Preprint. Not for redistribution. Contents are (c) IEEE.
IEEE Transactions on Power Systems, Vol. 27, No. 3, 2012

A "Random Chemistry" Algorithm for Identifying Multiple Contingencies that Initiate Cascading Failure

Margaret J. Eppstein and Paul D. H. Hines, *Member, IEEE*

***Abstract*—This paper describes a stochastic "Random Chemistry" (RC) algorithm to identify multiple $(n-k)$ contingencies that initiate large cascading failures in a simulated power system. The method requires only $O(\log(n))$ simulations per contingency identified, which is orders of magnitude faster than random search of this combinatorial space. We applied the method to a model of cascading failure in a power network with $n = 2896$ branches and identify $148,243$ unique, minimal $n-k$ branch contingencies $(2 \leq k \leq 5)$ that cause large cascades, many of which would be missed by using pre-contingency flows, linearized line outage distribution factors, or performance indices as screening factors. Within each $n-k$ collection, the frequency with which individual branches appear follows a power-law (or nearly so) distribution, indicating that a relatively small number of components contribute disproportionately to system vulnerability. The paper discusses various ways that RC generated collections of dangerous contingencies could be used in power systems planning and operations.**

***Index Terms*—Contingency screening, cascading failure, power systems reliability**

I. Introduction

Power systems are generally operated according to the $n-1$ security criterion, where any single one of n components can fail without violating bus voltage, branch flow or stability limits. This approach works well if the probability of k multiple, nearly simultaneous failures ($n-k$ contingencies) is vanishingly small. However, $n-k$ contingencies do occur and sometimes trigger sequences of cascading outages that result in large blackouts, such as the events of August 14, 2003 in North America [1] and November 4, 2006 in Europe [2]. Because blackout sizes have a power-law (scale-free) size distribution [3], the risk due to multiple, simultaneous outages is large, despite the relatively small probabilities of these events. As a result, North American reliability standards now require that, "Each Transmission Operator shall operate to protect against instability, uncontrolled separation, or cascading outages resulting from multiple outages" [4].

In order to meet current standards, there is a need for tools to support efficient $n-k$ analysis. However, the computational complexity of finding high-impact $n-k$ contingencies makes complete enumeration and random search infeasible. Exhaustive $n-k$ analysis requires $n!/(n-k)!$ simulations if we assume that sequence matters, and $n!/(k!(n-k)!)$ if sequence is ignored [5]. For a system with $n = 10^4$ components that could fail, simulating all $n-3$ contingencies would require more than 10^{11} simulations, which is computationally infeasible. Due to the size of the search space, random search methods such as Monte-Carlo simulation are not sufficient.

Numerous techniques have been proposed to reduce the computational burden associated with multiple contingency analysis. The earliest contingency screening methods involved developing a performance metric that correlates to system stress, and evaluating the sensitivity of this metric to various component outages (e.g., [6], [7]). For multiple contingency analysis, Ref. [8] proposes a method that uses event trees and substation configuration data to identify collections of probable multiple-failure combinations. Reference [9] describes a method to identify collections of $n-k$ contingencies that result in overloaded transmission lines using line outage distribution factors (see Sec. IV-F). Optimization methods have also been proposed to identify high-impact $n-k$ contingencies [10]-[12]. Such optimization approaches can identify small sets of the highest-impact vulnerabilities, but are not designed to identify large unbiased collections for statistical analysis.

Most existing screening and optimization methods are based on identifying disturbances that result in limit violations. However, the presence of a limit violation is not a sufficient condition for a cascading failure. In many cases, limit violations are eliminated after one or two subsequent component outages and do not result in significant loss of load (e.g., the small cascade of July 3, 1996 [13]).

A method that could efficiently identify large unbiased collections of $n-k$ contingencies that lead to cascading failure could be useful for a variety of planning and operations applications in power systems engineering. Such collections could be used to provide valuable insight into the risk associated with large blackouts (e.g., by serving as input to risk estimation methods [14], [15]), estimate a system's cascading outage propagation rate [16], or serve as test scenarios in assessing the efficacy of a given procedure or control system designed to protect against cascading outages. Along planning time horizons, knowing the relative contributions that individual components make to $n-k$ vulnerability could be used to prioritize components for upgrades or preventative maintenance (such as checking for hidden relay failures). Subsequently, the method could be used to estimate how upgrades to these components

This work was supported in part by the U.S. Dept. of Energy award DE-OE0000447, a workforce development sub-award from Sandia National Laboratories supported by the U.S. Dept. of Energy through Inter-Entity Work Order M610000767, and the Vermont Advanced Computing Center which is supported by the NASA award NNX 06AC88G.

The authors are with the College of Engineering and Mathematical Sciences, University of Vermont, Burlington, VT USA (e-mail: Maggie.Eppstein@uvm.edu, paul.hines@uvm.edu).

would impact the resulting dangerous contingency collections. For operational time scales, the method could be used to flag transmission lines that contribute disproportionately to multiple contingency vulnerability in order to assist operators with decisions that affect power flow on these lines.

Thus motivated, the goal of this paper is to describe a search method for efficiently identifying large collections of minimal $n-k$ contingencies (for k simultaneous outages) that lead to large cascading failures in a simulated power system. Our paper is organized as follows. Sec. II describes the proposed search method and the cascading failure simulator with which we test the search, and Sec. III describes the test grid used in the experiments. Sec. IV describes experimental results that illustrate the method and some patterns in the resulting contingency sets. We provide some discussion and conclusions in Sec. V.

II. Algorithms

A. Random chemistry (RC) algorithm

We have developed a stochastic approach for rapidly identifying a minimal simultaneous $n-k$ contingency that results in system failure. By "minimal", we mean that no smaller subset of the k outages result in system failure. Here we will use U to represent the universal set of all n possible outages, and we assume that the system is initially $n-1$ secure. "System failure" is defined as a sequence of cascading outages that exceeds some user-specified blackout size criterion. For brevity, we hereafter use the phrase "malignancy" or "malignant contingency" to mean a minimal set of k outages that result in system failure.

Our approach was originally inspired by Kauffman [17], who outlined a simple hypothetical nonlinear feature-set selection procedure (dubbed "Random Chemistry") for stochastically detecting small auto-catalytic sets of k nonlinearly interacting molecules out of a large number of n candidate molecules, in only $O(\log(n))$ steps. More recently, Eppstein *et al.* [18] adapted this idea into an algorithm for finding k epistatically-interacting genetic variations that predispose for disease in genome-wide association studies.

The paragraphs that follow describe a version of the Random Chemistry (RC) algorithm adapted to find malignant $n-k$ malignancies in a power system.

RC Step 1: Search for a (large, non-minimal) random set S of component outages that causes system failure. An initial target set size k_{init} is specified, and random combinations of outages of components in U, each of size k_{init}, are tested until a set S is found that causes the system to fail. If no such set is found within some constant number of tries T, the target set size k_{init} is doubled (although it is truncated to a maximum size of n) and the process is repeated. As long as k_{init} is sufficiently large, this step typically requires only one or a few tries, since it is trivial to find a large (non-minimal) set of outages that cause system failure. In the worst case, this step is guaranteed to terminate successfully when $S = U$, so is upper bounded by $O(log_2(n - k_{\text{init}}))$ trials. For the results in this paper, we used $k_{\text{init}} = 80$ and $T = 20$.

RC Step 2: Reduce the size of the discovered set S by stochastically reducing the set by a constant fraction.

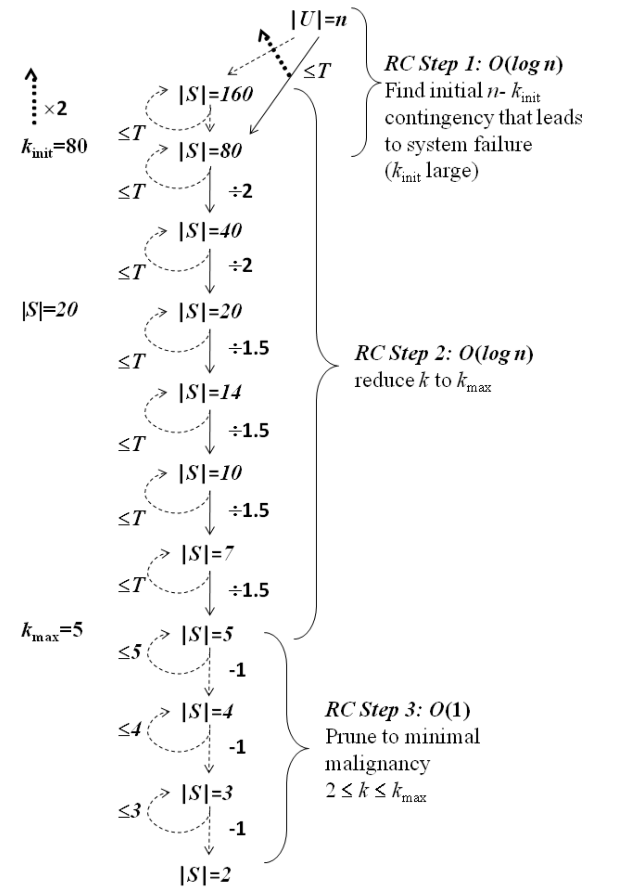

Figure 1. The set reduction process in the RC algorithm, illustrated for the RC parameter settings used in this work. In RC Step 1, a large random set of branch outages is identified that results in system failure. During RC Step 2, the candidate set of component outages S is repeatedly stochastically reduced by the fractional amount indicated and then tested with a simulator; as soon as a reduced set of outages is found that results in a system failure, the reduced set is accepted and the reduction process is repeated on the smaller set. If the maximum number of tries T is ever exceeded, the RC trial is aborted and a new trial is initiated. In RC Step 3, after the set size has been reduced to $k_{\max}$, the set is linearly pruned until a minimal $n-k$ malignancy is identified, for $2 \leq k \leq k_{\max}$. Dashed arrows denote optional paths, depending on stochastic results.

Up to T random subsets of S, each of size $|S|/C$ (truncated to not less than a final target set size $k_{\max}$) for some small constant C, are tested until a new subset S_2 is found that still causes system failure. S is then replaced by S_2 and the process is repeated until $|S| = k_{\max}$, where $k_{\max}$ is a small positive integer that provides an upper bound on the malignancy size k sought by the algorithm. If, at any point, the number of attempts surpasses the specified maximum number of tries T, this RC search run is aborted and is considered a "failed" RC trial. (Note, however, that a "failed" RC trial simply implies that the smallest minimal $n-k$ malignancy contained in S probably has $k > k_{\max}$, which is larger than what we are seeking). This set reduction step is upper bounded by $O(\log_C(k_{\text{init}} - k_{\max}))$ trials, where $k_{\max} < k_{\text{init}} \leq n$. To reduce the number of aborted runs, we use a larger (more aggressive) $C = 2$ when the set size is greater than 20, and then reduce C to 1.5 for $|S| \leq 20$.

RC Step 3: Prune individual outages from the remaining set S until a minimal $n-k$ malignancy is identified.

Individual components are tested for possible removal from S, until a minimal malignant subset is identified (that is, the resulting set of outages S results in system failure, but no smaller subset of S does). This step requires between $k_{\max}-2$ and $\sum_{i=3}^{k_{\max}} i$ simulations.

Each execution of the RC algorithm (comprising RC Steps 1-3, above) is thus upper bounded by $O(\log(n))$ grid simulations, as graphically depicted in Fig. 1. The specific number of trials required to achieve a "successful" RC run (i.e., when RC Step 2 succeeds in reducing $|S|$ to $k_{\max}$), and the number of simulations per RC run, both vary stochastically and are a function of the particular grid as well as the RC parameter settings. The runtime of the RC algorithm is relatively insensitive to the choice of the initial set size k_{init}. However the choice of the constant $k_{\max}$ is more important. If $k_{\max}$ is too large, then the runtime becomes dominated by RC Step 3 (which, although is technically upper bounded by a constant time, requires up to $\sum_{i=3}^{k_{\max}} i$ simulations). On the other hand, if $k_{\max}$ is too small, then there is a higher probability that RC Step 2 will fail, since it becomes increasingly likely that the smallest minimal $n-k$ malignancy contained in S is larger than $k_{\max}$. In practice, we observe that choosing $k_{\max}=5$ provides a good balance between these constraints.

B. Obtaining large collections of malignancies

The RC algorithm is repeated as many times as desired to sample the space of $n-k$ minimal malignancies (for $2 \leq k \leq k_{\max}$). Since RC trials are independent, the likelihood that a duplicate malignancy will be found increases as the collection of identified malignancies grows. If one has identified j out of m_k malignancies (for a given k), of which i are unique, then the probability p_j that the next $n-k$ malignancy found is unique is:

$$p_j = \frac{m_k - i}{m_k} \tag{1}$$

If a sufficient fraction of the complete collection of malignancies for a given k is identified, then the rate of change in p_j becomes observable. We can take advantage of this to estimate the size of the complete collection m_k, as follows. Taking the reciprocal of (1), the expectation is that the next unique $n-k$ malignancy will be found after a total of $\Delta j = p_j^{-1}$ additional $n-k$ malignancies have been found. Solving for m_k yields the expected value:

$$\hat{m_k} = \frac{i\Delta j}{\Delta j - 1} \tag{2}$$

As $i \to m_k$, the error in this estimate of m_k decreases. The error can be further damped by averaging over successive estimates of $\hat{m_k}$ as j is further increased (skipping the singularities in $\hat{m_k}$ where the observed $\Delta j = 1$).

The stochastic set reduction approach of the RC algorithm is inherently biased toward finding malignancies with lower k, by virtue of the fact that the number of contained subsets of size k is largest for $k = |S|/2$, in conjunction with the observation that the frequency with which a random subset of a given size k causes system failure decreases with increasing k (see Sec. IV). However, as long as a uniform random number generator is used for selecting subset components in RC Steps 1 and 2, the pruning of components in RC Step 3 is done in a uniformly randomly permuted order, and the search is terminated when the first minimal $n-k$ malignancy is identified during pruning, then repeated application of RC search will yield unbiased collections of $n-k$ malignancies, for a given k. Alternatively, if the aim is to further decrease the search time and increase the success rate of individual RC trials, one could elect to non-uniformly bias the selection of subsets (e.g., by selecting branch components with probabilities corresponding to pre-contingency flows) in RC Steps 1 and 2, and to exhaustively search for and return all minimal $n-k$ malignancies embedded in the final set S in RC Step 3. However, doing so would necessarily propagate this sampling bias into the resulting collection of $n-k$ malignancies.

C. Cascading failure simulator (CFS)

The modeling of cascading failure in power systems is a challenging problem. There are many mechanisms by which a small set of disturbances can propagate to become large blackouts, including cascading thermal overloads, relay failure, voltage collapse, dynamic instability, and operator error [5]. Different modeling approaches have advantages and disadvantages in terms of capturing subsets of these mechanisms [19]. High-level statistical models [20], [16] provide high-level information about system risk, but are not designed to identify specific components that contribute to risk. Some have presented work on purely topological models of cascading failure [21], however the flow patterns in these models differ substantially from those described by Kirchhoff's and Ohm's laws, and therefore need to be treated with caution [22]. Models of cascading failure that are based on a dc power flow, such as OPA [23]-[24], are computationally efficient and numerically stable and thus facilitate statistical observations from large numbers of simulations. However dc power flow models have known limitations [25], and cannot capture some aspects of cascading failure, such as dynamic instability, voltage collapse or distance relays. Sequential static cascading failure models based on the ac power flow exist [26], [27], but are more computationally intensive and require that one make assumptions to handle power-flow cases that do not converge. Dynamic, mid/long-term transient stability models of specific historical cascading failures exist (e.g., [28]), however these models tend to require extensive calibration and are very computationally expensive. Even a full dynamic model of cascading failure requires assumptions, since operator behavior is almost always a critical component to cascading failure, and many of the system parameters (such as area control error management procedures) are unknown.

Thus all cascading failure models simplify power system dynamics to some extent. Because this paper focuses on the evaluation of the RC algorithm, and because there is not yet a generally accepted and publicly available ac power flow model of cascading failure, the simulator used here is based on the dc power flow approximations (modified from [22]). The simulator described here can easily be replaced with a different one that is more appropriate to a particular type of cascading failure.

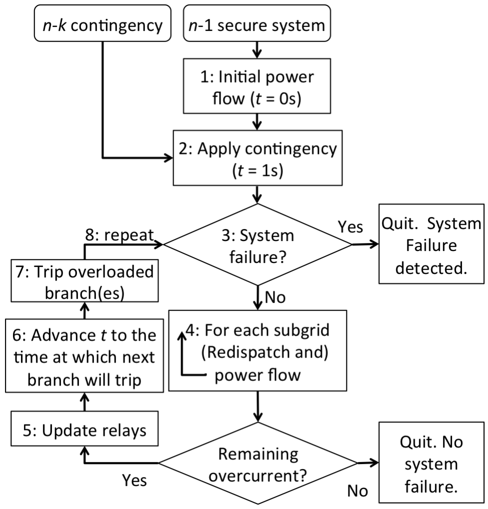

Figure 2. Algorithm flow diagram for the cascading failure simulator (CFS) used in this study.

Figure 2 illustrates our cascading failure simulator (CFS), which is designed to test whether an $n-k$ contingency results in system failure. The paragraphs that follow describe CFS in additional detail.

CFS Step 1: Initialize power flow. Initialize the simulation by calculating pre-contingency power flow on a power system that has been initially dispatched to ensure $n-1$ security with respect to the universal set (U) of n components that can fail.

CFS Step 2: Apply contingency. Apply the $n-k$ contingency to the system by changing the status of the k failed components, then update the system susceptance matrix **B** and identify the separated islands in the system.

CFS Step 3: Check for system failure. If "system failure" is detected, based on some user-defined criterion, terminate the simulation and mark this contingency as malignant.

CFS Step 4: Re-dispatch and recalculate power flow. If the results from Steps 2 or 7 indicate that the system has just separated into additional islands, re-dispatch the generation in each new island as follows. First, allow generators to ramp up or down to match supply and demand as closely as possible, constrained by limits $P_{g,\min}(t)$ and $P_{g,\max}(t)$. These limits are set to the most restrictive of either the upper and lower generation limits of the machine or the ramp rate of the machine times the amount of time since the last power flow calculation. Thus, for cases with longer time periods between power flow calculations, generators are allowed to increase/decrease output over a greater range of values. If, after re-dispatching generators, there remains more generation than load ($R = (\sum_{g\in G} P_g - \sum_{d\in D} P_d) > 0$, where G is the set of generator buses, and D is the set of load buses), generators in this island are tripped sequentially, starting with the smallest machine, until $R \leq 0$. The rationale for allowing generator tripping to correct for supply surplus cases is that generator ramping can be a slow process; when automatic generation controls cannot quickly correct frequency error, some generators will likely trip as a result of overspeed relays. If, after this, $R < 0$ load is shed by multiplying all loads in the island by a scalar $\lambda = \sum_{g\in G} P_g / \sum_{d\in D} P_d$. After achieving a balance between supply and demand in each island, a standard dc power flow calculation is used to find power/current flow on each transmission line. Because a supply/demand balance is achieved before this power flow calculation, the solution will be independent of the choice of a reference bus. If no branches are overloaded after the power flow calculation, the simulator terminates and indicates that this contingency did not result in system failure.

CFS Step 5: Update relays. Use a time-delayed overcurrent relay on each branch to determine when/if a branch trips due to overload. While it is decreasingly common for overcurrent relays to be used in the bulk power grid, these relays approximately represent a number of processes by which branches may shut down, such as the overheating of a transmission line and sagging into vegetation, the time an operator is willing to tolerate an overload before manually disconnecting the line, or backup relaying systems, such as Zone 2 or Zone 3 distance protection. In CFS branch j, with power flow f_j and flow limit $\overline{f_j}$, fails if its accumulated overload (o_j) exceeds a limit $\overline{o_j}$. The accumulated overload between time periods t and $t+\Delta t$ (Δo_j) is calculate from:

$$\Delta o_j(t,\Delta t) = \begin{cases} \int_t^{t+\Delta t} (f_j(t) - \overline{f_j})dt, & \text{if } f_j(t) > \overline{f_j} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

The threshold $\overline{o_j}$ is chosen such that a transmission line will trip after 5 seconds of being 50% above the branch flow limit. As defined, if branch j's flow is more than 150% of its limit, it will trip faster, and slower for less than 150%. To find Δt, the simulator finds the minimum time until the next branch fails. Practically this means that, unless there are two branches with identical state variables $o_j(t)$ and $f_j(t)$, not more than one branch will trip at each iteration of the algorithm.

CFS Step 6: Advance time until the next branch failure. Set $t = t + \Delta t$, with Δt as selected in CFS Step 5.

CFS Step 7: Trip branch(es). Switch the branch status to failed for each branch for which a relay has tripped and update **B** to reflect the change in network structure.

CFS Step 8: repeat from CFS Step 3. Continue until either there are no remaining overloaded branches, or until a system failure is detected.

It should be noted that CFS is similar to other quasi steady state cascading failure models. Our model differs from the short-term behavior of OPA [23] only in the method by which generation is re-dispatched after system separation and in that branch failure is treated deterministically. Whereas OPA frequently re-dispatches generators using an algorithm that seeks to satisfy the line flow constraints, here we assume that the system operators have limited ability to optimally re-dispatch generation during a cascading failure, and rely largely on automatic generation control to achieve a new power balance. The model is also similar to those described in [29] and [30], except that CFS allows for generator tripping to correct for extreme imbalances, whereas [29] and [30] allow generators to ramp quickly between $P_{\min}$ and $P_{\max}$.

III. Test network

We applied the RC search method to identify minimal $n-k$ contingencies that trigger system failure in a model of the 2004 peak winter Polish transmission system, which is available with MATPOWER [31]. The test case has $n = 2896$ branches (transmission lines and transformers), 2383 buses, and 24.6 GW of load. The pre-contingency dc branch power flows have a mean of 34.3 MW, median of 18.7 MW, standard deviation of 53.8 MW, and maximum of 918 MW. We increased 12 of the initial branch flow limits to achieve a pre-contingency power flow case such that no single branch outage could trigger cascading system failure, as defined below. After doing so, all of the pre-contingency flows were less than 95.2% of their long-term (Rate A) limits. With the case thus prepared we applied the RC method with U defined as the set of all 2896 branches in the network.

For the results presented here, "system failure" is defined as a state in which at least 10% of the buses are no longer connected to the largest island. While various definitions of system failure could be employed, we selected a network-separation-based measure of system disruption, rather than a load-shed-based one, for several reasons. Most importantly, the method chosen to rebalance supply and demand after the grid separates into islands becomes increasingly important as the network subdivides. Once significant network separation occurs, the system is clearly in a dangerous state, but the size of the blackout that will result depends heavily on automated control and operator actions, making the simulation results increasingly uncertain. To compare load shedding and network separation metrics, we measured the blackout sizes in MW and fraction of nodes separated for all $n-2$ contingencies in the test network that were either predicted by (4) to cause overloads or found by the RC method to cause separation of at least 10% of the buses. The histogram in Fig. 3 shows a tri-modal distribution, with a natural break between the lowest and middle modes at 15% separation (for clarity, we only show contingencies resulting it at least 3% separation). All of the disruptions larger than 15% separation caused a loss of at least 1727 MW of power (Fig. 3, right panel); by lowering the threshold to the more conservative 10% separation, we were able to detect all of the the disruptions in the middle and highest modes, as well as most (all but 5) of the disruptions in the lowest mode that also caused a power loss of at least 1727 MW (Fig. 3, right panel, above and to the right of the dashed lines).

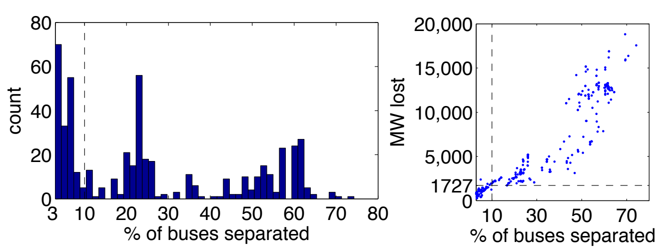

Figure 3. Left: histogram of disruption size, shown for $n-2$ contingencies causing at least 3% separation. Right: Scatter plot of the relationship between power loss and percent of buses separated. Dashed lines indicate the thresholds used for major disruptions defined as 'system failure'.

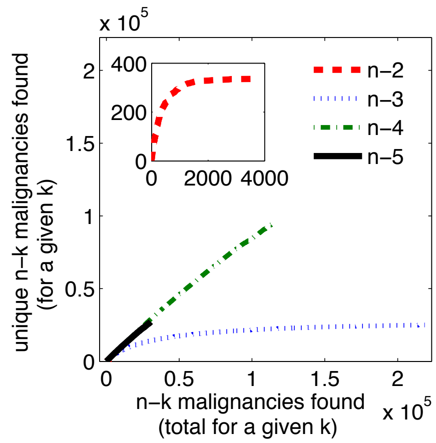

Figure 4. Comparing the number of unique versus total $n-k$ malignancies found by the RC algorithm, for a given $k \in \{2, 3, 4, 5\}$, gives an indication of how complete the collections are for each value of k. The inset shows a close-up near the origin showing the curve for all $n-2$ malignancies found.

IV. Results

A. Collections of malignancies identified with RC search

We report on 735,500 successful RC trials, in which we identified a total of $\{336, 25059, 95677, 27171\}$ unique $\{n-2, n-3, n-4, n-5\}$ malignancies, respectively. It is highly likely that the 336 unique $n-2$ malignancies found form a complete collection, since no new $n-2$ malignancies were found in over 7×10^5 successful RC trials after the last unique $n-2$ malignancy was identified (Fig. 4); this number also agrees closely with the estimate of $\hat{m}_2 = 338$ obtained using the method described in Sec. II-B. The asymptotic behavior of the $n-3$ curve (Fig. 4) implies that the identified collection of unique $n-3$ malignancies is approaching completion, and we estimate $\hat{m}_3 = 2.69 \times 10^4$. On the other hand, the identified collections of unique $n-4$ and $n-5$ malignancies are far from complete, as most newly identified $n-4$ and $n-5$ sets continued to be unique (Fig. 4).

In 1000 successful RC trials on this system and with these settings, we observed that it took an average of 1.57 RC trials to successfully find an $n-k$ malignancy, for $2 \le k \le 5$. Some of these RC trials required as few as 12 simulations, while others required up to 147 simulations, with a mean of 48 simulations per RC trial. On the test system with $k_{\max} = 5$, we identified minimal malignant sets with $k = \{2, 3, 4, 5\}$ in $\{50\%, 30\%, 16\%, 4\%\}$ of the successful RC trials, respectively. Thus, for this system, it typically requires an average of only about 151, 251, 471, and 1884 ($= 48 \times 1.57 \times \frac{100}{\{50,30,16,4\}}$) simulations to find an $n-\{2, 3, 4, 5\}$ malignancy, respectively.

B. Efficiency of RC search *vs.* random search

The expected number of simulations required to find one out of 336 $n-2$ malignancies using random search on this grid would be $12,476 = \binom{2896}{2}/336$. Using our estimate of $\hat{m}_3$, the expected number of simulations required to find an $n-3$ malignancy using random search on this grid is about $150,000 \approx \binom{2896}{3}/(2.69 \times 10^4)$. Thus, random search requires

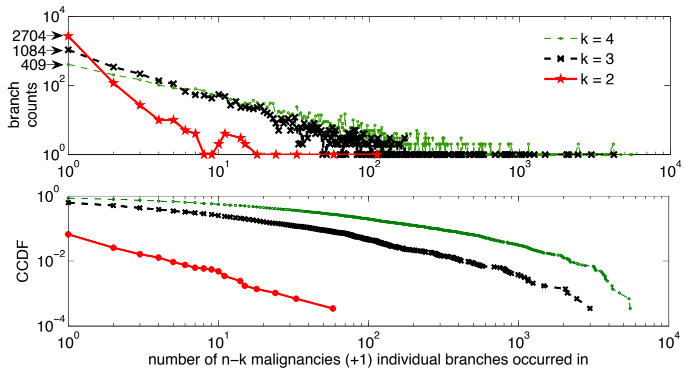

Figure 5. The number of times each of the 2896 individual branches occurred in the observed $n-k$ malignancies (top), and the complementary cumulative distribution functions of these data (bottom). Note that the x-axes have been shifted to the right by one so that branches occurring in zero malignancies are shown at the far left, with counts indicated.

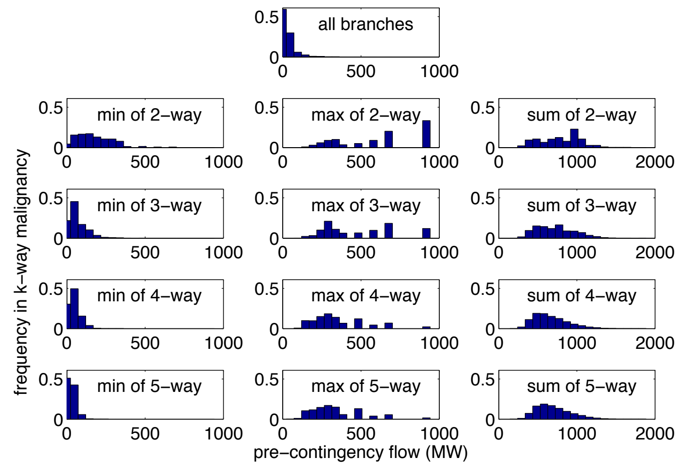

Figure 6. Frequency histograms for pre-contingency branch power flows. The top subplot shows the distribution of pre-contingency flows of all 2896 branches. The remaining four rows of subplots correspond to the distributions of pre-contingency flows in branches participating in $n-k$ malignancies, for $k = \{2, 3, 4, 5\}$ respectively, where the first column shows the minimum pre-contingency flow of the k branches, the second column shows the maximum pre-contingency flow of the k branches, and the third column shows the sum of flows in all k branches.

about 85 times as many simulations as RC search to find an $n-2$ malignancy in this grid model, and about 600 times as many simulations as RC search to find an $n-3$ malignancy. While it is impossible to reasonably estimate the number of $n-4$ or $n-5$ malignancies in this system (since these curves in Fig. 4 have not begun to asymptote), and empirical studies using random search yielded too few $n-4$ or $n-5$ malignancies to reliably estimate success rates, it is likely that the relative efficiency of RC search compared to random search continues to grow as k increases.

C. *Individual branch contributions to system vulnerability*

Out of the 2896 branches in the system, we observed a total of 192, 1812, and 2487 individual branches that occurred in at least one $n - \{2, 3, 4\}$ malignancy, respectively. (Since the $k - 5$ collection is so small (Fig. 4), we have omitted these data from this subsection.) The frequencies with which branches occurred in $n-k$ malignancies follow very heavy-tailed distributions with modes at 0 (Fig. 5). The (presumably complete) $n-2$ collection of malignancies exhibits a strong fit to a power law distribution ($R^2 = 0.96$). The frequency distributions for higher k malignancies span several orders of magnitude and appear to be approaching power law distributions (with shallower slopes at higher k), but show more truncated tails (Fig. 5), probably resulting from the fact that these collections are not complete.

Most branch outages never or rarely interacted with other branch outages to trigger cascading failures (Fig. 5, points near the left). In this study, 98.7%, 98.1%, and 96.7% of all branches occurred in $\leq 1\%$ of the $n - \{2, 3, 4\}$ malignancies, respectively. Conversely, a few branches appeared in very large numbers of malignancies (Fig. 5, points in the heavy tails on the right), and thus contributed disproportionately to system vulnerability. For example, in this system one branch occurred in 112 (33% of all) $n-2$ malignancies. As k increased, several individual branches were found to occur in thousands of malignancies. However, owing to the large numbers of malignancies at high k, no single branch occurred in more than 17% of the identified $n-3$ malignancies or in more than 6% of the identified $n-4$ malignancies. Practically speaking, this means that as k increases it becomes increasingly difficult to predict which line failures contribute to system vulnerability based on metrics designed for $n-1$ or $n-2$ contingency screening.

D. *Power flow characteristics of branches in malignancies*

When a transmission line or transformer fails, the current that was previously moving through that branch is immediately re-routed to parallel paths to satisfy Kirchhoff's and Ohm's laws. Thus, the more power that is interrupted by the outage, the more the flows on parallel transmission paths will increase, or decrease, which may result in an overloaded component and trigger a cascade. Therefore, it seems reasonable to conjecture that there is a relationship between the amount of power flowing through a transmission line and the frequency with which it occurs in $n-k$ malignancies. Consequently, we compared the distributions of pre-contingency branch flows in the system as a whole, with those of the branches that occur in various $n-k$ malignancies. While the maximum and sum of the pre-contingency flows in all k-branches participating in a given $n-k$ malignancy tends to be relatively high (Fig. 6, columns 2 and 3, respectively), the minimum pre-contingency flow of one of the k branches in a malignancy may be quite low, especially as k increases (Fig. 6, column 1). This indicates that in many cases the joint failure of a large transmission line and one or more small ones can interact to trigger large blackouts. If conventional contingency screening methods were used to initially prune out transmission lines and transformers with less than 50 MW from the set of possible component outages (U), 12% of the $n-2$ malignancies would not have been found, and the problem gets worse as k increases. Doing so would have hidden $\{49\%, 62\%, 83\%\}$ of the $n - \{3, 4, 5\}$ malignancies found in these experiments, respectively.

E. *Malignant* vs. *flow-matched benign pairs*

We selected 336 new pairs of branch outages, such that the distributions of the minimum, maximum, and sum of the pre-contingency flows in these 336 new branch pairs were statistically identical to those of the 336 $n-2$ malignancies (KS test, $p > 0.99$). We verified that none of these new pairs triggered system failure, and henceforth refer to them as "benign pairs". We then searched for network features that might distinguish malignant from benign pairs. To date, we have found no significant differences in the distributions of individual branch metrics, including pre-contingency flow and measures of network centrality, applied to the 115 branches that occur in malignant but not benign pairs *vs.* the 134 branches that occur in these benign pairs but not malignant pairs. We have found significant differences in the distributions of some pairwise metrics, including (a) the minimum electrical distances between branch pairs (KS test, $p < 2.5 \times 10^{-5}$) and (b) the minimum shortest path (in number of hops) between branch pairs (KS test, $p < 4.5 \times 10^{-12}$). In both cases, the distribution of malignant pairs was left-shifted relative to benign pairs, however there is still significant overlap in their distributions.

F. *Assessment of performance indices*

Finally, we examined to what extent performance indices can be used to detect malignant contingencies without simulation, or to reduce the search space so that fewer contingencies need simulation to identify those that result in large cascading failure. The change in flows $\Delta f_{i|\overline{r,s}}$ resulting from the simultaneous outage of branches r and s, can be computed using line outage distribution factors ($d_{i,j}$) [32] and the technique in [9] as:

$$\Delta f_{i|\overline{r,s}} = \begin{bmatrix} d_{i,r} & d_{i,s} \end{bmatrix} \begin{bmatrix} 1 & -d_{r,s} \\ -d_{s,r} & 1 \end{bmatrix}^{-1} \begin{bmatrix} f_r \\ f_s \end{bmatrix} \tag{4}$$

Computing $\Delta f_{i|\overline{r,s}}$ for all 4,191,960 branch pairs in the test system indicated that 186,428 pairwise branch outages would result in branch overloads. Of these, only 220 resulted in system failure in the simulator. On the other hand, because line outage distribution factors do not return values for cases that separate a network into islands, 116 of the $n-2$ malignancies found by RC search were not flagged as overloads using (4).

Eqn. 4 can be used to produce a performance index for double line outages, following the method of [6], as:

$$PI_{\overline{r,s}} = \sum_{i=1}^{n} w_i \left(\frac{f_i + \Delta f_{i|\overline{r,s}}}{\bar{f}_i} \right)^2 \tag{5}$$

where w_i is a branch weight, and f_i is a measure of flow (power or current) on branch i, with flow limit $\overline{f_i}$. For the results reported here we set the weights to normalized pre-contingency branch flows ($w_i = f_i / \sum_{j=1}^{n} f_j$), which provided improved results over the case with $w_i = 1, \forall i$. If one were to screen out $n-2$ contingencies with a post-contingency change in performance index of $\Delta PI_{\overline{r,s}} > 10^{-3}$, the search space is reduced by only 91%, but doing so would preclude the identification of 4% of the $n-2$ malignancies. At a higher threshold of $\Delta PI_{\overline{r,s}} > 10^{-2}$, the size of the search space would be reduced by 3 orders of magnitude, but 42% of the malignancies would be eliminated from the search set.

V. Discussion and Conclusions

This paper describes a stochastic "Random Chemistry" (RC) method to identify large sets of minimal $n-k$ contingencies that result in large blackouts in a cascading failure simulator. The RC algorithm requires only $O(\log(n))$ simulations per dangerous contingency (malignancy) found, which is orders of magnitude more efficient than random search for realistically sized grids.

We applied the RC method to a 2383 bus power system model and identified 148,243 unique, simultaneous $n-k$ branch malignancies ($2 \leq k \leq 5$) that result in large cascading failures. The results show a number of interesting patterns that illustrate the utility of the RC algorithm. First, the results indicate that the frequencies with which specific branches interact to trigger large blackouts follow power law (or nearly so) distributions. The heavy tails on these distributions indicate that a small number of branches occur in orders of magnitude more blackout scenarios than the majority of branches. Improvements at these locations may have (at least near-term) reliability benefits. Second, we found that outages in branches with seemingly insignificant amounts of pre-contingency power flow can interact with failures in larger transmission lines to initiate very large cascading failures, and that the probability of such interactions increases with increasing k. Due to these interactions, prescreening branches based on low pre-contingency power flow would preclude identification of many $n-k$ malignancies. Finally, our results indicate that, even when using a linearized power system model, direct linearized methods, such as line outage distribution factors and performance indices, do not predict which combination of outages will result in large cascading failures well enough to preclude the need for simulation.

Future work will investigate how this algorithm might be incorporated into power system applications. For example, importance sampling techniques can be used to combine prior information about the probability and importance of component failures to generate estimates of blackout risk [14], [15]; the RC method could be used to generate unbiased importance distributions of component outages (see Sec. IV-C) to be used as input to such risk estimation approaches. This approach could be extended for evaluation of the blackout risk effects of wide area control schemes or system upgrades by simulting these changes and using the RC method to re-estimate grid component importance distributions.

The results in this paper were produced assuming that the k branch outages occur simultaneously. While we acknowledge that order and timing can have important effects, we have not yet explored the sensitivity of system failure to the $k!$ possible orders (each with various timings) of the identified $n-k$ malignancies. Similarly, our results were produced from a power system model that captures only a subset of cascading failure mechanisms. Future work will employ more detailed cascading failure simulation models, explore the sensitivity of the results to modeling assumptions, and evaluate the impact of order and timing on cascading failures.

VI. Acknowledgement

The authors thank D.M. Rizzo and three anonymous reviewers for constructive suggestions that helped to improve this paper.

References

[1] USCA, "Final Report on the August 14, 2003 Blackout in the United States and Canada," US-Canada Power System Outage Task Force, Tech. Rep., 2004.

[2] UTCE, "Final Report System Disturbance on 4 November 2006," Union for the Co-ordination of Transmission of Electricity, Tech. Rep., 2007.

[3] D. Newman, B. Carreras, V. Lynch, and I. Dobson, "Exploring complex systems aspects of blackout risk and mitigation," *IEEE Transactions on Reliability*, vol. 60, no. 1, pp. 134 –143, 2011.

[4] NERC, "Standard TOP-004-2: Transmission Operations," North American Electric Reliability Corporation, Standard, 2007.

[5] M. Vaiman, K. Bell, Y. Chen, B. Chowdhury, I. Dobson, P. Hines, M. Papic, S. Miller, and P. Zhang, "Risk assessment of cascading outages: Methodologies and challenges," *IEEE Transactions on Power Systems*, vol. (in press), 2011.

[6] G. Ejebe and B. Wollenberg, "Automatic contingency selection," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-98, pp. 97 – 109, 1979.

[7] V. Brandwajn, A. Kumar, A. Ipakchi, A. Bose, and S. Kuo, "Severity indices for contingency screening in dynamic security assessment," *Power Systems, IEEE Transactions on*, vol. 12, no. 3, pp. 1136 –1142, aug 1997.

[8] Q. Chen and J. McCalley, "Identifying high risk n-k contingencies for online security assessment," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 823 – 834, 2005.

[9] C. M. Davis and T. J. Overbye, "Multiple element contingency screening," *IEEE Transactions on Power Systems*, vol. 26, no. 3, 2011.

[10] V. Donde, V. Lopez, B. Lesieutre, A. Pinar, C. Yang, and J. Meza, "Severe multiple contingency screening in electric power systems," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 406–417, May 2008.

[11] D. Bienstock and A. Verma, "The n-k problem in power grids: New models, formulations, and numerical experiments," *SIAM Journal on Optimization*, vol. 20, no. 5, pp. 2352–2380, 2010.

[12] C. Rocco, J. Ramirez-Marquez, D. Salazar, and C. Yajure, "Assessing the vulnerability of a power system through a multiple objective contingency screening approach," *IEEE Transactions on Reliability*, vol. 60, no. 2, pp. 394 –403, 2011.

[13] WSCC Operations Committee, "Western Systems Coordinating Council Disturbance Report For the Power System Outages that Occurred on the Western Interconnection on July 2, 1996 and July 3, 1996," Western Systems Coordinating Council, Tech. Rep., 1996.

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

[15] Q. Chen and L. Mili, "Risk-based composite power system vulnerability evaluation to cascading failures using importance sampling," in *Proc. of the IEEE Power and Energy Society General Meeting*, Detroit, 2011.

[16] 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, no. 4, pp. 650–662, Apr. 2010.

[17] S. Kauffman, *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*. Oxford Univ. Press, USA, 1996.

[18] M. Eppstein, J. Payne, B. White, and J. Moore, "Genomic mining for complex disease traits with 'random chemistry'," *Genetic Programming and Evolvable Machines (special issue on Medical Applications)*, vol. 8, pp. 395–411, 2007.

[19] IEEE Task Force on Understanding, Prediction, Mitigation and Restoration of Cascading Failures, "Survey of Tools for Risk Assessment of Cascading Outages," in *IEEE Power and Energy Society General Meeting*, Detroit, 2011.

[20] 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–II: Express Briefs*, vol. 55, no. 9, pp. 927–931, 2008.

[21] J.-W. Wang and L.-L. Rong, "Cascade-based attack vulnerability on the US power grid," *Safety Science*, vol. 47, pp. 1332–1336, 2009.

[22] P. Hines, E. Cotilla-Sanchez, and S. Blumsack, "Do topological models provide good information about vulnerability in electric power networks?" *Chaos: An interdisciplinary journal of non-linear science*, vol. 20, no. 3, 2010.

[23] B. A. Carreras, V. E. Lynch, I. Dobson, and D. E. Newman, "Critical points and transitions in an electric power transmission model for cascading failure blackouts," *Chaos: An interdisciplinary journal of non-linear science*, vol. 12, no. 4, pp. 985–994, 2002.

[24] R. Fitzmaurice, A. Keane, and M. O'Malley, "Effect of short-term risk-aversive dispatch on a complex system model for power systems," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 460–469, 2011.

[25] B. Stott, J. Jardim, and O. Alsac, "Dc power flow revisited," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1290 – 1300, 2009.

[26] D. P. Nedic, I. Dobson, D. S. Kirschen, B. A. Carreras, and V. E. Lynch, "Criticality in a cascading failure blackout model," *Electrical Power and Energy Systems*, vol. 28, no. 9, pp. 627–633, 2006.

[27] M. Bhavaraju and N. Nour, "TRELSS: A computer program for transmission reliability evaluation of large-scale systems," EPRI, Tech. Rep. EPRI-TR-100566, 1992.

[28] D. Kosterev, C. Taylor, and W. Mittelstadt, "Model validation for the August 10, 1996 WSCC system outage," *IEEE Transactions on Power Systems*, vol. 14, pp. 967 – 979, 1999.

[29] R. Pfitzner, K. Turitsyn, and M. Chertkov, "Controlled tripping of overheated lines mitigates power outages," Preprint: arXiv:1104.4558v2, 2011.

[30] D. Bienstock, "Optimal adaptive control of cascading power grid failures," Preprint: arXiv:1012.4025v1, 2010.

[31] R. Zimmerman, C. Murillo-Sánchez, and R. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12 –19, feb. 2011.

[32] J. Guo, Y. Fu, Z. Li, and M. Shahidehpour, "Direct calculation of line outage distribution factors," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1633–1634, 2009.

Margaret J. Eppstein is an Associate Professor in the Department of Computer Science at the University of Vermont. She received her B.S. in Zoology from Michigan State University in 1978, her M.S. in Computer Science in 1983, and her Ph.D. in Environmental Engineering in 1997, both from the University of Vermont. She has been a faculty member in Computer Science at the University of Vermont since 1983 (lecturer from 1983-2001; research assistant professor from 1997-2002; assistant professor from 2002-2008; associate professor since 2008), and was founding Director of the University of Vermont Complex Systems Center from 2006-2010. Her current research interests involve forward and inverse modeling and analysis of complex systems in a wide variety of application domains, including biological, environmental, technological, and social systems.

Paul Hines (S'96,M'07) is an Assistant Professor in the School of Engineering at the University of Vermont. He received the Ph.D. in Engineering and Public Policy from Carnegie Mellon U. in 2007 and the M.S. degree in Electrical Engineering from the U. of Washington in 2001. Formerly he worked at the US National Energy Technology Laboratory, the US Federal Energy Regulatory Commission, Alstom ESCA, and for Black and Veatch. His main research interests are in the areas of complex systems and networks, cascading failures in power systems, wind integration and energy security policy.