Thirty-ninth Hawaii International Conference on System Sciences, January 2006, Kauai, Hawaii. © 2006 IEEE

An estimator of propagation of cascading failure

Ian Dobson, Kevin R. Wierzbicki, Benjamin A. Carreras, Vickie E. Lynch, David E. Newman

***Abstract*— We suggest a statistical estimator to measure the extent to which failures propagate in cascading failures such as large blackouts. The estimator is tested on a saturating branching process model of cascading failure and on failure data generated by the OPA simulation of cascading blackouts. The estimator is a standard estimator for the branching process parameter modified so as to avoid saturation effects. We discuss the statistical effectiveness of the estimator and show how estimating the failure propagation and the initial failures leads to estimates of the distribution of the numbers of cascading failures. The estimator is derived from a simple and high-level branching process model of cascading failure. We discuss how branching process approximations and in particular the propagation parameter may arise in several different models, including models of interacting infrastructures.**

I. INTRODUCTION

Cascading failure is the main way that blackouts become widespread. For example, the August 2003 blackout spread to a sizable region of Northeastern America by cascading [28]. A traditional way to address cascading failure is to limit the failures that initiate cascades. To pursue the complementary approach of limiting the propagation of failures after they are initiated, we first need to be able to quantify how much failures propagate from data. This paper suggests such an estimator and tests it on a branching process model of cascading failure and on the OPA model [4] of cascading failure blackouts. The estimator may be applicable to cascading failure of other large, interconnected infrastructures.

We model the growth of blackout failures using a branching process and then estimate the branching process parameter λ that measures the extent to which failures propagate. Branching process models are an obvious choice of stochastic model to capture the gross features of cascading blackouts because they have been developed and applied to other cascading processes such as genealogy, epidemics and cosmic rays [21]. The first suggestion to apply branching processes to blackouts appears to be in [15] and subsequent applications to blackouts appear in [14], [16].

I. Dobson and K.R. Wierzbicki are with the ECE department, University of Wisconsin, Madison WI 53706 USA; email dobson@engr.wisc.edu. B.A. Carreras and V.E. Lynch are with Oak Ridge National Laboratory, Oak Ridge TN 37831 USA; email: carrerasba@ornl.gov. D.E. Newman is with the Physics department, University of Alaska, Fairbanks AK 99775 USA; email ffden@uaf.edu. We gratefully acknowledge support in part from NSF grants ECS-0214369 and ECS-0216053. This paper is an account of work sponsored in part by the Power Systems Engineering Research Center (PSERC). 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.

Thirty-ninth Hawaii International Conference on System Sciences, January 2006, Kauai, Hawaii. © 2006 IEEE.

We now explain how branching processes can be useful approximations to some of the gross features of cascading blackouts. Our idealized probabilistic model of cascading failure [17] describes a general cascading process in which component failures weaken and further load the system so that subsequent failures are more likely. We have shown that this cascade model and variants of it can be well approximated by a Galton-Watson branching process with each failure giving rise to a Poisson distribution of failures in the next stage [15], [13]. Moreover, some features of this cascade model are consistent with results from cascading failure simulations [6], [14], [25]. All of these models can show criticality and power law regions in the distribution of failure sizes or blackout sizes consistent with NERC data [8]. The distribution of the number of high voltage transmission lines lost in North American contingencies from 1965 to 1985 [1] also has a heavy tail distribution [9]. Initial work fitting branching process models to observed blackout data is in [16].

The Galton-Watson branching process model [21], [2] gives a way to quantify the propagation of cascading failures with a parameter λ. In the Galton-Watson branching process the failures are produced in stages. The process starts with Z_0 failures at stage zero to represent the initial disturbance. The failures in each stage independently produce further failures in the next stage according to an probability distribution with mean λ called the offspring distribution. That is, each failure in each stage produces an average of λ failures in the next stage.

The eventual behavior of the branching process is governed by the parameter λ. In the subcritical case of $\lambda < 1$, the failures will die out (that is, reach and remain at zero failures at some stage) and the mean number of failures in each stage decreases exponentially. In the supercritical case of $\lambda > 1$, although it possible for the process to die out, often the failures increase exponentially until the system size or saturation effects are encountered.

Saturation is thought to be a significant effect because many observed cascading blackouts do not proceed to the entire interconnection blacking out. The reasons for this may well include inhibition effects such as load shedding relieving system stress, or successful islanding, that apply in addition to the stochastic variation that will limit some cascading sequences. Understanding and modeling these inhibition or saturation effects is important, but in this paper we avoid describing the saturation process by only estimating the propagation of failures before saturation is encountered.

At the critical case of $\lambda = 1$, the branching process has a power law distribution of number of failures with exponent

-1.5. A corresponding power law region can be observed in the distributions of number of failures in the cascading failure model [17] and in the distribution of blackout size in blackout models [4], [10], [25] when the system has a particular loading called the critical loading. The implications for blackout risk of the power law region are that the risk of large blackouts is approximately the same or even exceeding the risk of small blackouts [5]. This observation justifies the study of large blackouts; an exponential tail in the distribution of blackout size would imply that large blackouts have negligible risk and that a risk-based analysis would ignore large blackouts. Moreover, at criticality the mean blackout size starts to increase more rapidly and above criticality there is an increasing risk of large blackouts. The terminology of criticality comes from statistical physics and does not, at this stage of knowledge of overall blackout risk, necessarily imply improper power system operation. Indeed there is some evidence that power systems may organize themselves to near critical loading in response to strong societal forces balancing economic use of the transmission system and reliability [7], [8].

One requirement on the failure data in order to estimate λ is that the failures be grouped into stages. The estimator we propose depends on the number of failures in the stages and particularly on the total number of failures and the number of failures in the initial and final stages. Many cascading failure simulations naturally produce failures in stages as the simulation iterates. However, if the method is applied to real data, the problem of grouping the data into stages must be addressed (see [16] for initial work grouping failures together according to closeness in time).

One direct way to estimate the probability distribution of number of line failures at a given level of system loading is simply to run the simulation or record real blackout data until sufficient data is accumulated to estimate the probability distribution of blackout sizes. This is straightforward but requires a large number of simulations or an impractically long observation time. If the distribution of line failures is near criticality and it has a power law character, the probability distribution requires many observations to determine its form for the larger blackouts. For example, it can take of the order of 1000 to 10000 real or simulated blackouts to accurately estimate the probability of the larger number of line failures in the near critical case. The near critical case is pertinent because there is some evidence and explanation indicating that the North American power transmission system is designed and operated near criticality [8], [7]. However, [8], [7] mainly apply to the distribution of energy unserved and the exact relationship between an "external", societal impact measure of blackout size such as energy unserved and an "internal" measure of blackout size such as number of lines failed remains unclear. On the other hand, empirical data for North American line outages [1] has a heavy tail that is fairly close to a power law [9].

If one assumes a branching process model, one can predict the probability distribution of the number of failures from the initial distribution of failures and the estimate of λ using an analytic formula (an example is shown in section III-D). We test the estimate for λ on line failure data from the OPA cascading failure simulation by predicting the probability distribution of the number of line failures using the estimate of λ and comparing this distribution to the probability distribution of line failures observed in OPA. This approach of estimating the distribution of line failures by first estimating λ has the potential to be much faster because λ is the mean of the offspring distribution that generates the branching process and the offspring distribution does not have heavy tails.

II. Branching process with saturation

This section describes the details of the main branching process model used in this paper. Suppose that there are N identical components and all components are initially unfailed. The process saturates when $S \leq N$ components fail. Component failures occur in stages with Z_n the number of failures in stage n and Y_n the total number of failures up to and including stage n.

$$Y_n = Z_0 + Z_1 + Z_2 + ... + Z_n \tag{1}$$

There are a deterministic number Z_0 of initial failures. Each of the Z_n failures in stage n independently causes a further number of failures in stage $n+1$ according to a Poisson distribution with mean λ, except that if the total number of failures exceeds S, then the total number of failures is limited to S. That is, the jth failure in stage n causes $Z_{n+1}^{[j]}$ failures in stage $n+1$ according to the Poisson distribution and the total number of failures in stage $n+1$ is

$$Z_{n+1} = \min\left\{Z_{n+1}^{[1]} + Z_{n+1}^{[2]} + \cdots + Z_{n+1}^{[Z_n]},\, S - Y_n\right\}, \tag{2}$$

where $Z_{n+1}^{[1]}, Z_{n+1}^{[2]}, \cdots, Z_{n+1}^{[Z_n]}$ are independent. (A different form of saturation is described in [15], [14].) The total number of failures Y is distributed according to a saturating Borel-Tanner distribution:

$$P[Y=r] = \begin{cases} Z_0\lambda(r\lambda)^{r-Z_0-1}\dfrac{e^{-r\lambda}}{(r-Z_0)!}; & Z_0 \leq r < S \\ 1 - \displaystyle\sum_{s=Z_0}^{S-1} Z_0\lambda(s\lambda)^{s-Z_0-1}\dfrac{e^{-s\lambda}}{(s-Z_0)!}; & r = S \end{cases} \tag{3}$$

Some simulations such as OPA may produce a random number Z_0 of initial failures. If the initial failures are independent and of equal probability p_0, then they are distributed according to a binomial distribution. If the number of components is large and p_0 is small, the distribution of initial failures is approximately Poisson with mean $\theta = Np_0$ and in this case the total number of failures is distributed according to a saturating generalized Poisson distribution:

$$P[Y=r] = \begin{cases} \theta(r\lambda+\theta)^{r-1}\frac{e^{-r\lambda-\theta}}{r!}; & 0 \leq r < S \\ 1 - \displaystyle\sum_{s=0}^{S-1}\theta(s\lambda+\theta)^{s-1}\dfrac{e^{-s\lambda-\theta}}{s!}; & r = S \end{cases} \tag{4}$$

We are interested in the total number of failures conditioned on there being a nonzero number of failures and this is distributed according to

$$P[Y=r]=\begin{cases}\theta(r\lambda+\theta)^{r-1}\frac{e^{-r\lambda-\theta}}{r!(1-e^{-\theta})}; & 1\le r<S\\ 1-\sum_{s=1}^{S-1}\theta(s\lambda+\theta)^{s-1}\dfrac{e^{-s\lambda-\theta}}{s!(1-e^{-\theta})}; & r=S\end{cases}\tag{5}$$

If there is an arbitrary distribution of nonzero initial failures $P[Z_0 = z_0]$ for $z_0 = 1, 2, 3, ...$, then the total number of failures is distributed according to a combination of the saturating Borel-Tanner distributions:

$$P[Y\!=\!r]=\begin{cases}\sum_{z_0=1}^{r}P[Z_0\!=\!z_0]z_0\lambda(r\lambda)^{r-z_0-1}\dfrac{e^{-r\lambda}}{(r-z_0)!}; r<S\\ 1-\sum_{s=1}^{S-1}P[Y\!=\!s]; \qquad r=S\end{cases}\tag{6}$$

III. Estimators of propagation $\widehat{\lambda}_s$ and mean initial failures $\widehat{\theta}$

A. Definition of $\widehat{\lambda}_s$

We suppose that the cascading failure simulation produces $Z_0 > 0$ initial failures in stage 0 and then iterates to produce further numbers of failures Z_1, Z_2, Z_3,... in stages 1,2,3,... respectively. The assumption of $Z_0 > 0$ implies that all statistics are conditioned on the start of a cascade. The simulation is run K times to produce K independent realizations of the cascade. The failures in the kth run are written as $Z_0^{(k)}$, $Z_1^{(k)}$, $Z_2^{(k)}$, $Z_3^{(k)}$,... . The simulation results can be tabulated as follows:

$$\begin{array}{cccccc} & \text{stage 0} & \text{stage 1} & \text{stage 2} & \text{stage 3} & \cdots\\ \text{run 1} & Z_0^{(1)} & Z_1^{(1)} & Z_2^{(1)} & Z_3^{(1)} & \cdots\\ \text{run 2} & Z_0^{(2)} & Z_1^{(2)} & Z_2^{(2)} & Z_3^{(2)} & \cdots\\ \text{run 3} & Z_0^{(3)} & Z_1^{(3)} & Z_2^{(3)} & Z_3^{(3)} & \cdots\\ . & . & . & . & . & \\ . & . & . & . & . & \\ \text{run } K & Z_0^{(K)} & Z_1^{(K)} & Z_2^{(K)} & Z_3^{(K)} & \cdots\end{array}\tag{7}$$

Define the cumulative number of failures in run k up to and including stage n as

$$Y_n^{(k)} = Z_0^{(k)} + Z_1^{(k)} + Z_2^{(k)} + ... + Z_n^{(k)}\tag{8}$$

Each run has a stage at which the number of failures is zero and remains zero for all subsequent stages, either because the cascade dies out, or all N of the components in the system have failed. The number of failures at which saturation occurs is $S \le N$. Define

$$s(k) = \max\{\, n \mid Y_n^{(k)} < S \ \text{ and } \ Z_{n-1}^{(k)} > 0\,\}\tag{9}$$

Then $s(k)$ is either the first stage at which there are zero failures or the last stage before saturation.

We define the estimator of λ as

$$\widehat{\lambda}_s = \frac{\sum_{k=1}^K\left(Z_1^{(k)}+Z_2^{(k)}+...+Z_{s(k)}^{(k)}\right)}{\sum_{k=1}^K\left(Z_0^{(k)}+Z_2^{(k)}+...+Z_{s(k)-1}^{(k)}\right)}\tag{10}$$

$$= \frac{\sum_{k=1}^K\left(Y_{s(k)}^{(k)}-Z_0^{(k)}\right)}{\sum_{k=1}^K Y_{s(k)-1}^{(k)}}\tag{11}$$

The appendix derives formulas for the mean and variance of $\widehat{\lambda}_s$.

B. The standard estimator $\widehat{\lambda}_n$

The estimator $\widehat{\lambda}_s$ is a variant of the well known estimator:

$$\widehat{\lambda}_n = \frac{\sum_{k=1}^K\left(Y_n^{(k)}-Z_0^{(k)}\right)}{\sum_{k=1}^K Y_{n-1}^{(k)}}\tag{12}$$

The difference is that $\widehat{\lambda}_n$ uses a fixed number of stages n for each run whereas $\widehat{\lambda}_s$ only uses information from stages before saturation. $\widehat{\lambda}_n$ is a maximum likelihood estimator [11], [24], [19]. Moreover, for fixed n and as $K \to \infty$,

(i) If $\lambda < \infty$, $\widehat{\lambda}_n \to \lambda$ almost surely and $E\widehat{\lambda}_n \to \lambda$; that is, $\widehat{\lambda}_n$ is an strongly consistent and asymptotically unbiased estimate of λ;

(ii) if the offspring distribution variance $\sigma^2 < \infty$, then $\widehat{\lambda}_n$ has an asymptotically normal distribution

$$N\left(\lambda,\sigma^2\left[K\sum_{i=0}^{n-1}\lambda^i\right]^{-1}\right)=N\left(\lambda,\frac{\sigma^2(\lambda-1)}{K(\lambda^n-1)}\right)$$

Note that the standard deviation of $\widehat{\lambda}_n \sim 1/\sqrt{K}$.

C. Estimator of mean initial failures $\widehat{\theta}$

K samples of the initial failures are given by $Z_0^{(1)}$, $Z_0^{(2)}$,..., $Z_0^{(K)}$. We fit the initial failures with a Poisson distribution with parameter θ conditioned on nonzero number of failures:

$$P[Z_0 = r] = \frac{e^{-\theta}}{1-e^{-\theta}}\frac{\theta^r}{r!}, \qquad r = 1, 2, 3, ...\tag{13}$$

Let the sample mean of the initial failures be

$$\overline{Z_0} = \frac{1}{K}\sum_{k=1}^K Z_0^{(k)}.\tag{14}$$

Then both maximum likelihood and method of moments estimation of θ in (13) yields an estimate $\widehat{\theta}$ satisfying

$$\overline{Z_0} = \frac{\widehat{\theta}}{1-e^{-\widehat{\theta}}}.\tag{15}$$

D. Estimating distribution of number of failures from λ

If we assume that the cascading process is approximated by a branching process with a Poisson offspring distribution, then we can obtain the probability distribution of the number of failures from λ. For example, assuming one initial failure, Figure 1 shows the probability distributions obtained for $S = N = 1000$ and three values of λ. For subcritical $\lambda = 0.6$ well below 1, the probability of large number of failures of size near S is exponentially small. The probability of exactly S failures is also very small. As λ increases in the subcritical range $\lambda < 1$, the mechanism by which there develops a significant probability of large number of failures near S is that the power law region extends towards S failures [14]. For near critical $\lambda \approx 1$, there is a power law region extending to S failures. For supercritical $\lambda = 1.2$, there is an exponential tail. This again implies that the probability of large number of failures $< S$ is exponentially small. However there is a significant probability of exactly S failures that increases with λ.

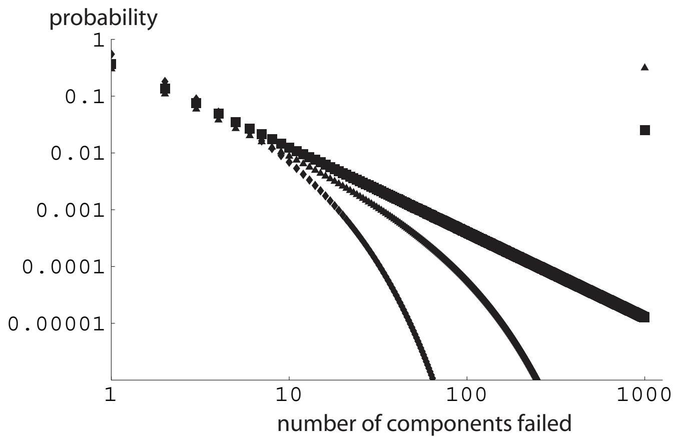

Fig. 1. Log-log plot of PDF of total number of failures in branching process model for three values of λ. $\lambda = 0.6$ is indicated by the diamonds. $\lambda = 1.0$ (criticality) is indicated by the boxes. $\lambda = 1.2$ is indicated by the triangles.

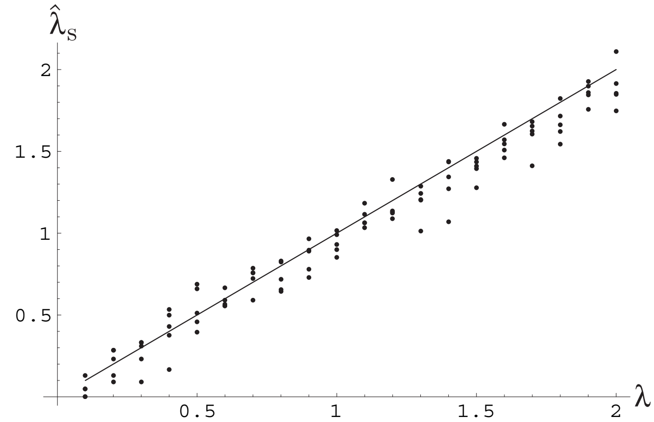

Fig. 2. $\widehat{\lambda}_s$ plotted against λ from 20 runs of saturating branching process with saturation $S = 20$. 5 computations of $\widehat{\lambda}_s$ are shown for each value of λ to indicate the scatter of the results around the line of slope one.

IV. Results

We first consider the statistical performance of the estimator $\widehat{\lambda}_s$ on an ideal saturating branching process and then test the use of analytic branching process formulas including the estimator $\widehat{\lambda}_s$ in predicting the distribution of line failures from the OPA simulation.

A. Testing on saturating branching process

The estimator $\widehat{\lambda}_s$ was tested on the saturating branching process with Poisson offspring distribution and $Z_0 = 1$ for various specified values of λ and saturation S. Figures 2 and 3 show how well $\widehat{\lambda}_s$ matches λ for saturations $S = 20$ and $S = 100$. The necessity for $\widehat{\lambda}_s$ taking account of saturation is shown by the corresponding results for $\widehat{\lambda}_n$ in Figure 4 (compare to Figure 2). For subcritical λ well below $\lambda = 1$, $\widehat{\lambda}_n$ and $\widehat{\lambda}_s$ give the same performance because it is likely that cascades die out before reaching saturation.

We determined the bias and variance of $\widehat{\lambda}_s$ numerically by computing $\widehat{\lambda}_s$ 1000 times and computing the sample mean and standard deviation of $\widehat{\lambda}_s$. The standard deviation $\sigma(\widehat{\lambda}_s)$ was found to decrease $\sim 1/\sqrt{K}$ as K increases.

For saturation $S = 20$, $0 < \lambda < 2$, and number of runs $10 \leq K \leq 1000$, $\widehat{\lambda}_s$ underestimated λ with a bias less than 0.1:

$$-0.1 < \mu(\widehat{\lambda}_s) - \lambda \leq 0$$

and

$$\sigma(\widehat{\lambda}_s) \leq \frac{0.6}{\sqrt{K}}.$$

For example, $K = 20$ runs gives $\sigma(\widehat{\lambda}_s) \leq 0.13$.

For saturation $S = 100$, $0 < \lambda < 2$, and number of runs $10 \leq K \leq 150$, $\widehat{\lambda}_s$ underestimated λ with a bias less than 0.07:

$$-0.07 < \mu(\widehat{\lambda}_s) - \lambda \leq 0$$

and

$$\sigma(\widehat{\lambda}_s) \leq \frac{0.5}{\sqrt{K}}.$$

For example, $K = 20$ runs gives $\sigma(\widehat{\lambda}_s) \leq 0.11$.

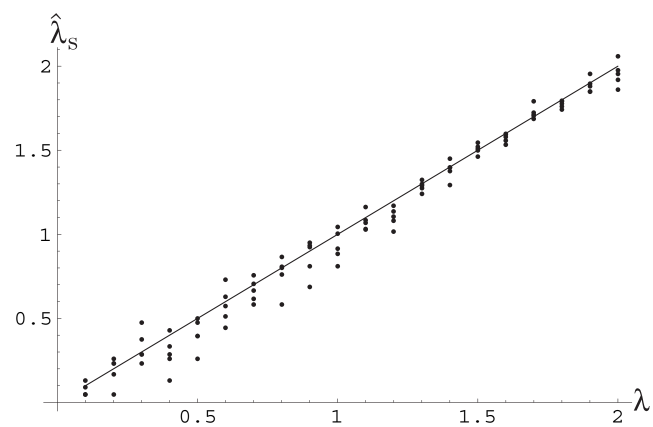

Fig. 3. $\widehat{\lambda}_s$ plotted against λ from 20 runs of saturating branching process with saturation $S = 100$. 5 computations of $\widehat{\lambda}_s$ are shown for each value of λ to indicate the scatter of the results around the line of slope one.

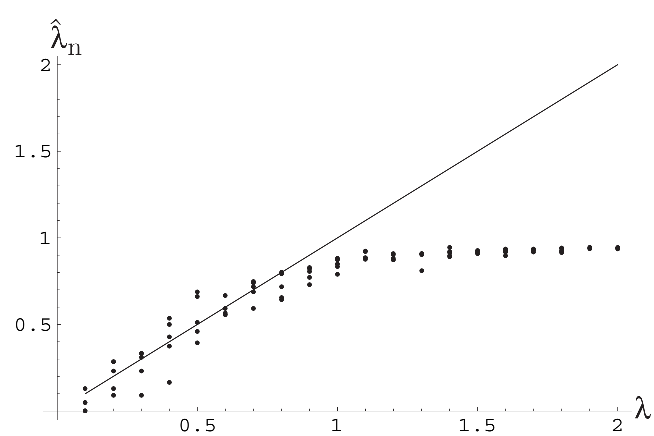

Fig. 4. Standard estimator $\widehat{\lambda}_n$ plotted against λ from 20 runs of saturating branching process with saturation $S = 20$. 5 computations of $\widehat{\lambda}_n$ are shown for each value of λ to indicate the scatter of the results around the line of slope one.

B. Computing $\widehat{\lambda}_s$ from OPA results and predicting distribution of line failures

The OPA model produces cascading line failures in stages resulting from a random initial set of line failures. (Here the intent of the term "line failure" includes line outages in which the line is intact but temporarily unavailable to transmit power as well as line outages in which the line is damaged.) The power transmission system is modelled using DC load flow and LP generator dispatch and cascading line overloads and failures are represented [4]. The power system is assumed to be fixed with no transmission line upgrade process.

For each case considered, OPA was run so as to produce 5000 cascading failures with a nonzero number of line failures. These 5000 runs yield line failure data in the form (7). All our statistics are conditioned on a nonzero number of line failures. Then $\widehat{\lambda}_s$ and $\widehat{\theta}$ were obtained using equations (11) and (15). The empirical distribution of the initial failures Z_0 was also estimated. The number of runs and the resulting data set are large enough that the standard deviation of $\widehat{\lambda}_s$ is negligible. The objective is to test an accurate $\widehat{\lambda}_s$ by evaluating its ability to predict the probability distribution of line outages. (The influence on statistical accuracy of the small number of runs desirable in practice is evaluated in section IV-A.)

For each case, the probability distribution of line failures was predicted using two methods and compared to the empirical probability distribution of line failures. The first method used the estimates $\widehat{\lambda}_s$ and $\widehat{\theta}$ in the saturating generalized Poisson distribution formula (5) to predict the distribution. This tests the validity of the branching process model of initial failures and propagation underlying (5) and the estimates $\widehat{\lambda}_s$ and $\widehat{\theta}$. The second method used the empirical distribution of initial failures and the estimate $\widehat{\lambda}_s$ in the formula (6) to predict the distribution. This tests the validity of the branching process propagation of failures and the estimate $\widehat{\lambda}_s$.

The first three cases used the IEEE 118 bus system at average load levels of 0.9, 1.0, and 1.3 times the base case loading. (The OPA parameters (explained in [4]) are $\gamma = 1.67$, $p_0 = 0.0001$ and $p_1 = 1$ and (11) is evaluated with saturation $S = 15$ lines.) The results are shown in Figures 5-7 and Table I. The matches in Figures 5 and 6 are very good. Note that the OPA simulation of cascading failure producing the empirical probability distributions of line failures is substantially more complicated than the simple analytic probability distribution formulas (5) and (6) used to predict the same distributions. Figure 7 shows a case with a large initial disturbance (the average number of lines initially failed is $\widehat{\theta} \approx 12$). In this case the match for the generalized Poisson distribution is good and the match for the formula (6) using the empirical distribution of initial failures is poorer, especially for the smaller number of failures. We suspect that the poorer match is due to some of the cascades reaching saturation because of the large initial disturbance.

The last two cases used the 190 bus tree-like test system [4] at average load levels of 1.0 and 1.2 times the base case loading. (The OPA parameters are $\gamma = 1.94$, $p_0 = 0.005$ and $p_1 = 0.15$, and (11) is evaluated with saturation $S = 15$ lines.) The results are shown in Figures 8 and 9 and Table I. The results in Figure 9 show saturation effects not well captured in the branching process model that make the prediction of the distribution poorer. (One could set $S = 15$ in (5) and (6) to model all cascades stopping when 15 lines fail, but this very crude model of saturation gives a spike at 15 lines that is also a poor fit to the empirical distribution since quite a few of the cascades proceed beyond 15 lines failed.) The position of the peak in the empirical distribution of Figure 9 indicates a concentration of number of line failures near 15, and we associate this with a tendency for the cascades to saturate at approximately 15 lines failed. There also seems to be a secondary peak and saturation at approximately 30 lines failed. At smaller loadings such as in Figure 8, fewer cascades encounter the saturation and, although the saturation effect is perhaps discernible, it is small. We expect larger saturation values in larger test systems. It is interesting to note that the July and August 1996 Western area blackouts saturated at over 30 high voltage line trips and the August 2003 Eastern interconnect blackout saturated at several hundred high voltage line trips [16].

The results suggest that good predictions of the probability distributions of the number of line failures can be obtained as long as saturation effects are not significant. We do not understand how to accurately model the saturation effects at present. The ability to predict the probability distribution of the number of line failures in non saturating cases supports the applicability of branching models to cascading failure before saturation is reached and the usefulness of the estimate $\widehat{\lambda}_s$ of failure propagation. $\widehat{\lambda}_s$ can be obtained much more efficiently than empirical probability distributions of line failures obtained by brute force.

TABLE I
ESTIMATORS FOR OPA CASES

power system	loading factor	$\widehat{\theta}$	$\widehat{\lambda}_s$
IEEE 118 bus	0.9	1.10	0.19
IEEE 118 bus	1	1.66	0.41
IEEE 118 bus	1.3	12.20	0.44
tree-like 190 bus	1.0	1.49	0.53
tree-like 190 bus	1.2	6.21	0.61

V. GENERALITY AND LIMITATIONS OF λ CONCEPT

This paper approximates cascading failure by a single-type Galton-Watson branching process with random or deterministic initial failures. The most immediate motivation is that this branching process approximates a loading dependent model of cascading failure [17] and captures some, but not all, qualitative features of cascading failure in blackout simulations [4], [10], [25]. Moreover, the single-type Galton Watson branching process also approximates other cascading processes, such as avalanches in idealized sandpiles [23], [27], initial spread of epidemics, growth of populations etc. Thus estimation of λ is already established in cascading failure in other fields [19]. This section shows two examples of more elaborate models that involve cascading failure to start to explore how the parameter λ may apply more generally.

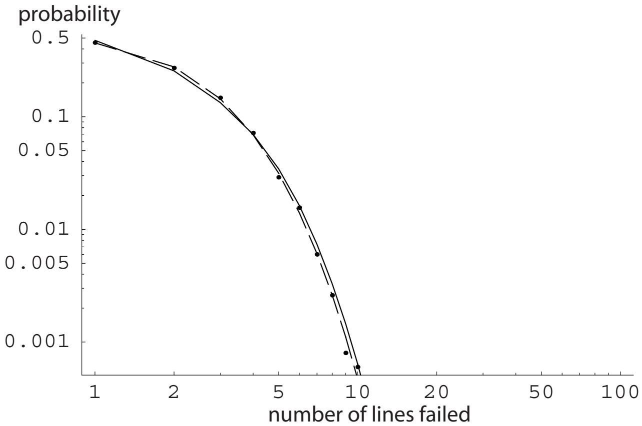

Fig. 5. IEEE 118 bus system with loading factor 0.9. PDF estimated with initial failure distribution and λ_s (solid line) and $\widehat{\theta}$ and $\widehat{\lambda}_s$ (dashed line) compared with OPA empirical PDF (dots). Note the log-log scales.

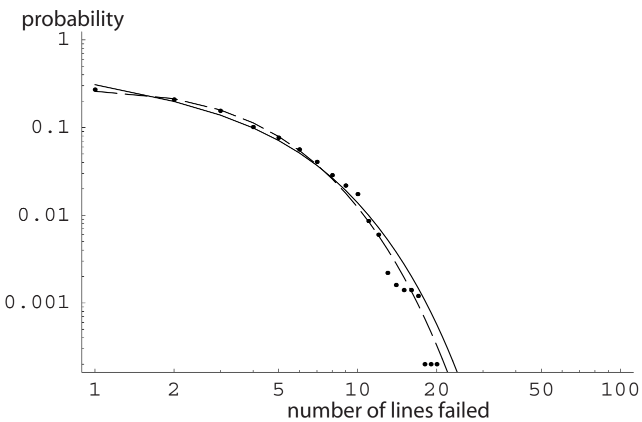

Fig. 6. IEEE 118 bus system with loading factor 1.0. PDF estimated with initial failure distribution and λ_s (solid line) and $\widehat{\theta}$ and $\widehat{\lambda}_s$ (dashed line) compared with OPA empirical PDF (dots).

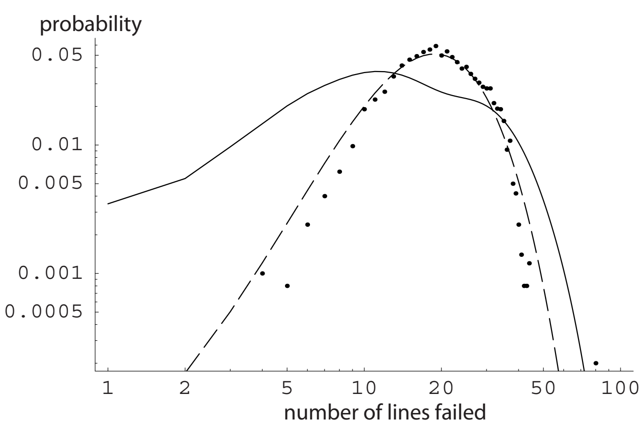

Fig. 7. IEEE 118 bus system with loading factor 1.3. PDF estimated with initial failure distribution and λ_s (solid line) and $\widehat{\theta}$ and $\widehat{\lambda}_s$ (dashed line) compared with OPA empirical PDF (dots).

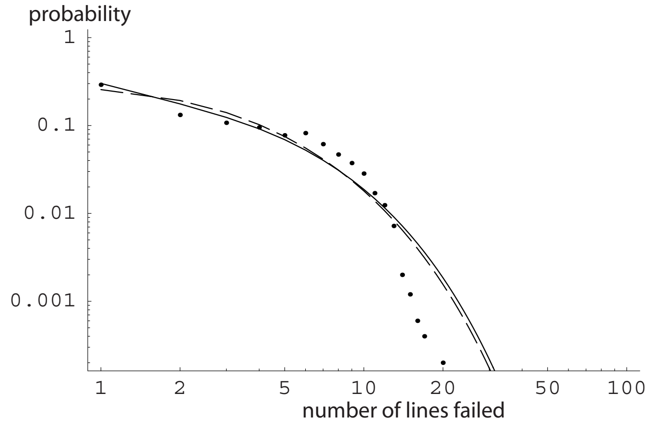

Fig. 8. Tree-like 190 bus system with loading factor 1.0. PDF estimated with initial failure distribution and λ_s (solid line) and $\widehat{\theta}$ and $\widehat{\lambda}_s$ (dashed line) compared with OPA PDF (dots).

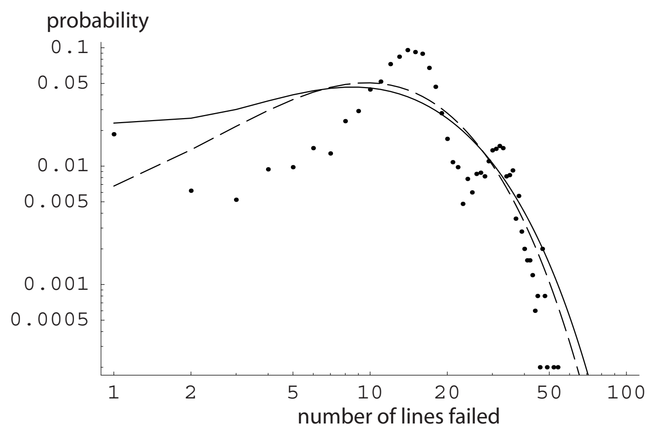

Fig. 9. Tree-like 190 bus system with loading factor 1.1. PDF estimated with initial failure distribution and λ_s (solid line) and $\widehat{\theta}$ and $\widehat{\lambda}_s$ (dashed line) compared with OPA PDF (dots).

A. Demon model of two coupled infrastructures

The Demon model [26] abstractly models cascading infrastructure failure as a dynamic complex system. The Demon model is based on the forest fire model of Bak, Chen and Tang [3] with modifications by Drossel and Schwabl [18]. For a single system, the model is defined on a 2-d network. Network nodes represent components of the infrastructure system and lines represent couplings between components. The components can be operating, failed or failing. The rules of the model are that for each time step:

1) Any operating component fails with a probability P_f. P_f is the spontaneous failure rate independent of other components.
2) An operating component fails with a probability P_n if at least one of the nearest components is failing. P_n describes the extent to which failures propagate to neighboring components.
3) A failing component becomes a failed one.
4) A failed component is repaired and becomes operational with probability P_r.

We can generalize this single system to two coupled infrastructure systems by taking two of these 2-d networks, making a fraction g of the nodes in each system correspond to the component in the corresponding 2-d position in the other system, and adding another rule:

5) An operating component in system 1 fails if the corresponding component in system 2 is failed or failing. An operating component in system 2 fails if the corresponding component in system 1 is failed or failing.

In this paper, we will assume that the two coupled systems have identical networks and identical parameters except that the probability $P_f^{(1)}$ of spontaneous failures in system 1 may differ from the the probability $P_f^{(2)}$ of spontaneous failures in system 2. We assume that the systems are coupled symmetrically with coupling coefficient c and that the fraction of components coupled to other system is g. These rules allow both analytic mean field and numerical calculations and we start with a mean field calculation that generalizes the calculation in [18]. The number of components in each system is N. Let $O^{(i)}(t)$ be the fraction of operating components in system i at time t; that is, the number of operating components at time t divided by N. In the same way, we can define the fraction of failed components, $F^{(i)}(t)$, and the fraction of failing components, $B^{(i)}(t)$.

The mean field equations for the coupled system are:

$$\begin{aligned}
B^{(1)}(t+1) &= P_f^{(1)}O^{(1)}(t) + P_n f O^{(1)}(t)B^{(1)}(t)\\
&\quad + cgO^{(1)}(t)(B^{(2)}(t)+F^{(2)}(t))\\
F^{(1)}(t+1) &= (1-P_r)F^{(1)}(t) + B^{(1)}(t)\\
O^{(1)}(t+1) &= (1-P_f^{(1)})O^{(1)}(t) + P_r F^{(1)}(t)\\
&\quad - P_n f O^{(1)}(t)B^{(1)}(t) - cgO^{(1)}(t)(B^{(2)}(t)+F^{(2)}(t))\\
B^{(2)}(t+1) &= P_f^{(2)}O^{(2)}(t) + P_n f O^{(2)}(t)B^{(2)}(t)\\
&\quad + cgO^{(2)}(t)(B^{(1)}(t)+F^{(1)}(t))\\
F^{(2)}(t+1) &= (1-P_r)F^{(2)}(t) + B^{(2)}(t)\\
O^{(2)}(t+1) &= (1-P_f^{(2)})O^{(2)}(t) + P_r F^{(2)}(t)\\
&\quad - P_n f O^{(2)}(t)B^{(2)}(t) - cgO^{(2)}(t)(B^{(1)}(t)+F^{(1)}(t))
\end{aligned} \tag{16}$$

In (16), f is the average number of neighboring nodes within a system that a failing node can propagate failure to. During failure propagation, since at least one of the nodes neighboring a failing node has failed, $f \approx k-1$, where the network degree k of each system is the average number of nodes within each system that are joined to a given node. Equations (16) are consistent with the condition $O^{(i)}(t)+B^{(i)}(t)+F^{(i)}(t)=1$.

In the limit with no spontaneous failures, $P_f^{(1)} = P_f^{(2)} = 0$, and for a steady state solution, (16) can be reduced to two coupled equations,

$$\begin{aligned}
\left(1-P_n f O^{(1)}\right)P_r(1-O^{(1)}) &= cg(1+P_r)(1-O^{(2)})O^{(1)},\\
\left(1-P_n f O^{(2)}\right)P_r(1-O^{(2)}) &= cg(1+P_r)(1-O^{(1)})O^{(2)}.
\end{aligned} \tag{17}$$

If $c \neq 0$, then $O^{(1)} = 1$ implies $O^{(2)} = 1$; that is, the systems are effectively decoupled. Therefore, to have truly coupled systems, system 1 must be in a supercritical state with $O^{(1)} < 1$. Here criticality occurs in passing from an average state with almost no failures to an average state with some failures.

For this mean field computation we have assumed identical system symmetrically coupled. In particular $P_f^{(1)} = P_f^{(2)} = 0$. Then $O^{(1)} = O^{(2)}$, and (17) lead to the following identical solutions for the two systems in steady state:

$$O^{(i)} = \begin{cases} 1, & \hat{g} \le 1 \\ \dfrac{1}{\hat{g}}, & \hat{g} > 1 \end{cases} \tag{18}$$

$$F^{(i)} = \begin{cases} 0, & \hat{g} \le 1 \\ \dfrac{\hat{g}-1}{\hat{g}(1+P_r)}, & \hat{g} > 1 \end{cases} \tag{19}$$

$$B^{(i)} = \begin{cases} 0, & \hat{g} \le 1 \\ \dfrac{\hat{g}-1}{\hat{g}(1+P_r)}P_r, & \hat{g} > 1 \end{cases} \tag{20}$$

Here, $\hat{g}$ is the control parameter given by

$$\hat{g} = P_n f + \frac{cg(1+P_r)}{P_r} \tag{21}$$

The critical point occurs at $\hat{g} = 1$.

Observe from (21) that if the systems are uncoupled so that $c = 0$, then $\hat{g} = P_n f$ both measures the average propagation of failures in each system and detects criticality when $\hat{g} = 1$. Thus for uncoupled systems (or a system considered in isolation) $\hat{g}$ has similar characteristics to λ in the branching process model. However, the situation is more complicated when the systems are coupled and then (21) shows that the criticality depends also on the coupling and the probability of repair P_r.

We have tested the results from the mean field computation by comparing them with numerical results for the identical coupled systems formed from the systems listed in Table II. The results for the average number of operating components are shown in Figure 10. These results were obtained for fixed $P_r = 0.001$, $c = 0.0005$, $P_f^{(1)} = 0.00001$, and $P_f^{(2)} = 0$ and we have varied the propagation parameter P_n. The results show very good agreement with the mean field theory results as the network degree k increases. For $k = 2$, the system are practically 1-d and the mean field theory is not applicable.

TABLE II

SYSTEMS TO BE COUPLED

system	k	number of nodes
Open 3-branch Tree	2	3070
Closed 3-branch Tree	3	3070
Square	3.96	10000
Hexagon	5.9	4681

The density of operating components is practically the same in both systems. This is expected because the systems are nearly identical with the only symmetry breaking feature being the probability of spontaneous failures $P_f^{(2)}$ that is zero in the second system.

Perhaps the most important point for the purposes of this paper is that the critical point moves as the coupling cg increases. Figure 11 shows the critical point as a function of

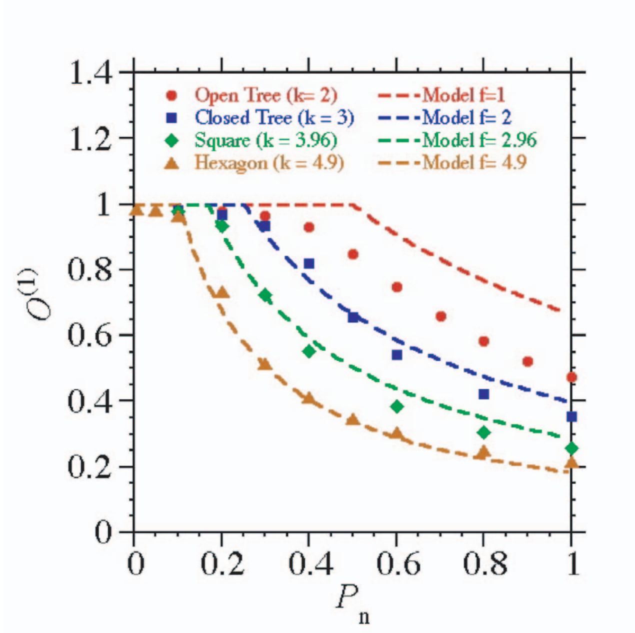

Fig. 10. Variation of fraction of operating components $O^{(1)}$ in system 1 as the propagation parameter P_n is varied.

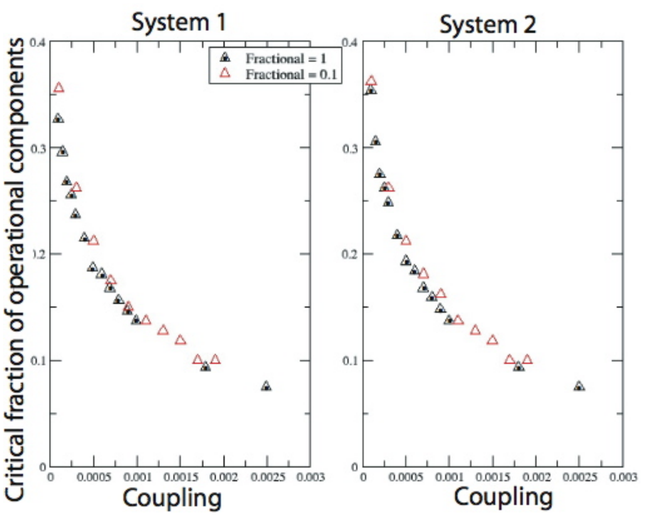

Fig. 11. Fraction of operating components $O^{(i)}$ at critical point as a function of coupling cg for system 1 and system 2. The triangles with dots inside are results obtained with $g = 1$ (all components coupled between the systems) and the open triangles are results obtained with $g = 0.1$ and a tenfold increase in c.

the coupling cg. A factor of four variation in the critical point is seen. Figure 11 shows the critical points of both system 1 and system 2 calculated from the change in the functional form of the operating node density (18). The two systems are virtually indistinguishable from each other.

Following [18], we computed a measure of propagation of failing components

$$\widehat{\lambda}_d = \left\langle \frac{B(4)}{B(3)} \right\rangle \tag{22}$$

where $\langle\ \rangle$ denotes averaging over all the cascades occurring in the model runs. While there is critical value for $\widehat{\lambda}_d$ at the critical point discussed above, the critical value of $\widehat{\lambda}_d$ at the critical point is typically less then 1 (up to 20% less for the cases examined) and it depends on the coupling and the details of the system. We are continuing to investigate measures of propagation and their interpretation and possible applicability to the coupled system.

B. *Two-type branching process model*

To show a more elaborate branching process model where the generalization of λ with all the characteristics of the single type case is not yet apparent, consider a two-type Galton-Watson branching process [2]. One motivation for the two types in blackout models is that type 1 could represent component failures within the network and type 2 could represent the amount of load shed or energy unserved. Another motivation is that the two types could represent failures in two coupled infrastructures [26]. One way to generalize λ is the matrix

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{pmatrix}$$

where λ_{ij} is the average number of type j failures caused by each type i failure in the previous stage. The coupled cascading processes still show criticality but the condition for criticality becomes $\lambda_{\max} = 1$ where $\lambda_{\max}$ is the largest eigenvalue of Λ [26]. $\lambda_{\max}$ is a quadratic root function of the entries of Λ. The criticality property remains a function of Λ, but the function is not trivial as in the single type case.

Suppose that the two-type branching process represents two coupled infrastructures. Then, if the first infrastructure is incorrectly modeled in isolation, it will be a single-type branching process with parameter λ_{11}, and its proximity to criticality will be computed as $1 - \lambda_{11}$. On the other hand, the model of the two coupled infrastructures will yield a proximity to criticality of $1 - \lambda_{\max}$. This illustrates the importance of mutual couplings with other infrastructures in determining the propensity for cascading failure.

VI. Discussion and Conclusion

In this paper, we approximate cascading failure by a single type Galton-Watson branching process with saturation in order to propose a method of quantifying the propagation of failures λ and hence estimate the probability distribution of the number of failures. The proposed estimator $\widehat{\lambda}_s$ requires K observations of cascades with the initial and final failures grouped in stages. Although $\widehat{\lambda}_s$ is not asymptotically unbiased, it does work in the presence of saturation effects that are thought to occur in blackouts.

Testing on a saturating branching process suggests that an order of $K = 10$ observations can be sufficient to determine $\widehat{\lambda}_s$ with worst-case bias and standard deviation of order 0.1. The standard deviation of $\widehat{\lambda}_s$ decreases like $1/\sqrt{K}$ as K increases and it may be possible to correct for the bias in future work. This compares favorably with the much larger number of observations needed to estimate the often heavy-tailed distribution of number of failures directly. The ability to estimate the propagation of failures and the distribution of failures with a modest number of observations would expand the opportunities for using cascading failure simulations to study the effect of transmission system upgrades on cascading failure and is crucial for the practicality of monitoring failures in the power grid to assess cascading failure.

Initial testing of the estimator $\widehat{\lambda}_s$ by predicting the empirical probability distributions of line failures from OPA using branching process models shows a good fit for cases in which saturation effects are not present. Similarly good fits for the non-saturating case are obtained by approximating the initial failures as a Poisson distribution and estimating the mean $\widehat{\theta}$ of the Poisson distribution. These results are also promising in showing the applicability of simple branching process models to cascading processes in blackouts before saturation. Further progress on saturating cases is likely to require a better understanding of the saturation effects.

We also examine quantities sharing some properties of λ in an abstract complex systems dynamical model of two interacting infrastructures. While there remains a useful parameter $\hat{g}$ governing the criticality of the model, the value of $\hat{g}$ at criticality varies with the amount of coupling between the two infrastructures. When the infrastructures are not coupled, $\hat{g}$ is analogous to λ in that it describes the average propagation of failures in each system and the system becomes critical when $\hat{g} = 1$. However, the situation is more complicated when the infrastructures become coupled. A two-type generalization of the branching process is a different model of coupled cascading systems and in this model λ can be generalized to a more complicated function of the model parameters that measures the dominant propagation and becomes 1 at criticality. The complex systems dynamical model and the two-type branching model are very different but both are simple models of coupled cascading processes that have the potential to capture some overall features of failures cascading in and between infrastructures. Both models show that the propensity for cascading failure is strongly influenced by the mutual couplings between the infrastructures. In particular, consideration of the infrastructures in isolation could give a misleading assessment of the propensity for cascading failure.

As more is learned about the universal or application-specific properties of cascading failure, the high level models of cascading failure may be elaborated. Each elaborated branching process has further parameters to fit to describe the cascade. Further work is needed to determine the best compromise between simplicity and elaboration, but this paper makes progress in estimating λ for a single-type Galton-Watson branching process with saturation. This branching process model or its elaborations could well be applicable to assessing cascading failure risk in large networked infrastructures and in interacting infrastructures.

References

[1] R. Adler, S. Daniel, C. Heising, M. Lauby, R. Ludorf, T. White, An IEEE survey of US and Canadian overhead transmission outages at 230 kV and above, *IEEE Transactions on Power Delivery*, vol. 9, no. 1, Jan. 1994, pp. 21 -39

[2] K.B. Athreya, P.E. Ney, *Branching Processes*, Dover NY 2004 (reprint of Springer-verlag Berlin 1972).

[3] P. Bak, K. Chen, C. Tang, A forest-fire model and some thoughts on turbulence, *Physics Letters A*, vol. 147, July 1990, pp. 297-300.

[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, *Chaos*, vol. 12, no. 4, December 2002, pp. 985-994.
Note: all the authors' papers are available at
http://eceserv0.ece.wisc.edu/~dobson/home.html

[5] B.A. Carreras, V.E. Lynch, D.E. Newman, I. Dobson, Blackout mitigation assessment in power transmission systems, *36th Hawaii International Conference on System Sciences*, Hawaii, 2003.

[6] B.A. Carreras, V.E. Lynch, I. Dobson, D.E. Newman, Dynamical and probabilistic approaches to the study of blackout vulnerability of the power transmission grid, *37th Hawaii International Conference on System Sciences*, Hawaii, 2004.

[7] B.A. Carreras, V.E. Lynch, I. Dobson, D.E. Newman, Complex dynamics of blackouts in power transmission systems, *Chaos*, vol. 14, no. 3, September 2004, pp. 643-652.

[8] B.A. Carreras, D.E. Newman, I. Dobson, A.B. Poole, Evidence for self organized criticality in a time series of electric power system blackouts, *IEEE Transactions on Circuits and Systems I*, vol. 51, no. 9, September 2004, pp. 1733-1740.

[9] J. Chen, J.D. McCalley, A cluster distribution as a model for estimating high-order event probabilities in power systems, *Eighth International Conference on Probability Methods Applied to Power Systems*, Ames Iowa, September 2004.

[10] J. Chen, J.S. Thorp, I. Dobson, Cascading dynamics and mitigation assessment in power system disturbances via a hidden failure model, *International Journal of Electrical Power and Energy Systems*, vol 27, no 4, May 2005, pp. 318-326

[11] J.-P. Dion, N. Keiding, Statistical inference in branching processes, in *Branching Processes*, editors A. Joffe, P. Ney, Marcel Dekker, New York 1978, pp. 105-140.

[12] I. Dobson, B.A. Carreras, V. Lynch, D.E. Newman, Complex systems analysis of series of blackouts: cascading failure, criticality, and self-organization, *Bulk Power System Dynamics and Control - VI*, August 22-27, 2004, Cortina dAmpezzo, Italy.

[13] I. Dobson, B.A. Carreras, D.E. Newman, Probabilistic load-dependent cascading failure with limited component interactions, *IEEE International Symposium on Circuits and Systems*, Vancouver Canada, May 2004.

[14] I. Dobson, B.A. Carreras, V.E. Lynch, B. Nkei, D.E. Newman, Estimating failure propagation in models of cascading blackouts, *Eighth International Conference on Probability Methods Applied to Power Systems*, Ames Iowa, September 2004. Also to appear in *Probability in the Engineering and Informational Sciences*, vol. 19, 2005.

[15] I. Dobson, B.A. Carreras, D.E. Newman, A branching process approximation to cascading load-dependent system failure. *37th Hawaii International Conference on System Sciences*, Hawaii, January 2004.

[16] I. Dobson, B.A. Carreras, D.E. Newman, Branching process models for the exponentially increasing portions of cascading failure blackouts, *38th Hawaii International Conference on System Sciences*, January 2005, Hawaii.

[17] I. Dobson, B.A. Carreras, D.E. Newman, A loading-dependent model of probabilistic cascading failure, *Probability in the Engineering and Informational Sciences*, vol. 19, no. 1, January 2005.

[18] B. Drossel, F. Schwabl, Forest-fire model with immune trees, *Physica A*, vol. 199, October 1993, pp. 183-197.

[19] P. Guttorp, *Statistical inference for branching processes*, Wiley, NY, 1991

[20] R. C. Hardiman, M. T. Kumbale, Y. V. Makarov, An advanced tool for analyzing multiple cascading failures, *Eighth International Conference on Probability Methods Applied to Power Systems*, Ames Iowa, September 2004.

[21] T.E. Harris, *Theory of branching processes*, Dover NY 1989.

[22] D.S. Kirschen, D. Jawayeera, D.P. Nedic, R.N. Allan, A probabilistic indicator of system stress, *IEEE Transactions on Power Systems*, vol. 19, no. 3, 2004, pp. 1650-1657.

[23] D.S. Lee, K.-I. Goh, B. Kahng, D. Kim, Branching process approach to avalanche dynamics, *Journal of the Korean Physical Society*, vol. 44, no. 3, March 2004, pp. 633-637.

[24] K. Nanthi, Statistical estimation for stochastic processes, *Queen's papers in pure and applied mathematics*, 62, 1983.

[25] D.P. Nedic, I. Dobson, D.S. Kirschen, B.A. Carreras, V.E. Lynch, Criticality in a cascading failure blackout model, *Fifteenth Power Systems Computation Conference*, Liege Belgium, August 2005.

[26] D. E. Newman, B. Nkei, B. A. Carreras, I. Dobson, V. E. Lynch, P. Gradney, Risk assessment in complex interacting infrastructure systems, *Thirty-eighth Hawaii International Conference on System Sciences*, Hawaii, January 2005.

[27] F. Slanina, Self-organized branching process for a one-dimensional rice-pile, *European Physical Journal B*, 209-216, 2002.
[28] U.S.-Canada Power System Outage Task Force, *Final Report on the August 14th blackout in the United States and Canada*. United States Department of Energy and National Resources Canada, April 2004.

Appendix

This appendix derives formulas for mean and variance of $\widehat{\lambda}_s$ similar to the formulas for $\widehat{\lambda}_n$ in [19, pp. 36-37].

Define the event

$$A = \bigcap_{k=1}^{K} \left\{ s(k) = s_k, Y^{(k)}_{s_k-2} = y_k, Z^{(k)}_{s_k-1} = z_k \right\}$$

$$\begin{aligned}
&\mathrm{E}\,\widehat{\lambda}_s = \mathrm{E}\,\frac{\sum_{k=1}^{K}(Y^{(k)}_{s(k)} - Z^{(k)}_0)}{\sum_{k=1}^{K} Y^{(k)}_{s(k)-1}} \\
&= \sum_{s_k,y_k,z_k} \mathrm{E}\left(\frac{\sum_{k=1}^{K}(Y^{(k)}_{s(k)} - Z^{(k)}_0)}{\sum_{k=1}^{K} Y^{(k)}_{s(k)-1}} \,\middle|\, A \right) P(A) \\
&= \sum_{s_k,y_k,z_k} \mathrm{E}\left(\frac{\sum_{k=1}^{K}(y_k + z_k + Z^{(k)}_{s_k} - Z^{(k)}_0)}{\sum_{k=1}^{K}(y_k + z_k)} \,\middle|\, A \right) P(A) \\
&= \sum_{s_k,y_k,z_k} \left(1 + \frac{\sum_{k=1}^{K}(\lambda z_k - Z^{(k)}_0)}{\sum_{k=1}^{K}(y_k + z_k)} \right) P(A) \\
&= 1 + \lambda \mathrm{E}\frac{\frac{1}{K}\sum_{k=1}^{K} Z^{(k)}_{s(k)-1}}{\frac{1}{K}\sum_{k=1}^{K} Y^{(k)}_{s(k)-1}} - \mathrm{E}\frac{\frac{1}{K}\sum_{k=1}^{K} Z^{(k)}_0}{\frac{1}{K}\sum_{k=1}^{K} Y^{(k)}_{s(k)-1}} \qquad (23)
\end{aligned}$$

Let

$$B(k) = \left(s(k), Y^{(k)}_{s(k)-2}, Z^{(k)}_{s(k)-1} \right)$$

$$\begin{aligned}
&\mathrm{Var}\widehat{\lambda}_s = \mathrm{Var}\frac{\sum_{k=1}^{K}(Y^{(k)}_{s(k)} - Z^{(k)}_0)}{\sum_{k=1}^{K} Y^{(k)}_{s(k)-1}} \\
&= \mathrm{Var}\left(1 + \frac{\sum_{k=1}^{K}(Z^{(k)}_{s(k)} - Z^{(k)}_0)}{\sum_{k=1}^{K}(Y^{(k)}_{s(k)-2} + Z^{(k)}_{s(k)-1})} \right) \\
&= \mathrm{Var}\left(\frac{\sum_{k=1}^{K}(Z^{(k)}_{s(k)} - Z^{(k)}_0)}{\sum_{k=1}^{K}(Y^{(k)}_{s(k)-2} + Z^{(k)}_{s(k)-1})} \right) \\
&= \mathrm{E}\,\mathrm{Var}\left(\frac{\sum_{k=1}^{K}(Z^{(k)}_{s(k)} - Z^{(k)}_0)}{\sum_{k=1}^{K}(Y^{(k)}_{s(k)-2} + Z^{(k)}_{s(k)-1})} \,\middle|\, B(k) \right) \\
&\qquad + \mathrm{Var}\,\mathrm{E}\left(\frac{\sum_{k=1}^{K}(Z^{(k)}_{s(k)} - Z^{(k)}_0)}{\sum_{k=1}^{K}(Y^{(k)}_{s(k)-2} + Z^{(k)}_{s(k)-1})} \,\middle|\, B(k) \right) \\
&= \\
&\sigma^2 \mathrm{E}\,\frac{\sum_{k=1}^{K} Z^{(k)}_{s(k)-1}}{\left(\sum_{k=1}^{K} Y^{(k)}_{s(k)-1}\right)^2} + \mathrm{Var}\frac{\sum_{k=1}^{K}(\lambda Z^{(k)}_{s(k)-1} - Z^{(k)}_0)}{\sum_{k=1}^{K} Y^{(k)}_{s(k)-1}} \\
&= \frac{\sigma^2}{K}\mathrm{E}\,\frac{\frac{1}{K}\sum_{k=1}^{K} Z^{(k)}_{s(k)-1}}{\left(\frac{1}{K}\sum_{k=1}^{K} Y^{(k)}_{s(k)-1}\right)^2} \\
&\qquad + \frac{1}{K}\mathrm{Var}\frac{\frac{1}{\sqrt{K}}\sum_{k=1}^{K}(\lambda Z^{(k)}_{s(k)-1} - Z^{(k)}_0)}{\frac{1}{K}\sum_{k=1}^{K} Y^{(k)}_{s(k)-1}}
\end{aligned}$$

This uses

$$\begin{aligned}
\mathrm{Var}X &= \mathrm{E}X^2 - (\mathrm{E}X)^2 \\
&= \mathrm{E}\mathrm{E}(X^2|Y) - \mathrm{E}(\mathrm{E}(X|Y))^2 + \mathrm{E}(\mathrm{E}(X|Y))^2 \\
&\qquad -(\mathrm{E}\mathrm{E}(X|Y))^2 \\
&= \mathrm{E}\mathrm{Var}(X|Y) + \mathrm{Var}\mathrm{E}(X|Y)
\end{aligned}$$