Complex Systems Techniques applied to Power Transmission Expansion Planning. Part I: Generating Random Networks that are Consistent with Power Transmission

Alice Patania, Jean-Gabriel Young, Sara Lumbreras, María Pereda, Ilaria Bertazzi, Daniel Citron, Masahiko Haraguchi

***Abstract*—Transmission Expansion Planning (TEP) could benefit from studying how network structure affects power grid performance. To accomplish this, we require large ensembles of power grid data. These are not available, so it is therefore necessary to generate random networks that have a structure that is consistent with power grids. We propose a generative model that is able to capture the main characteristics of the power grid. This model is based on an epsilon-disk algorithm that incorporates the spatial component of power systems while taking into account the particular characteristics of sources (power plants), sinks (demand) and edges (transmission lines) in the problem. We study the performance of the algorithm and apply it to a real case study based on the Spanish power system.**

***Index Terms*— Transmission Expansion Planning, Network Generation**

I. Introduction

Transmission Expansion Planning (TEP) can be viewed as a network design problem that presents some special characteristics. In particular, the evaluation of the performance of the network involves relatively sophisticated physical considerations (Kaltenbach, Peschon, & Gehrig, 1970) that can be expressed in mathematical formulations of diverse complexity. Additionally, the number of possible ways to expand the network grows exponentially with the system size - in principle; we could link each possible pair of nodes with any cable type. This means that the problem is, for real systems, unmanageably large.

In this context, it would be very valuable to guide the expansion process by restricting the network designs or some of their characteristics. For instance, if we knew that, in general, good designs had a given degree distribution, we could reduce the space of feasible solutions by constraining the degree distribution of the expanded network. We take inspiration from an earlier work by some of the authors, where bounds are imposed on some of the variables of an optimization problem, successfully reducing the feasible space and dramatically improving computation times (S. Lumbreras, Ramos, & Sánchez-Martin, 2014).

However, the kind of information needed to formulate these bounds is not available to the network planner; we do not know, in general, what characteristics make good power networks. In order to extract this knowledge it would be necessary to evaluate different network designs under a range of performance criteria, and then study whether there are any given network metrics that can predict, to some extent, the performance of a given design. If good network designs are found to have a range of values for a specific metric, then we can add that constraint to the design process. However, this kind of study is not directly possible, as the set of real power systems that can be studied is very limited.

We propose to circumvent this limitation by *generating random networks that are consistent with power transmission*. This considers the special characteristics of the problem so that the networks generated are consistent with the features of real power grids. Additionally, we require an efficient enough method so that a sufficient number of case studies can be generated in manageable computation times.

This paper describes this proposal. We first describe TEP as a design problem and the power grid from a network perspective (sections II and III). Then we present the algorithm (section V) and our results (section VI). Finally, section VII presents conclusions and outlines further lines of research.

II. TEP as a Network Design Problem

TEP can be understood as a network design problem (NDP) that minimizes the cost of constructing a network that links power generators and demand nodes, incorporating all

This working paper has been produced as a result of the work carried out at the 2015 Complex Systems Summer School at the Santa Fe Institute. The authors acknowledge support from the Spanish MICINN Project CSD2010-00034 (SimulPast CONSOLIDER-INGENIO 2010).

Jean-Gabriel Young is with Université Laval (email: jean-gabriel.young.1@ulaval.ca)

Alice Patania is with Politecnico di Torino and I.S.I. Foundation (email: alice.patania@polito.it)

Sara Lumbreras is with Universidad Pontificia Comillas (email: slumbreras@comillas.edu)

María Pereda is with Universidad de Burgos (e-mail: mpereda@ubu.es).

Ilaria Bertozzi with Università di Torino (e-mail ibertazz@unito.it)

Daniel Citron is with Cornell University (email: dtc65@cornell.edu).

Masahiko Haraguchi is with Columbia University (email.: haraguchi.masahiko@gmail.com). mh2905@columbia.edu).

constraints that describe the operation of power plants and the laws that govern the physical power flows in the system.

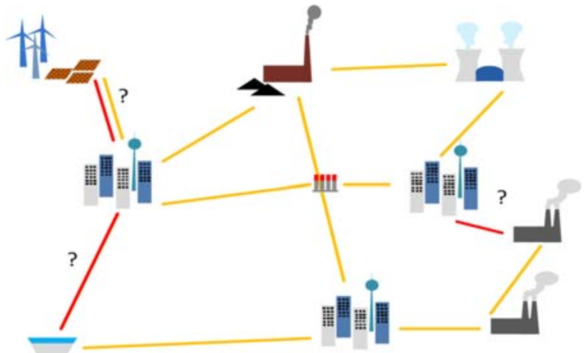


Figure 1. The Transmission Expansion Planning Problem decides which transmission links should be added to the network that links generation and demand in the power system (Lumbreras, 2012).

In general, NDPs consist of identifying an optimal subgraph *F* of an undirected graph *G* subject to some feasibility conditions. Well known NDPs are the Minimum Spannning Tree Problem, the Travelling Salesman Problem and the Shortest Path Problem. Reference (Feremans, Labbé, & Laporte, 2003) reviews the main properties and applications of these problems, which have important implications in fields such as transportation, telecommunications, project management or biology.

Power network design has some additional characteristics that distinguish it from the latter cases:

- The network is spatially embedded. That is, each node is located at a specific a position in space. Edges are more expensive to build when they link nodes that are far apart. We will assume that the cost of building scales linearly with geographical distance. This is especially important in the current context of the problem. While conventional fossil-fuel power plants used to be located close to the urban areas, we now have renewable generation that needs to be built where the resource (*e.g.* wind, sun) is abundant. For this reason, renewable sources are often located farther away from the main demand centers. In addition, distributed generation is another new factor to consider that is reflected in the spatial distribution of sources.
- There are different types of nodes and edges with different attributes. We have nodes that can be sources (generators), sinks (loads) or neither (substations, which are intermediate points of connection in the grid). Generators have attributes such as their technology (coal, solar, etc), maximum generation capacity (which cannot be surpassed) or their operation cost (the marginal cost of providing one extra MW of energy). Demand nodes have their load as their main parameter. Edges have a maximum flow capacity. Some other technical considerations can be considered by means of these attributes.
- The fact that the edges have a maximum flow makes TEP a special type of capacitated NDP (Raack, 2014).
- There are several performance objectives that we must take into account, which in a basic case should include:
  - Investment cost of the network, that is, the cost of the transmission lines that need to be built.
  - Operation cost of the network, which represents the base-case efficiency of a design. It is calculated as the sum of the costs of the power generated to serve demand. If some demand is not met, the operation cost should be increased to penalize for it.
  - Reliability penalties: the designs that are not robust with respect to node or edge failures should also be penalized. How to take this into consideration is a challenge on its own, which is studied in part III of this project.
- Calculating operation cost is a nontrivial problem. We should use cheap generators over more expensive ones, but we also need to consider transmission constraints (maximum flow capacities and Kirchhoff's Laws). This means that minimum operation cost cannot be immediately calculated and an optimization problem must be solved instead. This optimization problem is described in the appendix section of this paper.

III. Power grid as a network

Power grids differ from typical systems studied in network science because they are spatially embedded, undirected, multi-type graphs which support directed dynamics (Kirchhoff's circuit laws). In general, spatially embedded networks have a strongly constrained topology that translates to fractal scaling, heavily constrained degree distributions, few or no long distance connections, and trivial clustering-degree correlations ((Boccaletti, Latora, Moreno, Chavez, & Hwang, 2006), See §2.5). For example, some spatial graphs are planar and the upper bound for the number of links scales linearly with the number of nodes N instead of with N^2, as it happens with non-spatially embedded networks.

Power grid networks display peculiar topological characteristics that are not found in other spatial graphs: they are sparsely connected (Albert, Albert, & Nakarado, 2004; Wang, Thomas, & Scaglione, 2008), feature some long distance "random" links, and appear to be much-more clustered than their random equivalent counterparts (Solé, Rosas-Casals, Corominas-Murtra, & Valverde, 2008).

The characteristic degree distribution for power grids is generally exponential (Albert et al., 2004; Amaral, Scala, Barthelemy, & Stanley, 2000). However, in (Wang, Scaglione, & Thomas, 2010) it was stated that low degree nodes are not well captured by the exponential distribution and must therefore be specified explicitly (Wang et al., 2010). The clustering-degree correlation is trivial, *i.e.* $\mathrm{C(k)}$ is roughly independent of k, where $\mathrm{C(k)}$ is the clustering coefficient as a function of the degree (Ravasz & Barabási, 2003). In contrast, hierarchical networks roughly follow $\mathrm{C(k)}{\sim}k^1$ (Ravasz & Barabási, 2003).

Random graph models have been proven to be powerful tools that allow gaining insights about the features of real systems (Hébert-Dufresne, Allard, Young, & Dubé, 2013a; Hébert-Dufresne, Allard, Young, & Dubé, 2013b). In this paper we propose a new model for generating random spatially embedded networks with realistic power grid characteristics. By generating power grids at a whim, we open the way to the study of the interplay between network characteristics and the efficiency and robustness of power grids.

IV. Related work

There is substantial work reported in the literature on growth models for spatial graphs and random power grid networks. Most authors focus on the fact that nodes are largely connected to their geographical neighbors. For example, reference (Manna & Sen, 2002) introduces a variation on Barabasi-Albert's Preferential Attachment (PA) with spatially embedded nodes. The authors impose a diminishing return on long-range connections, with a probability of connection $p_{ij} \propto k_j {l_{ij}}^{-\alpha}$, where l_{ij} is the Euclidean distance between nodes i and j, and k_j is the degree of node j. This model builds on earlier ideas such as the ones presented in reference (Waxman, 1988), where the connection probability $p_{ij} \propto e^{-l_{ij}/\alpha}$ decreases exponentially with distance. Another popular approach is to focus not on the spatial aspect of power grids but rather on their statistical properties (Amaral et al., 2000). For example, adding rules to classical PA, such as node aging and edge cost efficiency yield. This produces models that correctly reproduce the degree distributions of power grids (Amaral et al., 2000). Incorporating limited subsets of accessible nodes leads to similar effects (Mossa, Barthelemy, Stanley, & Amaral, 2002). In the realm of power grids, the most important results are perhaps the ones of Wang et al. (Wang et al., 2008; Wang et al., 2010). They introduce a scalable model for generating random power-grid topologies. Their idea is to connect randomly placed nodes so that an edge-length distribution is respected. Sources are then randomly selected among the nodes and the electrical properties of the power grids are derived from the topology.

V. Epsilon-disk model

Our model explores a different connection mechanism. Rather than imposing an edge-length distribution, we connect nodes based on a local rule, such that the structural properties of the power grid arise solely as a consequence of node placements.

Our generation algorithm proceeds in 3 steps:

- We assign the nodal locations, types (source or sink) and attributes (maximal and minimal capacity, cost, demand). These features are either fixed or specified according to a probabilistic distribution function.
- Edges are placed. We use a deterministic process which probes the neighborhood of the sinks for potential connections, until their electrical demand is adequately satisfied.
- Transmission line capacities are assigned, taking into consideration the operation of the system and the investment cost of the lines.

This computationally efficient model produces power grids with realistic topology.

1) Node placement and attributes

Inside a fixed area, given the expected number of sinks K and sources S, $\mathrm{K} + \mathrm{S}$ nodal locations are selected according to a random distribution function. In this report, we use a uniform distribution, but we could also consider a Poisson distribution (Wang et al., 2008) and a spatially peaked distribution so that they cluster around a fixed location.

Nodal attributes are also drawn from random distribution functions. One must specify the demand of power of every sink i, as well as a minimum $\underline{G_i}$ and maximum $\overline{G_i}$ production capacity for each source i.

2) Growth of the disk and connectivity

The disk model introduced here is loosely based on the topological construction of the Vietoris–Rips complex (Vietoris, 1927). We take into consideration the physical length of the edges, and the electrical properties of the nodes to construct the edges of the graph.

The general idea of our connection algorithm is to consider growing disks of epsilon radius centered on the sinks. Every node that enters a disk becomes connected to the center node, regardless of its type.

The disk model is defined by a real valued parameter ε, $0 \le \varepsilon_{min} \le \varepsilon$. Starting with disks of radius ε_s centered on each node, we increase the radius of each sink by a value of ε_i at each step, until an adequacy condition is met. If the adequacy condition is is not verified for all the connected components then ε_i is increased for the sinks in the connected components that do not verify the condition.

3) Power adequacy

Given the power g_i generated by source i, as well as a minimum and maximum capacity ($\underline{G_i}$,$\overline{G_i}$) it is easy to see that the sum of all power demands D_i in a connected graph must satisfy the inequality:

$$\sum_i \overline{G_i} \ge \sum_i D_i \ge \sum_i \underline{G_i} \cdot x_i \qquad (1)$$

where x_i represents the commitment of a given power generator, that is, whether it is working or not.

There exists at least a set of g_i such that every source in the graph operates at a level that lies in its accessible range [$\underline{G_i}$,$\overline{G_i}$].

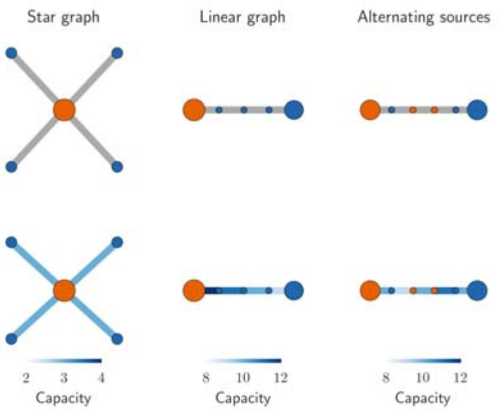
Setting some $x_i = 0$ for some of the sources amounts to not using some of the sources. Identifying the optimal set of committed sources Ω reduces to the subset-sum problem, which is known to be NP-complete (Moore & Mertens, 2011). We therefore require that every source produces power within its capacity range, a stricter but much less computationally demanding adequacy condition. This is equivalent to assuming that the unit-commitment problem has been determined beforehand, so that only active sources have a non-zero lower bound on their generation $\underline{G_i} > 0$.

4) *Edge capacities*

The previous steps have assigned electrical properties to nodes and adequacy has been enforced. We then assign carrying capacities to the edges. A simple solution consists of assigning edges based on optimal power flow. This process is carried out by solving a power flow model as described in the appendix section. This process is carried out by solving a simplified setting of the power flow model described in the appendix section, in which only the second-stage constraints are taking into consideration. It should be noted that this power flow model can be easily stated as a matrix equation involving the Laplacian of the graph ((Newman, 2010), See §6.14.1).

Figure 2 shows the edge capacities assigned by our method, for some very simple power grids. Notice how this simple algorithm correctly identifies and ignores redundant edges. The two edges that differentiate the star and bowtie graph are assigned a null capacity. A similar behavior is observed on the tadpole graph. The linear graphs showcase an important and vital feature of the capacity assignment algorithm: it accounts for information beyond the immediate neighborhood of a node. Even if the most demanding sinks are far from the most powerful source, intermediary edges correctly receive a higher capacity than imposed by their immediate neighborhood.

Very importantly, this process assumes that continuous values of capacity can be assigned, while real power grids can have only transmission lines with discrete capacity values. It should be understood that this step can be modified easily to get reasonable networks that abide the discreteness constraint, for instance, by rounding capacities up or down.



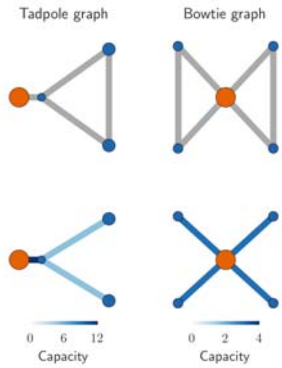


Figure 2. Five test graphs. (top) Backbone of the graphs. (bottom) Optimal capacity identified by our algorithm. Orange nodes represent sources, and blue nodes represent sinks. The size of a node is proportional to its demand (sinks) or maximum capacity (sources). Both the demands and maximum capacities sum to 12 MW.

VI. Results

1) *Properties of the algorithm*

The model introduced in the previous section is strongly influenced by the demand distribution. Indeed, for a fixed amount of available power (i.e. a fixed set of sources), connected components will normally need to be larger to meet the adequacy condition. To study the effect that this choice can have on the topology of the generated random networks, we create random power grids with both lognormal and uniform demand distributions. We then vary the expected total demand and characterize the structure of the network as a function of the total demand. In both case studies, we create networks of 10 sources of minimal capacity $\min(C_i) = 0$ and maximal capacity $\max(C_i) = 10$, and 100 sinks. The nodes are placed uniformly at random in a square. The density of the lognormal distribution is defined as

$$p(x)dx = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\ln(x)^2}{2\sigma^2}} dx \qquad (2)$$

and has an expected value of $e^{\frac{\sigma^2}{2}}$. We therefore expect the demand to match the capacity when

$$\sigma = \sqrt{2 \cdot \ln(10)} \approx 2.14\sqrt{2 \cdot \ln(10)} \approx 2.14$$

Similarly, the expected value of an uniform distribution defined on $[0,2d]$ equals d, such that we expect demand to match capacity when $d = 10$. The results of our simulations are shown in Fig. 2. In both cases, nodes are randomly placed in a square embedding space, and connected with our algorithm.

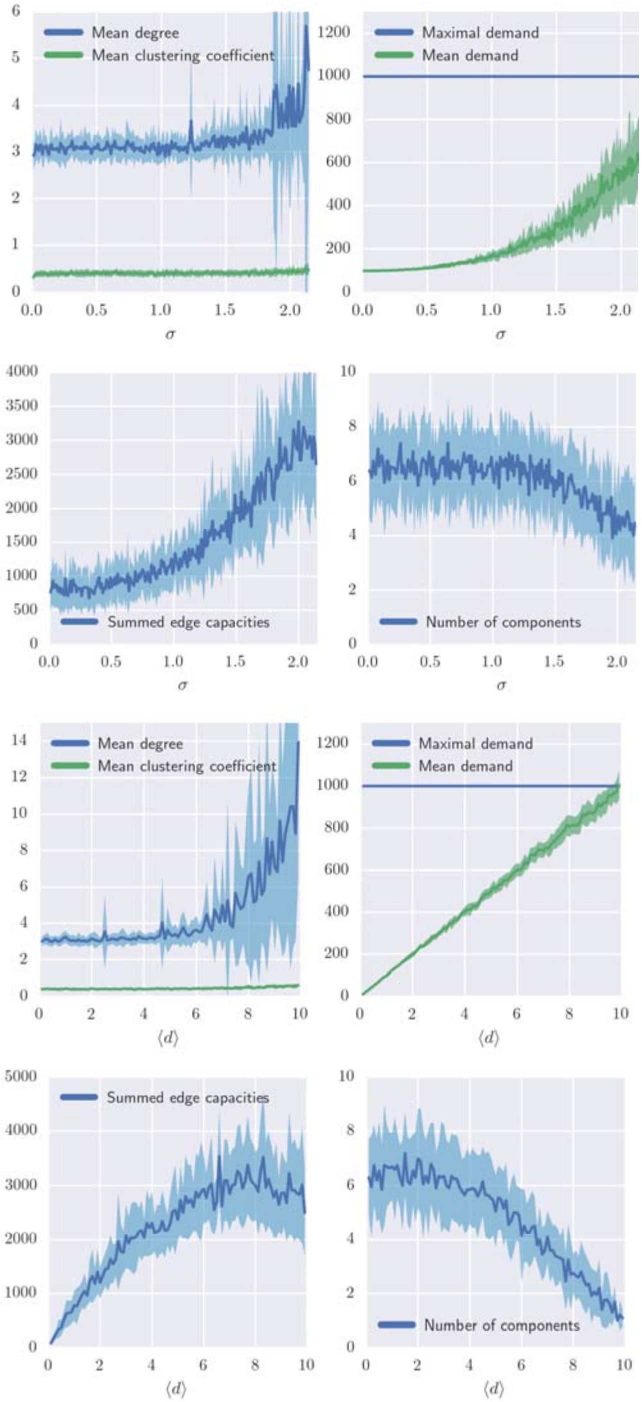


Figure 3. Effect of the demand distribution on the network structure. Network characteristics as a function of the scale parameter of a lognormal distribution of demands (top) or as the average of a uniform distribution of demands (bottom). In both case studies,

Figure 3 shows that demand increases the generated graphs enter more extreme regimes. The average degree increases and peaks when the demand matches the available capacity. This means that the disks become very large before the adequacy condition is met. As a result, a non-vanishing portion of the nodes lie in the disk of every sink. This not only leads to higher average degrees, but also to a reduction in the number connected components. Surprisingly, these additional connections do not seem to affect the clustering significantly. We can therefore conclude that the additional edges essentially act as bridges between the different regions of the graph. We also observe that the carrying capacities of the edges also increase with the demand. It's a straightforward consequence of the fact that more cables are needed to deliver the power required by the sinks. Figure 4 confirms the intuitions developed with the analysis of Figure 3. It shows typical outputs of the algorithm for extreme choices of parameters. When there is much more available power than required by the sinks (left), the graphs tend to be sparsely connected. Only the immediate neighborhood of each sink is explored before the adequacy condition is quickly met. However, when the demand roughly equals the capacity, the grids become densely connected. In comparing the results for the lognormal and uniform case, we also see that heterogeneity in the demand distribution translates to heterogeneity in the capacity of cables.

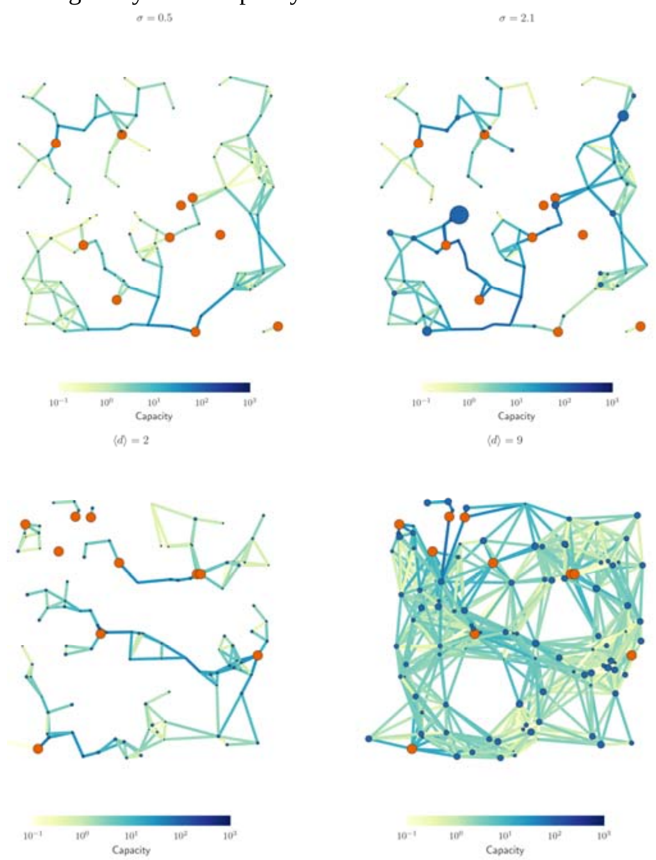


Figure 4. Typical power grids produced by our algorithm, with lognormal distribution of demands (top) and uniform distribution of demands (bottom). We use the parameters described in the caption of Fig. 2. The leftmost figures show power grids in a regime where the capacity of the sources far outweighs the total demand of the sinks. The rightmost figures show the same power grids in a regime where the demand is roughly equal to the maximum capacity.

In Figure 5, we study the effect of the shape of the embedding space on the random power grid structure. Again, we create graphs of 110 nodes, 10 of which are sources of maximal capacity $\max(C_i) =$, while the rest are sinks with demands drawn from the uniform distribution. Given the sizeable effects that the demand distribution has on the resulting power grid, we consider both of the extreme regimes highlighted in Fig. 3. The nodes are placed uniformly at random in a rectangle of constant area A, and the length x of one of the sides is varied to produce shapes ranging from a thin rectangle to a square. When x is very low, one side of the rectangle is disproportionately larger than the other. This can be used to represent different geographical conditions in power systems, with countries such as Germany being described by proportionate squares, and others such as Chile or Italy displaying very stretched geometries.

In both regimes (low and high demand), widely different side lengths lead to an increase in average degree, a reduction in overall edge capacity, and an increase in the number of connected components. Essentially, asymmetric grids break down in highly clustered components of sinks, connected by low capacity edges. This effect is somewhat similar to the one observed in Fig. 2. We conclude that demand increases and asymmetry are analogous effect. The root of this similarity is that both phenomena lead to reduced power availability in the immediate vicinity of sinks. The disks must therefore grow much longer before the adequacy condition is met.

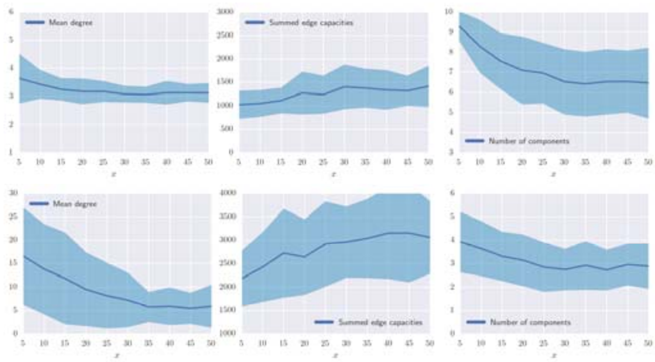


Figure 5. Effect of the shape of the underlying space on network characteristics. We use the parameters described in the caption of the previous figures to generate power grids with uniform distributions of demand, of mean 2 and 8 [MW]. The graphs are embedded in a rectangle of constant area A = 2500 [km] and the length x of one the side is varied to produce a range of different shapes.

2) *Computational complexity of the algorithm*

One of the key objectives of our algorithm is computational efficiency. Faster network generation allows one to study larger systems with better statistics. Figure 6, where we show the running time of the algorithm as a function of the number of nodes in the power grid leads us to believe that our algorithm runs roughly in $O(N+M)$, for a power grid of N nodes.

For the sake of simplicity, let us assume that we generate a network in a square grid of linear length L. This allows us to calculate the computational cost of each step. First, it is clear that assigning nodal parameters is linear in the number of nodes.

Second, we argue that the connection step runs $O(N+M)$, where M is the number of edges. Indeed, it is easy to see that it takes at worst $\sqrt{2}{\varepsilon_s}^{-1}L$ disk growth steps to find a partition of the nodes in adequate components ($\sqrt{2}L$ is the radius of the largest possible disk). Each of these steps involves adding a fraction of the M edges, and has a linear cost in the number of nodes because one must check every component for adequacy. Thus connecting the nodes requires $O(N+M)$ operations (with a potentially large multiplicative factor if ε_s is small).

Finally, assigning capacities takes $O(n^{\alpha})$ operations, where α depends on the efficiency of the matrix equation solver rather than the specifics of the graph. Recall that this step requires that we solve the power flow equations. The most straightforward method for this, matrix inversion, entails $\alpha \leq 2.377$. Moreover, since we solve this problem for every connected component, the average size of the component comes into play. If the size of the average component scales with the network size, $n = aN$ and each problem contains an extensive number of variables. Conversely, if the size of a component does not depend on network size (as it happens in the typical case), then n is the average component size and does not depend on N. Thus, for fragmented graphs, the algorithm is $O(N+M)$ because the connection step is the most costly, whereas it runs in $O(N^{2+\alpha})$for graphs with few connected components.

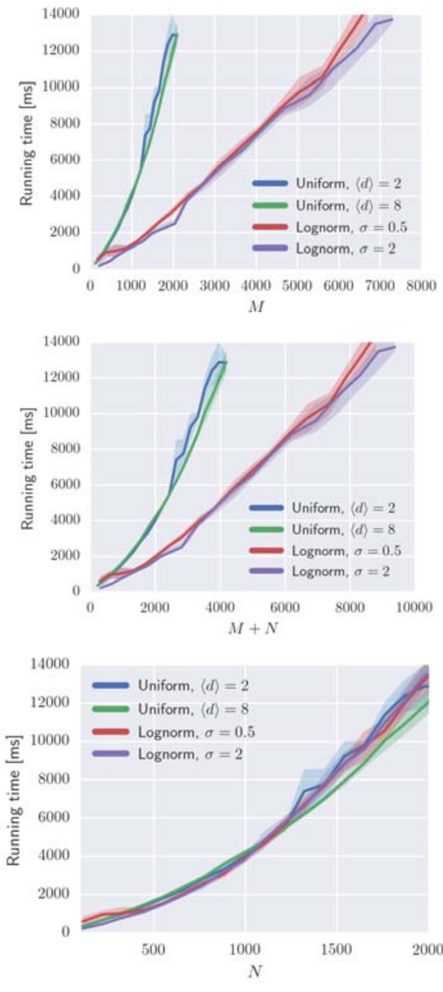
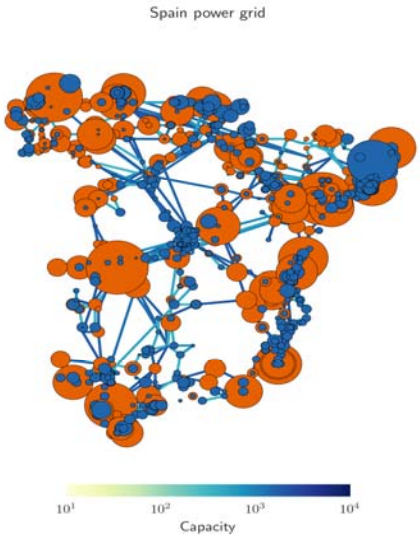
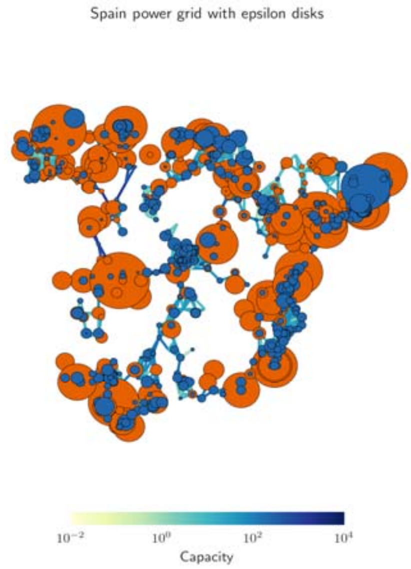


Figure 6. Complexity of the algorithm. Running time as a function of the number of nodes and. We generate random power grids with *10m* sources of a maximum capacity of 100 MW and *100m* sinks with demands drawn from various distributions. The nodes are uniformly placed in a square plane, and connected with our algorithm. The running time is averaged over 25 realizations, for each value of m=1,2,..,20.

B. *Case study: the Spanish power grid*

We test our model by analyzing the Spanish power grid. To do so, we study three versions of the grid. The first version is the actual power grid, reconstructed from reference (S. Lumbreras, 2014). In the second version, we use the same nodes but place the transmission lines using the epsilon-disk model and assign capacities by calculating optimal power flows. The third version is built with the same procedure but nodes are now randomly placed in a square embedding space. Figure 7 illustrates each version. The second version only depends on ε_s (we selected $\varepsilon = 0.5$ and a space of linear dimension $L = 1000$) and is therefore a deterministic graph, whereas the third version is sampled from random graph ensemble.





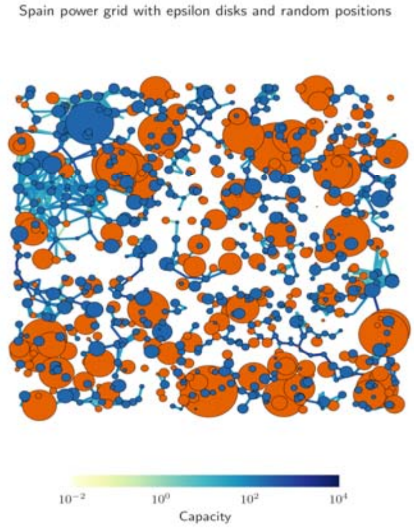
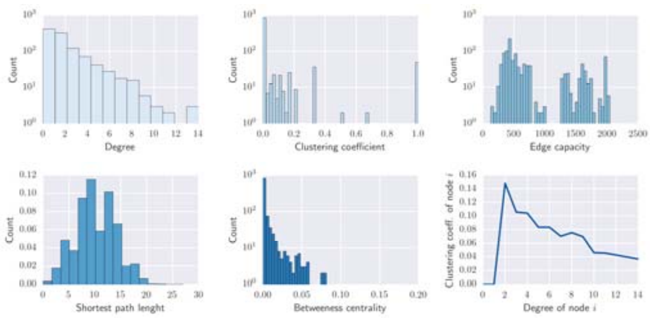


Figure 7. Actual power grid of Spain (top), Spanish power grid as connected with the epsilon-disks algorithm (center). Randomized version of Spain (bottom). The same set of nodes is used to build these three versions of the Spanish power grid.

In figure 7, we show the distribution of several network properties for the three ensembles. The actual power grid features an exponential distribution of degrees, a clustering coefficient distribution which is much more skewed towards high values than random graphs, a bimodal distribution of edge capacities, a roughly binomial distribution of shortest path length centered around 10, low centrality and clear degree-clustering correlations. The exponential degree distribution, low average centrality and long shortest path length are signatures of spatially embedded graphs. While the exponential degree distribution is expected (Albert et al., 2004; Amaral et al., 2000), the degree-clustering correlation stands in contrast to known results (Ravasz & Barabási, 2003). This could indicate that the grid is more hierarchical than it was previously believed (Ravasz & Barabási, 2003). However, given that most edges are not central (see betweenness centrality), we can conclude that the hierarchy is not dominant.



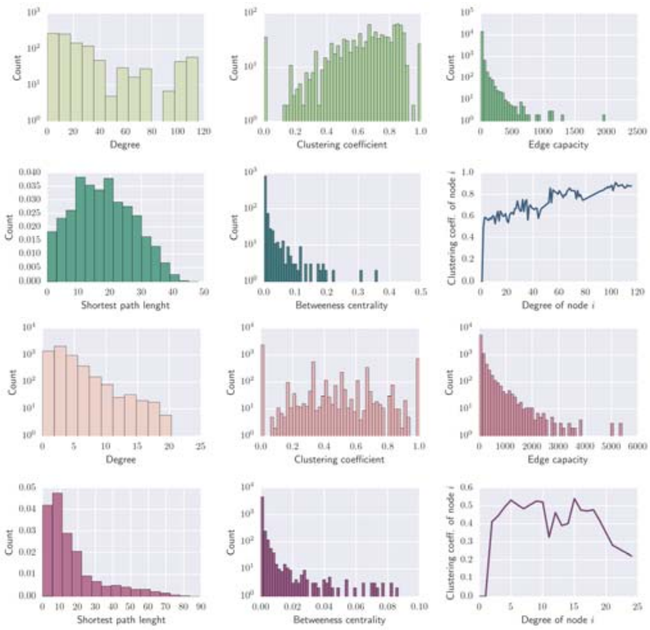


Figure 8. Structure of the power grid. Actual power grid of Spain (top). Power grid of Spain as connected with the-epsilon disks algorithm (center). Randomized version of Spain (bottom). The randomized results are averaged over 10 realization of the ensemble. The betweenness centrality and shortest path length distributions are computed for each component separately and averaged afterwards.

Our rewired version does not exhibit an exponential distribution of degrees, there is much more clustering than in the actual power grid and there is a wildly different degree-clustering correlation function. Yet, the spatial properties of the grid are very well captured by the epsilon-disk model (see, for instance, the shortest path length). The initial node placement is largely responsible for this agreement. Indeed, the actual grid is mostly connected through a series of local connections, and our model captures this idea very well. Unsurprisingly, we do not capture the bimodal distribution of edge capacities since the power flow method identifies the minimum necessary capacity. Only very peculiar node placements could naturally yield a bimodal distribution of edge capacities (e.g. a densely packed group of sinks far from a single source).

The random position ensemble is perhaps the one that yields the most surprising results. Where enforcing node placement negated our ability to model an exponential distribution of degrees, a random distribution in space recovers it. The betweenness centrality distribution is also better reproduced by the random ensemble. While the average clustering coefficient comes closer to reality, it is much higher in the random ensemble than in the actual system. In turn, this implies that the degree-clustering correlation function is not accurately reproduced. The same can be said about the edge capacities and shortest-path length. In short, while the random ensemble better captures some local and global features, it also incorporates too much clustering in the ensemble and fails to connect the network on a larger scale (long shortest path). A plausible cause for this problem lies in the capacity to demand ratio of Spain. In our dataset, the power production of Spain exceeds the average demand by a significant margin. As noted previously (See 1), this translates to a sparsely connected graph, which in turns leads to longer paths and reduced navigability. One could perhaps add so-called random links to the model to counteract the effect.

In Figure 9, we show the eigenvalue density of the Laplacian matrix of the three ensembles. The matrix plays an important role in determining power flows on power grids (see V.4), but also in capturing the important structural features of the network. Its eigenvalues are related to the degree distribution, the partition of nodes in subgroups and to the outcome of a plethora of dynamical processes on networks (Newman, 2010) . While we do not capture the spread of the spectrum of the real power grid in either of the two synthetic versions, we reproduce the proper overall shape with the random graph ensemble. This could mean that our algorithm could be used to generate artificial networks with similar dynamical properties, under proper renormalization.

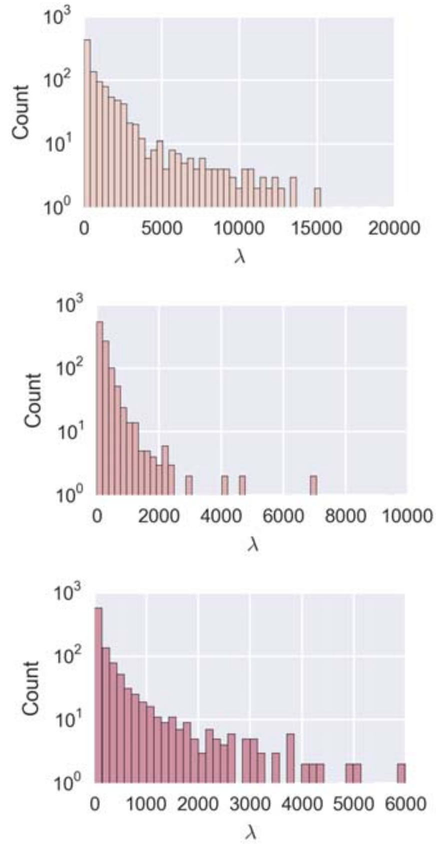


Figure 9. Spectrum of the Laplacian matrix (left) Actual power grid of Spain. (center) Power grid of Spain as connected with the epsilon-disks algorithm (right) Randomized version of Spain. The randomized results (bottom) are averaged over 10 realizations of the ensemble.

VII. Conclusions and further Work

In this paper, we proposed an efficient algorithm that generates random, realistic power grids. As highlighted by our case study of the Spanish power grid, this random graph ensemble reproduces many features of real systems as it is, and simple improvements could make it even more accurate. For example, we have seen how node placement is crucial for the reproduction of spatial properties. Moving from purely random node placement to correlated node placements could therefore improve our model. Adding long-range correlations our including preferential direction for epsilon disks growth could enhance its ability to account for the spatial aspect of power grids.

Further work could determine whether these networks can be used to evaluate different design types and therefore guide the optimization process.

VIII. References

Albert, R., Albert, I., & Nakarado, G. L. (2004). Structural vulnerability of the north american power grid. *Physical Review E, 69*(2), 025103.

Amaral, L. A., Scala, A., Barthelemy, M., & Stanley, H. E. (2000). Classes of small-world networks. *Proceedings of the National Academy of Sciences of the United States of America, 97*(21), 11149-11152.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D. (2006). Complex networks: Structure and dynamics. *Physics Reports, 424*(4), 175-308.

Feremans, C., Labbé, M., & Laporte, G. (2003). Generalized network design problems. *European Journal of Operational Research, 148*(1), 1-13.

Hébert-Dufresne, L., Allard, A., Young, J., & Dubé, L. J. (2013a). Global efficiency of local immunization on complex networks. *Scientific Reports, 3*

Hébert-Dufresne, L., Allard, A., Young, J., & Dubé, L. J. (2013b). Percolation on random networks with arbitrary k-core structure. *Physical Review E, 88*(6), 062820.

Kaltenbach, J. C., Peschon, J., & Gehrig, E. H. (1970). A mathematical optimization technique for the expansion of electric power transmission systems. *Power Apparatus and Systems, IEEE Transactions on, PAS-89*(1), 113-119.

Lumbreras, S. (2014). *Decision support methods for large-scale flexible transmission expansion planning.* Unpublished

Lumbreras, S., Ramos, A., & Sánchez-Martin, P. (2014). Offshore wind farm electrical design using a hybrid of ordinal optimization and mixed-integer programming. *Wind Energy,*

Manna, S. S., & Sen, P. (2002). Modulated scale-free network in euclidean space. *Physical Review E, 66*(6), 066114.

Moore, C., & Mertens, S. (2011). *The nature of computation* Oxford University Press.

Mossa, S., Barthelemy, M., Stanley, H. E., & Amaral, L. A. N. (2002). Truncation of power law behavior in "scale-free" network models due to information filtering. *Physical Review Letters, 88*(13), 138701.

Newman, M. (2010). *Networks: An introduction* Oxford University Press.

Raack, C. (2014). Capacitated network design. *Operations research proceedings 2013* (pp. 363-368) Springer.

Ravasz, E., & Barabási, A. (2003). Hierarchical organization in complex networks. *Physical Review E, 67*(2), 026112.

Solé, R. V., Rosas-Casals, M., Corominas-Murtra, B., & Valverde, S. (2008). Robustness of the european power grids under intentional attack. *Physical Review E, 77*(2), 026102.

Vietoris, L. (1927). Über den höheren zusammenhang kompakter räume und eine klasse von zusammenhangstreuen abbildungen. *Mathematische Annalen, 97*(1), 454-472.

Wang, Z., Scaglione, A., & Thomas, R. J. (2010). Generating statistically correct random topologies for testing smart grid communication and control networks. *Smart Grid, IEEE Transactions on, 1*(1), 28-39.

Wang, Z., Thomas, R. J., & Scaglione, A. (2008). Generating random topology power grids. *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual,* pp. 183-183.

Waxman, B. M. (1988). Routing of multipoint connections. *Selected Areas in Communications, IEEE Journal on, 6*(9), 1617-1622.