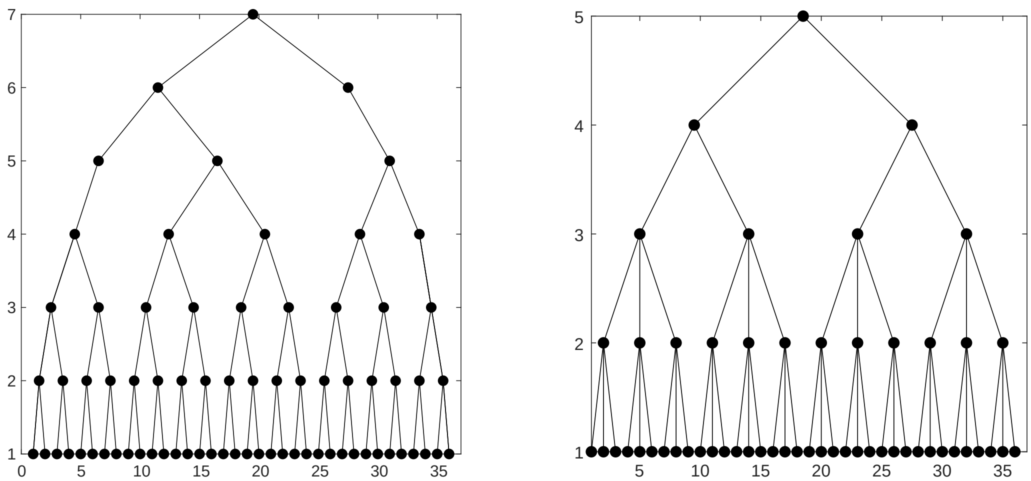

Figure 2: BCT with 2 (left) or 3 (right) brigades per team.

The two networks in fig. 2 show a slightly more realistic but more complicated example, motivated by the choice between 2-brigade or 3-brigade BCTs mentioned earlier (Salmerón and Appleget, 2014). Each hierarchy has 36 nodes (brigades) in its bottom row, but the left hierarchy has them grouped in 2-brigade units and the right hierarchy in 3-brigade units. The total number of nodes in the left and right hierarchies are 74 and 55, respectively. The left hierarchy uses branching ratios of 1 or 2 throughout, while the right hierarchy uses branching ratios of 2 or 3.

Decap i rows below top	Brigades lost in BCT2	Brigades lost in BCT3
0	36	36
1	12 or 24	18
2	8 or 12 or 16	9
3	4 or 8	3
4	4	1

Table 1: Number of “brigades” lost by decapitation i rows **below the top.**

Decap i rows above bottom	Brigades lost in BCT2	Brigades lost in BCT3
0	1	1
1	2	3
2	4	9
3	4 or 8	18
4	8 or 12 or 16	36

Table 2: Number of “brigades” lost by decapitation i rows **above the bottom.**

Table 1 shows the number of brigades (bottom-row nodes) lost by decapitation i rows **below the top** of the hierarchy. We see that the 2-brigade structure (BCT2) almost always loses more brigades than the 3-brigade structure (BCT3). It would seem that BCT3 is less vulnerable to decapitation.

Table 2 shows the number of brigades lost by decapitation i rows **above the bottom** of the hierarchy. We see that now the 3-brigade structure (BCT3) almost always loses more brigades (and never loses fewer brigades) than the 2-brigade structure (BCT2). It would seem that BCT2 is less vulnerable to decapitation.

In many situations one must choose between alternative hierarchical structures with unity of command. Part of the choice is based on addressing the question "Which topology is more vulnerable to decapitation?" Counting the number of nodes lost as a result of decapitation seems paradoxical. In section 3 we formulate our representation of hierarchical networks with unity of command. In section 4 we formalize the paradox in two propositions, and in sections 5–7 we explore 3 different approaches to resolving the paradox.

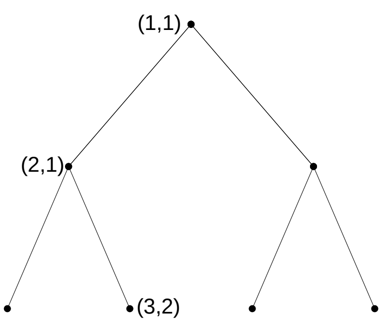

Figure 3: Hierarchical network with 3 rows and 2 branches at each node.

3 HIERARCHICAL NETWORKS WITH UNITY OF COMMAND

We now formulate our mathematical representation of hierarchical networks with unity of command.

3.1 Introduction

We study hierarchical networks as in the simple example of fig. 3, though generalizing to arbitrary branching ratio and number of rows. The single node in the highest row receives a message — a command — 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.

The lexical meaning of "command" is unambiguous. Nonetheless, in human networks it is important to recall Leighton's insight that the "commander's true function … was not coercion but decision, the final responsibility of choosing among various possible courses of action." (Leighton, 1952, p.425). Unity of command entails both the uniqueness of each specific command relationship, and the unity of purpose that encompasses all of the individual orders that are issued.

The concept of a message 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 message that propagates down the hierarchy. Furthermore, the message is intended to evolve in a pre-planned manner. For instance, a battlefield 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. Decapitation interrupts this process, with adverse impact on the overall functionality of the hierarchy. Central to our analysis is the uncertainty that accompanies the decapitation process.

3.2 Formulation of Network 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 now we do not consider varying branching ratio among the nodes, though examples in sections 6.2 and 7 will consider varying branching ratio.) Fig. 3 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:

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

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

3.3 Distinction from Random Networks

The properties of random networks have been of interest for a long time. For instance, Erdös and Rényi (1959) study “asymptotic statistical properties of random graphs”. A “scale free” network is one in which the nodal distribution of a random network follows a power law, at least asymptotically as the degree grows. That is, the fraction of nodes of degree k is proportional to $k^{-\gamma}$ (at least asymptotically in k) where γ is a positive constant. Grant *et al.* (2011) note that the 11 command and control networks that they studied were “closest in form to scale-free networks”. Barabasi and Albert (1999, p.509) demonstrate that the power-law nodal distribution in complex random networks is “a consequence of two generic mechanisms: (i) networks expand continuously by the addition of new vertices, and (ii) new vertices attach preferentially to sites that are already well connected.” These properties introduce an element of self-organization in the growth of the network, which is important in many networks that are neither “completely regular [nor] completely random” (Watts and Strogatz, 1998. p.440).

Our hierarchical networks with unity of command are not scale free, either at finite degree or asymptotically, as we now show.

The “degree” of a node is the number of connections that it has to other nodes. In the networks that we are considering the single top node is connected to B other nodes; the N_R nodes in the bottom row are each connected to only 1 other node; all other nodes are connected to $B+1$ nodes. Thus a node in the networks we are considering has degree 1, or B, or $B+1$. The fraction of nodes with these degrees are:

$$p(1) = \frac{N_R}{N} = \frac{(B-1)B^{R-1}}{B^R - 1} \tag{2}$$

$$p(B) = \frac{1}{N} = \frac{B-1}{B^R - 1} \tag{3}$$

$$p(B+1) = \frac{N - 1 - N_R}{N} = \frac{B^{R-1} - B}{B^R - 1} \tag{4}$$

From eqs.(2)–(4) we see that the hierarchical networks with unity of command that we are studying are not scale free.

Nor are our network models relevant as the branching ratio, B, grows asymptotically. Hierarchical networks with unity of command are fundamentally different from complex and partially random networks such as the world wide web, electrical power grids, citation patterns of scientific articles, and other such networks cited by Barabasi and Albert (1999). Unity of command imposes a non-random structure that distinguishes the network from random scale free networks.

Newman (2003) points out that it is "large-scale statistical properties" that are of interest in scale free and related networks. It is worth quoting Newman (2003) at length:

> "Many of the questions that might previously have been asked in studies of small networks are simply not useful in much larger networks. A social network analyst might have asked. 'Which vertex in this network would prove most crucial to the network's connectivity if it were removed?' But such a question has little meaning in most networks of a million vertices — no single vertex in such a network will have much effect at all when removed. On the other hand, one could reasonably ask a question like 'What percentage of vertices need to be removed to substantially affect network connectivity in some given way?' and this type of statistical question has real meaning even in a very large network." (p.169)

Size itself is not the essential distinction; a modern army may have several million "nodes", and yet removal of particular nodes can have enormous impact. The essential distinction is the strong structural constraint imposed by unity of command, as opposed to the random element in the structure of the complex networks that Newman examines. The two propositions discussed in section 4 show that removal of nodes under unity of command nonetheless entails a paradox, and introduces deep uncertainties, even though the topology is highly structured.

4 DECAPITATION PARADOX

In this section we explore the decapitation paradox in hierarchical networks with unity of command and constant branching ratio. If a single node of a network is removed or incapacitated, then the entire hierarchical sub-network below and including that node is lost: "decapitation" of a sub-network removes all of its nodes. The paradox arises from the following question regarding two networks with different numbers of rows and/or branches at each node. Which network is more vulnerable to decapitation? The paradox is manifested in two propositions. As before, the topology of a hierarchical network with unity of command is specified by the number of rows, R, and number of branches at each node, B, referred to as the branching ratio.

The paradox is embodied in the following two propositions, whose proofs appear in appendices A and B.

Proposition 1 ***Given:***

1. *Two hierarchical networks with unity of command, denoted T and T', whose topologies are (R, B) and (R', B'), respectively.*
2. *Network T has fewer branches at each node: $B < B'$.*
3. *A single node is decapitated in each network. In each case the decapitated node is i rows* ***above the bottom row*** *of its network, for $i = 0, \ldots, \min(R-1, R'-1)$.*

Let ℓ and ℓ' denote the total number of nodes lost in networks T and T', respectively, due to the decapitation.

Then: *The network with lower branching ratio* ***loses fewer nodes*** *than the network with greater branching ratio unless the decapitation is in the bottom row:*

$$\ell \leq \ell' \tag{5}$$

with equality if and only if $i = 0$.

Proposition 2 ***Given:***

1. *Two hierarchical networks with unity of command, denoted T and T', whose topologies are (R, B) and (R', B'), respectively.*
2. *Network T has fewer branches at each node: $B < B'$. The two networks have the same total number of nodes.*
3. *A single node is decapitated in each network. In each case the decapitated node is i rows* ***below the top row*** *of its network, for $i = 0, \ldots, \min(R-1, R'-1)$.*

Let L and L' denote the total number of nodes lost in networks T and T', respectively, due to the decapitation.

Then: *The network with lower branching ratio* ***loses more nodes*** *than the network with greater branching ratio unless the decapitation is in the top row:*

$$L \geq L' \tag{6}$$

with equality if and only if $i = 0$.

Roughly speaking, proposition 1 states that the network with lower branching ratio loses **fewer** nodes, while proposition 2 states that the network with lower branching ratio loses **more** nodes. Proposition 1 locates the decapitated node by counting from the **bottom** of the network, while Proposition 2 counts from the **top** and also assumes that the two networks have the same total number of nodes. The paradox is that, seemingly, the more vulnerable network depends on whether one counts from the bottom or from the top.

It is important to resolve this paradox because one might, mistakenly, answer the question "Which topology is more vulnerable to decapitation?" by counting from the top (or from the bottom), without recognizing that top- or bottom-counting yield different answers.

A paradox is not a contradiction; it results from a faulty perception or misunderstanding. One can unambiguously locate a node either by counting from the bottom or from the top. However, when counting from the bottom, both sub-networks have the same number of rows, which is not the case when counting from the top if the networks have the same number of nodes. Hence the difference between the two propositions.

Furthermore, a high-row node with low branching ratio commands a large fraction of all the nodes of the network. For instance, each node in the first row below the top of the left hand network in fig. 1 commands 15 of the 31 nodes, while each node 1 row below the top of the righthand network commands only 6 of the 31 nodes. Precisely the reverse holds for low rows, as seen again in fig. 1.

On a more conceptual level, we resolve the paradox by recognizing that the original question is not well posed. The question was: Given two networks with different numbers of rows and/or branches, which network is more vulnerable to decapitation? The answer is that vulnerability to decapitation is not a unique property of a hierarchical network with unity of command. Networks with low branching ratio will be relatively more vulnerable to decapitation high in the network, but relatively less vulnerable for decapitation low in the network. Conversely, networks with high branching ratio will be relatively less vulnerable to decapitation high in the network, but relatively more vulnerable to decapitation low in the network.

This conclusion is itself significant for its implications about differential vulnerability to low- or high-row decapitation. It demonstrates that hierarchical topology — R and B — by itself does not uniquely determine vulnerability to decapitation.

We now extend the exploration of assessing vulnerability to decapitation, revealing further ambiguities and their resolution.

In section 5 we compare the average loss due to decapitation in networks with different topologies, assuming equal probability for decapitation of each node. This yields unique comparisons of vulnerability, but depends on the assumed probability distribution. Section 6 looks at average losses with alternative probability distributions, resulting in different assessments of vulnerability. If one knows the probability distribution for decapitation, then the average loss is a unique measure of vulnerability. If the probability distribution is uncertain, then one can consider the robustness to uncertainty in the distribution, as explored in section 7.

5 AVERAGE LOSS WITH EQUAL PROBABILITIES

Consider a hierarchical network with R rows and constant branching ratio B. A node in the network will be chosen at random, with equal probability for each node. The selected node will be removed, which decapitates the sub-network in which this node is at the top. Let $\overline{\ell}$ denote the average number of nodes lost. In appendix C we show that, if $B = 1$:

$$\overline{\ell} = \frac{R+1}{2} \tag{7}$$

while if $B > 1$:

$$\overline{\ell} = \frac{RB^R}{B^R - 1} - \frac{1}{B-1} \tag{8}$$

	B= 1	2	3	4	5
R= 1	1.0	1.00	1.00	1.00	1.00
2	1.5	1.66	1.75	1.80	1.83
3	2.0	2.42	2.61	2.71	2.77
4	2.5	3.26	3.55	3.68	3.75
5	3.0	4.16	4.52	4.67	4.75
6	3.5	5.09	5.50	5.66	5.75
7	4.0	6.05	6.50	6.66	6.75
8	4.5	7.03	7.50	7.66	7.75
9	5.0	8.01	8.50	8.66	8.75
10	5.5	9.00	9.50	9.66	9.75

Table 3: Average number of nodes lost, $\overline{\ell}(R, B)$, with uniform probability distribution.

Table 3 shows that, on average, large networks lose more nodes from a decapitation than small networks. However, in terms of the average *fraction* of lost nodes, large networks are less vulnerable to decapitation than small networks. Once again, vulnerability to decapitation is in part a matter of definition: different assessments result from different definitions.

Table 3 also allows comparison of the vulnerabilities of networks with different (R, B) values. For instance, consider networks $(R, B) = (5, 2)$ and $(R, B) = (3, 5)$, which both have 31 nodes. Network (5,2) loses on average 4.16 nodes from a uniformly random decapitation, while network (3,5) loses only 2.77 nodes on average. In this respect, the latter network is less vulnerable to decapitation.

We must recall, however, that these averages presume a uniform probability distribution across all of the nodes. We explore the relaxation of this assumption in the next section, and we will see that vulnerability varies substantially as the probability distribution changes.

6 AVERAGE LOSS WITH ALTERNATIVE PROBABILITY DISTRIBUTIONS

The uniform probability distribution might be adopted in the absence of more specific information, motivated perhaps by the principle of indifference. However, many decapitations are anything but uniformly distributed. For example, political assassinations will generally be biased to high-ranking individuals. The Archduke Franz Ferdinand of Austria and his wife Sophie were assassinated in 1914 by Gavrilo Princip who chose them precisely because of their high social rank. Likewise, Steinbruner (1981–1982) emphasized that the nuclear capabilities of both the United States and the Soviet Union were highly vulnerable to elimination of a relatively small number of critical targets. In contrast, terrorist violence is usually directed explicitly at ordinary citizens in order to instill fear in the general populace.

The average loss due to decapitation is a useful measure of the vulnerability of a network. However, one must consider the choice of the probability distribution. In sections 6.1 and 6.2 we consider alternative probability distributions for networks with constant and variable branching ratios. However, the probability distribution may be uncertain, so in section 7 we consider uncertainty in the probability distributions themselves.

6.1 Networks with Constant Branching Ratio

Let p_{ij} denote the probability of removal of the ith node in the jth row counting from the top of a network with R rows and constant branching ratio B. Recall that $N(R, B)$ in eq.(1) is the number of nodes in a network with R rows and branching ratio B, and that $N_j = B^{j-1}$ is the number of nodes in the jth row. Generalizing eq.(36) in appendix C, the average loss is:

$$\begin{aligned} \bar{\ell}(R, B) &= \sum_{j=1}^{R} \sum_{i=1}^{N_j} p_{ij} N(R - j + 1, B) && (9) \\ &= \sum_{j=1}^{R} p_j N(R - j + 1, B) && (10) \end{aligned}$$

where p_j is the probability of a decapitation in the jth row counting from the top:

$$p_j = \sum_{i=1}^{N_j} p_{ij}, \quad j = 1, \ldots, R \tag{11}$$

For notational convenience we define the vector of sub-network sizes:

$$\nu^T = (N(R, B), \, N(R - 1, B), \, \ldots, \, N(1, B)) \tag{12}$$

Now the average loss in eq.(10) becomes:

$$\bar{\ell}(R,B) = p^T \nu \tag{13}$$

where p is the vector of row probabilities $(p_1, \ldots, p_R)$.

We will also be interested in the average fraction of nodes lost, which is defined as:

$$\bar{f}(R,B) = \frac{\bar{\ell}(R,B)}{N(R,B)} \tag{14}$$

The results in table 3 were calculated with the uniform distribution:

$$p_j^{\mathrm{uni}} = \frac{N_j}{N(R,B)}, \quad j = 1, \ldots, R \tag{15}$$

We now consider two different distributions. The first distribution gives greater weight to higher rows in the network, and the second distribution preferentially weights lower nodes:

$$p_j^{\mathrm{top}} = \frac{(R-j+1)^2 N_{R-j+1}}{\sum_{j=1}^R j^2 N_j}, \quad j = 1, \ldots, R \qquad \text{(top-heavy)} \tag{16}$$

$$p_j^{\mathrm{bot}} = \frac{j^2 N_j}{\sum_{j=1}^R j^2 N_j}, \quad j = 1, \ldots, R \qquad \text{(bottom-heavy)} \tag{17}$$

Note that p_j^{top} is the same as p_j^{bot} in reverse order:

$$p_j^{\mathrm{top}} = p_{R-j+1}^{\mathrm{bot}} \tag{18}$$

	B= 1	2	3	4	5
$R=1$	1.00	1.00	1.00	1.00	1.00
2	1.80	2.77	3.76	4.76	5.76
3	2.57	6.15	11.72	19.28	28.84
4	3.33	12.69	34.94	76.07	142.0
5	4.09	25.47	103.2	299.6	700.7

Table 4: Average number of nodes lost, $\bar{\ell}(R,B)$, with the top-heavy probability distribution, eq.(16).

	B= 1	2	3	4	5
$R=1$	1.00	1.00	1.00	1.00	1.00
2	1.20	1.22	1.23	1.23	1.23
3	1.42	1.48	1.51	1.52	1.52
4	1.66	1.77	1.80	1.82	1.83
5	1.90	2.07	2.11	2.14	2.15

Table 5: Average number of nodes lost, $\bar{\ell}(R,B)$, with the bottom-heavy probability distribution, eq.(17).

Tables 4 and 5 show the average number of nodes lost with the top-heavy and bottom-heavy probability distributions, respectively. The far greater losses with the top-heavy distribution reflect the fact that sub-networks below high nodes are much larger than sub-networks below low nodes.

More importantly, we note that one may get different assessments of the vulnerability to decapitation, depending on the probability distribution. Compare the hierarchies with $(R, B) = (5, 2)$ and $(R, B) = (3, 5)$, which both have 31 nodes. In the top-heavy case the average losses are 25.47 and 28.84 nodes, respectively, indicating slightly greater vulnerability for configuration (3,5) over (5,2). However, in the bottom-heavy case the average losses for (5,2) and (3,5) are 2.07 and 1.52, respectively, indicating greater vulnerability for (5,2). Which configuration is more vulnerable, as assessed by the average nodal loss? The answer may depend on what probability distribution is employed.

	B= 1	2	3	4	5
$R = 1$	1.000	1.000	1.000	1.000	1.000
2	0.900	0.925	0.942	0.952	0.960
3	0.857	0.879	0.901	0.918	0.930
4	0.833	0.846	0.873	0.895	0.910
5	0.818	0.821	0.853	0.878	0.897

Table 6: Average fraction of nodes lost, $\overline{f}(R, B)$, with the top-heavy probability distribution, eq.(16).

	B= 1	2	3	4	5
$R = 1$	1.0000	1.0000	1.0000	1.0000	1.0000
2	0.6000	0.4074	0.3077	0.2471	0.2063
3	0.4762	0.2127	0.1162	0.0725	0.0493
4	0.4167	0.1183	0.0452	0.0215	0.0118
5	0.3818	0.0668	0.0175	0.0063	0.0028

Table 7: Average fraction of nodes lost, $\overline{f}(R, B)$, with the bottom-heavy probability distribution, eq.(17).

Different assessments of vulnerability also arise when comparing fractional losses of configurations with different numbers of nodes, as seen in tables 6 and 7. Consider for instance (5,3) with 121 nodes, and (4,5) with 156 nodes. The top-heavy distribution results in average fractional losses of 0.853 for (5,3), and 0.910 for (4,5); thus (4,5) is more vulnerable. However, with the bottom-heavy distribution the fractional losses are 0.0175 for (5,3), and 0.0118 for (4,5); thus (5,3) is more vulnerable.

In short, the average loss of nodes due to random decapitation is useful for comparing the vulnerabilities of different hierarchical configurations, if one knows the probability for decapitation of the different nodes. However, different assessments of vulnerability may result from different distributions. In section 7 consider the assessment of vulnerability given uncertainty about this probability distribution.

6.2 Networks with Variable Branching Ratio: BCTs

We now evaluate the average loss due to random decapitation of 2-brigade, 3-brigade and 4-brigade hierarchies (fig. 2 in section 2 and fig. 4). As before, R is the number of rows, p_j is the probability that decapitation occurs somewhere in the jth row counting from the top, and we assume equal probability for decapitation of each node in the jth row. The total number of nodes in the network is N. Let N_j denote the number of nodes in the jth row from the top, for which the expression prior to

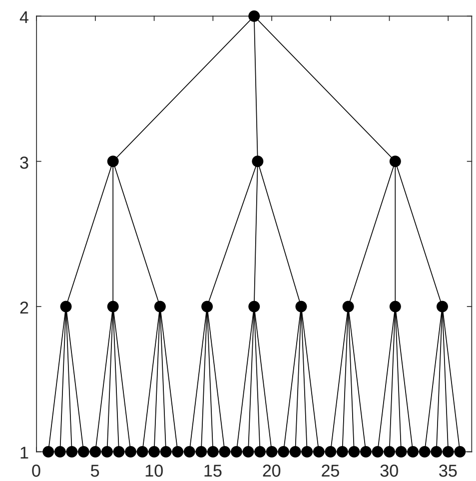

Figure 4: BCT with 4 brigades per team.

eq.(1) is no longer valid because the branching ratios in the BCT networks are not constant. Let S_{ij} denote the number of nodes in the subnetwork headed by the ith node in the jth row from the top. The average fractional loss due to decapitation is:

$$\overline{f} = \frac{1}{N} \sum_{j=1}^{R} \sum_{i=1}^{N_j} \frac{p_j}{N_j} S_{ij} \tag{19}$$

Recall that the average total loss, $\overline{\ell}$, is related to the average fractional loss, $\overline{f}$, as $\overline{\ell} = N\overline{f}$.

We evaluate $\overline{f}$ for each of the 3 networks in figs. 2 and 4, and for the top-heavy and bottom-heavy probability distributions in eqs.(16) and (17), $p_j^{\rm top}$ and $p_j^{\rm bot}$. Instead of the distribution $p_j^{\rm uni}$ in eq.(15), which has equal probability for all nodes, we use a distribution with equal probability for all rows. This row-uniform distribution is defined:

$$p_j^{\rm runi} = \frac{1}{R}, \quad j = 1, \ldots, R \tag{20}$$

We find that the 2-brigade network has lower average fractional losses, but higher total losses, than both the 3- and 4-brigade networks. Specifically, the average fractional losses for the 2-, 3- and 4-brigade networks, with the bottom-heavy, row-uniform, and top-heavy probability distributions, are:

$$\overline{f}_{\rm bot,2} = 0.0362, \quad \overline{f}_{\rm bot,3} = 0.0385, \quad \overline{f}_{\rm bot,4} = 0.0471 \tag{21}$$

$$\overline{f}_{\rm runi,2} = 0.3065, \quad \overline{f}_{\rm runi,3} = 0.3636, \quad \overline{f}_{\rm runi,4} = 0.3828 \tag{22}$$

$$\overline{f}_{\rm top,2} = 0.7894, \quad \overline{f}_{\rm top,3} = 0.8825, \quad \overline{f}_{\rm top,4} = 0.9025 \tag{23}$$

We see that the 3-brigade network is more vulnerable than the 2-brigade network to all three distributions, as assessed by the average fractional loss. Similarly, the 4-brigade network is more vulnerable than the 3-brigade network. For instance, with the top-heavy distribution, the 2-brigade network loses 78.94% of its nodes, while the 3- and 4-brigade networks lose 88.25% and 90.25% of their nodes.

Assessment of vulnerability with the average total nodal loses is reversed:

$$\overline{\ell}_{\rm bot,2} = 2.679, \quad \overline{\ell}_{\rm bot,3} = 2.117, \quad \overline{\ell}_{\rm bot,4} = 2.308 \tag{24}$$

$$\overline{\ell}_{\rm runi,2} = 22.68, \quad \overline{\ell}_{\rm runi,3} = 20.00, \quad \overline{\ell}_{\rm runi,4} = 18.75 \tag{25}$$

$$\overline{\ell}_{\rm top,2} = 58.41, \quad \overline{\ell}_{\rm top,3} = 48.54, \quad \overline{\ell}_{\rm top,4} = 44.22 \tag{26}$$

The 2-brigade network loses more nodes on average than either the 3- or 4-brigade networks. The 3-brigade network loses more nodes on average than the 4-brigade network with the row-uniform and top-heavy distributions, but not with the bottom-heavy distribution. Recall that the 2-brigade network has 74 nodes while the 3- and 4-brigade networks have 55 and 49 nodes, respectively:

In summary, eqs.(21)–(23) suggest that the 2-brigade network is less vulnerable per unit deployed than the 3- or 4-brigade networks, while eqs.(24)–(26) suggest that the 2-brigade network is more costly in absolute loss.

7 AVERAGE LOSS WITH UNCERTAIN PROBABILITY DISTRIBUTIONS

We now consider uncertainty in the probability distribution, p, where p_i is the probability that decapitation occurs in the ith row from the top. p determines the average number or fraction of lost nodes. We continue to consider networks in which the branching ratio is not necessarily the same at all nodes.

Let $\widetilde{p}$ denote the best estimate of the probability distribution. This may, for instance, be the row-uniform distribution in eq.(20), or the top-heavy or bottom-heavy probability distributions in eqs.(16) and (17). The true distribution, p, is uncertain because it depends on the behavior of adversaries whose goals and capabilities may be poorly known to us. We do not know how much each element of p deviates from the corresponding element of $\widetilde{p}$. Our uncertainty about p is unbounded in the space of mathematically legitimate distributions. That is, we do not know or have a reasonable estimate of the maximal deviation between p and $\widetilde{p}$. While $\widetilde{p}$ and p are probability distributions, our uncertainty about them is non-probabilistic: we know no probability distribution for the deviation between them. The absolute-error info-gap model represents this unbounded non-probabilistic uncertainty:

$$\mathcal{U}(h) = \left\{ p : \ \sum_{i=1}^{R} p_i = 1, \ \ p_i \geq 0, \ \ |p_i - \widetilde{p}_i| \leq h, \ i = 1, \ldots, R \right\}, \quad h \geq 0 \tag{27}$$

$\mathcal{U}(h)$ is the set of all mathematically legitimate probability distributions $(p_1, \ldots, p_R)$ whose elements p_i deviate from the corresponding elements $\widetilde{p}_i$ of the nominal distribution by no more than h. The value of h is unknown and unbounded, so the info-gap model is a family of nested sets of probability distributions, unbounded in the space of all probability distributions.

The average fractional loss due to decapitation, $\overline{f}$, is defined in eq.(19). We require that the average fraction of lost nodes not exceed a critical value, which we denote $\overline{f}_{\mathrm{c}}$. Alternatively, the value of $\overline{f}_{\mathrm{c}}$ might arise from the question: What is the greatest fractional loss, $\overline{f}_{\mathrm{c}}$, which we can be quite confident will not be exceeded? Formally, the performance requirement is:

$$\overline{f} \leq \overline{f}_{\mathrm{c}} \tag{28}$$

The robustness is the greatest horizon of uncertainty, h, up to which all probability vectors p in the uncertainty set $\mathcal{U}(h)$ satisfy the requirement in eq.(28). That is, the definition of the robustness function is:

$$\widehat{h}_f(\overline{f}_{\mathrm{c}}) = \max \left\{ h : \left(\max_{p \in \mathcal{U}(h)} \overline{f} \right) \leq \overline{f}_{\mathrm{c}} \right\} \tag{29}$$

Robustness curves for average fractional loss with 2-, 3- and 4-brigade BCTs are shown in figs. 5–7. The putative probability distributions in these 3 figures are bottom-heavy, row-uniform, and top-heavy respectively, eqs.(17), (20) and (16).

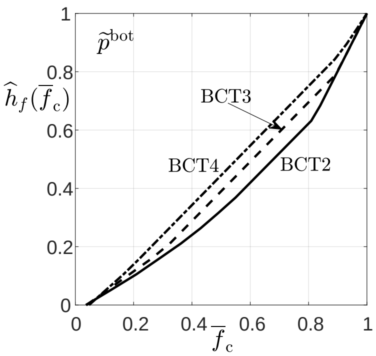

Figure 5: Robustness curves for 2-, 3-, and 4-brigade networks networks. Bottom-heavy putative distribution.

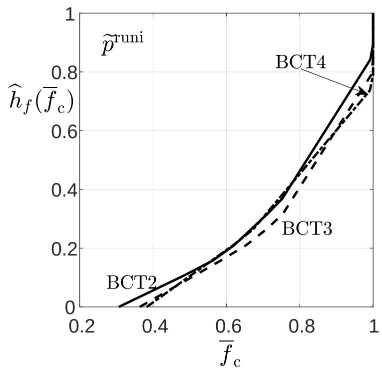

Figure 6: Robustness curves with row-uniform putative distribution.

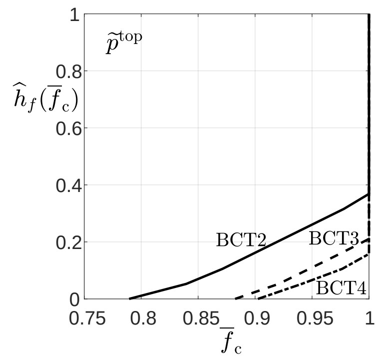

Figure 7: Robustness curves with top-heavy putative distribution.

Each robustness curve reaches the horizontal axis at the value of average fractional loss that is predicted with the putative probability distribution. For example, in fig. 7 the robustness curves for 2-, 3- and 4-brigade networks reach zero robustness at $\overline{f}_{\rm c} =$ 0.79, 0.88 and 0.90, respectively, as we expect from eq.(23). This is the "zeroing" property, characteristic of all info-gap robustness curves. It means that predicted outcomes — based on the best available information — have no robustness to error and uncertainty in that information. Consequently, prioritization of one's alternatives on the basis of predicted outcomes is unreliable. So, for instance, if the best estimate of the probability distribution is that it is top-heavy, as in fig. 7, one might be inclined to prefer the 2-brigade over the 3-brigade network, and the 3-brigade over the 4-brigade network because the predicted losses are lowest for BCT2 and highest for BCT4, as in eq.(23). However, those predictions have zero robustness to error in the probability distribution, so that prioritization is unreliable if the probability distribution is uncertain.

The positive slopes of the robustness curves in figs. 5–7 express the trade off between robustness and fractional loss due to decapitation: greater robustness against uncertainty, $\widehat{h}_f(\overline{f}_{\rm c})$, is obtained only by allowing greater fractional loss of nodes, $\overline{f}_{\rm c}$. For instance, in fig. 5 we see that BCT2 has no robustness for average fractional loss of 0.0326 (the zeroing property), but it has robustness of $\widehat{h}_f =$ 0.24 for $\overline{f}_{\rm c}$ = 0.4 as a result of the trade off: less demanding performance (greater $\overline{f}_{\rm c}$) has greater robustness against uncertainty.

Fractional loss is a proxy for loss of effectiveness, as explained in section 2. Thus the positive slope of the robustness curves expresses the trade off between robustness to uncertainty and the effectiveness of the organization: greater effectiveness is guaranteed only at lower robustness to uncertainty.

In fig. 5 we see that BCT4 is more robust than BCT3, which is more robust than BCT2, over nearly the entire range of fractional loss. Thus the robust preference is for BCT4 over BCT3, and for BCT3 over BCT2. However, the robustness increments between the adjacent options are fairly small, so these robust preferences are not strong, and the options seem roughly equivalent in terms of their average fractional loss with a putative (but uncertain) bottom-heavy probability distribution. In terms of fractional loss in the bottom-heavy case, one is roughly indifferent between the options (though less so in comparing BCT4 against BCT2).

Fig. 6 shows an even clearer case of equivalence between the 3 options in terms of robustness to uncertainty in a row-uniform probability distribution: the robustness curves strongly overlap one another.

Fig. 7 shows a clear robust preference for the 2-brigade option over BCT3, which in turn is weakly preferred over BCT4, in the case of an uncertain top-heavy probability distribution. We note, however, that the robustness is much lower than in the previous two figures at the corresponding $\overline{f}_{\rm c}$ values (note the different horizontal scales). This means that, while BCT2 is most preferred in this case, one has less confidence that the corresponding fractional losses will not be exceeded.

Summarizing figures 5–7 one could conclude that the robustness of BCT2 is roughly equal to or greater than the robustness of BCT3 or BCT4, though BCT2's robustness is low in the presence of uncertain top-heavy probability distributions.

8 DISCUSSION

We have studied decapitation in a hierarchical network with unity of command. By decapitation we mean the removal of a node and of the entire sub-network below that node. Unity of command means that each node receives a command or message from only one other node. We have focussed on the question of vulnerability to decapitation: given two different networks, which network is more vulnerable to decapitation?

One might approach this question by counting the number of nodes lost when decapitating a specific node. However, the 2 propositions in section 4 present a paradox. A different answer to the vulnerability question is obtained if one selects the decapitated node in each network by counting i rows from the top of each network, or i rows from the bottom. The resolution of the paradox hinges on recognizing that the original question — which of the two networks is more vulnerable to decapitation — is not well posed. Networks with low branching ratio will be relatively more vulnerable to decapitation high in the network, but relatively less vulnerable for decapitation low in the network. Conversely, networks with high branching ratio will be relatively less vulnerable to decapitation high in the network, but relatively more vulnerable to decapitation low in the network.

We refined the question by examining the average number (or fraction) of nodes lost, given a specified probability for decapitation of each node. However, in sections 5 and 6 we showed by example that one can get quite different assessments of vulnerability by using different probability distributions. This answer to the vulnerability question is satisfactory only if the probability distribution is well known. However, the probability distribution for decapitation is likely to be highly uncertain because it reflects goals and capabilities of an adversary.

Hence in section 7 we refined the question once again by exploring the robustness, to uncertainty in the decapitation probability, of average fractional loss of nodes. Specifically, for a given network topology, we ask the robustness question: how much error in the estimated probability distribution can be tolerated without the average fractional loss exceeding a critical value? A large value of robustness implies that the fractional loss will be acceptable even if the estimated probability distribution errs greatly. Conversely, low robustness implies low tolerance to error in the estimated distribution. This allows us to compare alternative topologies, as we illustrated by comparing 2-brigade, 3-brigade and 4-brigade BCTs.

To summarize, vulnerability to decapitation can be assessed by counting the number (or fraction) of lost nodes resulting from a specific decapitation if one knows which specific node is at risk. If the nodal vulnerabilities are uncertain, one can use a probability distribution for nodal decapitation to assess average loss, assuming knowledge of the probability distribution. If the probability distribution is uncertain, one can assess vulnerability to decapitation by evaluating the robustness to uncertainty in the probability distribution, while satisfying a performance requirement on the average loss.

Military and other organizations are exploring alternatives to strict hierarchy, such as networking sensors, commanders, and end-effectors, in order to relax the hierarchical constraint and to expedite action and enhance effectiveness. The methodology developed in this paper can be adapted to explore the vulnerabilities of these alternative organizational structures.

9 REFERENCES

Barabási, Albert-László and Réka Albert, 1999, Emergence of scaling in random networks, *Science,* vol. 286, pp.509–512, 15 October 1999.

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

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

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

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.

P. Erdös and A Rényi, 1959, On random graphs, *Publicationes Mathematicae* (Debrecen), vol. 6, pp.290–297.

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.

Grant, Tim J., B.C. Buizer, R.J. Bertelink, 2011, Vulnerability of C2 Networks to Attack: Measuring the topology of eleven Dutch Army C2 systems, 16th International Command and Control Research and Technology Symposium (ICCRTS 2011), Quebec City, Canada.

Leighton, Richard M., 1952, Allied unity of command in the second World War: A study in regional military organization, *Political Science Quarterly,* vol. 67(3), pp. 399–425.

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

Newman, M.E.J., 2003, The structure and function of complex networks, *SIAM Review,* vol. 45. no. 2., pp.167–256.

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.

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

Steinbruner, John D., Nuclear Decapitation, *Foreign Policy,* No. 45 (Winter, 1981-1982), pp. 16–28.

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

Watts, Duncan J. and Steven H. Strogatz, 1998, Collective dynamics of 'small-world' networks, *Nature,* vol. 393, pp.440–442, 4 June 1998.

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 PROOF OF PROPOSITION 1, SECTION 4

The sub-networks removed from networks T and T', by removal of the node, have $i+1$ rows and B or B' branches at each node, respectively, where $B' > B$.

The total number nodes removed by the decapitation, in each network, are:

$$\ell = \sum_{j=1}^{i+1} B^{j-1} \tag{30}$$

$$\ell' = \sum_{j=1}^{i+1} {B'}^{j-1} \tag{31}$$

If $i = 0$ then $\ell = 1 = \ell'$. If $i > 0$ then $\ell < \ell'$ because $B' > B \geq 1$. ∎

B PROOF OF PROPOSITION 2, SECTION 4

The following lemma will be instrumental in proving proposition 2. The basic idea of this lemma is that each node in a network is the top of a sub-network with the same number of branches at each node, but with fewer rows. The 1st row below the top of the network contains B nodes. If the total number of nodes in the network in N, then each of these B nodes is the top of a sub-network with $(N-1)/B$ nodes. The lemma generalizes this to the ith row below the top of the network.

Lemma 1 ***Given:***

1. *A hierarchical network with unity of command containing R rows and B branches at each node except in the bottom row.*
2. *The total number of nodes in the network is N.*

Then *the number of nodes below and including a node in the ith row below the top, for $i = 0, 1, \ldots, R-1$, is:*

$$N_{(i)} = \frac{N - \sum_{j=1}^{i} B^{j-1}}{B^i} \tag{32}$$

Proof of lemma 1.

1. If $B = 1$ then the network is a single path with N nodes. $N_{(i)}$ in eq.(32) equals $N - i$ which is the number of nodes below and including the ith node below the top. Thus eq.(32) is correct for $B = 1$.
2. Now consider $B > 1$.
3. A node in the ith row below the top of the network, for $i = 0, 1, \ldots, R-1$, is the top of a network with $R-i$ rows and B branches at each branching node. From eq.(1) we see that this node is the top of a sub-network with $(B^{R-i}-1)/(B-1)$ nodes.
4. From eq.(1) we see that $N = (B^R - 1)/(B-1)$ and that $\sum_{j=1}^{i} B^{j-1} = (B^i - 1)/(B-1)$.
5. Thus the righthand side of eq.(32) becomes:

$$\frac{(B^R-1)/(B-1) - (B^i-1)/(B-1)}{B^i} = \frac{B^R - B^i}{B^i(B-1)} = \frac{B^{R-i}-1}{B-1} \tag{33}$$

6. This agrees with the result in step 3. Thus eq.(32) is correct. ∎

Proof of proposition 2. By lemma 1:

$$L = \frac{N - \sum_{j=1}^{i} B^{j-1}}{B^i} \tag{34}$$

$$L' = \frac{N - \sum_{j=1}^{i} B'^{j-1}}{B'^i} \tag{35}$$

Hence, if $i = 0$, then $L = N = L'$. If $i \geq 1$, then $L > L'$ because $B' > B$. ∎

C DERIVATION OF EQS.(7) AND (8) IN SECTION 5

Let $N(R, B)$ denote the number of nodes in a network with R rows and branching ratio B, based on eq.(1). Let $N_j = B^{j-1}$ denote the number of nodes in the jth row, counting from the top of the network, for $j = 1, \ldots, R$.

The probability of removing a node in the jth row from the top is $N_j/N(R,B)$ because each node is equally likely to be removed.

A node in the jth row from the top is at the top of a sub-network with $R - j + 1$ rows. Thus, removal of this node removes $N(R - j + 1, B)$ nodes.

Thus the average number of nodes lost is:

$$\bar{\ell} = \frac{1}{N(R,B)} \sum_{j=1}^{R} N_j N(R - j + 1, B) \tag{36}$$

If $B = 1$ then $N_j = 1$ and $N(R, B) = R$ so eq.(36) becomes:

$$\bar{\ell} = \frac{1}{R} \sum_{j=1}^{R} (R - j + 1) = \frac{1}{R} \sum_{j=1}^{R} j = \frac{R+1}{2} \tag{37}$$

which is eq.(7).

If $B > 1$, eq.(36) becomes:

$$\bar{\ell} = \frac{B-1}{B^R - 1} \sum_{j=1}^{R} B^{j-1} \frac{B^{R-j+1}-1}{B-1} \tag{38}$$

$$= \frac{1}{B^R - 1} \sum_{j=1}^{R} \left(B^R - B^{j-1} \right) \tag{39}$$

$$= \frac{RB^R}{B^R - 1} - \frac{1}{B^R - 1} \underbrace{\sum_{j=1}^{R} B^{j-1}}_{N(R,B)} \tag{40}$$

$$= \frac{RB^R}{B^R - 1} - \frac{1}{B - 1} \tag{41}$$

which is eq.(8).