Towards Estimating the Statistics of Simulated Cascades of Outages with Branching Processes

Junjian Qi, *Student Member, IEEE,* Ian Dobson, *Fellow, IEEE,* Shengwei Mei, *Senior Member, IEEE*

***Abstract*—Branching processes can be applied to simulated cascading data to describe the statistics of the cascades and quickly predict the distribution of blackout sizes. We improve the procedures for discretizing load shed data so that a Galton-Watson branching process may be applied. The branching process parameters such as average propagation are estimated from simulated cascades and the branching process is then used to estimate the distribution of blackout size. We test the estimated distributions with line outage and load shed data generated by the improved OPA and AC OPA cascading simulations on the IEEE 118 bus system and the Northeast Power Grid of China.**

***Index Terms*—Branching process, cascading failure, power transmission reliability, simulation.**

I. Introduction

CASCADING blackouts of transmission systems are complicated sequences of dependent outages that lead to load shed. Of particular concern are the larger cascading blackouts; these blackouts are rare and high impact events that have substantial risk, and they pose many challenges in simulation, analysis, and mitigation [1].

Simulations of cascading outage produce samples of cascades, and it is useful to be able to statistically describe these simulated cascades with high-level probabilistic models such as branching processes. Branching processes have descriptive parameters that characterize the system resilience to cascading. For example, the propagation of outages is a key parameter for branching processes that quantifies the average extent to which outages cause further outages. Moreover, the branching process can be used to predict the probability distribution of blackout size (that is, the frequencies of blackouts as a function of blackout size) from much shorter simulation runs. It is much quicker [2] to estimate the parameters of a branching process from a shorter simulation run and then predict the distribution of blackout size using the branching process than it is to run the simulation for very long time to accumulate enough cascades (especially the rare large cascades) to empirically estimate the distribution of blackout size.

J. Qi and S. Mei are with the State Key Laboratory of Power Systems, Electrical Engineering Dept., Tsinghua University, Beijing 100084 China (e-mails: qjj08@mails.tsinghua.edu.cn, meishengwei@mail.tsinghua.edu.cn). J. Qi visited the Electrical and Computer Engineering Dept. at Iowa State University, Ames IA 50011 USA. I. Dobson is with the Electrical and Computer Engineering Dept., Iowa State University, Ames IA 50011 USA (e-mail: dobson@iastate.edu). We gratefully acknowledge support in part from the Tsinghua Scholarship for Overseas Graduate Studies, National High Technology Research and Development Program of China (863 Program, 2011AA05A118), NSFC grant-50525721, Special Funding for National Excellent Doctoral Dissertation Author-201153, DOE grant DE-SC0002283, and NSF grant CPS-1135825.

However, these advantages of branching processes can only be realized when the branching process is validated, and this needs to be done for a range of different simulations and test cases, and for real data. The progress so far in establishing branching process models for estimating cascade propagation and distributions of blackout size is:

1) Branching processes can match the distribution of number of line outages [2] and load shed [3] simulated by the OPA simulation on the IEEE 118 and 300 bus systems. There is also a match for the distribution of load shed for the TRELSS simulation in one case of an industrial system of about 6250 buses [3].
2) Branching processes can match the propagation and distribution of the number of cascading line outages in real data [4], [5].
3) Branching processes can approximate CASCADE, another high-level probabilistic model of cascading [6].

In matching load shed in [3] for item (1), two approaches and types of branching processes were tried to accommodate the continuously varying load shed data. (Note that line outages are easier than load shed since number of line outages is a nonnegative integer whereas load shed varies continuously.) The approach that worked best in [3] discretized the load shed data so that it became integer multiples of the chosen discretization unit. Then the discretized load shed data was processed with a Galton-Watson branching process, which works with nonnegative integers. However in [3], the choice of the discretization unit was ad hoc and no systematic approach was given. Analyzing the load shed data with the branching process is important because load shed is measure of blackout size that is of great significance to both utilities and society, whereas line outages are of direct interest only to utilities.

Instead of applying a Galton-Watson branching process to discretized load shed, it is also possible to directly analyze load shed using a continuous state branching process model [3]. Each generation is a continuously varying amount of load shed that propagates according to a continuous offspring probability distribution. One challenge with the continuous state branching process is determining the form to assume for the continuous offspring distribution. Reference [3] assumes a Gamma distribution for simplicity and computational convenience, but the problem of a justified choice of continuous offspring distribution for cascading load shed remains open. The calculations for continuous state branching processes are analogous to the calculations for Galton-Watson branching processes but more technically difficult.

In this paper, we advance the application of branching processes to cascading blackouts by improving the processing of load shed data and validating the branching process on other simulations and test systems. In particular,

1) We describe a new and systematic procedure to discretize load shed data so that it can be analyzed as a Galton- Watson branching process with a Poisson offspring distribution. This new procedure can be applied to any continuously varying measure used to track cascading.
2) We show how branching process models can match the distribution of number of cascading line outages and load shed simulated by two enhanced versions of the OPA simulation on the IEEE 118 bus system and the Northeast Power Grid of China (NPGC).

The rest of this paper is organized as follows. Section II briefly introduces Galton-Watson branching processes. Section III explains estimating branching process parameters and the distribution of outages. Section IV explains how to discretize the load shed data so that it can be analyzed with a Galton-Watson branching process. Section V introduces the improved OPA and AC OPA cascading simulations. Section VI tests the proposed method with line outage data and load shed data generated by the two simulations on IEEE 118 bus system and the NPGC system. Finally the conclusion is drawn in Section VII.

II. Branching processes

Branching processes have long been used in a variety of applications to model cascading processes [7], [8], but have been applied to the risk of cascading outage only recently [9], [4], [6], [2], [3], [5]. This section gives an informal overview of branching processes. For more detail, see [5] and [7], [8].

The Galton-Watson branching process gives a high-level probabilistic model of how the number of outages in a blackout propagate. The initial outages propagate randomly to produce subsequent outages in generations. Each outage in each generation (a "parent" outage) independently produces a random number 0, 1, 2, 3, ... of outages ("children" outages) in the next generation. The distribution of the number of children from one parent is called the offspring distribution. The children outages then become parents to produce another generation, and so on. If the number of outages in a generation becomes zero, then the cascade stops.

The mean of the offspring distribution is the parameter λ. λ is the average family size (the average number of children outages for each parent outage). λ quantifies the tendency for the cascade to propagate, since large average family sizes will tend to cause the outages to grow faster. In this paper, we have $\lambda < 1$, and the outages will always eventually die out. The branching process model does not directly represent any of the physics or mechanisms of the outage propagation, but, after it is validated, it can be used to predict the total number of outages. The parameters of a branching process model can be estimated from a much smaller data set, and then predictions of the total number of outages can be made based on the estimated parameters. While it is sometimes possible to observe or produce large amounts of data to make an empirical estimate of the total number of outages (indeed this is the way the branching process prediction is validated), the ability to do this via the branching process model with much less data is a significant advantage that enables practical applications. The simplicity of the branching process model also allows a high-level understanding of the cascading process without getting entangled in the various and complicated mechanisms of cascading. The branching process should be seen as complementary to detailed modeling of cascading outage mechanisms.

The intent of the branching process modeling is not that each parent outage in some sense "causes" its children outages; the branching process simply produces random numbers of outages in each generation that can statistically match the outcome of the cascading. For example, when used to track the number of transmission line outages, the branching process does not specify which lines outage, or where they are, or explain why they outaged. The branching process describes the statistics of the number of lines outaging in each generation, and the statistics of the total number of lines outaged.

Similarly, when used to track load shed, the branching process does not specify which load is shed, or where, or explain why load is shed. It is worth noting that the underlying cascading processes and interactions that we are tracking with a branching process are complicated and varied. For example, there are situations in which load shed tends to inhibit the chance of further cascading and there are other situations in which load shed can tend to increase the chance of further cascading. It is not obvious that branching processes can summarize this complexity, and it is a goal of this paper to show evidence that this can be done. In particular, we show that the branching process describes the statistics of the load shed in each generation for the purpose of summarizing the cascade propagation so that the distribution of the total load shed can be statistically estimated. The branching process analysis is a bulk statistical analysis that should be regarded as complementary to detailed causal analysis.

III. Estimating branching process parameters

This section explains how the branching process parameters are estimated from the simulated data and used to estimate the distribution of blackout size, in particular the distribution of the number of line outages or the distribution of load shed.

The simulation naturally produces the line outages or load shed in generations or stages; each iteration of the "main loop" of the simulation produces another generation. In the case of line outages, the number of line outages in each generation are counted. In the case of load shed, the continuously varying load shed amounts in each generation are processed and discretized as described in section IV to produce integer multiples of the chosen discretization unit. In either case, M cascades[1] are simulated to produce nonnegative integer data that can be arranged as

[1]Note that a single outage shedding load followed by no further load shed is regarded as a cascade with only one generation.

	generation 0	generation 1	generation 2	$\cdots$
cascade 1	$Z_0^{(1)}$	$Z_1^{(1)}$	$Z_2^{(1)}$	$\cdots$
cascade 2	$Z_0^{(2)}$	$Z_1^{(2)}$	$Z_2^{(2)}$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
cascade M	$Z_0^{(M)}$	$Z_1^{(M)}$	$Z_2^{(M)}$	$\cdots$

where $Z_j^{(m)}$ is the number of outages produced by the simulation in generation j of cascade number m. Since the average propagation $\lambda < 1$, each cascade eventually terminates with a finite number of generations when the number of outages in a generation becomes zero. Each cascade has a nonzero number of outages in generation 0. The shortest cascades stop in generation 1 by having no outages in generation 1 and higher generations, but some of the cascades will continue for several or occasionally many generations before terminating. The assumption of a positive number of outages in generation 0 implies that all statistics assume (or, in statistical terminology, are conditioned on) a cascade starting.

Note that all the outages are parent outages, and all the outages in generations 1 and higher are children outages. It follows that the average propagation λ (the average family size) can be estimated as the total number of children divided by the total number of parents:

$$\hat{\lambda} = \frac{\sum_{m=1}^{M}(Z_1^{(m)} + Z_2^{(m)} + ...)}{\sum_{m=1}^{M}(Z_0^{(m)} + Z_1^{(m)} + ...)} \tag{1}$$

This is the standard Harris estimator of propagation [8], [10]. We will also estimate the variance of the offspring distribution as explained in Appendix A and use this variance in section IV to choose the load shed discretization.

The empirical probability distribution of the number of initial outages Z_0 can be obtained as

$$P[Z_0 = z_0] = \frac{1}{M}\sum_{m=1}^{M} I[Z_0^{(m)} = z_0], \tag{2}$$

where the notation $I[\text{event}]$ is the indicator function that evaluates to one when the event happens and evaluates to zero when the event does not happen.

The average number of initial outages θ is estimated as

$$\hat{\theta} = \frac{1}{M}\sum_{m=1}^{M} Z_0^{(m)}. \tag{3}$$

We assume that the offspring distribution is a Poisson distribution with mean λ. There are general arguments suggesting that the choice of a Poisson offspring distribution is appropriate [3], based on the offspring outages being selected from a large number of possible outages that have small probability and are approximately independent.

We are most interested in statistics of the total number of outages Y_∞ produced by the cascades, since Y_∞ indicates either the total number of line outages or the total load shed. Given the probability distribution of Z_0 and the average propagation $\hat{\lambda}$, the formula for calculating the probability distribution of Y_∞ is the following mixture of Borel-Tanner distributions [4]:

$$P[Y_\infty = r] = \sum_{z_0=1}^{r} P[Z_0 = z_0] z_0 \hat{\lambda}(r\hat{\lambda})^{r-z_0-1}\frac{e^{-r\hat{\lambda}}}{(r-z_0)!}.$$

IV. Processing and discretizing load shed

This section discusses the processing and discretization of the load shed data to produce integer counts of the discretization unit of load shed. Our Galton-Watson branching process assumes a Poisson offspring distribution, which always has its variance equal to its mean. The key idea is to choose the discretization so that the offspring distribution also has its variance equal to its mean and is therefore consistent with the Poisson distribution.

A. Initial Processing

Very small load shed amounts (less than 0.5% of total load) are considered negligible and rounded to zero. The cascades with no load shed are discarded. For those cascades that have no load shed in initial generations and non-negligible load shed in subsequent generations, the initial generations with no load shed are discarded to guarantee that generation zero always starts with a positive amount of load shed. Moreover, when any intermediate generation with zero load shed is encountered, the current cascade will be ended and a new cascade started at the next generation with nonzero load shed. There are M cascades in total and $X_k^{(m)}$ denotes the load shed at generation k of cascade m.

B. Discretization

To apply a Galton-Watson branching process to load shed data, we need to discretize the continuously varying load shed data in MW to integer multiples of a unit of discretization Δ. In particular, we use the following discretization to convert the load shed $X_k^{(m)}$ MW to integer multiples $Z_k^{(m)}$ of Δ MW:

$$Z_k^{(m)} = \text{int}\left[\frac{X_k^{(m)}}{\Delta} + 0.5\right] \tag{4}$$

where $\text{int}[x]$ is the integer part of x. We add 0.5 before taking the integer part to ensure that the average values of $Z_k^{(m)}$ and $X_k^{(m)}/\Delta$ are equal. The discretization (4) is straightforward except that the choice of the discretization unit Δ matters. Here we give a justifiable way to choose Δ. A previous paper [3] made a heuristic and subjective choice of Δ.

We first discuss how the choice of the discretization unit Δ affects the mean and variance of the offspring distribution. Consider a second choice of discretization unit Δ' and corresponding discretized data ${Z'}_k^{(m)}$ so that

$${Z'}_k^{(m)} = \text{int}\left[\frac{X_k^{(m)}}{\Delta'} + 0.5\right].$$

Neglecting the effects of rounding, so that $Z_k^{(m)} \approx X_k^{(m)}/\Delta$ and ${Z'}_k^{(m)} \approx X_k^{(m)}/\Delta'$, yields

$${Z'}_k^{(m)} \cong \frac{\Delta}{\Delta'} Z_k^{(m)}. \tag{5}$$

The mean λ of the offspring distribution is a dimensionless ratio of load shed that does not depend on the scaling or units of the load shed (see (1)). To understand this, we recall that the offspring distribution is defined as the distribution of the number of units of load shed in generation $k+1$ *assuming one unit of load shed in generation* k. When the discretization is changed from Δ to Δ', both $Z_k^{(m)}$ and $Z_{k+1}^{(m)}$ are multiplied by Δ/Δ', so that the mean of the offspring distribution does not change. The variance of $Z_{k+1}^{(m)}$ is multiplied by $(\Delta/\Delta')^2$, but since $Z_k^{(m)}$ is also multiplied by Δ/Δ', the offspring distribution variance σ^2, which is the variance in generation $k+1$ arising from one unit of load shed in generation k, is only multiplied by Δ/Δ'. To explain these scalings with an example, suppose that in the first cascade there are 2 units of load shed in generation k with discretization Δ so that $Z_k^{(1)}=2$. Then, according to the principles of branching processes, the distribution of load shed $Z_{k+1}^{(1)}$ in generation $k+1$ is the sum of two independent copies of the offspring distribution for discretization Δ, and therefore has mean 2λ and variance $2\sigma^2$. Now change the discretization to $\Delta'=2\Delta$ so that $Z_{k+1}'^{(1)} \cong Z_{k+1}^{(1)}/2$ and $Z_k'^{(1)} = Z_k^{(1)}/2 = 1$. Then the distribution of $Z_{k+1}'^{(1)}$ is the offspring distribution for discretization Δ', which has mean $EZ_{k+1}'^{(1)} = EZ_{k+1}^{(1)}/2 = 2\lambda/2 = \lambda$, and variance $E[(Z_{k+1}'^{(1)}-\lambda)^2] = E[(Z_{k+1}^{(1)}-2\lambda)^2]/4 = 2\sigma^2/4 = \sigma^2/2$. Thus changing the discretization unit Δ strongly affects σ^2, and in particular increasing Δ decreases σ^2 proportionally.

When we consider the influence of rounding on these scaling approximations, the mean of the offspring distribution λ is only slightly affected by Δ while the variance of the offspring distribution σ^2 has an overall strong tendency of decreasing proportionally with the increase of Δ (strict monotonicity for small changes in Δ is not guaranteed). We notate these strong and weak dependencies by writing $\sigma^2(\Delta)$ and $\lambda(\Delta)$ respectively and we can get $\sigma^2(\Delta) \cong \sigma^2(\Delta')\,\Delta'/\Delta$ and $\lambda(\Delta) \cong \lambda(\Delta')$ from the above analysis.

Our calculations assume a Poisson offspring distribution, and it is well known that the variance of a Poisson distribution is equal to its mean. Therefore, to be consistent with the Poisson distribution, we need to choose a discretization so that the variance of the offspring distribution is equal to its mean. That is, we need to choose Δ so that $\sigma^2(\Delta)=\lambda(\Delta)$.

Specifically, we discretize the data for $\Delta = 1\,\mathrm{MW}$ and estimate $\sigma^2(1)$ and $\lambda(1)$ and then the Δ that satisfies $\sigma^2(\Delta) = \lambda(\Delta)$ is approximately $\sigma^2(1)/\lambda(1)$. Calculated values of Δ are shown in Table II. This procedure requires σ^2 to be estimated and this calculation is explained in the Appendix. We have found that the distribution of load shed is not very sensitive to the exact value of Δ, so that the estimate of σ^2 and the chosen Δ need not be very accurate.

C. *Processing after Discretization*

After discretization the initial load shed in some cascades may become zero. These cascades are discarded.

V. Improved OPA and AC OPA Model

This section briefly summarizes the main features of the improved OPA and AC OPA cascading outage simulations. Since these are enhanced versions of the OPA model, we first summarize the OPA model. For detailed descriptions, see [11], [12], [13] for OPA, [14], [15] for improved OPA, and , [15], [16], [17] for AC OPA. This paper uses the forms of the improved OPA and AC OPA simulations in which the power grid is fixed and does not evolve or upgrade.

A. *OPA*

The OPA model represents transmission lines, loads, and generators and computes the network power flows with a DC load flow. Each simulation run starts from a solved base case solution for the power flows, generation, and loads that satisfy circuit laws and constraints. To obtain diversity in the runs, the system loads at the start of each run are varied randomly around their mean values by multiplying by a factor uniformly distributed in $[2-\gamma,\,\gamma]$. Initial line outages are generated randomly by assuming that each line can fail independently with probability p_0. This crudely models initial line outages due to a variety of causes including lightning, wildfires, bad weather, and operational errors. Whenever a line fails, the generation and load is redispatched to satisfy the transmission line and generation constraints using standard linear programming methods. The optimization cost function is weighted to ensure that load shedding is avoided where possible. If any lines were overloaded during the optimization, then these lines are those that are likely to have experienced high stress, and each of these lines fails independently with probability p_1. The process of redispatch and testing for line outages is iterated until there are no more outages. OPA has been validated against WECC data [18].

B. *Improved OPA*

Compared with OPA, the improved OPA considers in the cascading process the misoperation of protective relays and the failure of EMS or communications in the dispatching center. The maloperation of relays is simulated by tripping lines that are not overloaded with probability $\xi \times |F/F^{\max}|^a$. ξ is the base probability of unwanted operation of relays and $|F/F^{\max}|$ is the load ratio of the transmission line. The standard linear programming problem in OPA is calculated with probability η and $1-\eta$ is the failure rate of the dispatching center, which can be caused by breakdown of the EMS or communications.

C. *AC OPA*

Both OPA and improved OPA use DC power flow and only consider active power. Bus voltages are considered constant. In contrast, the AC OPA model uses AC power flow and thus can consider reactive power and voltage. The operation mode of the system is first determined by AC OPF and load shedding and will be readjusted by AC OPF until there is no further outage or failure once outages happen. The AC OPF takes into account both reactive power and voltage constraints.

VI. Results

This section presents results giving the branching process parameters computed from simulated cascades, the distributions of outages predicted from these parameters, and the comparison with the simulated distributions of outages. The cascading outage data is produced by the improved OPA simulation [14] on the NPGC system and the AC OPA simulation [16], [17] on the IEEE 118 bus test system.

The NPGC system consists of Heilongjiang, Jilin, Liaoning, and the northern part of Inner Mongolia. It has about 500 transmission lines, more than 300 transformers, 250 substations, and about 200 generators. We consider 500kV and 220kV transmission lines and substations which correspond to a 568 bus system. The NPGC system data includes the line limits.

The IEEE 118 bus system data is standard, except that line flow limits are determined by running the fast dynamics of the improved OPA and the slow dynamics of OPA that selectively upgrades lines in response to their participation in blackouts [12], starting from an initial guess of the line limits. This procedure results in a coordinated set of line limits.

For the improved OPA model, we use the same parameters as those in [14] and $p_1 = 0.999$, $\xi = 0.001$, $a = 10$, and $\eta = 0.95$. For the NPGC system $p_0 = 0.0007$, which is the same as [14], and for the IEEE 118 bus system $p_0 = 0.0001$, which is the same as [2]. For both models the load variability $\gamma = 1.67$ as in [2].

For testing the branching process model, the simulation is run so as to produce 5000 cascading outages with a nonzero number of line outages or non-negligible load shed.

A. Line Outage Results

We test the branching process model for estimating the distribution of total line outages on cascading data from the improved OPA simulations on the NPGC system under different load levels. The estimated branching process parameters $\hat{\theta}$ and $\hat{\lambda}$ are shown in Table I. The distributions of line outages are shown in Figs. 1–2. The distributions of total line outages (dots) and initial line outages (squares) are shown, as well as a solid line indicating the total line outages predicted by the branching process. The branching process data is also discrete, but is shown as a line for ease of comparison. The branching process prediction of the distribution of total line outages matches the empirical distribution quite well.

The branching process model is also tested with cascading data generated by the AC OPA model on the IEEE 118 bus system. The results are shown in Figs. 3–5. The branching process prediction of the distribution of total line outages matches the empirical distribution well.

TABLE I
ESTIMATED PARAMETERS FOR LINE OUTAGE DATA

system	model	load level	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\lambda}_{500}$
NPGC	Im OPA	1.15	1.69	0.02	0.02
NPGC	Im OPA	1.3	2.14	0.14	0.14
NPGC	Im OPA	1.6	2.54	0.19	0.20
IEEE 118	AC OPA	1.0	4.38	0.40	0.40
IEEE 118	AC OPA	1.2	7.01	0.52	0.52
IEEE 118	AC OPA	1.4	10.08	0.63	0.63

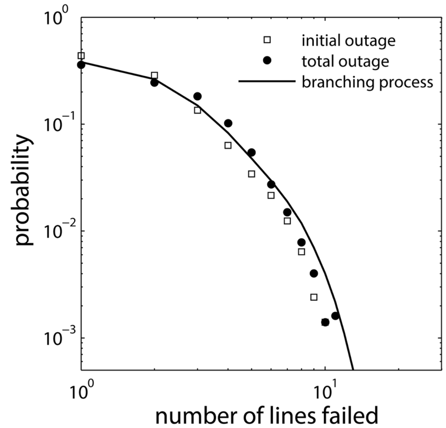


Fig. 1. Probability distributions of number of line outages by improved OPA on NPGC system at load level 1.3 times the base case. Dots indicate total numbers of line outages and squares indicate initial line outages; both distributions are empirically obtained from the simulated cascades. The solid line indicates the distribution of total numbers of line outages predicted with the branching process.

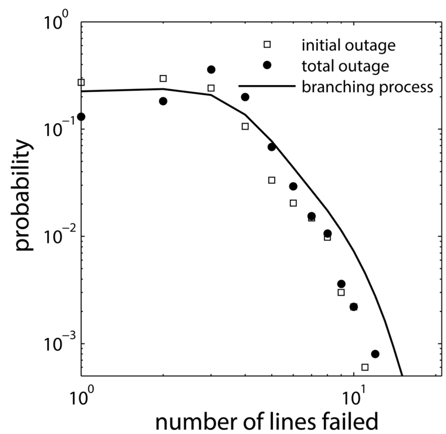


Fig. 2. Probability distributions of number of line outages by improved OPA on NPGC system at load level 1.6 times the base case.

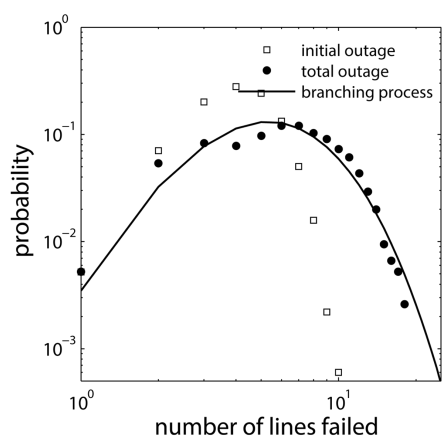


Fig. 3. Probability distributions of number of line outages by AC OPA on IEEE 118 bus system at base case load level.

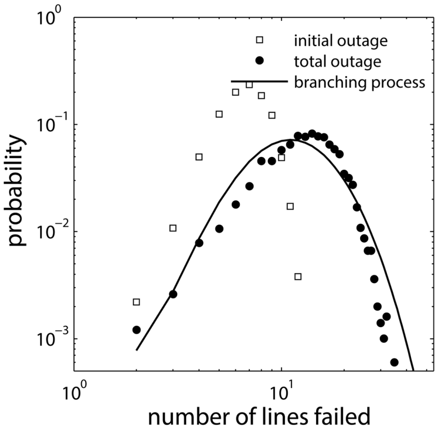


Fig. 4. Probability distributions of number of line outages by AC OPA on IEEE 118 bus system at load level 1.2 times the base case.

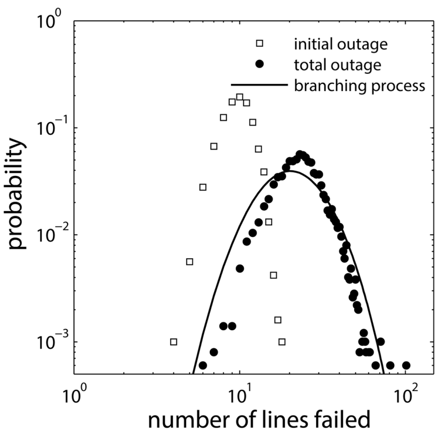


Fig. 5. Probability distributions of number of line outages by AC OPA on IEEE 118 bus system at load level 1.4 times the base case.

B. Load Shed Results

We also test the branching process model for estimating the distribution of total load shed on cascading data from the improved OPA and AC OPA simulations on the NPGC and IEEE 118 bus systems. The estimated branching process parameters $\hat{\theta}$ and $\hat{\lambda}$ are listed in Table II. In contrast with the line outages, the average propagation $\hat{\lambda}$ does not always increase with load level. When the load level increases from 1.3 times base case to 1.6 times base case for improved OPA simulation on NPGC system, $\hat{\lambda}$ decreases from 0.29 to 0.22.

TABLE II
ESTIMATED PARAMETERS FOR LOAD SHED DATA

system	model	load level	$\hat{\theta}$ (MW)	$\hat{\lambda}$	Δ (MW)	$\hat{\lambda}_{500}$
NPGC	Im OPA	1.15	177	0.10	198	0.11
NPGC	Im OPA	1.3	441	0.29	388	0.35
NPGC	Im OPA	1.6	5404	0.22	2465	0.23
IEEE 118	AC OPA	1.0	160	0.42	90	0.44
IEEE 118	AC OPA	1.2	224	0.58	167	0.58
IEEE 118	AC OPA	1.4	407	0.67	425	0.68

The results comparing the distributions of load shed for improved OPA simulations are shown in Figs. 6–8 and those for AC OPA are shown in Figs. 9–11. The branching process predictions of the distribution of load shed match the empirical distributions well, with a less good match for the high stress case of Figure 11.

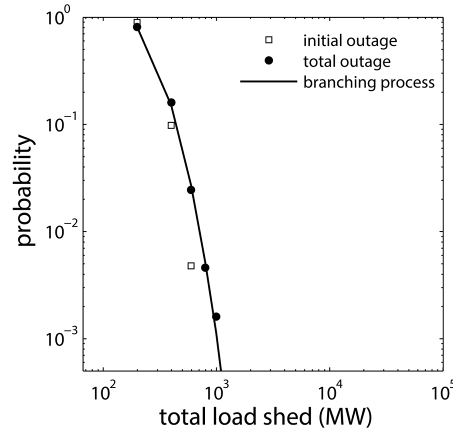


Fig. 6. Probability distribution of total load shed by improved OPA on NPGC system at load level 1.15 times the base case.

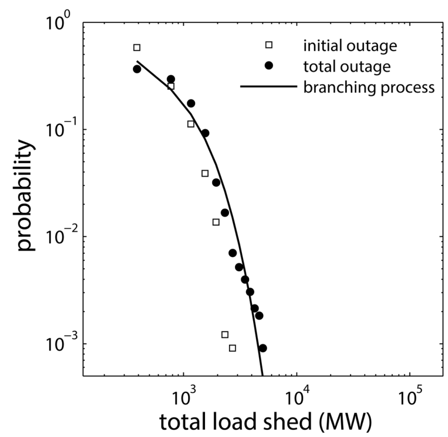


Fig. 7. Probability distribution of total load shed by improved OPA on NPGC system at load level 1.3 times the base case.

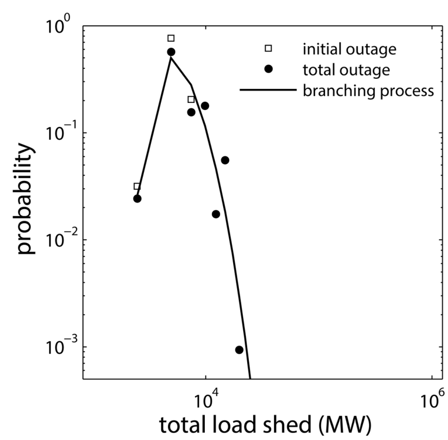


Fig. 8. Probability distribution of total load shed by improved OPA on NPGC system at load level 1.6 times the base case.

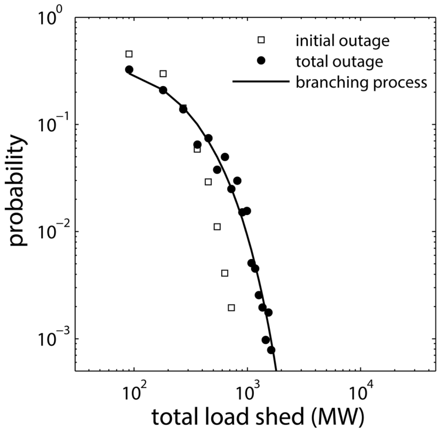


Fig. 9. Probability distribution of total load shed by AC OPA on IEEE 118 bus system at base case load level.

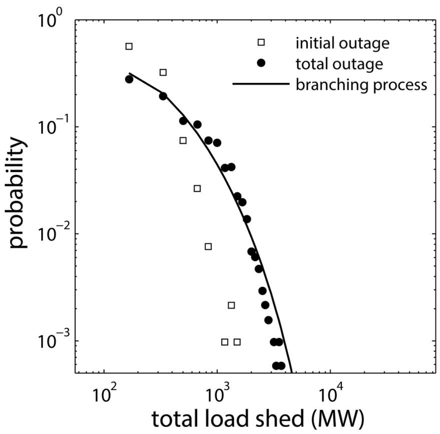


Fig. 10. Probability distribution of total load shed by AC OPA on IEEE 118 bus system at load level 1.2 times the base case.

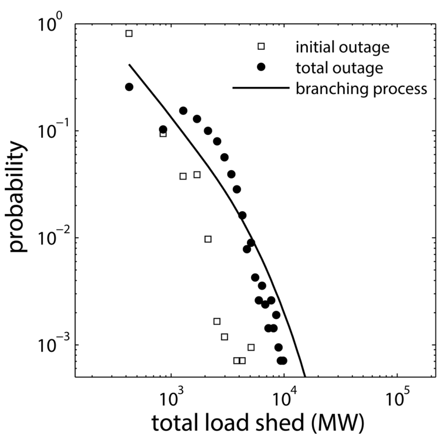


Fig. 11. Probability distribution of total load shed by AC OPA on IEEE 118 bus system at load level 1.4 times the base case.

A reasonable objection to the comparison in section VI-A and VI-B is that the same data is used both to estimate the distribution and to obtain the empirical distribution. To address this objection we divided the data into separate fitting and validation sets. Specifically, we estimate the distribution from the odd numbered cascades and compare with the empirical distribution for the even numbered cascades, and vice verse. The resulting matches, which are shown in Appendix B, are also satisfactory.

C. Efficiency

For testing and validating the branching process model, we use 5000 cascades, but the analysis and results of [2] and [3] suggest that, once validated, the approach can be applied for much fewer cascades. Tables I and II show the propagation $\hat{\lambda}_{500}$ estimated using the first 500 cascades. $\hat{\lambda}_{500}$ is close to $\hat{\lambda}$ in all cases except for load shed in NPGC at loading level 1.3.

Previous work has demonstrated the efficiency gained by estimating the distribution of discretized load shed by first estimating the propagation and then using a branching process model. According to [3], "if the initial load shed distribution is known accurately, then accurately estimating the distribution of the total amount of load shed via discretization and the Galton-Watson branching process requires substantially fewer cascades."

In particular, let $p_{\rm branch}$ be the probability of shedding total load $S\Delta$ MW, computed via estimating λ from $K_{\rm branch}$ simulated cascades with non-negligible load shed and then using the branching process model. $p_{\rm branch}$ is conditioned on a non-negligible amount of load shed. Let $p_{\rm empiric}$ be the probability of shedding total load $S\Delta$ MW, computed empirically by simulating $K_{\rm empiric}$ cascades with non-negligible load shed. If we require the same standard deviation for both methods, then section IV of [3] derives the following approximation of the ratio R of the required number of simulated cascades as

$$R = \frac{K_{\rm empiric}}{K_{\rm branch}} = \frac{p_{\rm empiric}(1-p_{\rm empiric})\theta}{\lambda(1-\lambda)\Delta}\left(\frac{dp_{\rm branch}}{d\lambda}\right)^{-2} \qquad (6)$$

R is a ratio describing the gain in efficiency when using branching process rather than empirical methods. To obtain numerically a rough estimate of R, we evaluate (6) for almost largest total load shed $S\Delta$ MW for all six load shed cases. Here we sort the total load shed of $K_{\rm empiric}$ cascades in descending order and choose $S\Delta$ to be the ith one, where $i = {\rm int}[5\% K_{\rm empiric} + 0.5]$. $dp_{\rm branch}/d\lambda$ is estimated by numerical differencing. The results in Table III show that $K_{\rm empiric}$ exceeds $K_{\rm branch}$ by an order of magnitude or more. That is, if the initial load shed distribution is known accurately, then accurately estimating the distribution of the total amount of load shed via discretization and the Galton-Watson branching process requires substantially fewer cascades.

TABLE III
EFFICIENCY GAIN R FOR LOAD SHED DATA

system	model	load level	$S\Delta$ (MW)	R
NPGC	Im OPA	1.15	419	28
NPGC	Im OPA	1.3	1851	27
NPGC	Im OPA	1.6	12759	15
IEEE 118	AC OPA	1.0	830	28
IEEE 118	AC OPA	1.2	1581	111
IEEE 118	AC OPA	1.4	3946	91

Similarly, we can also confirm the efficiency gains for line outage data by evaluating (6) with $\Delta = 1$ for almost largest total line outages S, which can be determined in a similar way as choosing $S\Delta$. The results in Table IV show that $K_{\rm empiric}$ exceeds $K_{\rm branch}$ by an order of magnitude or more. Similar or better efficiency gains for estimating the distribution of number of lines outages have also been observed in [2].

TABLE IV
EFFICIENCY GAIN R FOR LINE OUTAGE DATA

system	model	load level	S	R
NPGC	Im OPA	1.15	4	114
NPGC	Im OPA	1.3	6	68
NPGC	Im OPA	1.6	6	21
IEEE 118	AC OPA	1.0	13	32
IEEE 118	AC OPA	1.2	23	41
IEEE 118	AC OPA	1.4	46	36

The efficiency gains are not surprising because estimating the parameters of a distribution is generally expected to be more efficient than estimating the distribution empirically.

VII. CONCLUSION

In this paper, we advance the application of branching processes to cascades of power grid outages by improving the processing of load shed data and testing the capability of the branching process to estimate distributions of blackout size.

Load shed data needs to be discretized before applying a Galton-Watson branching process with a Poisson offspring distribution. We describe and justify a new way to select the unit of discretization so that it is compatible with the Poisson offspring distribution. The capability to describe load shed cascading data is useful, and the method should also apply to any continuously varying quantity that is used to track the progression and impact of cascading outages.

We used enhanced versions of the OPA simulation on the IEEE 118 bus system and NPGC systems to produce cascading outage data for the transmission lines outaged and the load shed. After discretizing the load shed data, we estimated the average propagation and the average size of the initial outage that are the parameters of a branching process model. We then used the branching process model to estimate the distributions of lines outaged and load shed. The estimated distributions are close to the empirical distributions, suggesting that the Galton-Watson branching process model with an average propagation can capture some overall statistical aspects of the cascading of line outages and load shed. While previous work on other simulations and test cases has reached similar conclusions [5], [3], it is necessary to test branching processes with many different simulations and test cases to establish the use of branching processes for power system cascading outages. In addition to providing a useful summary describing cascade statistics, the branching process enables the average propagation and the distribution of blackout size to be estimated with much fewer simulated cascades [2]. This is a significant advantage since simulation time is a limiting factor when studying cascading blackouts.

APPENDIX A
ESTIMATING VARIANCE OF THE OFFSPRING DISTRIBUTION

This appendix explains estimating the variance of the offspring distribution that is used to determine the discretization of the load shed data in section IV-B. Useful background is in [7] or [8].

Let $Y_t^{(m)} = Z_0^{(m)} + Z_1^{(m)} + \cdots + Z_t^{(m)}$ be the total outages up to and including generation t of cascade m, and let $Y_\infty^{(m)}$ be the total outages in all generations of cascade m. Group the cascades into K groups so that each group has at least n initial outages. The choice of K and n is discussed below. Let $Y_\infty^{[k]}$ be the sum of the outages in group k and let $\hat\lambda_\infty^{[k]}$ be the Harris estimator (1) evaluated for the outages in group k.

Given $\hat\lambda$, which is the Harris estimator (1) of the mean of the offspring distribution for all the cascades, we estimate the variance of the offspring distribution as

$$\hat\sigma^2 = \frac{1}{K}\sum_{k=1}^{K} Y_\infty^{[k]}(\hat\lambda_\infty^{[k]} - \hat\lambda)^2. \tag{7}$$

We discuss the choice of n and K. It is desirable to have large n and large K, but there is a tradeoff between n and K. If there are a total of p initial outages in the simulated cascades, then $nK \cong p$. To determine values of n and K, we simulated 5000 realizations of an ideal Galton-Watson branching process with Poisson offspring distribution of known mean and variance for a range of values of n. We found that for small n and large K, (7) underestimates the variance whereas for large n and small K, (7) is a noisy estimate. We chose $n = 20\,p/5000$ and $K \cong 250$ based on this testing.[2]

There are other estimators of offspring variance in the literature [10], [19], and while (7) seemed to perform well enough for our purposes here, the question of the most effective variance estimator remains open, especially for our cascading data in which the offspring mean and variance may vary somewhat with generation.

Now we give some justification for (7), mainly by following Yanev [20]. Consider a branching process X_0, X_1, X_2, ... with $X_0 = n$ and offspring mean λ and offspring variance σ^2. We assume $0 < \lambda < 1$ and $\sigma^2 < \infty$. (A branching process with $X_0 = n$ is equivalent to n independent branching process cascades whose initial outages are all 1. Thus X_0, X_1, X_2, ... corresponds to one of the K groups of cascades discussed above.) Let $Y_t(n) = X_0 + X_1 + \cdots + X_t$ denote the total outages up to and including generation t. Let $\hat\lambda_t(n) = (Y_t(n) - X_0)/Y_{t-1}(n)$ be the Harris estimator for the offspring mean computed from X_0, X_1, X_2, ..., X_t. Write $X_0^{(i)}$, $X_1^{(i)}$, $X_2^{(i)}$, ... for the branching process starting from the ith initial outage only. Let $W_t^{(i)} = \sum_{k=1}^{t}\left(X_k^{(i)} - \lambda X_{k-1}^{(i)}\right)$.

From [20, equation (8)]:

$$Y_{t-1}(n)(\hat\lambda_t(n) - \lambda)^2 = \frac{\frac{1}{n}\left(\sum_{i=1}^{n} W_t^{(i)}\right)^2}{\frac{1}{n}Y_{t-1}(n)} \tag{8}$$

[2]When load shed data is discretized, the number of initial outages p is inversely proportional to the discretization unit Δ. Choosing n proportional to p ensures that the grouping of the cascades into K groups will remain unchanged when the unit of discretization Δ changes.

We have

$$\frac{1}{n}\left(\sum_{i=1}^{n} W_t^{(i)}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(W_t^{(i)}\right)^2 + \frac{1}{n}\sum_{\substack{1\le i,j\le n\\ i\neq j}} W_t^{(i)}W_t^{(j)}$$

[20] states that $EW_t^2 = \sigma^2(1+\lambda+...+\lambda^{t-1})$ and $EW_t = 0$. Then the strong law of large numbers implies

$$\frac{1}{n}\left(\sum_{i=1}^{n} W_t^{(i)}\right)^2 \overset{\text{a.s.}}{\to} \sigma^2(1+\lambda+...+\lambda^{t-1}).$$

Moreover, $Y_{t-1}(n)/n \overset{\text{a.s.}}{\to} EY_{t-1} = 1+\lambda+...+\lambda^{t-1}$, a nonzero constant, as $n\to\infty$. Hence from (8),

$$Y_{t-1}(n)(\hat{\lambda}_t(n)-\lambda)^2 \overset{\text{a.s.}}{\to} \sigma^2 \quad \text{as } n\to\infty.$$

Now letting $t\to\infty$, we get $Y_\infty(n)(\hat{\lambda}_\infty(n)-\lambda)^2 \overset{\text{a.s.}}{\to} \sigma^2$ as $n\to\infty$, and $E\left[Y_\infty(n)(\hat{\lambda}_\infty(n)-\lambda)^2\right]\to\sigma^2$ as $n\to\infty$.

Write $\hat{\lambda}(p)$ to show the dependence of the Harris estimator $\hat{\lambda}$ for all the cascades on the total number of initial outages p. Define

$$\Delta = Y_\infty(n)\left[(\hat{\lambda}_\infty(n)-\lambda)^2 - (\hat{\lambda}_\infty(n)-\hat{\lambda}(p))^2\right] = \Delta_1+\Delta_2$$

where $\Delta_1 = Y_\infty(n)(\hat{\lambda}_\infty(n)-\lambda)(\hat{\lambda}(p)-\lambda)$

and $\Delta_2 = Y_\infty(n)(\hat{\lambda}_\infty(n)-\hat{\lambda}(p))(\hat{\lambda}(p)-\lambda)$.

By Cauchy-Schwartz,

$$|E\Delta_1|^2 \le E\left[\left(Y_\infty(n)(\hat{\lambda}_\infty(n)-\lambda)\right)^2\right]E\left[(\hat{\lambda}(p)-\lambda)^2\right] \quad (9)$$

Since $Y_\infty(n)(\hat{\lambda}_\infty(n)-\lambda) = (1-\lambda)\left[Y_\infty(n)-\frac{n}{1-\lambda}\right]$, from [20, equation (6) and Lemma 1] we obtain

$$E\left[\left(Y_\infty(n)(\hat{\lambda}_\infty(n)-\lambda)\right)^2\right] = (1-\lambda)^2\frac{n\sigma^2}{(1-\lambda)^3} = \frac{n\sigma^2}{1-\lambda}$$

From [20, Theorem 2(ii)],

$$E\left[(\hat{\lambda}(p)-\lambda)^2\right]\to\sigma^2(1-\lambda)/p \quad \text{as } p\to\infty.$$

Choose $p\ge n^{1+\delta}$ for some $\delta>0$. Then from (9),

$$|E\Delta_1|^2 \le \frac{n\sigma^2}{1-\lambda}\frac{\sigma^2(1-\lambda)}{n^{1+\delta}} = \frac{\sigma^4}{n^\delta}\to 0 \quad \text{as } n\to\infty,$$

and $E\Delta_1\to 0$ as $n\to\infty$.

$$\Delta_1-\Delta_2 = Y_\infty(n)(\hat{\lambda}(p)-\lambda)^2 = \frac{Y_\infty(n)}{n}n(\hat{\lambda}(p)-\lambda)^2$$

For $p\ge n^{1+\delta}$, and as $n\to\infty$,

$$nE\left[(\hat{\lambda}(p)-\lambda)^2\right] = n\frac{\sigma^2(1-\lambda)}{p} \le n\frac{\sigma^2(1-\lambda)}{n^{1+\delta}}\to 0.$$

That is, $E\left[n(\hat{\lambda}(p)-\lambda)^2\right] \overset{\text{L}_1}{\to} 0$. Moreover $Y_\infty(n)/n \overset{\text{L}_1}{\to} 1/(1-\lambda)$, and hence $\Delta_1-\Delta_2 \overset{\text{L}_1}{\to} 0$ and $E\left[\Delta_1-\Delta_2\right]\to 0$ as $n\to\infty$. Since we have already shown $E\Delta_1\to 0$, we conclude that $E\Delta = E\left[\Delta_1+\Delta_2\right]\to 0$. Hence for $p\ge n^{1+\delta}$,

$$E\left[Y_\infty(n)(\hat{\lambda}_\infty(n)-\hat{\lambda}(p))^2\right]\to\sigma^2 \quad \text{as } n\to\infty.$$

Since $\hat{\sigma}^2\to E\left[Y_\infty(n)(\hat{\lambda}_\infty(n)-\hat{\lambda}(p))^2\right]$ as $K\to\infty$, $\hat{\sigma}^2\to\sigma^2$ as $K,n\to\infty$.

Appendix B
Computing Estimated and Empirical Distributions from Separate Data

This appendix presents results by dividing the data into separate fitting and validation sets. Typical results for both line outage data and load shed data are given.

In Figs. 12–13, we estimate the distribution of the total outages using the branching process from the odd numbered cascades (solid line) and compare with the empirical distribution from the even numbered cascades (dots). The squares are empirically obtained distribution of initial outages from the even numbered cascades.

In Figs. 14–15, we estimate the distribution using the branching process from the even numbered cascades (solid line) and compare with the empirical distribution from the odd numbered cascades (dots). The squares are empirically obtained distribution of initial outages from the odd numbered cascades.

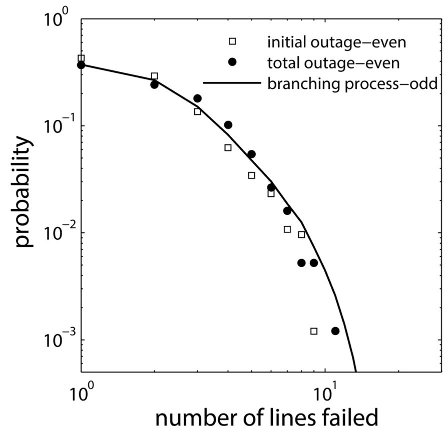


Fig. 12. Probability distribution of total line outages by improved OPA on NPGC system at load level 1.3 times the base case.

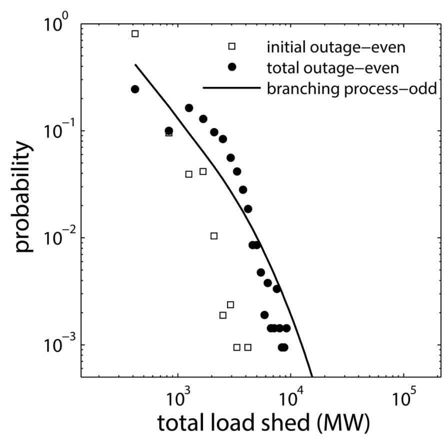


Fig. 13. Probability distribution of total load shed by AC OPA on IEEE 118 bus system at load level 1.4 times the base case.

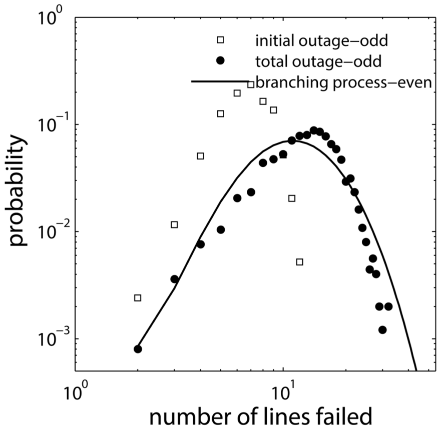


Fig. 14. Probability distribution of total line outages by AC OPA on IEEE 118 bus system at load level 1.2 times the base case.

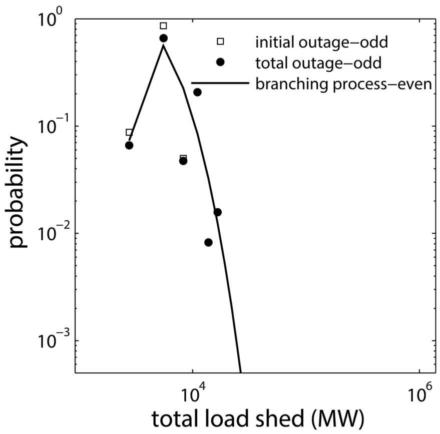


Fig. 15. Probability distribution of total load shed by improved OPA on NPGC system at load level 1.6 times the base case.

REFERENCES

[1] IEEE PES CAMS Task Force on Cascading Failure, "Initial review of methods for cascading failure analysis in electric power transmission systems," IEEE PES General Meeting, Pittsburgh PA USA, July 2008.

[2] I. Dobson, J. Kim, and K. R. Wierzbicki, "Testing branching process estimators of cascading failure with data from a simulation of transmission line outages," *Risk Analysis*, vol. 30, pp. 650–662, 2010.

[3] J. Kim, K. R. Wierzbicki, I. Dobson, and R. C. Hardiman, "Estimating propagation and distribution of load shed in simulations of cascading blackouts," *IEEE Systems Journal*, vol. 6, no. 3, Sept. 2012, pp. 548-557.

[4] H. Ren and I. Dobson, "Using transmission line outage data to estimate cascading failure propagation in an electric power system," *IEEE Transactions on Circuits and Systems Part II*, vol. 55, pp. 927–931, 2008.

[5] I. Dobson, "Estimating the propagation and extent of cascading line outages from utility data with a branching process," *IEEE Trans. Power Systems*, vol. 27, no. 4, Nov. 2012, pp. 2146-2155.

[6] J. Kim, I. Dobson, "Approximating a loading-dependent cascading failure model with a branching process," *IEEE Trans. Reliability*, vol. 59, no. 4, December 2010, pp. 691-699.

[7] K. B. Athreya and P. E. Ney, *Branching Processes*, Dover NY, 2004.

[8] T. E. Harris, *Theory of Branching Processes*, Dover NY, 1989.

[9] I. Dobson, B. A. Carreras, and D. E. Newman, "A branching process approximation to cascading load-dependent system failure," *37th Hawaii Intl. Conference on System Sciences*, Hawaii, HI, Jan. 2004.

[10] P. Guttorp, *Statistical Inference for Branching Processes*, Wiley NY, 1991.

[11] I. Dobson, B. A. Carreras, and V. E. Lynch, "An initial model for complex dynamics in electric power system blackouts," *34th Hawaii Intl. Conference on System Sciences*, Maui, HI, pp. 710–718, Jan. 2001.

[12] 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.

[13] H. Ren, I. Dobson, B.A. Carreras, "Long-term effect of the n-1 criterion on cascading line outages in an evolving power transmission grid," *IEEE Trans. Power Systems*, vol. 23, no. 3, August 2008, pp. 1217-1225.

[14] S. Mei, F. He, X. Zhang, S. Wu, and G. Wang, "An improved OPA model and blackout risk assessment," *IEEE Trans. Power Systems*, vol. 24, pp. 814–823, May. 2009.

[15] S. Mei, X. Zhang, M. Cao, *Power Grid Complexity*, Tsinghua University Press, Beijing and Springer, Berlin 2011.

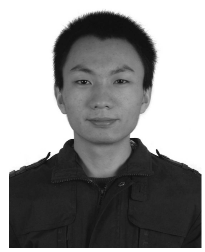
[16] S. Mei, Yadana, X. Weng, and A. Xue, "Blackout model based on OPF and its self-organized criticality," *Proceedings of the 25th Chinese Control Conference*, pp. 7–11, 2006.

[17] S. Mei, Y, Ni. Weng, G. Wang, and S. Wu, "A study of self-organized criticality of power system under cascading failures based on AC-OPA with voltage stability margin," *IEEE Trans. Power Systems*, vol. 23, pp. 1719–1726, Nov. 2008.

[18] B.A. Carreras, D.E. Newman, I. Dobson, N.S. Degala, "Validating OPA with WECC data," 46th Hawaii Intl. Conference on System Sciences, Maui, HI, Jan. 2013.

[19] J. P. Dion, "Estimation of the variance of a branching process," *Ann. Statist*, vol. 3, pp. 1183–1187, 1975.

[20] N. M. Yanev, "On the statistics of branching processes," *Theory of Probability and its Applications*, vol. 20, pp. 612–622, 1975.



Junjian Qi (S'12) received the B.E. degree in Electrical Engineering from Shandong University in 2008 and is currently a Ph.D. candidate of the Department of Electrical Engineering of Tsinghua University. His research interests include blackout model, cascading failure, and power system state estimation.

Ian Dobson (F'06) received the BA in Mathematics from Cambridge University and the PhD in Electrical Engineering from Cornell University.

He is currently Sandbulte professor of engineering at Iowa State University. His interests include cascading failure risk, blackouts, and synchrophasors.

Shengwei Mei (SM'05) received the B.E. degree in mathematics from Xinjiang University, Urumqi, China, the M.S. degree in operations research from Tsinghua University, Beijing, China, and the Ph.D degree in automatic control from the Chinese Academy of Sciences, Beijing, in 1984, 1989, and 1996, respectively.

He is currently a Professor at Tsinghua University. His research interests include power system analysis and control.