Hawaii International Conference on System Sciences, January 2003, Hawaii.

A probabilistic loading-dependent model of cascading failure and possible implications for blackouts

Ian Dobson
ECE Department
University of Wisconsin
Madison WI 53706 USA
dobson@engr.wisc.edu

Benjamin A. Carreras
Oak Ridge National Laboratory
Oak Ridge TN 37831 USA
carrerasba@ornl.gov

David E. Newman
Physics Department
University of Alaska
Fairbanks AK 99775 USA
ffden@uaf.edu

Abstract

Catastrophic disruptions of large, interconnected infrastructure systems are often due to cascading failure. For example, large blackouts of electric power systems are typically caused by cascading failure of heavily loaded system components. We introduce the CASCADE model of cascading failure of a system with many identical components randomly loaded. An initial disturbance causes some components to fail by exceeding their loading limit. Failure of a component causes a fixed load increase for other components. As components fail, the system becomes more loaded and cascading failure of further components becomes likely. The probability distribution of the number of failed components is an extended quasibinomial distribution. Explicit formulas for the extended quasibinomial distribution are derived using a recursion. The CASCADE model in a restricted parameter range gives a new model yielding the quasibinomial distribution. Some qualitative behaviors of the extended quasibinomial distribution are illustrated, including regimes with power tails, exponential tails, and significant probabilities of total system failure.

1. CASCADE model

Cascading failure is a standard cause of catastrophic failure in large, interconnected infrastructure systems. For example, loading-dependent cascading failure occurs in large blackouts of electric power transmission systems. The importance of these infrastructures to society motivates the construction and study of models that capture salient features of cascading failure.

We define the CASCADE model of probabilistic cascading failure with the following general features:

1. Multiple identical components, each of which has a random initial load and an initial disturbance.
2. When a component overloads, it fails and transfers some load to the other components.

Property 2 can cause cascading failure: a failure additionally loads other components and some of these other components may also fail, leading to a cascade of failure. The components become progressively more loaded as the cascade proceeds. An initial version of CASCADE was used to examine power transmission system critical loading and power tails in probability distributions of blackout size [8].

1.1. Description of model

The CASCADE model has n identical components with random initial loads. For each component the minimum initial load is $L^{\min}$ and the maximum initial load is $L^{\max}$. For j=1,2,...,n, component j has initial load L_j that is a random variable uniformly distributed in $[L^{\min}, L^{\max}]$. $L_1, L_2, \cdots, L_n$ are independent.

Components fail when their load exceeds L^{fail}. When a component fails, a fixed amount of load P is transferred to each of the components.

To start the cascade, we assume an initial disturbance that loads each component by an additional amount D. Other components may then fail depending on their initial loads L_j and the failure of any of these components will distribute an additional load $P \geq 0$ that can cause further failures in a cascade. The CASCADE model can be defined more precisely in algorithmic form:

Algorithm for CASCADE model

0. All n components are initially unfailed and have initial loads $L_1, L_2, \cdots, L_n$ determined as independent random variables uniformly distributed in $[L^{\min}, L^{\max}]$.

1. Add the initial disturbance D to the load of component j for each $j = 1, ..., n$. Initialize iteration counter i to one.

2. Test each unfailed component for failure: For $j = 1, ..., n$, if component j is unfailed and its load $> L^{\text{fail}}$ then component j fails. Suppose that m_i components fail in this step.

3. If $m_i = 0$, stop; the cascading process ends.

4. If $m_i > 0$, then increment the component loads according to the number of failures m_i: Add $m_i P$ to the load of component j for $j = 1, ..., n$.

5. Increment iteration counter i and go to step 2

A simple example of the CASCADE model with 5 components producing a cascade is shown in Table 1. In this cascade, two components fail in iteration 1, one component fails in iteration 2, and one component fails in iteration 3, and then the cascade ends with a total of 4 components failed. That is, $m_1 = 2, m_2 = 1, m_3 = 1$, and then the cascade ends with a total of $m_1 + m_2 + m_3 = 4$ components failed. Fig. 1 shows the succession of load increases in this cascade labelled by their iteration number.

Table 1. Component loads increasing in a cascade example with 5 components

initial load range $[L^{\min}, L^{\max}] = [0.5, 8.5]$
failure load $L^{\text{fail}} = 9.5$
initial disturbance $D = 3$
load increment $P = 1$
iteration counter i

1	2	3	4	5	component number	i
8	6	7	4	1	initial random load L_j	0
11	9	10	7	4	initial disturbance D added	1
13	11	12	9	6	1 and 3 fail; $2P$ added	2
14	12	13	10	7	2 fails; P added	3
15	13	14	11	8	4 fails; P added	4
		cascade ends with 1,3,2,4 failed				

– integer loads used for convenience only.
– loads above $L^{\text{fail}} = 9.5$ indicate failure.

Since the cascade of failures is only determined by the component loads *relative* to L^{fail}, an equivalent way to visualize the cascade shows the component loads as if fixed at their initial values and shifts the L^{fail} horizontal line downwards as shown in Fig. 2. In this point of view, the line at first shifts downwards by the initial disturbance D and then shifts down by $m_i P$ in subsequent iterations. In Fig. 2, m_i, the previous number of component failures, conveniently appears as the previous number of zeros that the line passes.

Consider the event that exactly r components fail in a cascade governed by the algorithm for the CASCADE

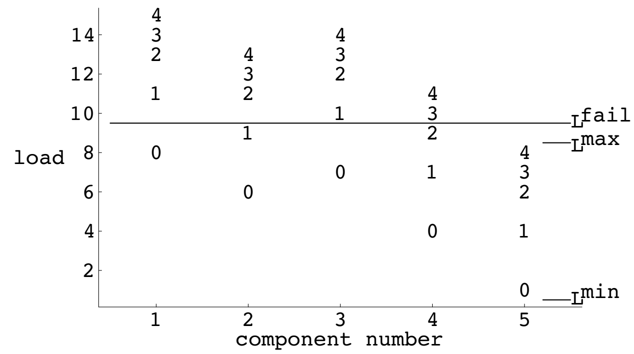

Figure 1. Component loads increasing in cascade example with 5 components. 0 indicates initial load and 1,2,3,4 indicate load at iterations 1,2,3,4 respectively. A component fails when its load exceeds the L^{fail} horizontal line.

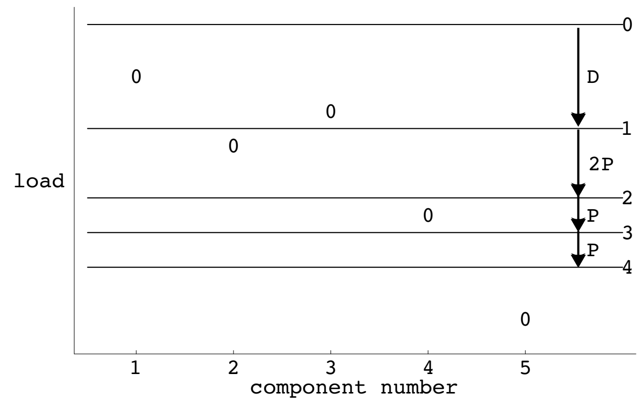

Figure 2. Alternative visualization of cascade example with loads (zeros) shown as if fixed at their initial values and the L^{fail} horizontal line shifting downwards. The line first shifts down by $D = 3$ and subsequently shifts down proportionally to the previous number of failures (number of zeros the line just passed).

model. For convenience of description, we renumber the components in the order they failed. Then an equivalent description of the event that exactly r components fail is that there are positive integers k and $m_1, m_2, \cdots, m_k$ such that equations (1–6) on page 4 apply. We give some examples to explain (2–6). Since component 1 failed on the first iteration after its load L_1 was increased by the initial disturbance D, $L_1 + D > L^{\text{fail}}$. This is an example of (2). After the first iteration, all the loads are increased by $m_1 P$. Since the $m_1 + 1$ component failed on the second iteration, $L_{m_1+1} + m_1 P + D > L^{\text{fail}}$. Moreover, since the $m_1 + 1$ component did not fail on the first iteration, $L_{m_1+1} + D \leq L^{\text{fail}}$. This is an example of (3). The component $r+1$ never fails and the load of component $r+1$ was increased by $D + rP$ by the end of the cascade. Therefore $L_{r+1} + D + rP \le L^{\rm fail}$. This is an example of (6).

The CASCADE model is summarized in table 2.

Table 2. CASCADE model with load range $[L^{\min}, L^{\max}]$ and failure load $L^{\rm fail}$

L_j	initial load of component j; uniformly distributed in $[L^{\min}, L^{\max}]$
P	load increase at each component when failure
D	initial disturbance
n	number of components
$L^{\max}$	max component load
$L^{\min}$	min component load
$L^{\rm fail}$	component failure load

Table 3. Normalized CASCADE model with load range $[0, 1]$ and failure load 1

		formula
ℓ_j	initial load of component j; uniformly distributed in $[0, 1]$	$\frac{L_j - L^{\min}}{L^{\max} - L^{\min}}$
p	load increase at each component when failure	$\frac{P}{L^{\max} - L^{\min}}$
d	initial disturbance	$\frac{D + L^{\max} - L^{\rm fail}}{L^{\max} - L^{\min}}$
n	number of components	
1	max component load	
0	min component load	
1	component failure load	

1.2. Normalized model

Suppose that the initial disturbance D is changed to $D+K$ and that the failure load $L^{\rm fail}$ is changed to $L^{\rm fail}+K$. Adding K to the initial disturbance has the effect of increasing all the loads by K at the beginning of the cascade, but since the failure load is also increased by K, the failure of components is unchanged, and this holds throughout the entire cascading process. That is, the cascade of component failures is unchanged, except that all loads and failure loads are increased in value by K. This conclusion holds for both positive and negative K. Thus we have

Principle 1 *The cascade process is unchanged if the initial disturbance D and the component failure load $L^{\rm fail}$ are incremented by the same amount K.*

Principle 1 also follows from the expressions in (2–6) only depending on the difference $L^{\rm fail} - D$.

It is convenient to normalize the model in two steps: The first step is to change the initial disturbance D to $D + L^{\max} - L^{\rm fail}$ and change the failure load $L^{\rm fail}$ to $L^{\rm fail} + L^{\max} - L^{\rm fail} = L^{\max}$. According to Principle 1 with $K = L^{\max} - L^{\rm fail}$, these changes have no effect on the component failures in the cascading process. These changes are chosen to make $L^{\rm fail}$ coincide with $L^{\max}$.

The second step defines the normalized initial load

$$\ell_j = \frac{L_j - L^{\min}}{L^{\max} - L^{\min}} \tag{12}$$

Then ℓ_j is a random variable uniformly distributed on $[0, 1]$. Moreover, the failure load is $\ell_j = 1$. Let

$$p = \frac{P}{L^{\max} - L^{\min}}, \qquad d = \frac{D + L^{\max} - L^{\rm fail}}{L^{\max} - L^{\min}} \tag{13}$$

Then p is the amount of load increase on any component when one other component fails expressed as a fraction of the load range $L^{\max} - L^{\min}$. Similarly d is the initial disturbance expressed as a fraction of the load range. The CASCADE model and its normalized parameters are summarized in table 3. Also, the normalized conditions for r components to fail are shown in (1) and (7–11).

2. Distribution of number of failed components

This section derives recursive and explicit formulas for the probability distribution of the number of failed components.

Definition: $f(r, d, p, n)$ is the probability that r components fail in the CASCADE model with normalized initial disturbance d, normalized load transfer amount p, and n components.

2.1. Cases $d \le 0$ and $d \ge 1$

When the initial disturbance $d \le 0$, no components fail and

$$f(r, d, p, n) = \left\{ \begin{array}{ll} 1 & ; r = 0 \\ 0 & ; 0 < r \le n \end{array} \right\}, \quad d \le 0 \tag{14}$$

When the initial disturbance $d \ge 1$, all n components fail immediately and

$$f(r, d, p, n) = \left\{ \begin{array}{ll} 0 & ; 0 \le r < n \\ 1 & ; r = n \end{array} \right\}, \quad d \ge 1 \tag{15}$$

2.2. Deriving the recursion for $0 < d < 1$

This subsection assumes throughout that $0 < d < 1$. The initial disturbance d causes immediate failure of the components that have initial load ℓ in $(1 - d, 1]$. Therefore the

Condition for r components to fail

$$r = m_1 + m_2 + \cdots + m_k \text{ with } m_i > 0,\ i = 1, 2, \cdots, k \tag{1}$$

$$L^{\text{fail}} - D < L_1, \cdots, L_{m_1} \tag{2}$$

$$L^{\text{fail}} - D - m_1 P < L_{m_1+1}, \cdots, L_{m_1+m_2} \leq L^{\text{fail}} - D \tag{3}$$

$$L^{\text{fail}} - D - (m_1 + m_2)P < L_{m_1+m_2+1}, \cdots, L_{m_1+m_2+m_3} \leq L^{\text{fail}} - D - m_1 P \tag{4}$$

$$\ldots$$

$$L^{\text{fail}} - D - (m_1 + m_2 + \cdots + m_{k-1})P < L_{m_1+\cdots+m_{k-1}+1}, \cdots, L_r \leq L^{\text{fail}} - D - (m_1 + m_2 + \cdots + m_{k-2})P \tag{5}$$

$$L_{r+1}, \cdots, L_n \leq L^{\text{fail}} - D - rP \tag{6}$$

Normalized version of (2–6):

$$1 - d < \ell_1, \cdots, \ell_{m_1} \tag{7}$$

$$1 - d - m_1 p < \ell_{m_1+1}, \cdots, \ell_{m_1+m_2} \leq 1 - d \tag{8}$$

$$1 - d - (m_1 + m_2)p < \ell_{m_1+m_2+1}, \cdots, \ell_{m_1+m_2+m_3} \leq 1 - d - m_1 p \tag{9}$$

$$\ldots$$

$$1 - d - (m_1 + m_2 + \cdots + m_{k-1})p < \ell_{m_1+\cdots+m_{k-1}+1}, \cdots, \ell_r \leq 1 - d - (m_1 + m_2 + \cdots + m_{k-2})p \tag{10}$$

$$\ell_{r+1}, \cdots, \ell_n \leq 1 - d - rp \tag{11}$$

probability of any component immediately failing is d and the probability of any component not immediately failing is $1-d$. Since the initial loads are independent, the probability that $r = 0$ components fail is

$$f(0, d, p, n) = (1 - d)^n \tag{16}$$

Also, in the case of one component ($n = 1$),

$$f(1, d, p, 1) = d \tag{17}$$

For the rest of the subsection we assume that $1 \leq r \leq n$. If $r \geq 1$ components fail, then a certain number of components k with $1 \leq k \leq r$ must have failed immediately due to the initial disturbance d only. Let E_k be the event that k components fail immediately. Then, since the initial component loads are independent,

$$P[E_k] = \binom{n}{k} d^k (1 - d)^{n-k} \tag{18}$$

Since $E_1, E_2, \cdots, E_r$ are mutually exclusive and collectively exhaustive events, the law of total probability gives

$$\begin{aligned} f(r, d, p, n) &= P[r \text{ components fail}] \\ &= \sum_{k=1}^{r} P[r \text{ components fail}|E_k]\, P[E_k] \end{aligned} \tag{19}$$

We claim that

$$P[r \text{ components fail}|E_k] = f(r - k, \frac{kp}{1 - d}, \frac{p}{1 - d}, n - k) \tag{20}$$

To establish claim (20), consider the $n - k$ components that did not immediately fail under the condition that E_k occurred. Since none of the $n - k$ components failed immediately, their loads ℓ must lie in $[0, 1 - d]$ and are uniformly distributed in $[0, 1 - d]$ (that is, the distribution conditioned on E_k is uniform in $[0, 1 - d]$.) For the $n - k$ components, $L^{\min} = 0$ and $L^{\max} = 1 - d$.

After E_k, each of the $n-k$ components has had a load increase d from the initial disturbance and an additional load increase kp from the immediate failure of k components. Therefore, after E_k, the total initial disturbance for each of the $n - k$ components is $D = kp + d$.

To summarize, after E_k, the failure of the $n - k$ initially unfailed components is governed by the CASCADE model with initial disturbance $D = kp + d$, load transfer $P = p$, $L^{\min} = 0$, $L^{\max} = 1 - d$, $L^{\text{fail}} = 1$, and $n - k$ components.

Normalizing using (13) yields that the failure of the $n-k$ initially unfailed components is governed by the CASCADE model with normalized initial disturbance $\frac{kp}{1-d}$, normalized load transfer $\frac{p}{1-d}$ and $n-k$ components. Therefore the probability that r components fail given E_k is the probability that $r - k$ of the $n - k$ components fail and this probability is given by (20).

Combining (18), (19), (20) yields the recursion

$$\begin{aligned} &f(r, d, p, n) = \\ &\sum_{k=1}^{r} \binom{n}{k} d^k (1 - d)^{n-k} f(r - k, \frac{kp}{1 - d}, \frac{p}{1 - d}, n - k) \\ &;\ 0 < r \leq n, \qquad 0 < d < 1 \end{aligned} \tag{21}$$

2.3. Extended quasibinomial formula for f

This subsection establishes the extended quasibinomial formulas for the distribution f of the number of failed components. The special case of the quasibinomial distribution is also discussed.

We summarize formulas (15–17) and recursion (21):

$$f(r,d,p,n) = \begin{cases} 0 & ;\ 0 \le r < n, \quad d \ge 1 \quad (22) \\ 1 & ;\ r = n, \quad d \ge 1 \quad (23) \\ (1-d)^n & ;\ r = 0, \quad 0 < d < 1 \quad (24) \\ d & ;\ r = 1,\ n = 1, \quad 0 < d < 1 \quad (25) \\ \displaystyle\sum_{k=1}^{r} \binom{n}{k} d^k (1-d)^{n-k} f(r-k, \frac{kp}{1-d}, \frac{p}{1-d}, n-k) & \\ & ;\ r > 0,\ n > 1, \quad 0 < d < 1 \quad (26) \end{cases}$$

Equations $(22-25)$ and recursion (26) define $f(r,d,p,n)$ for all $n \ge 1$ and $d > 0$. It is straightforward to prove by induction that $f(r,d,p,n)$ is a probability distribution for all $n \ge 1$ as detailed in Appendix A.

Consider the expressions

$$f(r,d,p,n) = \binom{n}{r} d(rp+d)^{r-1}(1-rp-d)^{n-r}$$
$$; \qquad 0 \le r \le (1-d)/p,\ r < n \quad (27)$$
$$f(r,d,p,n) = 0\ ; \quad (1-d)/p < r < n,\ r \ge 0 \quad (28)$$
$$f(n,d,p,n) = 1 - \sum_{s=0}^{n-1} f(s,d,p,n) \quad (29)$$

Appendix B proves that (27-29) satisfy $(22-26)$ and hence are explicit formulas for the extended quasibinomial probability distribution.[1]

For $np+d \le 1$ the probability distribution given by (27-29) reduces to (27) alone:

$$f(r,d,p,n) = \binom{n}{r} d(rp+d)^{r-1}(1-rp-d)^{n-r} \quad (30)$$

and this is the quasibinomial distribution introduced by Consul [5] to model an urn problem in which a player makes strategic decisions. Thus (27-29) is an extended quasibinomial distribution in that its range of parameters is extended to allow $np + d > 1$. We will see below that this extended parameter range often describes regimes with a high probability of all components failing.

[1] We comment on how (27-29) apply for $np+d > 1$ or $n > (1-d)/p$. If $(1-d)/p < n < 1 + (1-d)/p$, (27) applies for $r < n$, (28) does not apply, and (29) gives the probability of $r = n$. If $n > 1 + (1-d)/p$, (27) applies only for $r < (1-d)/p$, (28) gives zero probability for $(1-d)/p < r < n$, and (29) gives the probability of $r = n$.

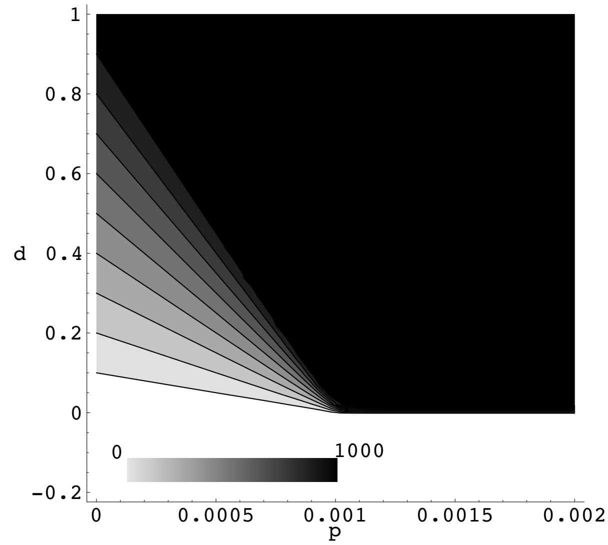

Figure 3. Average number $<r>$ of components failed as a function of p and d for $n = 1000$. Lines are contours of constant $<r>$. White indicates < 100 failures and black indicates > 900 failures.

Consul [5] has derived the mean of the quasibinomial distribution (30) as

$$nd \sum_{r=0}^{n-1} \frac{(n-1)!}{(n-r-1)!} p^r \quad (31)$$

3. Examples of CASCADE distribution

This section shows examples of the qualitative behavior of the CASCADE distribution for $n = 1000$ components and how modeling choices lead to parameterizations of the normalized CASCADE model.

3.1. Average number of failures

One way to summarize the behavior of the CASCADE distribution as the parameters vary is to examine the average number $<r>$ of components failed. A contour plot of $<r>$ as a function of parameters d and p is shown in Fig. 3. For the case $n = 1000$ considered in Fig. 3, the line $np + d = 1$ joins $(p,d) = (1/n, 0) = (0.001, 0)$ to $(p,d) = (0,1)$. The region of Fig. 3 to the right of the line $np + d = 1$ and above the line $d = 0$ is mostly black due to the high probability of all components failing making the average number of failures near 1000 (but note that the corner of the black region near $(p, d) = (0.001, 0)$ is rounded so that a neighborhood of $(p, d) = (0.001, 0)$ is white).

In the region to the left of the line $np + d = 1$ and above the line $d = 0$, the quasibinomial distribution (30) applies and the average number $< r >$ of failures is governed by (31). In particular, $<r>$ is proportional to d.

The following subsections examine the extended quasibinomial distribution in more detail along lines in Fig. 3 such as p constant and d varying, or d constant and p varying.

3.2. Binomial distribution for $p = 0$

One limiting case of the CASCADE model occurs when $p = 0$ and the distribution becomes binomial:

$$f(r, d, 0, n) = \binom{n}{r} d^r (1 - d)^{n-r} \tag{32}$$

It is well known that for large n and small d, (32) is well approximated by the Gaussian distribution with mean nd and variance $nd(1 - d)$.

3.3. Behavior for constant small d

We assume that the normalized initial disturbance is fixed at the small value $d = 0.0001$ and examine how the distribution changes as p increases. For $p = 0$, the distribution is binomial as explained above. The distribution as p increases from zero is shown in Fig. 4. The distribution for $p = 0.0001$ has a tail slightly heavier than binomial but still approximately exponential. The tail becomes heavier as p increases and the distribution for $p = 0.001$ has an approximate power tail over a range of r. Note that $p = 0.001$ approximately satisfies the condition $np + d = 1$. The distribution for $p = 0.002$ has an approximately exponential tail for small r, zero probability of intermediate r, and a probability of 0.08 of all 1000 components failing. (If an intermediate number of components fail, then the cascade always proceeds to all 1000 components failing.)

A previous, restricted version[2] of CASCADE obtained qualitatively similar results by assuming $p = d$ and increasing p. These results were qualitatively similar to simulation results obtained by increasing loading in models of blackouts caused by cascading failure in electric power transmission systems [8].

Increasing the normalized power transfer p may be thought of as strengthening the component interaction that causes cascading failure. The progression in the pdf as p increases from exponential tail to power tail and then to an exponential tail together with a significant probability of total failure is of interest. When the power tail or the significant probability of total failure occur, they have an important impact on the risk of cascading failure [3].

[2]The CASCADE model of [8] in effect assumed $P = D$ and $L^{\max} = L^{\text{fail}} = 1$. Average loading $L = (1 + L^{\min})/2$ was increased by increasing $L^{\min}$, leading to increasing $p(L) = d(L) = D/(2 - 2L)$.

The exponent of the power tail can be approximated as follows. Suppose that $np + d \leq 1$. By writing $\lambda = np$ and $\theta = nd$, and letting $n \to \infty$ and $p \to 0$ and $d \to 0$ in such a way that $\lambda = np$ and $\theta = nd$ are fixed, it can be shown [6] using Stirling's formula that the quasibinomial distribution $f(r, d, p, n)$ may be approximated by

$$g(r, \theta, \lambda) = \theta(r\lambda + \theta)^{r-1} \frac{e^{-r\lambda - \theta}}{r!} \tag{33}$$

which is the generalized Poisson distribution [6].

Now for $r >> 1$, use Stirling's formula to get

$$g(r, \theta, \lambda) \approx \frac{\theta}{(r\lambda + \theta)\sqrt{2\pi r}} \lambda^r e^{(\lambda-1)\left(-r-\frac{\theta}{\lambda}\right)}$$

In the limit, the condition $np + d = 1$ becomes $\lambda = 1$ and setting $\lambda = 1$ gives

$$g(r, \theta, \lambda) \approx \frac{\theta}{(r + \theta)\sqrt{2\pi r}} \tag{34}$$

which has gradient on a plot of log$g$ against log$r$ of $-1.5 + \frac{\theta}{r+\theta} \to -1.5$ as $r \to \infty$.

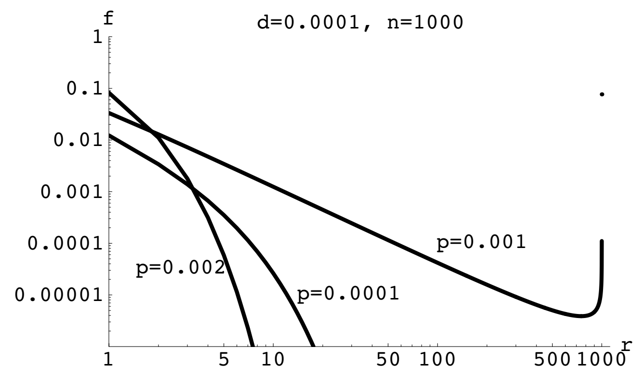

Figure 4. pdf for constant small d and varying p. Probability of no failures is 0.90 for all p. Lines are drawn through the plotted points. $p = 0.002$ has isolated point at $(1000, 0.08)$.

3.4. Effect of reducing maximum load

One way to try to mitigate cascading failure in the CASCADE model is to reduce the maximum component load $L^{\max}$ below L^{fail}. Let $K = L^{\text{fail}} - L^{\max}$ be the amount of the maximum load reduction. Then the system has no failures for any initial disturbance D smaller than K. The normalized amount of maximum load reduction is

$$k = \frac{K}{L^{\max} - L^{\min}} \tag{35}$$

and we will observe the effect of increasing k from zero.

The normalized initial disturbance $d(k)$ and the normalized power transfer $p(k)$ are now functions of k. In particular, according to (13),

$$d(k) = \frac{D - K}{L^{\max} - K - L^{\min}} = \frac{D - (L^{\max} - L^{\min})k}{(L^{\max} - L^{\min})(1-k)}$$

$$p(k) = \frac{P}{L^{\max} - K - L^{\min}} = \frac{P}{(L^{\max} - L^{\min})(1-k)}$$

Hence

$$d(k) = \frac{d(0) - k}{1-k}, \qquad p(k) = \frac{p(0)}{1-k} \tag{36}$$

and the probability of r components failing is

$$f(r, d(k), p(k), n) = f(r, \frac{d(0)-k}{1-k}, \frac{p(0)}{1-k}, n) \tag{37}$$

This shows the effect of reducing the maximum load. If $k > d(0)$ then $d(k) < 0$ and there are no failures (14). For small k, $d(k) \approx d(0) - k(1 - d(0))$ so that the normalized initial disturbance is decreased. However, the normalized load transfer amount $p(k) = \frac{p(0)}{1-k}$ increases. An example is shown in Fig. 5. The probability of no failures ($r = 0$; not shown in Fig. 5) is 0.37 for $k = 0$ and 0.90 for $k = 0.0009$. In this case the mitigation increases the probability of no failures and decreases the probability of failures over a range of $r \geq 1$. However, the increase in $p(k)$ makes the tail of the distribution slightly heavier.

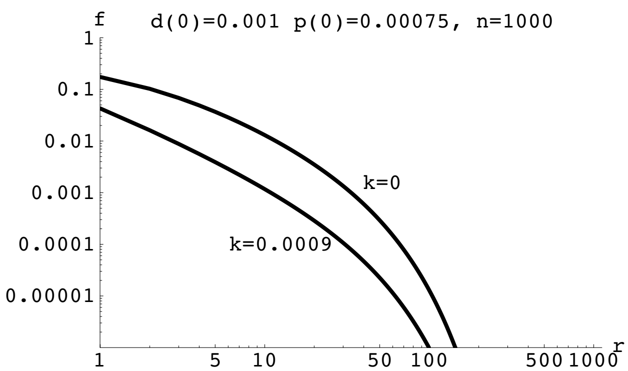

Figure 5. Effect on pdf of increasing k from 0.

3.5. Effect of changing average load

We assume that the initial disturbance $D = 0$ and examine the effect of changing the initial load range by changing the average initial load L defined by

$$L = \tfrac{1}{2}(L^{\max} + L^{\min}) \tag{38}$$

The range of initial load $W = L^{\max} - L^{\min}$. Then the initial load range can be parameterized by L and W:

$$[L^{\min}, L^{\max}] = [L - W/2, L + W/2] \tag{39}$$

We choose to let W be constant and examine the effect of varying L. Then, according to (13), the normalized initial disturbance $d(L)$ becomes an affine function of L and the power transfer p is a constant:

$$d(L) = \frac{L + W/2 - L^{\text{fail}}}{W}, \qquad p = \frac{P}{W} \tag{40}$$

The probability of r components failing is

$$f(r, d(L), p, n) = f(r, \frac{L + W/2 - L^{\text{fail}}}{W}, \frac{P}{W}, n) \tag{41}$$

If $L < L^{\text{fail}} - W/2$, then $d(L) < 0$ and there are no failures (14). If $L > L^{\text{fail}} + W/2$, then $d > 1$ and all components fail (15).

Now consider $L^{\text{fail}} - W/2 \leq L \leq L^{\text{fail}} + W/2$ so that $0 \leq d(L) \leq 1$. In particular, we choose $L^{\text{fail}} = 1$ and $W = 0.2$ so that $0.9 \leq L \leq 1.1$ and choose $p = 0.00075$ and $n = 1000$. Fig. 6 shows the pdf for $d = 0.0005$, 0.05, 0.2, 0.25 and, respectively, $L = 0.9001$, 0.91, 0.94, 0.95. For this case, $np + d = 1$ occurs at $d = 0.25$ and the pdf quickly becomes all components fail with probability one for $d > 0.25$. In the range $0 \leq d \leq 0.25$ the quasibinomial formulas (30) and (31) apply and hence the mean number of failed components increases linearly with d and L.

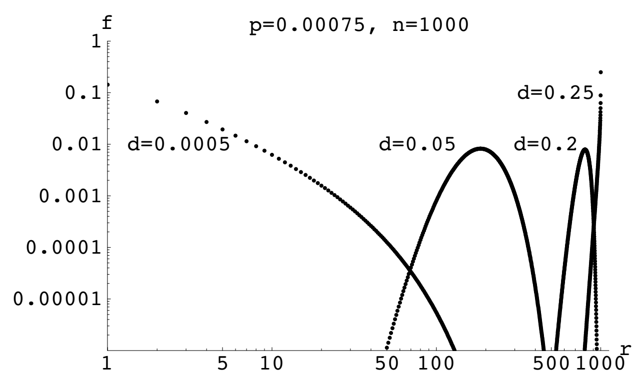

Figure 6. pdf for increasing d.

4. Application to blackouts

CASCADE is much too simple to represent with realism the detailed aspects of a large scale electric power system. However, it is plausible that the general features of CASCADE can be present in large blackouts involving cascading failure of transmission lines and generation.

There are many ways in which power system component failures interact, including via the protection system, redistribution of power flow, dynamics, and operator or planning errors. However, in accordance with the CASCADE model, it is generally true that these interactions are loading dependent and failure of one component tends to stress other components and make their failure more likely. If we focus on cascading failure with interactions due to line overloads and outages via redistribution of power flow (DC power flow with LP dispatch), then results from the OPA model [1, 2, 7] suggest that a variety of operational regimes are possible [1, 2, 4]. Some of these regimes yield distributions of blackout sizes that are qualitatively similar to those observed in CASCADE in the examples of Section 3.

The examples of Section 3 also illustrate how different models and parameterizations can be defined in the CASCADE model and then normalized so that the formulas for the CASCADE distribution apply. For example, the maximum load reduction in Section 3.4 roughly models some of the overall features of applying an n-k security criterion to a system with cascading failure. (If we regard initial component failures as providing an initial disturbance to n remaining components, then the reduction in maximum load roughly approximates in CASCADE the effect of the n-k criterion in that the system can survive a certain number k of initial component failures proportional to K. The analysis in Section 3.4 considers only the n remaining components where $n =$ n-k.)

The CASCADE model can be used to test ideas about cascading failure in complicated contexts such as large power system blackouts. Indeed, as sketched in section 3.3, a restricted version of the CASCADE model has already been used to show power tail behavior qualitatively similar to that obtained in blackout models [8]. Further application of CASCADE to better understand cascading failure in the OPA blackout model [1, 2, 4] and the power system itself is promising future work.

We have started to explore the relation of CASCADE to fiber bundle models of material failure [9]. Instead of regarding initial loads L_j as random and the failure load L^{fail} as fixed, one can regard L^{fail} as random and L_j as fixed. This makes CASCADE more comparable to a model of a single cascade of failure in an increasingly loaded fiber bundle. Global load sharing fiber bundle models redistribute the load of a failed fiber equally to all unfailed components so that, in contrast to CASCADE, the amount of load transferred to each unfailed component tends to increase as the cascade proceeds. However, Kloster et al. hold load transfer fixed when analyzing a single cascade of failure and obtain the generalized Poisson distribution (33) in equation (11) of [9].

5. Conclusions

We define and explain the CASCADE model of loading-dependent cascading failure of a system of identical components. Formulas for the pdf of the number of failed components are derived using a recursion. These formulas quantify the effect on the pdf of the initial disturbance and the amount of load transfer when a component fails. Thus features of loading-dependent cascading failure are captured in a probabilistic model with an analytic solution.

The pdf of the number of failed components is an extended quasibinomial distribution; that is, a quasibinomial distribution with an extended parameter range. CASCADE appears to be a new model that yields as a special case the quasibinomial distribution. The recursion offers a simple way to derive the distribution that avoids complicated algebra or combinatorics.

The tail of the pdf can range from exponential to an approximately power tail, or there can be a high probability of total failure (all components fail). Compared to systems with no cascading dependency, the power tail and total failure regimes show greatly increased probabilities of most of the components failing. Indeed under these regimes, the risk of catastrophic failure from cascading rare events can be comparable to or in excess of the risk of the more frequent smaller disruptions. (Although catastrophic failures are rarer, they have huge costs and their risk is the product of frequency and cost.) For a more detailed discussion of this claim in the context of blackout risk due to power tail regimes in distributions of blackout size see [3].

For future work, we look forward to systematic description of CASCADE model properties and its application to understand cascading failure in large interconnected complex systems. We hope that CASCADE can contribute to the goal of approximate quantification of the risks of catastrophic infrastructure failure due to cascading rare events and the mitigation of these risks.

6. Acknowledgements

I. Dobson and B.A. Carreras gratefully acknowledge coordination of this work by the Consortium for Electric Reliability Technology Solutions and funding in part by the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Power Technologies, Transmission Reliability Program of the U.S. Department of Energy under contract 9908935 and Interagency Agreement DE-A1099EE35075 with the National Science Foundation. I. Dobson, D.E. Newman, and B.A. Carreras gratefully acknowledge support in part from NSF grants ECS-0214369 and ECS-0216053. Part of this research has been carried out at Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U.S. Department of Energy under

contract number DE-AC05-00OR22725.

References

[1] B.A. Carreras, V.E. Lynch, M. L. Sachtjen, I. Dobson, D.E. Newman, Modeling blackout dynamics in power transmission networks with simple structure, 34th Hawaii Intl. Conf. System Sciences, Maui, Hawaii, Jan. 2001.

[2] B.A. Carreras, V.E. Lynch, I. Dobson, D.E. Newman, Dynamics, criticality and self-organization in a model for blackouts in power transmission systems, 35th Hawaii Intl. Conf. System Sciences, Hawaii, Jan. 2002.

[3] B.A. Carreras, V.E. Lynch, D.E. Newman, I. Dobson, Blackout mitigation assessment in power transmission systems, 36th Hawaii Intl. Conf. System Sciences, Hawaii, Jan. 2003.

[4] B.A. Carreras, V.E. Lynch, I. Dobson, D.E. Newman, Critical points and transitions in an electric power transmission model for cascading failure blackouts, to appear in Chaos, scheduled for December 2002.

[5] P.C. Consul, A simple urn model dependent upon predetermined strategy, Sankhyā: The Indian Journal of Statistics, Series B, vol. 36, Pt. 4, 1974, pp. 391-399.

[6] P.C. Consul, Generalized Poisson distributions, Dekker, NY 1989.

[7] I. Dobson, B.A. Carreras, V. Lynch, D.E. Newman, An initial model for complex dynamics in electric power system blackouts, 34th Hawaii Intl. Conf. System Sciences, Maui, Hawaii, Jan. 2001.

[8] I. Dobson, J. Chen, J.S. Thorp, B.A. Carreras, D.E. Newman, Examining criticality of blackouts in power system models with cascading events, 35th Hawaii Intl. Conf. System Sciences, Hawaii, Jan. 2002

[9] M. Kloster, A. Hansen, P.C. Hemmer, Burst avalanches in solvable models of fibrous materials, Physical Review E, vol. 56, no. 3, September 1997.

A. f is a probability distribution

This appendix proves by induction that f defined by (22−26) defines a probability distribution for $n \geq 1, d > 0$.

If $d \geq 1$, then (22), (23) define a probability distribution.

Now assume $0 < d < 1$. f nonnegative in (22–25) and recursion (26) imply that f is nonnegative for all $n \geq 1$.

In the case $n = 1$, (24), (25) give $\sum_{r=0}^{1} f(r, d, p, 1) = 1$. If $n > 1$ and $\sum_{r=0}^{m} f(r, d, m, p) = 1$ for $m < n$, then, according to (24) and (26),

$$\sum_{r=0}^{n} f(r,d,p,n) = (1-d)^n + \sum_{r=1}^{n}\sum_{k=1}^{r}\binom{n}{k} d^k(1-d)^{n-k} f(r-k, \frac{kp}{1-d}, \frac{p}{1-d}, n-k)$$
$$= (1-d)^n + \sum_{k=1}^{n}\sum_{r=k}^{n}\binom{n}{k} d^k(1-d)^{n-k} f(r-k, \frac{kp}{1-d}, \frac{p}{1-d}, n-k)$$
$$= \sum_{k=0}^{n}\binom{n}{k} d^k(1-d)^{n-k} = 1$$

In the special case $np + d \leq 1$, f reduces to the quasibinomial distribution (30) and we see that the recursive definition of f allows a proof that (30) is a probability distribution that is more elementary than the proof in [5].

B. Extended quasibinomial formulas satisfy the recursion

This appendix proves that extended quasibinomial formulas (27–29) satisfy recursion (22−26) for $d > 0$.

If $d > 1$, (27) does not apply (because $d > 1$ and $r < (1-d)/p \Rightarrow r < 0$) and (28), (29) satisfy (22), (23).

If $d = 1$, (27) (for $r = 0$) and (28) (for $0 < r < n$) verify (22). Then (29) verifies (23).

Now assume $0 < d < 1$. For $r = 0$, (27) satisfies (24). For $r = 1$ and $n = 1$, (29) satisfies (25). Moreover, for $r > 0$ and $n > 1$, we claim that (27–29) satisfies (26).

To prove this claim, we first assume that $0 < r \leq (1-d)/p$ and $n > 1$. Since

$$r \leq \frac{1-d}{p} \Rightarrow r-k \leq \frac{1-\frac{kp}{1-d}}{\frac{p}{1-d}}, \tag{42}$$

(27) applies to all the instances of f in the right hand side of (26) so that these instances of f may be written as

$$f(r-k, \frac{kp}{1-d}, \frac{p}{1-d}, n-k) = \binom{n-k}{r-k}\frac{kp}{1-d}\Big(\frac{rp}{1-d}\Big)^{r-k-1}\Big(1-\frac{rp}{1-d}\Big)^{n-r} \quad k = 1, 2, \cdots, r \tag{43}$$

(Note that $(r-k)\frac{p}{1-d} + \frac{kp}{1-d} = \frac{rp}{1-d}$.) Substitute (43) into the right hand side of (26) to get

$$\begin{aligned}
&\sum_{k=1}^{r}\binom{n}{k} d^k(1-d)^{n-k} \binom{n-k}{r-k}\frac{kp}{1-d}\Big(\frac{rp}{1-d}\Big)^{r-k-1}\Big(1-\frac{rp}{1-d}\Big)^{n-r}\\
&= \binom{n}{r}\sum_{k=1}^{r}\binom{r}{k}\frac{k}{r} d^k (rp)^{r-k}(1-rp-d)^{n-r}\\
&= \binom{n}{r}(1-rp-d)^{n-r}\sum_{k=1}^{r}\binom{r-1}{k-1} d^k (rp)^{r-k}\\
&= \binom{n}{r}(1-rp-d)^{n-r} d(rp+d)^{r-1}
\end{aligned} \tag{44}$$

According to (27), the left hand side of (26) also evaluates to (44).

Now we assume that $(1-d)/p < r < n$. Since

$$\frac{1-d}{p} < r < n \Rightarrow \frac{1-\frac{kp}{1-d}}{\frac{p}{1-d}} < r-k < n-k,$$

(28) implies that

$$f(r-k, \frac{kp}{1-d}, \frac{p}{1-d}, n-k) = 0\,, \quad k = 1, 2, \cdots, r$$

and hence (26) is verified.

Now we assume that $r = n$. Since $r = n \Rightarrow r - k = n - k$, (29) implies that

$$f(n-k, \frac{kp}{1-d}, \frac{p}{1-d}, n-k) = \\ 1 - \sum_{s=0}^{n-k-1} f(s, \frac{kp}{1-d}, \frac{p}{1-d}, n-k), \quad k = 1, 2, \cdots, r$$

Substitution into the right hand side of (26) gives

$$1 - (1-d)^n - \sum_{s=1}^{n-1} \sum_{k=1}^{s} \binom{n}{k} d^k (1-d)^{n-k} \\ f(s-k, \frac{kp}{1-d}, \frac{p}{1-d}, n-k) = 1 - \sum_{s=0}^{n-1} f(s, d, p, n) \quad (45)$$

where the last step uses the result established above that (27) satisfies (26). According to (29), the left hand side of (26) also evaluates to (45).