Robustness of Trans-European Gas Networks: The Hot Backbone

Rui Carvalho,[1,*] Lubos Buzna,[2,3,†] Flavio Bono,[4,‡] Eugenio Gutiérrez,[4] Wolfram Just,[1] and David Arrowsmith[1]

[1] *School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, U.K.*
[2] *ETH Zurich, UNO C 14, Universitätstrasse 41, 8092 Zurich, Switzerland*
[3] *University of Zilina, Univerzitna 8215/5, 01026 Zilina, Slovakia*
[4] *European Laboratory for Structural Assessment, Joint Research Centre, Institute for the Protection and Security of the Citizen (IPSC), Via. E. Fermi, 1 TP 480, Ispra 21027 (VA), Italy*

Here we uncover the load and vulnerability backbones of the Trans-European gas pipeline network. Combining topological data with information on inter-country flows, we estimate the global load of the network and its vulnerability to failures. To do this, we apply two complementary methods generalized from the betweenness centrality and the maximum flow. We find that the gas pipeline network has grown to satisfy a dual-purpose: on one hand, the major pipelines are crossed by a large number of shortest paths thereby increasing the efficiency of the network; on the other hand, a non-operational pipeline causes only a minimal impact on network capacity, implying that the network is error-tolerant. These findings suggest that the Trans-European gas pipeline network is robust, i.e. error-tolerant to failures of high load links.

PACS numbers: 05.10.-a, 89.75.-k, 89.75.Fb

I. INTRODUCTION

The world is going through a period when research in energy is overarching [1, 2]. Oil and gas prices are volatile because of geopolitical and financial crises. The rate of world-wide energy consumption has been accelerating, while gas resources are dwindling fast. Concerns about national supply and security of energy are on top of the political agenda, and global climate changes are now believed to be caused by the release of greenhouse gases into the atmosphere [1].

Although physicists have recently made substantial progress in the understanding of electrical power-grids [3, 4, 5, 6, 7, 8, 9, 10], surprisingly little attention has been paid to the structure of gas pipeline networks. Yet, natural gas is often the energy of choice for home heating, it is increasingly being used instead of oil for transportation [2], and gas-fired power plants are responsible for the production of approximately 20% of European electrical power [11]. Although renewable energy sources offer the best cuts in overall CO_2 emissions, the generation of electricity from natural gas instead of coal can significantly reduce the release of carbon dioxide to the atmosphere.

As demand for natural gas rises in Europe, it becomes more important to gain insights into the global transport properties of the European gas network. Unlike electricity, with virtually instantaneous transmission, the time taken for natural gas to cross Europe is measured in days. This implies that the coordination among transport operators is less critical than for power grids. Therefore, commercial interests of competing operators often lead to incomplete or incorrect network information, even at the topological level. Until now, modelling has typically been made in small systems by the respective operators, who have detailed knowledge of their own infrastructures. Nevertheless, Ukraine alone transits approximately 80% of Russian gas exports to Europe [12], suggesting the presence of a strong transportation backbone crossing several European countries.

Historically, critical infrastructure networks have evolved under the pressure to minimize local rather than global failures [13]. However, little is known on how this local optimization impacts network robustness and security of supply on a global scale. The failure of a few important links may cause major disruption to supply in the network, not because these links connect to degree hubs, but because they are part of major transportation routes that are critical to the operation of the whole network. Here we adopt the view that a *robust infrastructure network* is one which has evolved to be error-tolerant to failures of high load links. Our method is slightly different from previous work on real world critical infrastructure networks with percolation theory [6, 10, 14, 15], which assume the simultaneous loss of many unrelated network components. The absence of historical records on the simultaneous failure of a significant percentage of components in natural gas networks implies that the methods of percolation theory are of little practical relevance in our case.

Hence, our approach to the challenge of characterising the robustness of global transport on the European gas network was to characterise the *hot transportation backbone* which emerges when measuring network load and error tolerance.

II. TRANS-EUROPEAN GAS NETWORKS

We have extracted the European gas pipeline network from the Platts Natural Gas geospatial data [17].

*Electronic address: rui@maths.qmul.ac.uk
†Electronic address: lbuzna@ethz.ch
‡Electronic address: flavio.bono@jrc.it

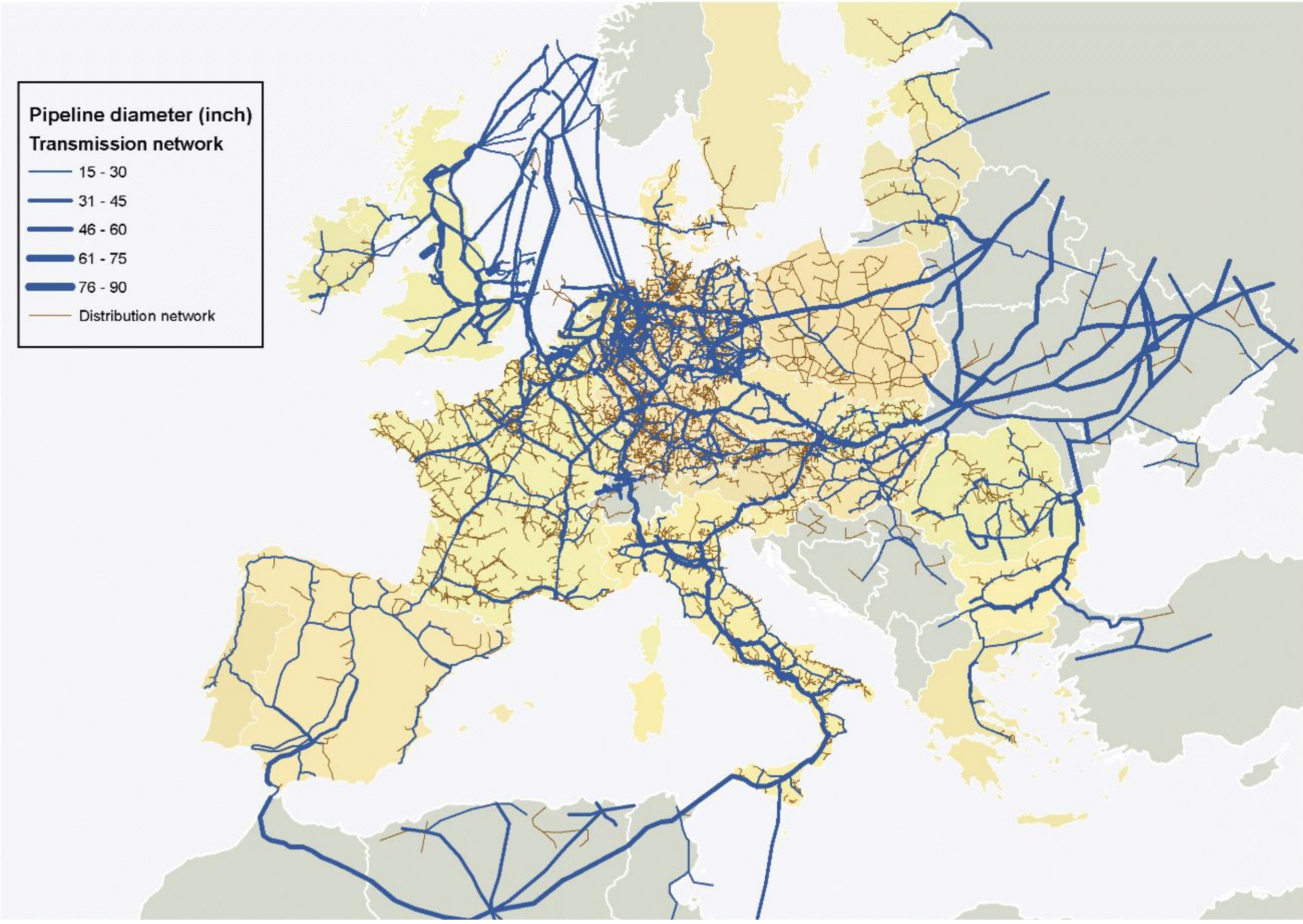

FIG. 1: (Colour online) European gas pipeline network. We show the transmission network (blue pipelines) overlaid with the distribution network (brown pipelines). Link thickness is proportional to pipeline diameter. We projected the data with the Lambert Azimuthal Equal Area projection [16].

TABLE I: Basic network statistics for the transmission and complete European gas pipeline networks. The complete network is the union of the transmission and distribution networks.

Statistics	**Gas Network (transmission)**	**Gas Network (complete network)**
Number of nodes	2207	24010
Number of edges	2696	25554
Total length [km]	119961	436289

The dataset covers all European countries (including non EU countries such as Norway and Switzerland), North Africa (main pipelines from Morocco and Tunisia), Eastern Europe (Belarus, Ukraine, Lithuania, Latvia, Estonia, Turkey) and Western Russia (see Fig. 1).

Similarly to electrical power grids, gas pipeline networks have two layers: transmission and distribution. The transmission network transports natural gas over long distances, whereas pipelines at distribution level cover urban areas and deliver gas directly to end customers. The gas pipeline transmission network is the connected component of the complete natural gas network, which is composed of all the important pipelines with diameter $D \geq 15$ inches, plus all other pipelines interconnecting major branches. We treated the resulting network as undirected due to lack of information on the direction of flows. However, network links are weighted according to pipeline diameter and length.

The European gas pipeline infrastructure is a continent-wide sparse network which crosses 38 countries, has about 2.4×10^4 nodes (compressor stations, city gate stations, liquefied natural gas (LNG) terminals and storage facilities) connected by approximately 2.5×10^4 pipelines (including urban pipelines), spanning more than 4.3×10^5 km (see Table I). The trans-European gas pipeline network is, in fact, an union of national infrastructure networks for the transport and delivery of natural gas over Europe. These networks have grown under different historical, political, economic,

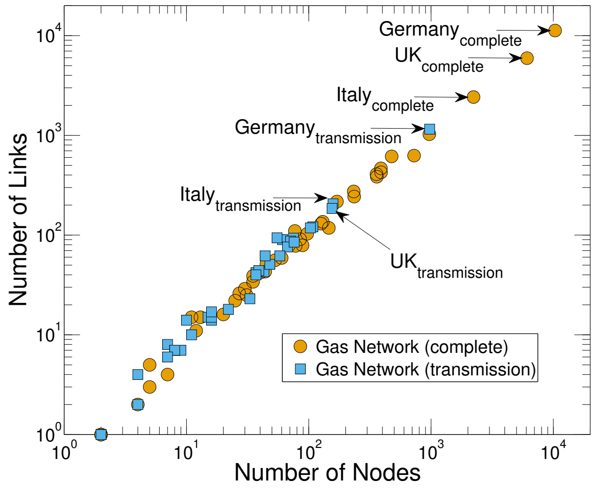

FIG. 2: (Colour online) Plot of the number L of links versus the number N of nodes for the transmission and complete (i.e. transmission and distribution) natural gas networks of the countries analysed. The three largest national transmission and complete networks are also labeled (Germany, UK and Italy).

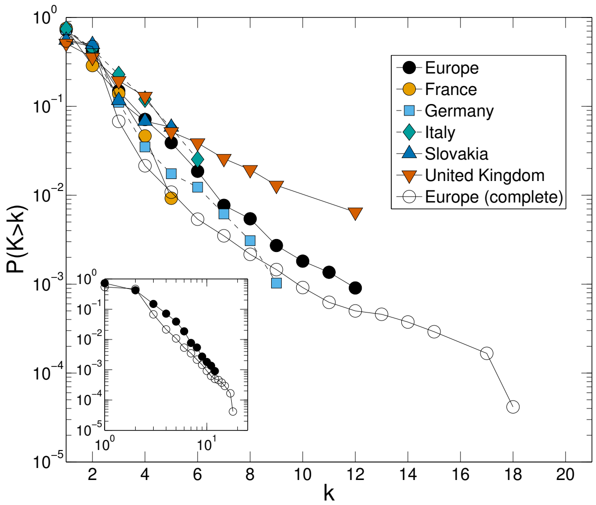

FIG. 3: (Colour online) Degree distribution of the European gas transmission and complete networks, together with national transmission gas networks larger than 100 pipelines. The inset shows the degree distribution of the European complete and transmission networks on a double-logarithmic scale, highlighting the presence of heavy tails on the degree distribution of the complete gas pipeline network.

technological and geographical constraints, and might be very different from each other from a topological point of view. To investigate similarities among the national gas networks, we first plotted in Fig. 2 the number N of nodes, versus the number L of links for each country. Figure 2 suggests that both the transmission and the complete (i.e., transmission and distribution) networks have approximately the same average degree because all points fall approximately along a straight line. Indeed, we found $\langle k_{\text{transmission}} \rangle = 2.4$ and $\langle k_{\text{full}} \rangle = 2.1$ [41]. Surprisingly, the size of the complete European national gas networks ranges over three orders of magnitude from 2 nodes (former Yugoslav Republic of Macedonia) to 10334 nodes (Germany). Further, the German transmission network is considerably larger than the Italian network. Germany has a long history of industrial usage of gas and is a major hub for imports from Russia and the North Sea [11, 18].

Given that the national networks have very different sizes, but approximately the same average degree, we looked for regularities in the probability distribution of degree of the European gas networks (see Fig. 3). In accordance with previous studies of electrical transmission networks [10, 19], the degree distribution of the natural gas transmission network decays exponentially. Unexpectedly, we found that the degree distribution of the complete gas network is heavy tailed, as can be seen in the inset of Fig. 3. This suggests that the gas network may be approximated by an exponential network at transmission level, but not when the distribution level is considered as well.

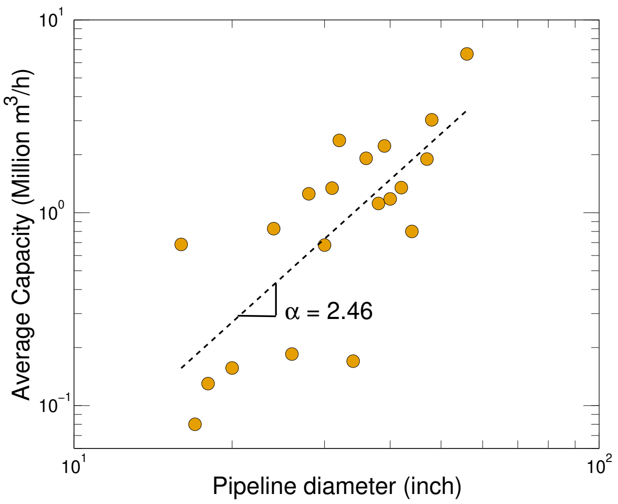

FIG. 4: (Colour online) Plot of digitized Gas Transmission Europe (GTE) pipeline capacity versus pipeline diameter on a double logarithmic scale. We digitized the European Gas Network map from GTE, and assigned the GTE capacities to pipelines in the Platts dataset. The straight line is a regression to the data, which corresponds to $C = aD^{\alpha}$ with $\alpha = 2.46$.

The July 2007 release of the Platts dataset, which we analysed, did not include information on the capacity of pipelines. To estimate pipeline capacity we compared cross-border flows based on capacity estimates whereby capacity $C \sim d^{\gamma}$ (see appendix), to reported cross-border flows extracted from the digitized Gas Transmission Europe (GTE) map [21]. Figure 4 is a plot of averaged pipeline capacity versus pipeline diameter in a double

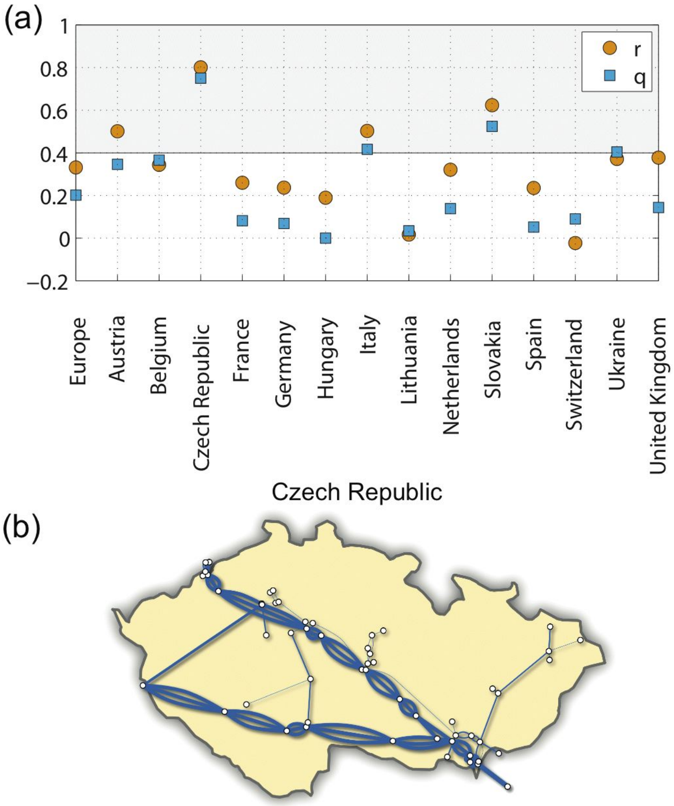

FIG. 5: (Colour online) (a) Pearson correlation coefficient, r, for the degrees of two linked nodes and the capacity of the linking pipeline; and percentage of capacity on parallel pipelines, q. We considered only countries with more than 50 pipelines. (b) Network layout for the Czech Republic. Link thickness is proportional to pipeline capacity.

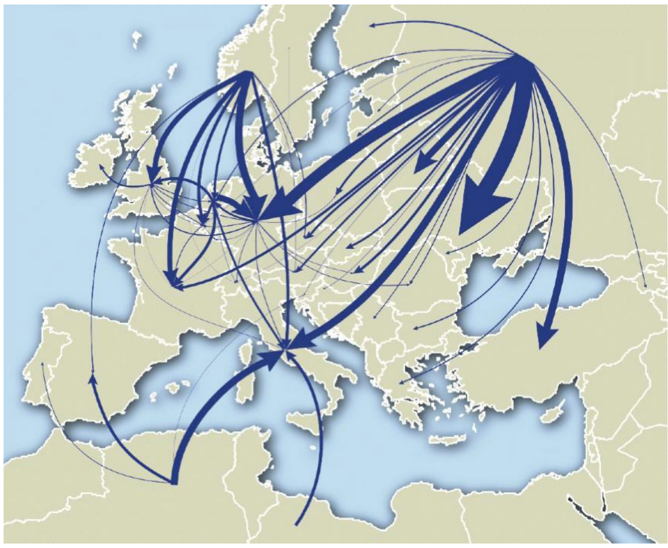

FIG. 6: (Colour online) Network of international gas trade movements by pipeline [20]. Link thickness is proportional to the annual volume of gas traded.

logarithmic scale for pipelines in the GTE dataset. We found a good match between the theoretical prediction of $C \sim d^{\gamma}$ with $\gamma \simeq 2.6$, and the capacity of major pipelines as made evident by the regression to the data $C \sim d^{\alpha}$ with $\alpha = 2.46$. Hence, we approximated the capacity of pipelines by $C \sim d^{5/2}$.

To understand the national structure of the network, we investigated the tendency of highly connected nodes to link to each other over high capacity pipelines. Figure 5(a) shows the Pearson correlation coefficient between the product of the degrees of two nodes connected by a pipeline and the capacity of the pipeline,

$$r = \frac{\sum_{e_{ij}} \left(k_i k_j - \overline{k_i k_j}\right)\left(c_{e_{ij}} - \overline{c_{e_{ij}}}\right)}{\sqrt{\left(k_i k_j - \overline{k_i k_j}\right)^2}\sqrt{\left(c_{e_{ij}} - \overline{c_{e_{ij}}}\right)^2}}. \tag{1}$$

where k_i and k_j are the degrees of the nodes at the ends of pipeline e_{ij}, and $c_{e_{ij}}$ is the capacity of the pipeline. Countries with high values of r have a gas pipeline network where degree hubs are interconnected by several parallel pipelines. Further, we plotted the percentage q of capacity on parallel pipelines for each national network. Typically, countries with high values of r also have high values of q.

Figure 5(a) shows relatively large differences among countries: Austria, the Czech Republic, Italy and Slovakia have both high values of r and q. A visual inspection of the network in these countries uncovers the presence of many parallel pipelines organized along high capacity corridors (see Fig. 5(b)). Taken together, these results suggest that some European national networks have grown structures characterised by chains of high capacity (parallel) pipelines over-bridging long distances. This has the consequence of improving local error-tolerance because the failure of one pipeline implies only a decrease in flow.

Motivated by the finding of error-tolerance in the gas pipeline networks, we then asked the question of whether there are global topological properties of the European network which could characterize network robustness.

The Platts dataset did not include volume or directionality of flows. Hence, we assessed the global structure of the European gas network under the availability of incomplete information on flows. To reduce uncertainty on flows, we combined the physical infrastructure network with the network of international natural gas trade movements by pipeline for 2007 [20] (see Fig. 6).

III. ANALYSIS OF MAN-MADE DISTRIBUTION NETWORKS WITH INCOMPLETE FLOW INFORMATION

To gain insights into the overlaid infrastructure and aggregate flow networks, we propose two complementary approaches which aim at identifying the backbones of the overlaid network. In both approaches, the flow network allows an approximate estimate of the volume of directed flows as we will detail in III A.

In the first approach, we assume that transport occurs along the shortest paths in geographical space. We search for a global backbone characterised by the presence of flow corridors where individual components were designed to sustain high loads.

In the second approach, we look at network vulnerability when single components fail. Recent studies of network vulnerability in infrastructure networks suggest that, although these networks have exponential degree distributions, under random errors or attacks the size of the percolation cluster decreases in a way which is reminiscent of scale-free networks [10, 22]. Typically, these studies presume that a large percentage of nodes or links may become non-operational simultaneously, i.e. the time scale of node or link failure is much faster than the time scale of repair. Whereas the underlying scenario of a hacker or terrorist attack on an infrastructure network causing large damage is certainly worth studying [23], the consequences of such attacks are overestimated when measuring damage by the relative size of the largest component. Here, we estimate the loss of flow when a single link is non-operational. We search for a global backbone characterised by corridors of interconnected nodes, where the removal of one single link causes a high loss of flow from source to sink nodes.

A. Generalized Betweenness Centrality

Many networks are in fact substrates, where goods, products, substances or materials flow from sinks to sources through components laid out heterogeneously in geographical space. Examples range from supply networks [24], spatial distribution networks [25] and energy networks [6] to communication networks [26]. Node and link stress in these networks is often characterised by betweenness centrality. The betweenness centrality of link e_{ij} is defined as the relative number of shortest paths between all pairs of nodes which pass through e_{ij}. The concept of betweenness centrality was originally developed to characterise the influence of nodes in social networks [27, 28] and, to our knowledge, was used for the first time in the physics literature in the context of social networks by Newman [29] and in the context of communication networks by Goh *et al.* [30][42].

Betweenness centrality is relevant in man-made networks which deliver products, substances or materials as cost constraints on these networks condition transportation to occur along shortest paths. However, nodes and links with high betweenness in spatial networks are often near the network barycentre [31], whose location is given by $\mathbf{x}_G = \sum_i x_i/N$, whereas the most important infrastructure elements are frequently along the periphery, close to either the sources or the sinks. Although flows are conditioned by a specific set of sources and sinks, traffic between these nodes may by highly heterogeneous and

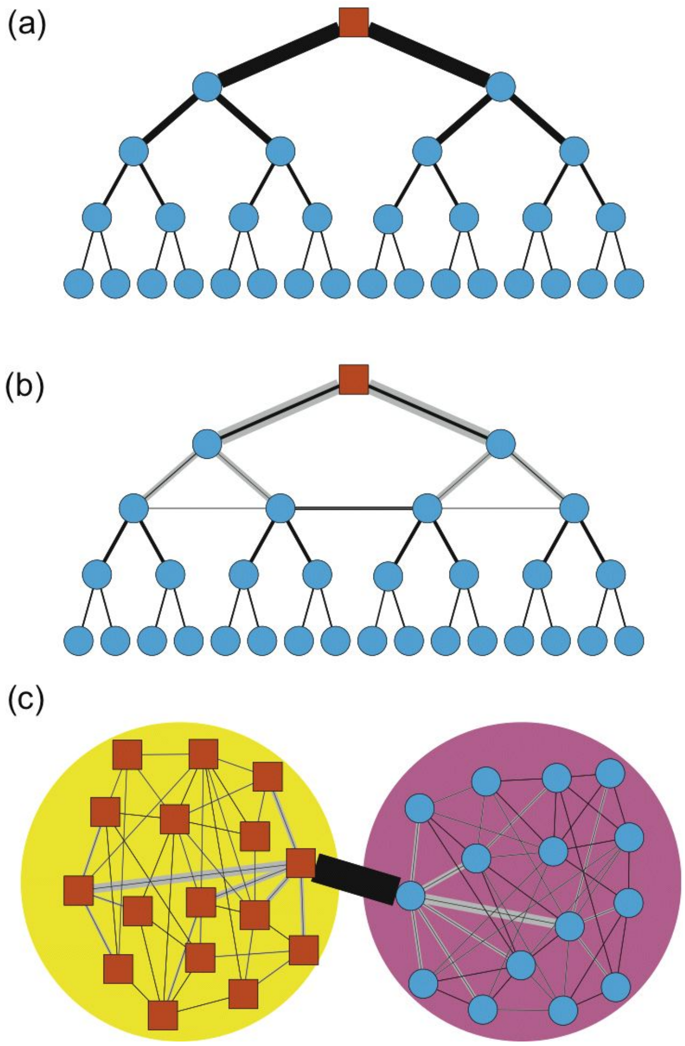

FIG. 7: (Colour online) Generalized betweenness (grey) and vitality (black) measures on: (a) a rooted tree, (b) a modified rooted tree with interconnections at a chosen level and (c) two communities connected by one link. Nodes are shaped according to their function: source nodes are squares (orange), and sink nodes are circles (blue). Both generalized betweenness and vitality depend on $T^{K,L}$, which is a constant for all examples. The smaller value of the two quantities is always drawn on the foreground so that both measures are visible.

one may have only access to aggregate transport data, but not to the detailed flows between individual sources and sinks (e.g. competition between operators may prevent the release of detailed data). Here we propose a generalization of betweenness centrality in the context of flows taking place on a substrate network, but where flow data is available only at aggregate level. We then show in the next section how the generalized betweenness centrality can help us to gain insights into the structure of trans-European gas pipeline networks.

The substrate network is often composed of sets of nodes which act like aggregate sources and sinks. The aggregation can be geographical (e.g. countries, regions

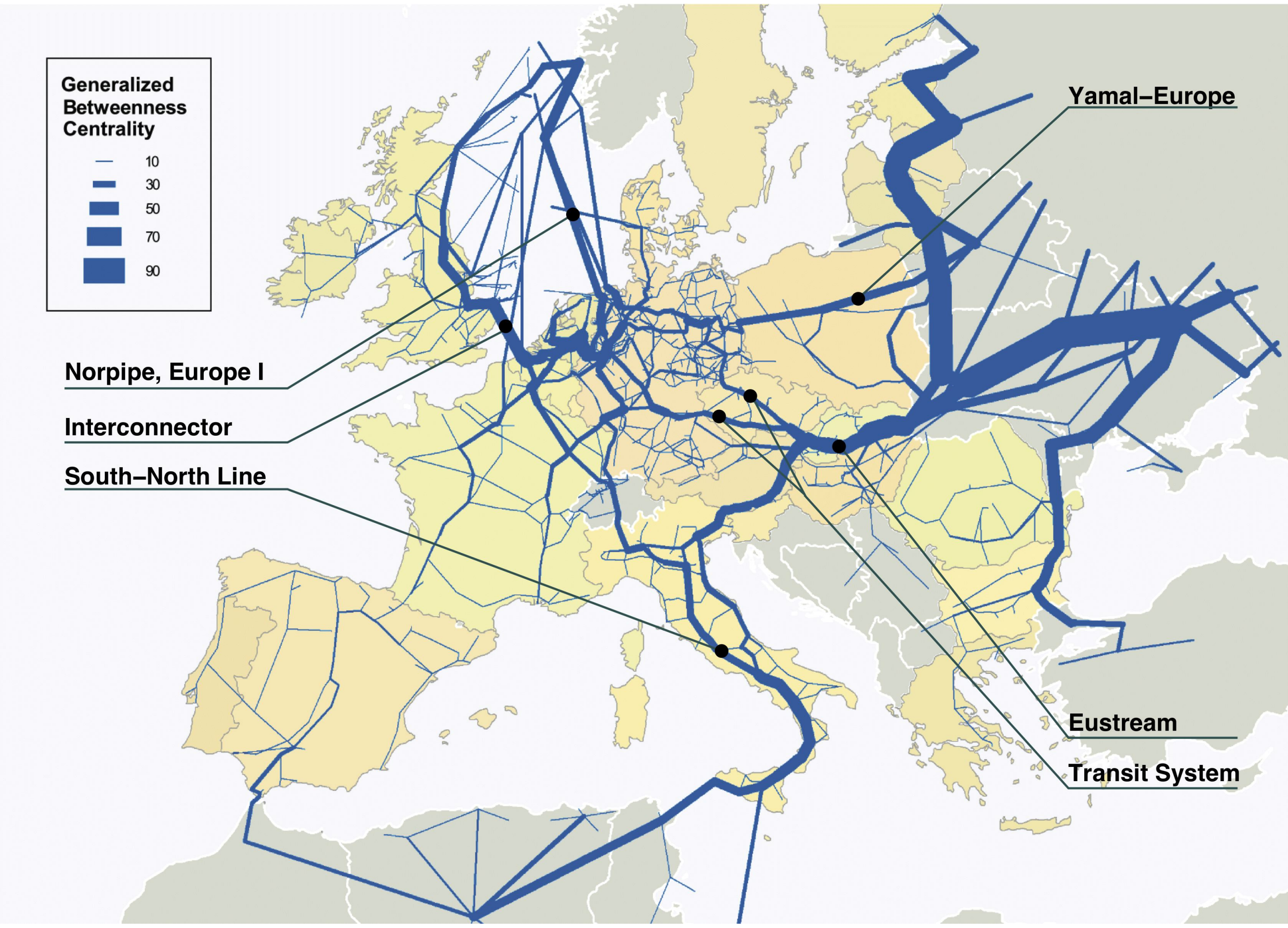

FIG. 8: (Colour online) Trans-European natural gas network. Link thickness is proportional to the generalized betweenness centrality (see Eq. 2). We labeled several major EU pipeline connections. The large difference between the generalized betweenness on these pipelines and the rest of the network suggests that the network has grown, to some extent, to transport natural gas with minimal losses along the shortest available routes.

or cities), or organizational (e.g. companies or institutions). If flow information is incomplete at microscopic scale for the substrate network and only available at aggregate level, then a possible extension of betweenness centrality for these networks is to weight the number of shortest paths between pairs of source and sink nodes by the amount of flow which is known to go through the network between aggregated pairs of sources and sinks. This can be accomplished by partitioning the substrate network into disjoint sets, $\mathcal{M} = \{\mathcal{M}_1, \cdots, \mathcal{M}_M\}$ which correspond to the nodes of the flow network, connected by links which are weighted by values of known aggregate flows.

Formally, the *generalized betweenness centrality* (generalized betweenness) of link e_{ij} of the substrate network was computed as follows. Let T be the adjacency matrix of $\mathcal{F}$, the flow network. Take each link e_{KL} in the flow network and compute the number of shortest paths from all pairs of nodes (s,t) in the substrate network $s \in K$ and $t \in L$ (where $K, L \in \mathcal{M}$ are sets of source and sink nodes, respectively), which pass through link e_{ij} of the substrate network. The contribution of that flow network link is then weighted by $T^{K,L}$, the amount of flow between K and L, and normalized by the maximum possible number of connections on the substrate network between sets K and L,

$$C_{ij} = \sum_{e_{KL} \in \mathcal{F}} \sum_{s \in K, t \in L} \frac{T^{K,L}}{|K||L|} \frac{\sigma_{st}(e_{ij})}{\sigma_{st}}. \tag{2}$$

Here, the first sum is over all links on the flow network. The second sum extends over all nodes s in the source set K and all nodes t in the sink set L, where K and L are nodes of the flow network but s and t are nodes of the substrate network. $T^{K,L}$ is the weighted adjacency matrix of flows between all K and L. $\sigma_{st}(e_{ij})$ is the number of shortest paths from source node s to sink node t which pass through link e_{ij}, and σ_{st} is the total number of shortest paths from node s to node t.

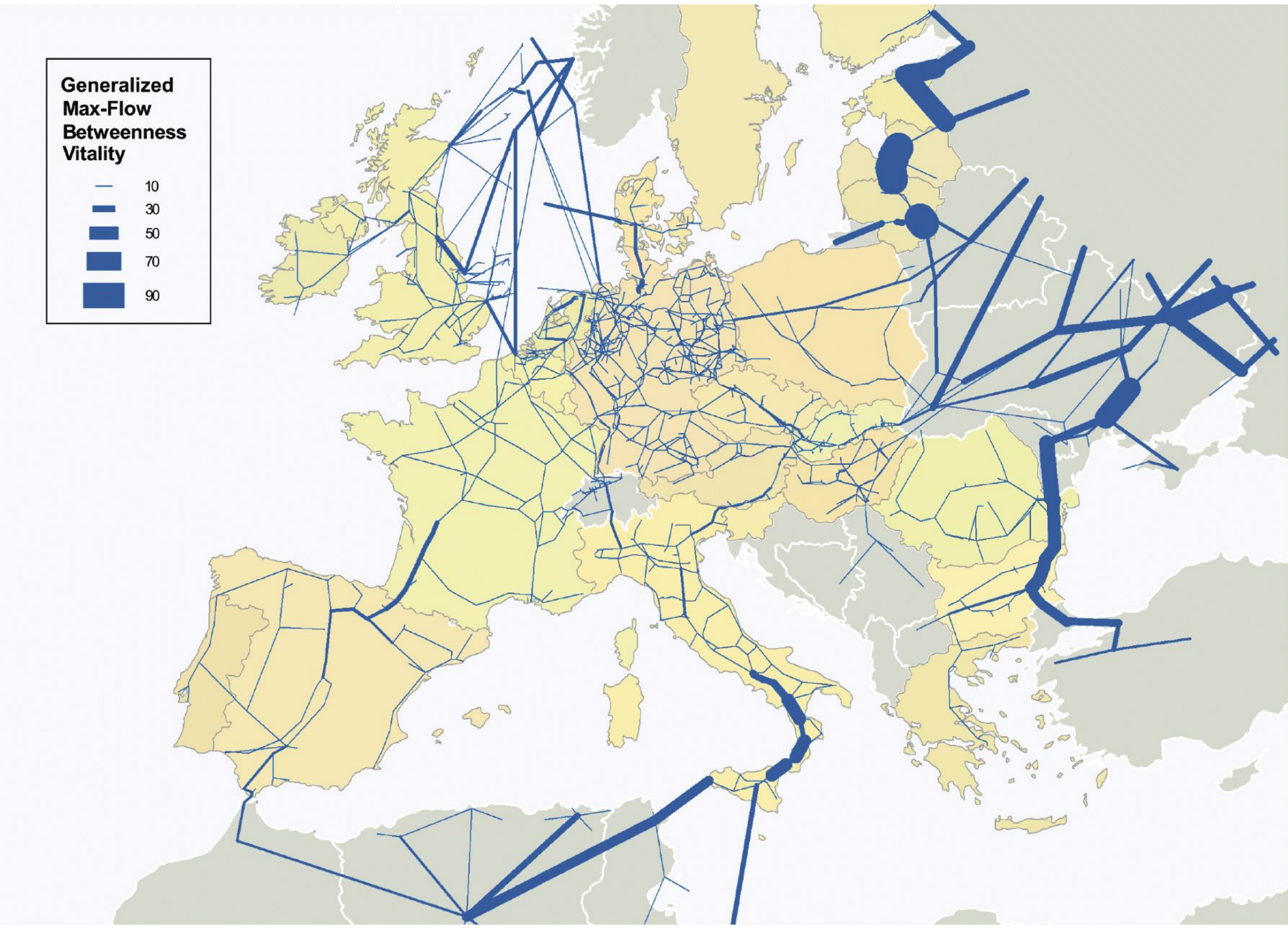

FIG. 9: (Colour online) Trans-European natural gas network. Link thickness is proportional to the generalized max-flow betweenness vitality (see Eq. 3). Pipelines close to the major sources tend to have high values of generalized vitality because this is where the network bottleneck is located. Pipelines along sparse interconnections between larger parts of the network (e.g., the Spanish-French border) also tend to have high values of generalized vitality, when compared to neighbouring pipelines.

B. Generalized Max-flow Betweenness Vitality

The maximum flow problem can be stated as follows: In a network with link capacities, we wish to send as much flow as possible between two particular nodes, a source and a sink, without exceeding the capacity of any link [32, 33]. The maximum flow from source s to sink t, f_{st}, is found with the max-flow min-cut theorem [32]. A cut is a partition of the nodes in the network into two disjoint subsets $\mathcal{S}$ and $\overline{\mathcal{S}}$, such that the source node fulfils $s \in \mathcal{S}$ and the sink node fulfils $t \in \overline{\mathcal{S}}$. The edge cut set includes all links which have an endpoint in $\mathcal{S}$ and another endpoint in $\overline{\mathcal{S}}$. If these links are removed from the network, no flow can be sent from the source to the sink. A *minimum edge cut set* is one for which the sum of capacities has the minimum possible value. The max-flow min-cut theorem states that the sum of the capacities of a minimum edge cut set is the maximum flow. In other words, the maximum flow is determined by the network's bottleneck.

The maximum flow problem and the shortest paths problem are complementary, as they capture different aspects of minimum cost flow. Shortest path problems capture link lengths, but not capacities; maximum flow problems model link capacities, but not lengths.

We are now interested in the answer to the question: How does the maximum flow between all sources and sinks change, if we remove a link e_{ij} from the network? In the absence of a detailed flow model, we calculated the flow that is lost when a link e_{ij} becomes non-operational assuming that the network is working at maximum capacity. In agreement with Eq. (2), we define the *generalized max-flow betweenness vitality* [28, 34] (generalized vitality):

$$V_{ij} = \sum_{e_{KL} \in \mathcal{F}} \sum_{s \in K, t \in L} \frac{T^{K,L}}{|K||L|} \frac{\Delta_{st}(e_{ij})}{f_{st}}, \qquad (3)$$

where the amount of flow which must go through link e_{ij} when the network is operating at maximum capacity is given by the vitality of the link [28]: $\Delta_{st}(e_{ij}) = f_{st}(G) - f_{st}(G \setminus \{e_{ij}\})$, and f_{st} is the maximal s-t flow in G. As in Eq. (2), K and L are disjoint sets of nodes on the substrate network, which we consider as the nodes of the flow network.

C. Generalized Betweenness Centrality versus Max-flow Betweenness Vitality

A close inspection of Eq. (2), generalized betweenness, and Eq. (3), generalized vitality, reveals that both measures have the physical units of gas flow given by $T_{K,L}$. Further, the relative number of shortest paths crossing a link e_{ij} is bounded by $0 \leq \frac{\sigma_{st}(e_{ij})}{\sigma_{st}} \leq 1$, and the relative quantity of flow which must go through the same link e_{ij} is also bounded by $0 \leq \frac{\Delta(e_{ij})}{f_{st}} \leq 1$. Thus, the generalized betweenness (2) and generalized vitality (3) can be compared for each link.

To examine the relationship between these two quantities, we considered three simplified illustrative networks: a rooted tree where the root is the source node and all other nodes are sinks, the same rooted tree with additional links interconnecting children nodes at a selected level, and two communities of source and sink nodes connected by one single link. We chose these particular examples because they resemble subgraphs which appear frequently on the European gas pipeline network, and thus they may help us to gain insights into the structure of the real world network.

Both the generalized betweenness and vitality have the same values on the links of trees where the root is the source node and the other nodes are sinks. To see this, consider without loss of generality the case when $T_{K,L} = 1$. Then, the generalized betweenness of a link is the proportion of sinks reachable (along shortest paths) over the link; and the generalized vitality is the proportion of sinks fed by the link. The two quantities have the same value on the links of a tree and we illustrated this in Fig. 7(a), where we drew link thickness proportional to the generalized betweenness (grey) and generalized vitality (black).

Figure 7(b) shows a modified tree network where we have connected child nodes at a chosen level. Here, the shortest paths between the root and any other node are unchanged from the example of the tree, but removing a link e_{ij} situated above the lateral interconnection does not cut all connections between the source (root) and sink nodes. As a consequence the values of generalized vitality are significantly smaller in the upper part of the graphs. Figure 7(c) shows two communities connected by one link, where source nodes are on one community and sink nodes on the other. This example is interesting for two reasons. First, the arguments used to explain why generalized betweenness and vitality take the same values on trees are also valid in this example. Second, the link connecting the two communities has a much higher value of generalized betweenness and vitality than the links inside the communities, which led us to expect that these two measures could hint at the presence of modular structure in the real world network.

IV. RESULTS

The generalized betweenness measure, defined by Eq. (2), assumes that gas is transported from sinks to sources along the shortest paths. To investigate whether this hypothesis is correct, we plotted the European gas pipeline network and drew the thickness of each pipeline proportional to the value of its generalized betweenness (see Fig. 8). We found that major loads, predicted by the generalized betweenness centrality, were on the well-know high capacity transmission interconnections such as the "Transit system" in the Czech Republic, the "Eustream" in Slovakia, the "Yammal-Europe" crossing Belarus and Poland, the "Interconnector" connecting the UK with Belgium or the "Trans-Mediterranean" pipeline linking Algeria to mainland Italy through Tunisia and Sicily. The dramatic difference between the values of generalized betweenness of all the major European pipelines and the rest of the network suggests that the network has grown to some extent to transport natural gas with minimal losses along the shortest available routes between the sources and end consumers. These major pipelines are the transportation backbone of the European natural gas network.

During the winter season, cross-border pipelines are used close to their full capacity [21]. In this situation, the generalized vitality of a pipeline, Eq. (3), can be interpreted as the network capacity drop, or the amount of flow that cannot be delivered, if that pipeline becomes non-operational.

The obvious drawback of generalized vitality is that it takes into account the overall existing network capacity without considering the length of paths. Conversely, the generalized betweenness considers the length of shortest paths, but not the capacity of pipelines. Since we assess the network from two complementary view points, we expect that the results will allows us to get a more complete picture of the general properties of the European gas pipeline network.

Figure 9 shows the values of generalized vitality in the European gas pipeline network. We found several relatively isolated segments with high generalized vitality located in Eastern Europe, close to the Spanish-French border, as well as on the south of Italy. The high values of generalized vitality in Eastern Europe can be explained by two factors. On one hand, the generalized vitality of pipelines close to the sources is higher than elsewhere simply because these pipelines are the bottleneck of the network. On the other hand, our approximation that a directed link in the flow matrix implies gas flowing from all nodes in the source to all nodes in the sink countries was clearly coarse-grained for flows between Russia and the Baltic states, as it would imply that pipelines in southern Russia would also supply the Baltic countries. This highlights boundary effects on the calculation of betweenness vitality, as the dataset excluded most ot the Russian gas pipeline network. The case of the Spanish-French border was different, though. The link with high vitality separates the Iberian Peninsula from the rest of mainland Europe. If this link was to be cut, then Portugal and Spain would only be linked to the pipeline network through Morocco. Finally, the south of Italy highlights an interesting example of two communities (Europe and North Africa) separated by the Trans-Mediterranean pipeline, which is reminiscent of the example in Fig. 7(c).

Perhaps surprisingly, we found that the generalized vitality is more or less homogeneous in most of mainland Europe. This result suggests that the EU gas pipeline network has grown to be error-tolerant and robust to the loss of single links.

Distribution networks originate from the need for effective connectivity among sources and sinks [35, 36]. For example, a spanning tree is highly efficient as it transports goods from sinks to sources in a way that shortens the total length of the network, thereby increasing its efficiency and viability. If the European gas pipeline network had been built as a spanning tree, its links would have very similar values of generalized betweenness and max-flow vitality (see Fig. 7).

The values of generalized betweenness are considerably higher than the corresponding values of generalized vitality for the most important pipelines in the European Union. In other words, the major pipelines are crossed by many shortest paths, but a non-operational pipeline causes only a minor capacity drop in the network. This dramatic contrast between the two measures reveals a *hot backbone* [37] showing that the Trans-European gas pipeline network is robust, i.e. error-tolerant to failures of high load links.

V. CONCLUSIONS

We analysed the Trans-European gas pipeline network from a topological point of view. We found that the European national gas pipeline networks have approximately the same value of average node degree, even if their sizes vary over three orders of magnitude. Like the electrical power grid, the degree distribution of the European gas transmission network decays exponentially. Unexpectedly, the degree distribution of the complete (transmission and distribution) gas pipeline network is heavy tailed. In some countries which are crucial for the transit of gas in Europe (Austria, the Czech Republic, Italy and Slovakia), we found that the main gas pipelines are organized along high capacity corridors, where capacity is split among two or more pipelines which run in parallel, over-bridging long distances. This implies that the network is error-tolerant because the failure of one pipeline causes only a decrease in flow. Motivated by the finding of error-tolerance in national networks, we then addressed the problem of capturing the topological structure of the European gas network.

At a global scale, the growth of the European gas pipeline network has been determined by two competing mechanisms. First, the network has grown under cost and efficiency constraints to minimize the length of transport routes and maximize transported volumes. Second, the network has developed error-tolerance by adding redundant links. The combination of the two mechanisms guarantees that the European gas pipeline network is robust, i.e. error-tolerant to failures of high load links. To reveal the network robustness, we analysed two measures, the generalized betweenness and generalized vitality, which highlight global backbones of transport efficiency and error-tolerance, respectively. Finally, we proposed that the hot backbone of the network is the skeleton of major transport routes where the network is robust, in other words, where values of generalized betweenness are high and values of generalized vitality are low. Our method is of potential interest as it provides a detailed geographical analysis of engineered distribution networks.

Further research in continent-wide distribution networks could proceed along several directions. Planned and under-construction pipelines may change the robustness of the network, in particular within their geographical vicinity. Liquefied natural gas (LNG), which is nowadays transported at low cost between continents, is increasingly supplying the pipeline network. The combined effect of LNG and storage facilities throughout the European coastline has the potential to reduce the dependency on one single exporting country, such as Russia. Last, but not least, the dispute between Russia and Ukraine in January 2009 has brought supply security to the top of the European political agenda and highlighted how the European gas network is robust to engineering failures, yet fragile to geopolitical crises.

Acknowledgments

We wish to thank Dirk Helbing, Sergi Lozano, Amin Seyyed Mazloumian and Russel Pride for valuable comments, and gratefully acknowledge the support of EU projects MANMADE (No. 043363) and IRRIIS (No. 027568).

APPENDIX: PIPELINE CAPACITY

The capacity of a pipeline can be schematically derived as follows: It is known that the flow of an incompressible viscous fluid in a circular pipe can be described in the laminar regime (with a parabolic velocity profile) by the Hagen-Poiseuille equation [38], which states that the volume of fluid passing per unit time is

$$dV/dt = \pi \Delta p r^4 / (8 \eta l) \,, \qquad \text{(A.1)}$$

where Δp is the pressure difference between the two ends of the pipeline, l is the length of the pipeline (thus, $-\Delta p/l$ is the pressure gradient), η is the dynamic viscosity, and r is the radius of the pipeline. However, the gas network operates in the turbulent regime, and the Hagen-Poiseuille equation is no longer valid. Therefore, we apply the Darcy–Weisbach equation for pipeline head loss, h_f. This is a phenomenological equation which describes the loss of energy due to friction within the pipeline and is valid in the laminar and turbulent regimes [39]:

$$h_f = f\frac{lv^2}{2gd}, \tag{A.2}$$

where f is called the Darcy friction factor, d is the pipeline diameter, v is the average velocity, and g is the acceleration of gravity.

Equation (A.2) can be written as a function of the volumetric flow rate $dV/dt = \pi\left(\frac{d}{2}\right)^2 v$ (which is the capacity of the pipeline [40]), as

$$h_f = \frac{16 f\, l\, (dV/dt)^2}{\pi^2 d^5}. \tag{A.3}$$

In general, the friction factor f and the pipeline loss h_f depend on the pipeline diameter d, so the capacity of the pipeline is given by $C = dV/dt = \frac{\pi}{4}\left(\frac{h_f}{f\,l}\right)^{1/2} d^{5/2} \sim d^\gamma$, where typically $\gamma \simeq 2.6$ for gas pipelines [40].

[1] http://www.aps.org/energyefficiencyreport/.
[2] *The greener grid*, Nature **454**, 551 (2008).
[3] D. J. Watts and S. H. Strogatz, Nature **393**, 440 (1998).
[4] M. L. Sachtjen, B. A. Carreras, and V. E. Lynch, Physical Review E **61**, 4877 (2000).
[5] A. E. Motter and Y.-C. Lai, Physical Review E **66**, 065102 (2002).
[6] R. Albert, I. Albert, and G. L. Nakarado, Physical Review E **69**, 025103 (2004).
[7] R. Kinney, P. Crucitti, R. Albert, and V. Latora, The European Physical Journal B **46**, 101 (2005).
[8] V. Rosato, S. Bologna, and F. Tiriticco, Electric Power Systems Research **77**, 99 (2007).
[9] I. Simonsen, L. Buzna, K. Peters, S. Bornholdt, and D. Helbing, Physical Review Letters **100**, 218701 (2008).
[10] R. V. Sole, M. Rosas-Casals, B. Corominas-Murtra, and S. Valverde, Physical Review E **77**, 026102 (2008).
[11] http://epp.eurostat.ec.europa.eu/.
[12] M. Reymond, Energy Policy **35**, 4169 (2007).
[13] D. H. Kim and A. E. Motter, in *Workshop on Complex Networks - from Biology to Information Technology* (Iop Publishing Ltd, Pula, Italy, 2007).
[14] P. Crucitti, V. Latora, and M. Marchiori, Physical Review E **69**, 045104 (2004).
[15] A. E. Motter, Physical Review Letters **93**, 098701 (2004).
[16] A. H. Robinson, J. L. Morrison, P. C. Muehrcke, A. J. Kimerling, and S. C. Guptill, *Elements of Cartography* (John Wiley and Sons, 1995).
[17] http://www.platts.com.
[18] http://www.eia.doe.gov/.
[19] L. A. N. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley, Proceedings of the National Academy of Sciences of the United States of America **97**, 11149 (2000).
[20] *BP statistical review of world energy* (2008).
[21] http://www.gie.eu.com.
[22] M. Rosas-Casals, S. Valverde, and R. V. Sole, International Journal of Bifurcation and Chaos **17**, 2465 (2007).
[23] T. Williams, Pipeline & Gas Journal pp. 56–58 (November 2007).
[24] D. Helbing, S. Lammer, U. Witt, and T. Brenner, Physical Review E **70**, 056118 (2004).
[25] M. T. Gastner and M. E. J. Newman, Physical Review E **74**, 016117 (2006).
[26] S. Mukherjee and N. Gupte, Physical Review E **77**, 036121 (2008).
[27] L. C. Freeman, Sociometry **40**, 35 (1977).
[28] D. Koschützki, K. A. Lehmann, L. Peeters, S. Richter, D. Tenfelde-Podehl, and O. Zlotowski, in *Network Analysis*, edited by U. Brandes and T. Erlebach (Springer, 2005).
[29] M. E. J. Newman, Physical Review E **64**, 016132 (2001).
[30] K. I. Goh, B. Kahng, and D. Kim, Physical Review Letters **87**, 278701 (2001).
[31] A. Barrat, M. Barthelemy, and A. Vespignani, Journal of Statistical Mechanics-Theory and Experiment, P05003 (2005).
[32] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications* (Prentice Hall, 1993).
[33] M. E. J. Newman, Physical Review E **70**, 056131 (2004).
[34] V. Latora and M. Marchiori, Physical Review E **71**, 015103 (2005).
[35] J. R. Banavar, A. Maritan, and A. Rinaldo, Nature **399**, 130 (1999).
[36] M. T. Gastner and M. E. J. Newman, Journal of Statistical Mechanics-Theory and Experiment, P01015 (2006).
[37] E. Almaas, B. Kovacs, T. Vicsek, Z. N. Oltvai, and A. L. Barabasi, Nature **427**, 839 (2004).
[38] L. D. Landau and E. Lifshitz, *Fluid Mechanics: Volume 6 (Course of Theoretical Physics)* (Butterworth-Heinemann, 1987).
[39] F. M. White, *Fluid Mechanics* (WCB/McGraw-Hill, 1999).
[40] Gas Transmission Europe, *Definition of available capacities at interconnection points in liberalized markets (Tech. Rep. 03CA043)* (2003).
[41] Sole *et al.* found similar results for European power grids [10]
[42] Goh *et al.* define load similarly to betweenness centrality, but the two measures are different in the presence of more than one shortest path between two nodes.