Correlated Failures of Power Systems: The Analysis of the Nordic Grid

M. Andreasson*, S. Amin†, G. Schwartz †, K. H. Johansson*, H. Sandberg* and S. S. Sastry†
*ACCESS Linnaeus Center, Royal Institute of Technology, Stockholm, Sweden. { mandreas, kallej, hsan }@kth.se
†TRUST Center, University of California, Berkeley. { saurabh, schwartz, sastry }@eecs.berkeley.edu

***Abstract*—In this work we have analyzed the effects of correlations between failures of power lines on the total system load shed. The total system load shed is determined by solving the optimal load shedding problem, which is a TSOs best response to a network failure. We have introduced a Monte Carlo framework for estimating the statistics of the system load shed as a function of stochastic network parameters, and provide explicit guarantees on the sampling accuracy. This framework has been applied to a 470 bus model of the Nordic power system and a correlated Bernoulli failure model. It has been found that increased correlations between Bernoulli failures of power lines can dramatically increase the expected value as well as the variance of the system load shed.**

I. Introduction

Power systems are among the largest and most complex dynamical systems created by mankind. Although power systems are carefully designed by experts, the dynamics of these systems is not yet fully understood as a consequence of the complexity and scale of the systems. Security and vulnerability of power grids is a increasing concern raised by several researchers [1], [2]. Since many vital functions in society, such as transportation, health care and communications rely on electricity supply, the security of the power grid is at east as important as these vital functions. Core functions of the power transmission system, such as the energy management system (EMS), are controlled by SCADA systems. SCADA systems in general are known to possess security flaws, making them vulnerable to deliberate attacks [3], [4]. These security flaws are inherited by any system which deploys SCADA, in particular power transmission systems. To mention one particular example, it has been discovered that stealthy deception attacks can be performed against state estimators of power systems [5]. While the consequences of such deception attacks is not known yet, it clearly justifies concerns about the security of power systems. Research on characterizing optimal attack and defense strategies has gained momentum over the past years. In [6], [7] the optimization problem of maximizing the power outage, for a given number of power transmission lines that an adversary is capable of disconnecting is considered. The system operator is assumed to take the best action to minimize the damage in form of compulsory load shedding. The problem formulation naturally gives rise to a minimax optimization problem. Because of the non-convexity and the existence of integer variables, the problem is inherently hard to solve for large systems.

This work was supported in part by the European Commission through the VIKING project, the Swedish Research Council, the Swedish Foundation for Strategic Research, and the Knut and Alice Wallenberg Foundation

Due to the unpredictable nature of power grids, many events governing the reliability of these systems are by nature stochastic, e.g. demands and generation. Various stochastic and sampling based methods for evaluating the reliability of power systems have been developed to deal with the stochastic nature of modern power systems [8], [9], [10], [11]. However, to our best knowledge, none of these methods attempt to model correlated failures of power system components, nor consider the security aspects of such failure distributions.

This work aims at studying the effects of increasing correlations between failures of power system components. In particular, we consider a Bernoulli model of correlated power line failures. We evaluate the impact of correlations between failures by the covariances between the failures. From a security perspective, our research is motivated by the belief that deliberate attacks against key components of power systems will typically target multiple components simultaneously, thus causing correlation among failures [12], [13], [14]. We measure the impact of a system failure by the minimal system load shed required to restore the system to a safe state. This formulation naturally gives rise to an optimization problem, which under some conditions can be made linear. For different values of the correlations of the failure distribution, we compute the sampled statistics of the total system load shed by Monte Carlo techniques, and provide guarantees on the convergence rate of these. In particular, we use a weighted sum of the mean and variance of the total system load shed as a risk measure of the failure statistics. To obtain statistical data from a realistic power system, we apply our concept to a full-scale model of the Nordic power system, acquired from publicly available sources. We have found that increasing correlations between Bernoulli failures of power lines lead to increased expected value and variance of the system load shed.

The rest of the paper is organized as follows. In Section II the optimal load shedding problem is presented and the total system load shed is defined. In Section III a Monte Carlo sampling technique is presented. A novel model of the Nordic power system is presented and evaluated in section IV. In Section V we analyze the impact of correlated failure statistics on the Nordic power system, followed by concluding remarks in Section VI.

II. Optimal load shedding

When a power system suffers a fault or any other contingency, power flows have to be reallocated to withstand the fault. If the fault is severe, some loads will have to be reduced in order to safely operate the power system. The procedure of reducing loads in a power grid is referred to as load shedding, and used extensively in the control of power systems [15], [16]. The related optimal load shedding problem seeks to minimize the necessary load shed, and is a well-studied problem [17].

In this section we derive the linear optimal load shedding problem, which is a form of an optimal power flow problem. For simplicity and for the sake of illustrating our methods, we consider only active power flows without energy losses in the power lines. The real power flow equations are [18]:

$$P^{line} = V^{line} B \sin\left(A\theta\right) \tag{1}$$

where P^{line} is a vector of active power flows in the transmission lines, $V^{line} = \mathrm{diag}\left(\left[V_i V_j\right]\right)$ where V_i is the voltage of bus i, $B = \mathrm{diag}(b_{ij})$ where b_{ij} is the admittance of the power line from node i to node j, θ is the vector of bus phase angles and A is the vertex-edge incidence matrix of the graph of the power system, defined as $A_{ij} = 1$ iff $e_i = (v_j, u) \in E$, $A_{ij} = -1$ iff $e_i = (u, v_j) \in E$ and $A_{ij} = 0$ otherwise. Here $\sin(x) = \left[\sin(x_i), \ldots, \sin(x_n)\right]^T$ for a vector x. By only considering sufficiently small phase angle differences, i.e. $\Delta\theta_{max} = \|A\theta\|_\infty$ being sufficiently small, we may linearize equation (1) around $A\theta = 0$;

$$P^{line} = V^{line} B A \theta \tag{2}$$

By summing the power flows to each bus, we get the linearized equation for the net power flows into the buses;

$$P = A^T V^{line} B A \theta =: L_B \theta$$

where P is a vector of real power injections to the buses. Note that L_B can be interpreted as a weighted Laplacian matrix of the graph associated with the power system, with weights corresponding to the line admittances times the bus voltages. We may assume, wlog, that the buses of the power system are partitioned such that $P = \left[P^g, P^l\right]^T$, where P^g are generation buses and P^l are load buses. We consider the optimization problem of the transmission system operator (TSO) of minimizing the total load shed of the system. The optimal load shedding problem can, for the linearized power flow equations, be formulated as a linear Program (LP);

$$\min_{\theta} \quad c^T \theta \tag{3}$$

$$\text{s.t.} \quad C\theta \preceq d \tag{4}$$

where

$$C = \begin{bmatrix} V^{line}BA \\ -V^{line}BA \\ H_g L_B \\ -H_g L_B \\ H_l L_B \\ -H_l L_B \\ A \\ -A \end{bmatrix} \quad d = \begin{bmatrix} P^{line}_{max} \\ P^{line}_{max} \\ P^g_{max} \\ 0^{n_g \times 1} \\ 0^{n_l \times 1} \\ -P^l_d \\ \Delta\theta_{max} \cdot 1^{n_p \times 1} \\ \Delta\theta_{max} \cdot 1^{n_p \times 1} \end{bmatrix} \tag{5}$$

and

$$c = \begin{bmatrix} 0^{1 \times n_g} & 1^{1 \times n_l} \end{bmatrix} L_B \tag{6}$$

$$H_g = \begin{bmatrix} I^{n_g \times n_g} & 0^{n_g \times n_l} \end{bmatrix} \quad H_l = \begin{bmatrix} 0^{n_l \times n_g} & I^{n_l \times n_l} \end{bmatrix} \tag{7}$$

and where n_g, n_l and n_p denotes the number of generator buses, load buses and power lines respectively, and $0 < \Delta\theta_{max} \leq \frac{\pi}{2}$ is a sufficiently small real number for which the linearized power flow equation (2) is a valid approximation. The matrix inequality in equation (4) combines line capacity constraints $\left(|P^{line}| \preceq P^{line}_{max}\right)$, power generation constraints $(\mathbf{0} \preceq P^g \preceq P^g_{max})$, power load constraints $\left(P^l_d \preceq P^l \preceq 0\right)$ and phase angle constraints $\left(|A\theta| \preceq \theta_{max} \cdot \mathbf{1}\right)$. The objective function $c^T\theta$ is the sum of the power injections from the demand buses, which is a linearly affine function of the total system load shed. It is easily shown that the minimum total load shed $S^*(C, d)$, which is the difference between the total power demand and total delivered power, is given by

$$S^*(C, d) \quad = \min_{\theta} \left\{ c^T\theta | C\theta \preceq d \right\} - 1^{1 \times n_l} \cdot P^l_d \tag{8}$$

III. Statistics of power system failures

In this paper we will consider stochastic failures of the power system, as in e.g. [8], [9]. To demonstrate the generality of our methods, we will not yet make any prior assumptions about these failures. Consider the matrices C and d as random variables, endowed with a probability measure $\mu^C \times \mu^d$. Since both the topology and load parameters of the power system are determined by C and d, such a probability measure can represent any type of failures of the power system. It can be seen that for any probability measure $\mu^C \times \mu^d$, the minimum total load shed $S^*(C, d)$ is also a random variable. Since the optimal load shedding is the best response to any contingency, $S^*(C,d)$ can be interpreted as a minimax cost in the case of adversarial attacks. The total load shed is a commonly used measure of the severeness of a power system outage [10], [11].

Because $0 \le S^*(C,d) \le -\sum P_d^l$, as seen from equation (4), the mean $\bar{S}^*$ and variance $\sigma^2_{S^*}$ of $S^*(C,d)$ always exist and are finite. We will use $\bar{S}^* + \alpha \cdot \sigma_{S^*}$ as a risk measure for the distribution $\mu^C \times \mu^d$. To see that this risk measure makes sense, we will show that it is closely related to the commonly used risk measure value at risk (VaR), which for a random variable X is defined as follows [19]:

$$\mathrm{VaR}_\alpha(X) = \inf\{l \in \mathbb{R} : \Pr(X > l) \le 1 - \alpha\}$$

The intuition of the expression $\mathrm{VaR}_\alpha(X)$ is that the maximum loss, in our case system load shed, is bounded by $\mathrm{VaR}_\alpha(X)$ with probability $1-\alpha$. One serious computational drawback with using VaR on sampled probability distributions, is that it requires knowledge of the full probability distribution of the random variable X. When dealing with samples of random variables, estimating VaR becomes hard since it requires estimation of the tail of the distribution X. The following proposition shows that one can always bound $\mathrm{VaR}_\alpha(X)$ with a linear combination of the mean and the variance of X:

Proposition 1. *The risk measure value at risk (*VaR_α*) satisfies*

$$\mathrm{VaR}_\alpha(X) \le \bar{X} + \frac{1}{\sqrt{\alpha}} \cdot \sigma$$

The proof is given in the appendix. Even though we have showed the existence and finiteness of $\bar{S}^*$ and $\sigma^2_{S^*}$, obtaining analytical closed form expressions for these quantities in terms of C and d is in general not possible. We will therefore use Monte Carlo techniques to estimate these quantities. By drawing N samples from the distribution $\mu^C \times \mu^d$, we obtain the following approximations of $\bar{S}^*$ and $\sigma^2_{S^*}$;

$$\bar{S}^* \approx \hat{S}^* = \frac{1}{N}\sum_{i=1}^{N} S^*(C_i, d_i) \tag{9}$$

$$\sigma^2_{S^*} \approx \hat{\sigma}^2_{S^*} = \frac{1}{N-1}\sum_{i=1}^{N}\left(S^*(C_i,d_i) - \hat{S}^*\right)^2 \tag{10}$$

Due to $S^*(C,d)$ being bounded, $\hat{S}^*$ and $\hat{\sigma}_{S^*}$ are guaranteed to converge to S^* and σ_{S^*} respectively.

Proposition 2. *Given $\epsilon > 0$, $\delta > 0$, the number of samples N_1 and N_2 which assure that*

$$\Pr\left[\left|\hat{S}^{*,N_1} - \bar{S}^*\right| \ge \epsilon\right] \le \delta$$
$$\Pr\left[\left|\hat{\sigma}_{S^*,N_2} - \sigma_{S^*}\right| \ge \epsilon\right] \le \delta$$

are

$$N_1 \ge \left\lceil \frac{\hat{S}^2}{4\delta\epsilon^2} \right\rceil \qquad N_2 \ge \left\lceil \frac{\hat{S}^4}{8\delta\epsilon^4} \right\rceil$$

Proof: Follows by lemma 3 and lemma 5 in the appendix, and the fact that $0 \le S^*(C,d) \le -\sum P_d^l$. ■

With proposition 2 we have guaranteed bounds of the estimation error of both the sampled expected value and the sampled variance of the load shed. The proposition can of course be used in the reverse direction. For given numbers N_1 and N_2, we can obtain bounds on the numbers δ and ϵ. With these explicit bounds on the error of the estimated mean and variance of the load shed, the number of samples can be chosen according to given accuracy requirements, and trade-offs between accuracy and the number of samples can be easily made a priory. The above propositions are the main theoretical foundation for the deployment of Monte Carlo techniques to estimate the statistics of the load shed.

IV. Model of the Nordic power system

In this paper, we will construct a model of the Nordic power transmission network to demonstrate our techniques. While IEEE standard power systems offer great transparency for research and can function as benchmark systems, we believe that it is important to demonstrate our concepts on real power systems. A problem with real power systems however is that TSOs in general are very restrictive with releasing information about their systems, or their internal network models, due to concerns that the data might be used for attack synthesis. To overcome the unavailability of data from the Nordic grid, we have built an approximate model of the Nordic power system, using only publicly available data. While there have been similar efforts to model other interconnected power systems, such as the main European power grid [20], there are no known complete models of the Nordic power system that are publicly available.

A. Collecting network data

The topology of the power system. i.e. the geographical positions of the main 400, 300, 220 and 132 kV power lines and HVDC links in the Nordic countries were obtained from the respective TSOs websites. The data obtained included the coordinates of the power buses, the connectivity of the buses through power lines, the voltage of the power lines and the number of parallel power lines, if applicable. The complete model has a total of 470 buses and 717 power lines. Information about all power generation facilities with capacity of at least 100 MW were also acquired from public sources. We have made the assumption that the remaining thermal power generation of power plants with generation less than 100 MW is located in populated areas, and hence proportional to the population in the demand buses. For the remaining wind power capacity we have assumed that the wind power generation is uniformly distributed over the land surface, and hence over the nodes.

B. Estimating power demand data

There is no available electricity demand data, other than cumulative data for the Nordic countries. This data is much to rough to be useful for our 470 node model. Following [20] we have used population census data to estimate the power demand. This methodology relies on the basic assumptions that household power demand is proportional to the population connected to a substation, as well as Industry power demand, since the workforce will settle relatively close to industries.

We collected population statistics from the Bureau of Statistics of the respective countries. We have collected population statistics for the major administrative regions of each country, and assumed that the demands are distributed uniformly over the non-generator buses within each region. The number of administrative regions in each country was between 12 and 21. Using smaller regions would introduce difficulties in assigning the right population to each substation. To estimate the power demands, both the yearly average and yearly maximum of the daily maximum power consumption were used to create two different load situations. The power lines and buses of the Nordic power grid are illustrated in figure 1.

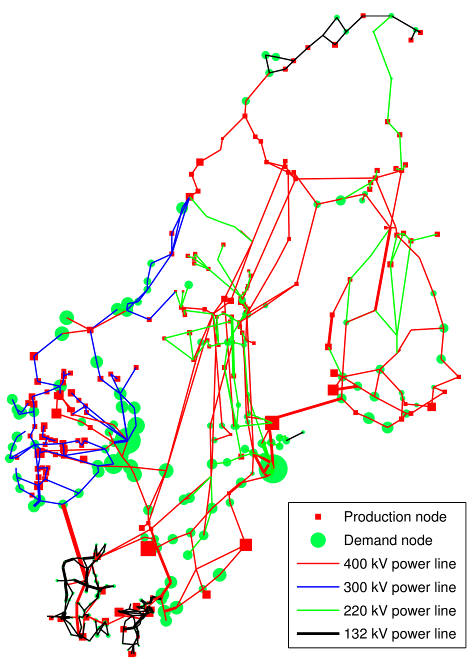

Figure 1. Model of the Nordel high voltage power transmission network. The circle area of the production and demand nodes are proportional to the production and demand respectively.

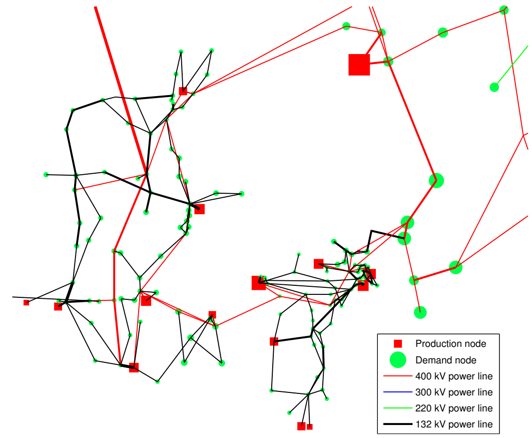

Figure 2. Model of the Nordel high voltage power transmission network in Denmark.

C. Estimating power line parameters

The only known parameters of the power lines obtained from public sources are the line voltages. To solve the optimal load shedding problem, also the line admittances as well as the maximum transmission capacities of each line need to be known. Neither of these are however in any form published by the TSOs. However, the admittances of power lines can be estimated by the length of the power line. Typically the reactance of high voltage power transmission lines is approximately $0.20\ \Omega/\mathrm{km}$ [21]. The lengths of a power line from a bus with coordinates x to a bus with coordinates y is estimated by the euclidean 2-norm as $l = \mathrm{dist}(x, y) = \|x - y\|_2$ which is always an underestimate of the actual line lengths. As for estimating transmission capacities, only cross-border transmission line capacity constraints are available from the Nordic TSOs. The transmission capacity of each power line of equal voltage is assumed to be the average capacity of the cross-border lines, which are shown in the table below.

Voltage	Capacity
400 kV	1030 MW
300 kV	650 MW
220 kV	415 MW
132 kV	143 MW

Figure 3. Estimated transmission line capacities.

D. Evaluating the model

The optimal load shedding problem was applied to the above derived model of the Nordic power transmission grid. By using the YALMIP [22] interface with the GLPK [23] LP solver in MATLAB [24], the optimal load shedding problem was solved. When solving the linear optimal load shedding problem with the yearly maximum loads, the total system load shed is found to be only 2 % of the total power demand. All buses with load shedding could be deduced to buses connected to 200 kV lines in the area of Stockholm, and to buses in northern Norway connected to 132 kV lines. To overcome these transmission bottlenecks, the capacity of the two power lines in Norway, and five power lines in Sweden, connected to the buses with load shedding were increased. The capacity of the critical lines was increased in steps of 5 MW equally for all the critical lines within each region, until no load shedding was necessary. By these adjustments, the capacity of the power lines in northern Norway was increased from 143 MW to 300 MW and the capacity of the power lines in Stockholm was increased from 415 MW to 600 MW. With these measures taken, no load shedding was necessary. To evaluate the model further, we examine if it satisfies the $N-1$ criterion, i.e. if the network is fully functional after the disconnection of any single power line. For the evaluation of the $N-1$ criterion, we used the yearly average of the daily maximum loads. For each $N-1$ contingency, the optimal load shedding problem was solved, and the resulting load shed was calculated. There were 51 power line failures that resulted in system load shed, but none which result in a total load shed of more than 1.6 %. Most of these contingencies were dead-end lines, who's removal necessarily results in system load shed of the end bus. Also, since no single line contingency caused considerable damage to the system, other than local, the $N-1$ criterion was considered to be essentially satisfied.

V. Simulations of correlated system failures

In this section we examine the effects of correlated system faults on the statistics of the minimum total system load shed. Our results are motivated by the belief that adversarial faults in general have stronger correlations than reliability faults [12], [13], [14]. Hence we will examine the effects of increasing correlations on the mean and variance of the system load shed. In the following empirical study we will consider failures in the form of power line disconnections. We model the disconnection of power line i as a binary random variable $X_i \in \{0,1\}$ where $X_i = 0$ represents to line i being fully functional with all parameters set to default, and $X_i = 1$ represents line i being disconnected, i.e. the admittance of line i being 0. Thus, the failure statistics of the whole power system are given by

$$P(X_1 = Y_1, \dots, X_{n_p} = Y_{n_p}) \;\forall\, Y_i \in \{0,1\}$$

Since parameterizing the full joint Bernoulli distribution would require $n^{n_p} \approx 10^{216}$ variables, we will only consider the joint Bernoulli distribution with the first two central moments given explicitly, i.e.

$$\bar{X}_i = \mathrm{E}[X_i] \quad \forall i \in \{1, \dots, n_p\}$$
$$\sigma_{ij} = \mathrm{E}\left[(X_i - \bar{X}_i)(X_j - \bar{X}_j)\right] \;\forall\, (i,j) \in \{1, \dots, n_p\}^2$$

which requires only $n_p^2 + n_p = 519120$ variables. To consider the effects of increasing correlations $\bar{X}_i = 0.02$ is kept constant, while σ_{ij} is increased. First we consider the scenario where σ_{ij} is increased equally for all $(i,j) \in \{1, \dots, n_p\}^2$. σ_{ij} is increased from 0 to 0.016 in steps of 0.004. For each step, 1000 Monte Carlo simulations are performed with a Bernoulli sampling algorithm described in [25] to acquire sampled approximations of $\bar{S^*}$ and σ_{S^*}. By proposition 2 the relative error of $\hat{S}^*$ is guaranteed to be less than 7 % with certainty 95 %. The result of the sampling are shown in figure 4. Clearly σ_{X^*} is increasing in σ_{ij}, but also $\bar{S}^*$ is increasing in σ_{ij}. We

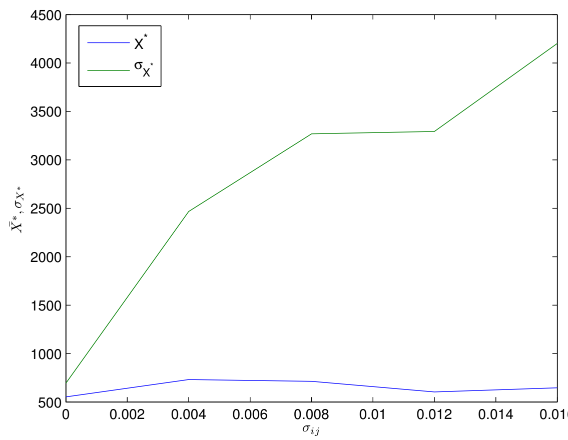

Figure 4. Expected value and standard deviation (MW) of the total load shed for different covariances for all power lines.

also consider the case where only the failure statistics of incident power lines are correlated. Let σ_{ij} be increased from 0 to 0.016 in steps of 0.004 if and only if the power lines i and j are connected to the same bus, otherwise $\sigma_{ij} = 0$. $\bar{X}_i = 0.02$ is again kept constant. The results of the sampling are shown in figure 5. Clearly both $\bar{S}^*$ and σ_{S^*} are increasing in σ_{ij}, and the effect of increasing correlations in X on $\bar{S}^*$ is much higher than when the correlation is increased for all power lines. Thus, by the common belief that security failures are more correlated than other failures, security induced failures will typically have higher expected cost and higher variance, even though the expected value of the failure probabilities are constant. Nevertheless, it is easy to find counterexamples where increased correlations between power line failures decrease the expected load shed. The simplest possible counterexample is a 3-node and 2-line power network shown in figure 6. Let the demand node have demand $-\bar{d}$, and the generation node a capacity $\bar{g} \geq \bar{d}$, and the line parameters such that the demand is satisfied under normal operation. It can be shown that the expected load shed of the system is $1 - \sigma_{12}$, where σ_{12} is the covariance between the failures of node 1 and 2. The intuition behind this counterexample is

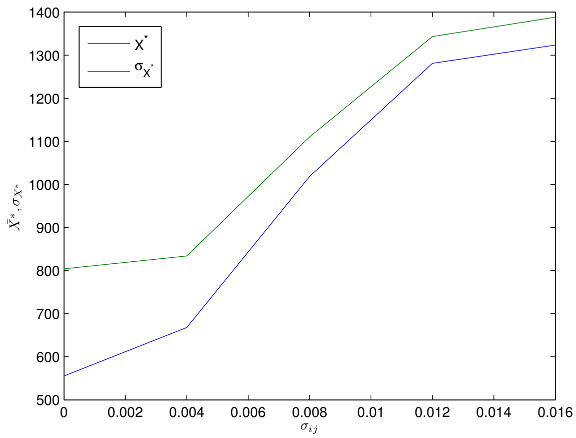

Figure 5. Expected value and standard deviation (MW) of the total load shed for different covariances for power lines connected to the same bus.

that while the probability of both lines failing increases, the probability of each failing individually decreases, with the result that the total probability of any line failing decreases. We here state sufficient conditions

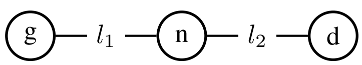

Figure 6. Topology of a 3-node, 2-line power network where increased correlations between system failures result in decreased system load shed. g is a generation bus, d a demand bus and n neither a demand nor generation bus.

under which increased correlations of power line failures imply increased expected load shed.

Proposition 3. *Let the power system satisfy the $n-k$ criterion, i.e. the disconnection of any k power lines does not induce any necessary load shedding. Furthermore, assume that all contingencies with at least $k+1$ line failures induce a total system load shed $\bar{c}$. Assume, wlog, that the moment $\phi_{1,\ldots,k+1} = \mathrm{E}[X_1 \cdot \ldots \cdot X_{k+1}]$ increases by $\Delta\phi$, but all other moments $\mathrm{E}[X_{i_1} \cdot \ldots \cdot X_{i_l}]$ are constant. Then*

1) *The central moment $\sigma_{1,\ldots,k+1} = \mathrm{E}\left[(X_1 - \bar{X}_1) \cdot \ldots \cdot (X_{k+1} - \bar{X}_{k+1})\right]$ also increases by $\Delta\phi$.*
2) *All other central moments of order less than or equal to $k+1$ remain constant.*
3) *The expected load shed $\bar{S}^*$ increases by $\bar{c} \cdot \Delta\phi$.*
4) *If the probability of necessary load shedding always remains less than $1/2$ when ϕ increases, the variance of the load shed, σ_{S^*}, increases by $\hat{c}^2 \cdot \Delta\phi$, where $0 < \hat{c} \leq \bar{c}$.*

A proof is given in the appendix. The following corollary follows directly from proposition 3.

Corollary 1. *Assume that all conditions of proposition 3 still hold, except that all contingencies with at least $k+1$ line failures induce a system load shed of* at least *$\bar{c}$. In this case the results of proposition 3 hold instead for the lower bound $\underline{S}^* \leq S^*$ of the system load shed, which is the total system load shed assuming all line failures result in the same system load shed $\bar{c}$.*

VI. Conclusions

In this work we have demonstrated that increased correlations between power line failures can dramatically increase the expected costs in terms of system load shed, although the expected value of the failure probabilities is kept constant. Furthermore we have demonstrated that increased correlations between power line failures can also increase the variance of the system load shed, thus increasing the risk of large system load sheds. We have demonstrated our results by simulating correlated power line failures in a model of the Nordic high-voltage power grid. We have furthermore provided sufficient conditions under which the mean and the variance of the total system load shed increase with increasing correlation between line failures.

In many situations, cascading failures further aggravate the state of a partially failing power system, leading to even larger losses. In future work, it would be of interest to consider the impact of correlated failures under power system dynamics, to also capture possible cascading failures. Also, it would be interesting to study if the conditions under which the expected value and the variance of the load shed are increasing in the correlations, can be relaxed.

References

[1] M.T.O. Amanullah, A. Kalam, and A. Zayegh. Network security vulnerabilities in scada and ems. In *Transmission and Distribution Conference and Exhibition: Asia and Pacific, 2005 IEEE/PES*, pages 1 –6, 2005.

[2] M. Amin. Security challenges for the electricity infrastructure. *Computer*, 35(4):8 –10, April 2002.

[3] David Watts. Security & vulnerability in electric power systems, 2003.

[4] Vinay M. Igure, Sean A. Laughter, and Ronald D. Williams. Security issues in scada networks. *Computers & Security*, 25(7):498 – 506, 2006.

[5] Yao Liu, Michael K. Reiter, and Peng Ning. False data injection attacks against state estimation in electric power grids. In *Proceedings of the 16th ACM conference on Computer and communications security*, CCS '09, pages 21–32, New York, NY, USA, 2009. ACM.

[6] J. Salmeron, K. Wood, and R. Baldick. Analysis of electric grid security under terrorist threat. *Power Systems, IEEE Transactions on*, 19(2):905 – 912, May 2004.

[7] Ali Pinar, Juan Meza, Vaibhav Donde, and Bernard Lesieutre. Optimization strategies for the vulnerability analysis of the electric power grid. *SIAM Journal on Optimization*, 20(4):1786–1810, 2010.

[8] W Li Roy Billinton, Billinton. Reliability assessment of electrical power systems using monte carlo methods. 1994.

[9] R. Allan and R. Billinton. Probabilistic assessment of power systems. *Proceedings of the IEEE*, 88(2):140 –162, February 2000.

[10] R. Billinton and P. Wang. Distribution system reliability cost/worth analysis using analytical and sequential simulation techniques. *Power Systems, IEEE Transactions on*, 13(4):1245 –1250, November 1998.

[11] R. Billinton and E. Khan. A security based approach to composite power system reliability evaluation. *Power Systems, IEEE Transactions on*, 7(1):65 –72, February 1992.

[12] Alfonso Valdes and Keith Skinner. Probabilistic alert correlation. RAID '00, pages 54–68, London, UK, 2001. Springer-Verlag.

[13] Hervé Debar and Andreas Wespi. Aggregation and correlation of intrusion-detection alerts. In *In Recent Advances in Intrusion Detection, LNCS 2212*, pages 85–103. Springer-Verlag, 2001.

[14] M. Sudit, A. Stotz, and M. Holender. Situational awareness of a coordinated cyber attack. In B. V. Dasarathy, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 5812 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 114–129, March 2005.

[15] J. Zhu. *Optimization of Power System Operation()*. 2009.

[16] J.A. Momoh, M.E. El-Hawary, and R. Adapa. A review of selected optimal power flow literature to 1993. ii. newton, linear programming and interior point methods. *Power Systems, IEEE Transactions on*, 14(1):105 –111, February 1999.

[17] M. Huneault and F.D. Galiana. A survey of the optimal power flow literature. *Power Systems, IEEE Transactions on*, 6(2):762–770, 1991.

[18] Ali Abur and Antonio Gomez Exposito. *Power System State Estimation: Theory and Implementation; electronic version*. Dekker, Abingdon, 2004.

[19] Darrell Duffie and Jun Pan. An overview of value at risk. *The Journal of Derivatives*, 4(3):7 – 49, 1997.

[20] Qiong Zhou and J.W. Bialek. Approximate model of european interconnected system as a benchmark system to study effects of cross-border trades. *Power Systems, IEEE Transactions on*, 20(2):782 – 788, May 2005.

[21] Brugg. High voltage xlpe cable systems, technical user guide. 2006.

[22] J. Löfberg. Yalmip : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[23] Andrew Makhorin. *GNU Linear Programming Kit, Version 4.9*. GNU Software Foundation, http://www.gnu.org/software/glpk/glpk.html, 2006.

[24] MATLAB. *version 7.11.0 (R2010b)*. The MathWorks Inc., Natick, Massachusetts, 2010.

[25] Jakob H. Macke, Philipp Berens, Alexander S. Ecker, Andreas S. Tolias, and Matthias Bethge. Generating spike trains with specified correlation coefficients. *Neural Computation*, 21(2):397–423, 2009. PMID: 19196233.

[26] Yaakov Bar-Shalom, Thiagalingam Kirubarajan, and X.-Rong Li. *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., New York, NY, USA, 2002.

APPENDIX

Proof: (of proposition 1) Note that

$$\mathrm{VaR}_\alpha(X) \leq \bar{X} + \frac{1}{\sqrt{\alpha}}\sigma \quad \Leftrightarrow$$
$$\Pr\left(X < \bar{X} + \frac{1}{\sqrt{\alpha}} \cdot \sigma\right) \geq 1 - \alpha$$

which is easily shown using Chebyshev's inequality

$$\Pr\left(X < \bar{X} + \frac{1}{\sqrt{\alpha}} \cdot \sigma\right) \geq$$
$$\Pr\left(\left|X - \bar{X}\right| < \frac{1}{\sqrt{\alpha}} \cdot \sigma\right) =$$
$$1 - \Pr\left(\left|X - \bar{X}\right| \geq \frac{1}{\sqrt{\alpha}} \cdot \sigma\right) \geq 1 - \alpha$$

■

Lemma 2. *The variance of a random variable X with compact support $[a, b]$ is bounded by:*

$$\mathrm{Var}[X] \leq \frac{(b-a)^2}{4}$$

Proof: Consider the random variable defined by $Y = X - \frac{a+b}{2}$. By basic probability theory we have

$$\mathrm{Var}[X] = \mathrm{Var}[Y] = \mathrm{E}[Y^2] - \left(\mathrm{E}[Y]\right)^2$$
$$\leq \left(\frac{b-a}{2}\right)^2 = \frac{(b-a)^2}{4}$$

■

Lemma 3. *Let $\hat{X}_N = \frac{1}{N}\sum_{i=1}^N X_i$ be the sampled mean of the RV X with N samples. Let X have compact support on $[a, b]$, then for any given $\epsilon > 0,\ \delta > 0$*

$$\Pr\left[\left|\hat{X}_N - \bar{X}\right| \geq \epsilon\right] \leq \delta$$

for

$$N \geq \left\lceil \frac{(b-a)^2}{4\delta\epsilon^2} \right\rceil$$

Proof: Since the samples are iid, we have

$$\mathrm{Var}[\hat{X}_N] = \mathrm{Var}\left[\frac{1}{N}\sum_{i=1}^N X_i\right] = \frac{1}{N^2}\mathrm{Var}\left[\sum_{i=1}^N X_i\right]$$
$$= \frac{N\sigma^2(X)}{N^2} \leq \frac{(b-a)^2}{4N}$$

By Chebyshev's inequality we have

$$\Pr\left\{\left|\hat{X_N} - \bar{X}\right| \geq \epsilon\right\} \leq \frac{\mathrm{Var}[\hat{X}_N]}{\epsilon^2} \leq \frac{(b-a)^2}{4N\epsilon^2} \leq \delta$$

■

Lemma 4. *Let $\hat{\sigma}_{X_N} = \frac{1}{N}\sum_{i=1}^N \left(X_i - \hat{X}_N\right)^2$ be the sampled variance of the RV X with N samples. Let X have compact support on $[a, b]$, then for any given $\epsilon > 0,\ \delta > 0$*

$$\Pr\left[\left|\hat{\sigma}^2_{X_N} - \sigma^2_X\right| \geq \epsilon\right] \leq \delta$$

for

$$N \geq \left\lceil \frac{(b-a)^4}{8\delta\epsilon^2} \right\rceil$$

Proof: By e.g. [26], the variance of the sampled variance is given by

$$\mathrm{Var}\left[\hat{\sigma}^2_{X_N}\right] = \frac{2\sigma^4_X}{N}$$

By lemma 2, the variance $\mathrm{Var}[X] = \sigma^2_X$ is bounded by

$$\sigma^2_X \leq \frac{(b-a)^2}{4}$$

Thus, by Chebyshev's inequality

$$\Pr\left[\left|\hat{\sigma}^2_{X_N} - \sigma^2_X\right| \geq \epsilon\right] \leq \frac{\mathrm{Var}[\sigma^2_{X_N}]}{\epsilon^2} \leq \frac{(b-a)^4}{8N\epsilon^2} \leq \delta$$

■

Lemma 5. *Given $\epsilon > 0$, $\delta > 0$, we have for a random variable X with compact support on $[a, b]$*

$$\Pr\left[\left|\hat{\sigma}_{X_N} - \sigma_X\right| \geq \epsilon\right] \leq \delta$$

for

$$N \geq \left\lceil \frac{(b-a)^4}{8\delta\epsilon^4} \right\rceil$$

Proof: By concavity of $\sqrt{\cdot}$, Chebyshev's inequality and lemma 4

$$\Pr\left[\left|\hat{\sigma}_{X_N} - \sigma_X\right| \geq \epsilon\right] \leq \Pr\left[\left|\hat{\sigma}^2_{X_N} - \sigma^2_X\right| \geq \epsilon^2\right]$$
$$\leq \frac{\mathrm{Var}[\sigma^2_{X_N}]}{\epsilon^4} = \frac{2\sigma^4_{X_N}}{N\epsilon^4} \leq \frac{(b-a)^4}{8N\epsilon^4} \leq \delta$$

■

Proof: (of proposition 3) The first part follows directly, since:

$$\sigma_{1,\ldots,k+1} = \mathrm{E}\left[(X_1 - \bar{X}_1) \cdot \ldots \cdot (X_{k+1} - \bar{X}_{k+1})\right] = \mathrm{E}\left[X_1 \cdot \ldots \cdot X_{k+1}\right] - \bar{X}_1 \cdot \mathrm{E}\left[X_2 \cdot \ldots \cdot X_{k+1}\right] + \ldots + (-1)^{k+1} \cdot \bar{X}_1 \cdot \ldots \cdot \bar{X}_{k+1}$$

The second part also follows directly by calculation. Wlog, consider

$$\sigma_{1,\ldots,l} = \mathrm{E}\left[(X_1 - \bar{X}_1) \cdot \ldots \cdot (X_l - \bar{X}_l)\right] = \mathrm{E}\left[X_1 \cdot \ldots \cdot X_l\right] - \bar{X}_1 \cdot \mathrm{E}\left[X_2 \cdot \ldots \cdot X_l\right] + \ldots + (-1)^l \cdot \bar{X}_1 \cdot \ldots \cdot \bar{X}_l$$

which is constant for all $l \leq k$ by the assumption that all other moments $\mathrm{E}[X_{i_1} \cdot \ldots \cdot X_{i_l}]$ are constant.

As for the third part, let $k = 1$ and consider a network with 3 power lines. The expected value of the system load shed is:

$$\bar{c} \cdot (\Pr[X_1 = 1, X_2 = 1] + \Pr[X_1 = 1, X_3 = 1] + \Pr[X_2 = 1, X_3 = 1] - \Pr[X_1 = 1, X_2 = 1, X_3 = 1]) = \bar{c} \cdot (\mathrm{E}[X_1X_2] + \mathrm{E}[X_1X_3] + \Pr[X_2X_3] - \Pr[X_1X_2X_3])$$

which increases by $\bar{c} \cdot \Delta\phi$ as $\mathrm{E}[X_1X_2]$ increases by $\Delta\phi$. One can generalize this and show that for arbitrary k and network size, the expected system load shed will still be proportional to $\bar{c} \cdot \mathrm{E}[X_1 \cdot \ldots \cdot X_{k+1}]$. By the assumption that all other moments $\mathrm{E}[X_{i_1} \cdot \ldots \cdot X_{i_l}]$ are constant.

To prove the last claim, we denote $p_s = \Pr[S^* = \bar{c}]$, and $p_n = 1 - p_s = \Pr[S^* = 0]$. By assumption $p_s < \frac{1}{2}$. For this case, the variance of the load shed is simply:

$$\sigma_{S^*} = p_n(0 - \bar{S}^*)^2 + p_s(\bar{c} - \bar{S}^*)^2 = p_n(p_s\bar{c})^2 + p_s(\bar{c} - p_s\bar{c})^2 = p_s\bar{c}^2(1 - p_s)$$

The latter expression is increasing since

$$\frac{\partial}{\partial p_s} \, p_s\bar{c}^2(1 - p_s) = \bar{c}^2(1 - 2p_s) > 0$$

for $p_s < \frac{1}{2}$. Since $0 < \frac{\partial \sigma_{S^*}}{\partial p_s} \leq \bar{c}^2$, $\bar{S}^* = \bar{c}p_s$, $\bar{S}^* = constant + \bar{c}\Delta\phi$ and

$$\Delta\sigma_{S^*} = \int_{\phi_0}^{\phi_0 + \Delta\Phi} \frac{\mathrm{d}\sigma_{S^*}}{\mathrm{d}\phi_{1,\ldots,k+1}} \; \mathrm{d}\phi_{1,\ldots,k+1}$$
$$= \int_{\phi_0}^{\phi_0 + \Delta\Phi} \frac{\partial \sigma_{S^*}}{\partial p_s} \frac{\partial p_s}{\partial \phi_{1,\ldots,k+1}} \; \mathrm{d}\phi_{1,\ldots,k+1}$$
$$= \int_{\phi_0}^{\phi_0 + \Delta\Phi} \frac{\partial \sigma_{S^*}}{\partial p_s} \; \mathrm{d}\phi_{1,\ldots,k+1}$$

it holds that

$$\Delta\sigma_{S^*} = \int_{\phi_0}^{\phi_0 + \Delta\Phi} \frac{\partial \sigma_{S^*}}{\partial p_s} \; \mathrm{d}\phi_{1,\ldots,k+1} > 0$$
$$\Delta\sigma_{S^*} = \int_{\phi_0}^{\phi_0 + \Delta\Phi} \frac{\partial \sigma_{S^*}}{\partial p_s} \; \mathrm{d}\phi_{1,\ldots,k+1}$$
$$\leq \int_{\phi_0}^{\phi_0 + \Delta\Phi} \bar{c}^2 \; \mathrm{d}\phi_{1,\ldots,k+1} = \bar{c}^2\Delta\phi$$

■