

PAPER • OPEN ACCESS

## Detours around basin stability in power networks

To cite this article: Paul Schultz *et al* 2014 *New J. Phys.* **16** 125001

View the [article online](#) for updates and enhancements.

### Related content

- [Deciphering the imprint of topology on nonlinear dynamical network stability](#)
- [Community consistency determines the stability transition window of power-grid nodes](#)
- [Comparative analysis of existing models for power-grid synchronization](#)

### Recent citations

- [Network-induced multistability through lossy coupling and exotic solitary states](#)  
Frank Hellmann *et al*
- [Monte Carlo basin bifurcation analysis](#)  
Maximilian Gelbrecht *et al*
- [Impact of network topology on the stability of DC microgrids](#)  
J. F. Wienand *et al*

## Detours around basin stability in power networks

Paul Schultz<sup>1,2</sup>, Jobst Heitzig<sup>1</sup> and Jürgen Kurths<sup>1,2,3,4</sup>

<sup>1</sup> Potsdam Institute for Climate Impact Research, D-14412 Potsdam, Germany

<sup>2</sup> Department of Physics, Humboldt University Berlin, D-12489 Berlin, Germany

<sup>3</sup> Institute for Complex Systems and Mathematical Biology, University of Aberdeen, AB24 3UE Aberdeen, UK

<sup>4</sup> Department of Control Theory, Nizhny Novgorod State University, Gagarin Avenue 23, 606950 Nizhny Novgorod, Russia

E-mail: [pschultz@pik-potsdam.de](mailto:pschultz@pik-potsdam.de)

Received 3 September 2014, revised 13 October 2014

Accepted for publication 27 October 2014

Published 2 December 2014

*New Journal of Physics* **16** (2014) 125001

doi:[10.1088/1367-2630/16/12/125001](https://doi.org/10.1088/1367-2630/16/12/125001)

### Abstract

To analyse the relationship between stability against large perturbations and topological properties of a power transmission grid, we employ a statistical analysis of a large ensemble of synthetic power grids, looking for significant statistical relationships between the single-node basin stability measure and classical as well as tailor-made weighted network characteristics. This method enables us to predict poor values of single-node basin stability for a large extent of the nodes, offering a node-wise stability estimation at low computational cost. Further, we analyse the particular function of certain network motifs to promote or degrade the stability of the system. Here we uncover the impact of so-called detour motifs on the appearance of nodes with a poor stability score and discuss the implications for power grid design.

Keywords: power grid, weighted complex networks, network motifs, logistic regression, single node basin stability



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

## 1. Introduction

### 1.1. Motivation

Following a sequence of large-scale blackout events [2, 17, 21] within the last decades, it became obvious that a deeper understanding of power grids from the complex system perspective is necessary. Still, there are the achievements and the knowledge that has been developed in about a century of power systems engineering, which is the main base of current power grid research. We expect, however, that the perspective of complex systems science adds important notions to help understanding power grids better, in particular their stability.

The physical grid itself—the transmission lines connecting various power stations, substations, consumers, etc—constitutes a complex coupling structure of the dynamical system. This can be well-described within the framework of complex networks theory. Our main intention is to uncover links between the topology of such a network and the stability of the dynamical system's stationary state of operation.

It is important to note that the network topology itself undergoes steady changes, driven by modernisation, grid expansion and climate change adaptation (including renewable energy production and distributed generation). There are, however, two distinct time scales. On the one hand, the real-time control of the system is affected by tripping of lines and the consecutive rearrangement of the power flow after a failure. On the other hand, structural changes—e.g. due to distributed generation of renewable energy—and their implications to the system's stability need to be considered for long-term planning.

We propose a strategy to directly estimate the power grid's stability, even on short time scales, using statistical network characteristics to omit the need of costly simulations. The focus lies on the identification of grid nodes that appear critical for stability.

A systematic analysis of the functional role network motifs play within complex dynamical systems started in biological systems [13, 20]. The authors studied directed transcriptional regulation networks and found reappearing small-sized structures, having a specific function, e.g. serving as feedbacks.

Beside the functional role of network motifs, their implications for linear dynamical stability have been analysed as well. In biological networks, a relatively high abundance of certain motifs appears to be correlated with dynamical stability against small perturbations of the system [18]. Further, motifs containing less edges are more likely to be linearly stable. Especially in [8] investigated directed feedback loops and found that motifs containing fewer nodes are more stable. Taken together, these findings indicate that small cycles in directed networks can be regarded as more stable and less prone to oscillatory behaviour. In addition [6] define the notion of reliability of information processing in close relation to stability. They find that there are certain motifs suppressing fluctuations and tending to synchronize the dynamics of single elements.

In the context of power grids, it has been shown by Menck *et al* [10] that there are network motifs in power grid networks that degrade the overall dynamical stability of the system. These are termed *dead trees*. Nodes adjacent to dead trees, referred to as *dead tree gateways*, show a significantly reduced single-node basin stability. Witthaut and Timme [24] otherwise highlighted the particular role of cycles in power grid networks. By the *addition* of a single line, the whole grid might be destabilized due to geometric frustration. We will show later, however, that a particular motif, namely a three-cycle, is helpful to stabilize the system.

In the following, we investigate the question if—in contrast to destabilising dead trees—there are also network motifs promoting the dynamical stability of a power grid. Further, we propose a logistic regression model aiming to predict the individual single-node basin stability using only a small set of network characteristics.

## 1.2. Ensemble of synthetic power grid networks

We make use of a random growth model for power grids and other spatially embedded infrastructure networks, which we recently published in [19], to create an ensemble of randomly generated synthetic power grid topologies as a basis for a statistical analysis. This allows us a more general analysis of a wide range of network topologies, i.e. guarantees a significant frequency of observed network motifs, and also avoids the unavailability problem with power grid data. Note that basic network characteristics like the degree distribution, average shortest-path length or global clustering coefficient, measured for synthetic network topologies, reproduce values published for real-world power grids quite well. Thus we conjecture that our results obtained with the model data can easily be translated to real-world power grids.

We use a large ensemble of 570 medium-sized network topologies with  $N = 100$  nodes and a mean degree of  $\bar{k} \approx 2.7$  (see [16]). The model parameters are  $\{N_0, p, q, r, s\} = \{1, 0.3, 0.1, 1, 0.2\}$ , see [19] for details.

In this random growth model, spatially embedded sparse networks are generated. An initial set of  $N_0$  nodes is connected by a minimum spanning tree and further subject to a growth process. The network growth is given by a repeated connection of new nodes to the grid according to an attachment rule which has four parameters  $p, q, r, s$ . The latter set of parameters mainly determines the final network characteristics that we have chosen in a way to yield topologies in accordance to common data sets.

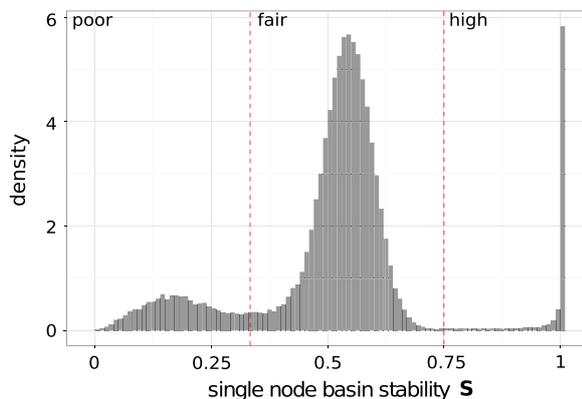
The line reactance is chosen to be proportional to the spatial line length  $X_{ij} = X' d_{ij}$ , where  $d_{ij}$  is the Euclidean distance between nodes  $i$  and  $j$  given by the link-length distribution.  $X'$  is the so-called specific reactance per length, a reasonable value for the transmission grid is  $X' = 0.265 \Omega \text{ km}^{-1}$  [9]. The spatial embedding of the synthetic networks gives rise to a heavy-tailed reactance distribution as observed for real data [22].

This setup serves as an input to the following swing equation in a co-rotating reference frame [1, 3, 4]

$$\ddot{\phi}_i = P_i - \alpha_i \dot{\phi}_i - \sum_{j=1}^N K_{ij} \sin(\phi_i - \phi_j), \quad (1)$$

$P_i$  is the net injected power at generator  $i$  and  $\alpha_i$  is a dissipation constant assigned to this generator.  $K_{ij}$  is determined by the voltage amplitudes, generator constants and depends on the line length, i.e.  $K_{ij} \propto X_{ij}^{-1}$ . Hence we treat power grid networks as weighted networks where the link weights  $X_{ij}^{-1}$  are a function of the entries of the reactance matrix.

Hence, this model operates on an aggregated grid model (see [4]) and there are some inherent simplifications. The networks, we consider, need to be thought of as reduced power grids after a removal of passive nodes (transformers, substations, loads), containing only generators busses with a net injected power that can also be negative (local demand is higher than local supply). Here we choose a symmetric power dispatch distribution with  $P_i = \pm 1$ . We



**Figure 1.** Basin stability histogram: distribution of single-node basin stability scores within the ensemble. The dashed lines delimit the classes of poor (14%), fair (79%) and high (7%) basin stability, where the percentages give the share of realisations in each class.

further restrict our analysis to the transmission part of the power grid, neglecting ohmic line losses.

The system (1) for a single node, on the one hand, is bistable in the relevant parameter regime with the coexistence of a stable focus and a stable limit cycle. The whole high-dimensional dynamics, on the other hand, thus gives rise to a multitude of attractors of which there is a dominant (concerning the basin size, see [10]) stationary state in which all generators are synchronized, i.e. they uniformly rotate with the grid's rated frequency and constant phase differences between the nodes. This state is a solution to the load-flow problem and referred to as the stationary state in the following.

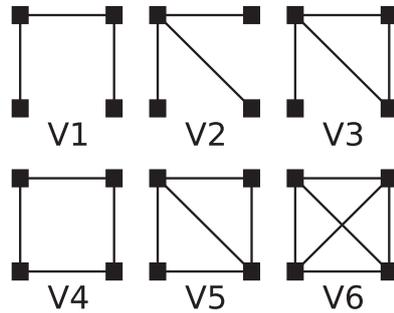
## 2. Stability analysis

As it has been pointed out in earlier studies, the assessment of dynamical stability in systems like the power grid requires non-local and large perturbations to be taken into account. Especially the analysis of a stationary state's basin of attraction [11, 23] in multistable systems gained recently importance. The numerical estimation of single-node basin stability using a Monte-Carlo rejection method has been proposed in [10]. The value  $S_i \in [0; 1]$  of single-node basin stability refers to the probability of the *whole system* to return to a specified stationary state after a (large) perturbation at node  $i$ .

Integrating (1), we estimate the single-node basin stability of the stationary state using a uniform distribution of initial conditions from the product set  $[-\pi, \pi] \times [-100, 100]$  with 100 trials at each node. This yields an upper bound for the standard error  $e$  of the estimated single-node basin stability of  $e = 0.05$ .

Note, that typical local deviations from the grid's rated frequency do not exceed 1 Hz. Our choice of initial conditions hence contains perturbations that are one order of magnitude larger, which we refer to as *large perturbations*. We observed that choosing a larger interval of frequency deviations does not alter the distribution of single node basin stability qualitatively.

From the resulting histogram (cf figure 1), we find that single-node basin stability is not unimodally distributed within the ensemble but shows rather well-separated three classes of



**Figure 2.** Four-sized motifs: sketch of the six possible connected network motifs containing four nodes. Further links, connecting the nodes within the motifs with the remaining network are omitted.

‘poor’, ‘fair’ and ‘high’ values. The majority of the nodes (79%) have a fair single-node basin stability, i.e. it is almost equally likely for the power grid to return to the stationary state or not after a large perturbation. The other two groups stand out as they represent the opposite cases of low (14%) and high (7%) probability of returning to a fixed point.

The observed single-node basin stability distribution coincides well with the distribution reported in [12], although the authors used Gilbert’s random graph model [5] and a uniform coupling  $K_{ij} \equiv K$ . Hence we conjecture that the shape of the distribution is rather set up by the parameters of the swing equation than by the network topology.

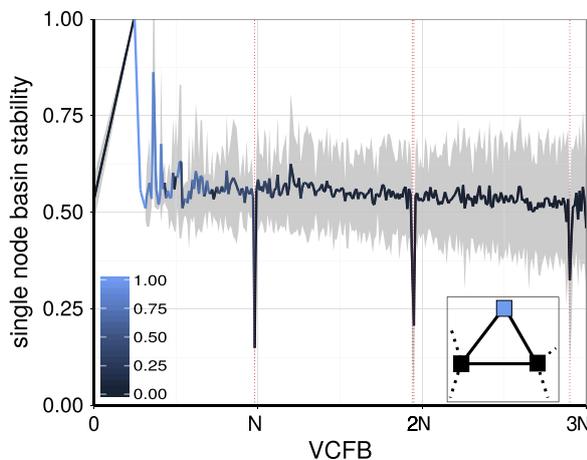
The statistical relation between single-node basin stability and network characteristics is typically nonlinear with no trivial physical explanations; thus the discretisation of basin stability to a categorical variable simplifies the analysis. We will especially focus on the critical class  $C_p$  of nodes with poor single-node basin stability which are easily destabilising the power grid. The class of non-poor nodes will be referred to as  $C_n$ .

### 2.1. Network motifs

Network motifs are isomorphism classes (‘types’) of small connected induced sub-graphs of a network, inheriting all links from the network. The motifs V1–V6 (cf figure 2) are identified using a simple brute-force algorithm that checks for each node the membership in one of the motifs. A single node can be contained in several motifs but a set of four nodes can only belong to one of them. In general, the function of the motifs also depends on the network context. However, we focus on the question which motifs are suited to predict poor single-node basin stability ( $C_p$ ) for the comprised nodes, without analysing their particular function.

The motifs we consider are all four-sized motifs, dead tree gateways and detours.

Dead tree gateways are identified using a set of specific values of the shortest-path betweenness that are typical for nodes adjacent to a dead tree ( $N-2$ ,  $2N-6$ ,  $2N-5$ ,  $3N-10$ , ...; see [10]). Nodes on detours are in turn characterized by a degree of two and a local clustering coefficient value of one. Furthermore, as pointed out in [15], detour nodes are nodes in triangles for which the shortest-path betweenness is very low because almost no shortest path passes through. However if the network is a resistance network, then the node may still share a significant amount of the power flow. This deviation is captured by a measure Newman termed *current flow betweenness*.



**Figure 3.** Detection of detours using  $VCFB_i$ ; plotted is the single-node basin stability  $S$  against the vertex current flow betweenness  $VCFB$ . The red vertical dotted lines with a vertex current flow betweenness of 98, 194, 195 and 290 refer to dead tree gateways. The shading of the line refers to the share of nodes in detours with the respective betweenness value, light colour indicates a higher number. The grey-shaded region marks one standard deviation. The inset contains a sketch of the detour motif, where detour node is marked in blue.

In analogy, we define an *edge current flow betweenness*  $ECFB_{ij}$ :

$$ECFB_{ij} = \left\langle I_{ij}^{st} \right\rangle_{s \neq t, s, t \in V}$$

$$I_{ij}^{st} = X_{ij}^{-1} \left| (R_{is} - R_{js})I_s - (R_{it} - R_{jt})I_t \right|, \quad (2)$$

where  $R$  is the Moore–Penrose pseudo-inverse of the matrix containing the inverse reactances  $X_{ij}^{-1}$ . Note that, in contrast to the original definition by Newman, we do not use uniform line resistance but the heterogeneous link weights  $X_{ij}^{-1}$ . Hence, the  $ECFB$  takes not only a node's position in the network into account but dynamical properties as well.

Especially, if one unit of in-feed and one unit of consumption are randomly distributed over all nodes of the network, the expected flow on the link  $i - j$  is at most the value given by  $ECFB_{ij}$ . Thus, instead of solving the load flow equations, it is more convenient and computationally less expensive to use the  $ECFB$  as an indicator. Summing up  $ECFB$  over all edges adjacent to a node  $i$  yields the node-wise characteristic *vertex current flow betweenness*  $VCFB_i$  which we make use of later on.

$$VCFB_i = \left\langle \frac{1}{2} \sum_j I_{ij}^{st} \right\rangle_{s \neq t, s, t \in V}. \quad (3)$$

As shown by [10], the shortest-path betweenness of a node can be used to detect dead tree gateways with a clear tendency towards poor single-node basin stability. Using instead  $VCFB_i$ , there are downward and upward peaks, relating to betweenness values with significantly lower or higher stability (cf figure 3). Firstly, we reproduce the results from the analysis using shortest-path betweenness, namely the four downward peaks at 98, 194, 195 and 290 (dashed vertical lines) that are present in the weighted analysis as well. Secondly, as a new feature, we

also identify pronounced upward peaks for low values of  $VCFB_i$ . The colour indication visualizes that this is the particular betweenness range found for the nodes in detour motifs, i.e. we uncover that these detour nodes in all cases fall into  $C_n$ .

We perform a two-by-two test, yielding the ratios of nodes falling into the four possible combinations of  $C_p/C_n$  and motif/not-motif.

As expected, dead tree gateways are a very good indicator of  $C_p$  as 90% have a poor single-node basin stability. Detour nodes, however, seem to be better suited to predict non-poor single-node basin stability, i.e., not a single detour node has a poor stability score. This is remarkable as this simple network motif of three nodes in a triangle seems to be sufficient to prevent the appearance of poor single-node basin stability in a power grid.

In case of a failure hitting a detour node, a stable power flow can still be maintained via the shorter opposite network path, i.e. we locally observe *redundant* routes for the power flow. Hence it is unlikely for a perturbation at a detour node to destabilize the whole system.

The role the four-sized motifs V1–V6 play for power grid stability is less pronounced. As a general feature we observe that the ratio of nodes with poor single-node basin stability decreases with the link density within a motif. This again supports the importance of local redundancy for the stability of the system.

### 3. Prediction of instability

In the previous section we investigated network motifs and found that two types of them—the detour nodes and dead end gateways—play a particular role to set up the overall system's stability. Hence, they are candidates to predict the single-node basin stability of the respective nodes.

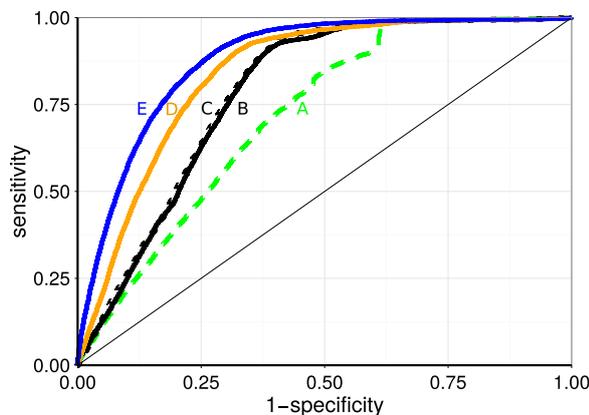
Now we derive a general model to predict whether a node belongs to  $C_p$  or  $C_n$ , using only a small set of network measures that require much less computationally effort than simulation-based estimators for basin stability. As we argued above, the predicted variable is binary, namely either 1 for  $S_i \in C_p$  or 0 for  $S_i \in C_n$ . Here,  $S_i$  is the single-node basin stability at node  $i$ . The aim is to derive a statistical model that predicts a probability  $p_i$  for each node to have a poor single-node basin stability.

Some of our explanatory variables are continuous, others are binary, namely the information whether a node is a detour node or a dead tree gateway. As we have seen the strong indication of the basin stability class given by these motifs in the last section, we employ the following modelling strategy. First we perform a pre-classification of the nodes in three subsets. These are

- (0) detour nodes,
- (i) nodes being neither detours nor dead tree gateways,
- (ii) dead tree gateways,

keeping in mind that a detour node cannot be a dead tree gateway. From the previous analysis we know that  $p_i = 0$  for nodes in set (0). For the remaining two sets (i) and (ii) we perform a logistic regression to estimate  $p_i$ . Consequently, the complete model comprises two logistic regressions with potentially different regression coefficients and thresholds.

To assess the predictive power, we use the cross-validated area-under-curve criterion AUC [7] applied to the receiver operating characteristic (ROC) curve of the individual regression



**Figure 4.** Receiver operating characteristic for set (i): ROC of predictions based on the five different logistic regression models A (dashed green), B (solid black), C (dots black), D (solid orange) to E (solid blue) applied to set (i) with an increasing number of explanatory variables. The models are described in detail in section 3.

**Table 1.** Stability of network motifs: frequency  $\nu_m$  of a node comprised in one of the motifs (cf figure 2), share of these nodes that belong into  $C_p$ .

Motif	$\nu_m$	Motif nodes in $C_p$
Dead tree gw.	5%	90%
Detour node	4%	0%
V1	82%	15%
V2	47%	16%
V3	20%	12%
V4	14%	11%
V5	3%	10%
V6	0.1%	13%

models. Reported values for the sensitivity or specificity of a certain model are always related to the prediction using the threshold closest to the point of perfect prediction.

We start with set (i) containing 91% of the nodes (cf table 1). In model A, using  $V_{CFB_i}$  as an explanatory variable (cf the discussion above), the corresponding ROC curve (cf figure 4) yields (figure 4 A, AUC = 0.717, std.err. = 0.003). The whole curve lies above the diagonal, thus any prediction is better than a random guess. The point closest to a perfect prediction corresponds to a sensitivity and specificity of 71% and 61%. Further improving this result, we enlarge our set of explanatory variables.

A common and simple characteristic for weighted networks is the strength  $\tilde{k}_i$  alias a weighted generalisation of the degree.

$$\tilde{k}_i = \sum_{j \in \mathcal{N}(i)} X_{ij}^{-1}, \quad (4)$$

where  $\mathcal{N}(i)$  denotes the set of nodes adjacent to node  $i$ . The rationale behind this definition is that a high line reactance (lowered line capacity) acts as a bottleneck for the power flow in the system, effectively reducing the strength of a node. Note, however, that a large strength on the one hand promotes the dispersion of a perturbation at this node but, on the other hand,

facilitates the spreading of perturbations on the network as well. Adding strength to the set of predictive variables (figure 4 B, AUC = 0.784, std.err. = 0.002) consequently improves the predictor because of the added information that is not contained in the flow-based betweenness, still the sensitivity and specificity do not go beyond 84% and 66%. Using the average neighbour's strength

$$\tilde{k}_i^N = \tilde{k}_i^{-1} \sum_{j \in \mathcal{N}(i)} \tilde{k}_j, \quad (5)$$

we also add the average neighbourhood line capacity to the local information contained in the strength, yielding only a small improvement of the model (figure 4 C, AUC = 0.790, std. err. = 0.002). Still we keep it, as it improves the final predictor by more than this intermediate improvement.

Adding the strength as an explanatory variable mainly improved the sensitivity whereas the specificity remained smaller. To get an increased specificity, we need a further predictor exposing truly non-poor nodes. Among the yet unused weighted characteristics, a candidate for this is the weighted local clustering coefficient. We define it as

$$\tilde{C}_i = \frac{\sum_{j,k,j \neq k} \min(X_{ij}^{-1}, X_{jk}^{-1}) \min(X_{ik}^{-1}, X_{jk}^{-1})}{\sum_{j,k,j \neq k} X_{ij}^{-1} X_{ik}^{-1}}, \quad (6)$$

so that  $\tilde{C}_i \in [0; 1]$  just as the unweighted local clustering coefficient. The addition of this explanatory variable further improves our model (figure 4 D, AUC = 0.839, std.err. = 0.002). High values of  $\tilde{C}_i$  close to 1 correspond to a high number of cycles in the proximity of a node  $i$  where the lines have a low reactance. Hence, we conjecture that such setup promotes the quick annihilation of perturbations spreading on the network. Indeed we mainly improve the specificity of the model to 73%, while the sensitivity remains at 84%.

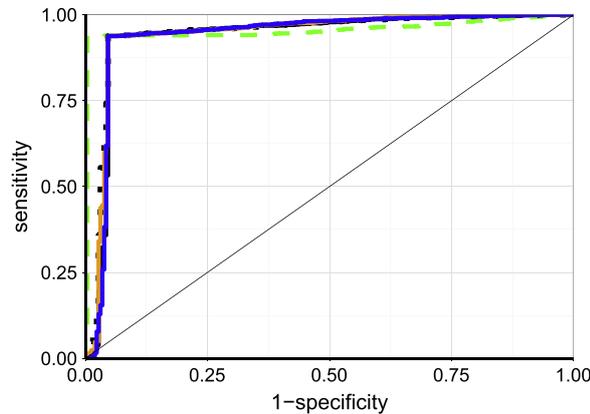
Furthermore, we anticipate that the addition of a centrality measure improves our predictor. In analogy to a piano string, nodes being more central are less likely to excite many perturbational modes than nodes with a low centrality. In other words, if the string is plucked close to one of the endpoints, more modes (overtones) are being excited than when the hammer acted in the centre of the string. Following this intuition, we expect perturbations hitting nodes with a high centrality to be less likely to destabilize the network.

A natural choice of a centrality measure for power grid networks is the *effective resistance closeness centrality* ERCC<sub>*i*</sub>, where the length of a network path between to nodes  $i$  and  $j$  in the standard unweighted closeness centrality is replaced by the effective resistance between the two nodes. The effective resistance ER<sub>*ij*</sub> is defined as the resistance a single edge that replaces all lines between two nodes  $i$  and  $j$  in an equivalent circuit would have. We define ERCC<sub>*i*</sub> as

$$\text{ERCC}_i = 1 / \langle \text{ER}_{ij} \rangle_j$$

$$\text{ER}_{ij} = R_{ii} - R_{ij} - R_{ji} + R_{jj}. \quad (7)$$

We finally have a model (figure 4 E, AUC = 0.915, std.err. = 0.002) for set (i) using the explanatory variables  $\tilde{k}_i$ ,  $\tilde{k}_i^N$ ,  $\tilde{C}_i$  and ERCC<sub>*i*</sub> where the best prediction reaches a sensitivity and specificity of 84% and 77%. In other words we are able to predict the single-node basin stability to be poor or non-poor correctly in about 80% of the cases using only (weighted) network information without any need of costly numerical simulations.



**Figure 5.** Receiver operating characteristic for set (ii): ROC of predictions based on the five different logistic regression models A to E applied to set (ii). The colours and line types are identical to figure 4.

Note that we indeed find that the weighted explanatory variables contain different information compared to the equivalent unweighted characteristics improving the prediction. This can be seen, using for instance shortest-path betweenness instead of  $VCFB_i$  (E', AUC = 0.876, std.err. = 0.002) or the standard closeness centrality instead of  $ERCC_i$  (E'', AUC = 0.839, std.err. = 0.002) in the final predictor, showing no improvement of the predictor compared to the previous step and less predictive performance than the predictor using the weighted measures.

Repeating the same procedure, adding explanatory variables step by step, for set (ii) we find model (figure 5 A, AUC = 0.959, std.err. = 0.005) with a sensitivity and specificity of 94% and 99% to have the best ROC (cf figure 5). This is a surprising result, as the addition of further explanatory variables to  $VCFB_i$  actually *decreases* the AUC and hence the predictory performance. This can typically happen in the case of over-fitting. Note that among all individual candidates for explanatory variables, a model using only  $VCFB_i$  for set (ii) still has the best characteristic. It also improves on the result from the motif analysis where we find dead tree gateways to have a 90% probability of falling into  $C_p$ .

The regression coefficients of the final subset models are listed in table 2, each coefficient is highly significant with a p-value less than 1‰. Taken that the logistic function is monotonically increasing with its argument we see that the negative sign of a regression coefficient relates to the corresponding variable decreasing the probability of a poor single-node basin stability.

From the coefficient of  $ERCC_i$  we see that a high centrality indeed corresponds to a low probability of a node having poor single-node basin stability, supporting our reasoning based on the piano metaphor from above.

#### 4. Discussion

In conclusion, the two main results of our analysis are the identification of the detour motif being important for enhancing power grid stability and a statistical model to predict poor single-node basin stability using only a small set of standard and novel network characteristics as explanatory variables.

**Table 2.** Coefficients of the logistic regression models: coefficients of the logistic regressions for poor single-node basin stability after the primary classification, the explanatory variables are strength  $\tilde{k}_i$ , average neighbour's strength  $\tilde{k}_i^N$ , weighted local clustering coefficient  $\tilde{C}_i$ , vertex current flow betweenness  $VCFB_i$  and effective resistance closeness centrality  $ERCC_i$ . Values are given for both sets (i) no detour/no dead tree gateway and (ii) no detour/dead tree gateway. Further we state the best values of the sensitivity and specificity at thresholds (i)  $p_i = 0.125$  and (ii)  $p_i = 0.913$ .

Coefficient	Value (i)	Value (ii)
Intercept	3.766625	5.6446
$VCFB_i$	-0.098590	-2.5557
$\tilde{k}_i$	0.074374	—
$\tilde{k}_i^N$	-0.395878	—
$\tilde{C}_i$	-7.588657	—
$ERCC_i$	-1.512116	—
sensitivity	84%	94%
specificity	77%	99%

Specifically, we find that the appearance of a detour in the network seems to prevent the detour node from having a poor stability in contrast to dead trees which are known to have the opposite effect. Using these two motifs that have a pronounced effect on the single-node basin stability of the comprised nodes for a primary node classification into three subsets, we are able to predict the stability class of about 80% of the nodes being neither dead tree gateways nor detour nodes.

There are at least two applications of our results, related to the two time scales of network topology evolution that were mentioned in the beginning.

Firstly, from the study of network motifs, general design principles for the long-term planning might be derived. Especially detours seem to be suited to promote stability, whereas dead tree gateways are possible weak points. It has been shown by [10] that reconnecting dead trees to the remaining networks, thus creating a new cycle, drastically improves the distribution of single-node basin stability. From our new findings we can interpret that in most of the cases, however, the authors transformed dead ends into detours.

Secondly, with a fast predictor based on information about the network structure and the line parameters, we developed a very fast tool to predict the appearance of nodes with poor single-node basin stability during live operation. This particular feature might prove useful to system operators, who usually have very limited computational resources that are already occupied by simulations of so-called 'N-1'-cases.

The novel weighted local clustering coefficient and the ERCC have been developed within the context of this work and we will give further details in a forthcoming publication.

We have chosen a small set of explanatory variables based on assumptions how these measures relate to the spreading of perturbations on the network. There is already literature on this field [14], however actual dynamics of this process in the power grid network is yet not fully understood. This would be an important point for consecutive research, building a theory to justify or falsify our assumptions on the physical meanings of the characteristics.

## Acknowledgment

This work was partially supported by the CoNDyNet project which is funded by the German Federal Ministry of Education and Research.

## References

- [1] Anderson P and Fouad A 2003 *Power System Control and Stability* 2nd edn (New York: Wiley)
- [2] Crucitti P, Latora V and Marchiori M 2004 A topological analysis of the Italian electric power grid *Phys. A: Stat. Mech. Appl.* **338** 92–97
- [3] Dörfler F and Bullo F 2010 Synchronization and transient stability in power networks and non-uniform Kuramoto oscillators *SIAM J. Control Optim.* **50** 1616–42
- [4] Filatrella G, Nielsen A and Pedersen N 2008 Analysis of a power grid using a Kuramoto-like model *Eur. Phys. J. B* **61** 485–91
- [5] Gilbert E 1959 Random graphs *Ann. Math. Stat.* **30** 1141–4
- [6] Klemm K and Bornholdt S 2005 Topology of biological networks and reliability of information processing *Proc. Natl. Acad. Sci. USA* **102** 18414–9
- [7] LeDell E, Petersen M L and van der Laan M J 2012 Computationally efficient confidence intervals for cross-validated area under the ROC curve estimates *U.C. Berkeley Division of Biostatistics Working Paper Series*
- [8] Ma'ayan A, Cecchi G A, Wagner J, Rao A R, Iyengar R and Stolovitzky G 2008 Ordered cyclic motifs contribute to dynamic stability in biological and engineered networks *Proc. Natl. Acad. Sci. USA* **105** 19235–40
- [9] Machowski J, Bialek J and Bumby J 2008 *Power System Dynamics: Stability and Control* 2nd edn (New York: Wiley)
- [10] Menck P J, Heitzig J, Kurths J and Schellnhuber H J 2014 How dead ends undermine power grid stability *Nat. Commun.* **5** 3969
- [11] Menck P J, Heitzig J, Marwan N and Kurths J 2013 How basin stability complements the linear-stability paradigm *Nat. Phys.* **9** 89–92
- [12] Menck P J and Kurths J 2012 Topological identification of weak points in power grids *Nonlinear Dyn. Electron. Syst. Proc. NDES 2012* pp 1–4 VDE
- [13] Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D and Alon U 2002 Network motifs: simple building blocks of complex networks *Science* **298** 824–7
- [14] Mones E, Araújo N A M, Vicsek T and Herrmann H J 2014 Shock waves on complex networks *Sci. Rep.* **4** 4949
- [15] Newman M J 2005 A measure of betweenness centrality based on random walks *Soc. Netw.* **27** 39–54
- [16] Pagani G A and Aiello M 2013 The Power Grid as a complex network: a survey *Phys. A: Stat. Mech. Appl.* **392** 2688–700
- [17] Pourbeik P, Kundur P and Taylor C 2006 The anatomy of a power grid blackout—root causes and dynamics of recent major blackouts *IEEE Power Energy Mag.* **4** 22–29
- [18] Prill R J, Iglesias P a and Levchenko A 2005 Dynamic properties of network motifs contribute to biological network organization *PLoS Biology* **3** 1881–92
- [19] Schultz P, Heitzig J and Kurths J 2014 A random growth model for power grids and other spatially embedded infrastructure networks *Eur. Phys. J.* **223** 2593
- [20] Shen-Orr S S, Milo R, Mangan S and Alon U 2002 Network motifs in the transcriptional regulation network of *Escherichia coli* *Nat. Genetics* **31** 64–8
- [21] UCTE 2006 Final report—system disturbance on 4 November 2006, *Union for the Coordination of Transmission of Electricity*

- [22] Wang Z, Scaglione A and Thomas R J 2010 Generating statistically correct random topologies for testing smart grid communication and control networks *IEEE Trans. Smart Grid* **1** 28–39
- [23] Wiley D A, Strogatz S H and Girvan M 2006 The size of the sync basin *Chaos* **16** 015103
- [24] Witthaut D and Timme M 2012 Braess's paradox in oscillator networks, desynchronization and power outage *New J. Phys.* **14** 083036