Message Passing for Distributed Optimisation of Power Allocation with Renewable Resources

Elizabeth Harrison
School of Engineering and Applied Sciences
Aston University
Birmingham, UK
Email: harrisem@aston.ac.uk

David Saad
School of Engineering and Applied Sciences
Aston University
Birmingham, UK
Email: d.saad@aston.ac.uk

K. Y. Michael Wong
Department of Physics
The Hong Kong University of Science and Technology
Hong Kong, China
Email: phkywong@ust.hk

***Abstract*—The increase in renewable energy generators introduced into the electricity grid is putting pressure on its stability and management as predictions of renewable energy sources cannot be accurate or fully controlled. This, with the additional pressure of fluctuations in demand, presents a problem more complex than the current methods of controlling electricity distribution were designed for. A global approximate and distributed optimisation method for power allocation that accommodates uncertainties and volatility is suggested and analysed. It is based on a probabilistic method known as message passing [1], which has deep links to statistical physics methodology. This principled method of optimisation is based on local calculations and inherently accommodates uncertainties; it is of modest computational complexity and provides good approximate solutions. We consider uncertainty and fluctuations drawn from a Gaussian distribution and incorporate them into the message-passing algorithm. We see the effect that increasing uncertainty has on the transmission cost and how the placement of volatile nodes within a grid, such as renewable generators or consumers, effects it.**

***Keywords*—Message passing, optimisation, power flow, distribution, renewable energy, uncertainty, electricity, networks**

I. Introduction

Due to the environmental effects and the non renewable nature of mainstream power sources such as coal and oil, there have been legislations and international agreements to curb CO_2 emission. Electricity producers must incorporate more renewable energy sources into the power grid and reduce their reliance on non-renewable energy sources. Technologies to harness renewable energy sources such as wind and solar are advanced enough that they can be used in electrical power grids. However, renewable sources are uncertain and fluctuating, reducing the reliability of power grids making them increasingly hard to manage. Power grids in the US, Europe and Asia have all experienced large blackouts [2] from not being able to fully control or manage the volatility of power grids, a trend that is bound to increase in years to come. This paper proposes a principled, probabilistic and computationally efficient Message Passing (MP) algorithm which can inherently consider the uncertainties and fluctuations of generators and consumers within a grid at the resource dispatch stage.

978-1-4673-8473-5/16$31.00 2016 IEEE

The current method of optimal power flow (OPF) [3] works by alternately repeating two steps until the overall optimum is found. The first step uses the Newton Raphson method to estimate initial values for distributing the electricity in order to achieve zero excess, enforcing Kirchoff's law, the second finds a better solution which reduces power loss and generating costs, while ensuring that all of the constraints are met, using gradient-descent or similar techniques. OPF is capable of considering small scale fluctuations from consumers, but not the more substantial fluctuations from renewable sources.

The power grid requires a dispatch method that is more robust to accidental damage or deliberate attack and can effectively consider bandwidth to eliminate tripping power lines. Fluctuations and uncertainties in generators and consumers need to be better incorporated to accommodate the growth in volatile components of the grid to ensure that power grids run smoothly and reliably. A recent report [4] shows that the UK is relying less on renewable energies than other European countries, arguably, due to the difficulty in incorporating them into the existing grid.

There are alternative methods of optimal dispatch; Chance Constrained OPF develops upon the already used OPF method but changes the constraints so that they are satisfied within a certain probability [5], [6], Interior Point method uses matrices to satisfy the constraints and uses a heuristic strategy of predictor-corrector to minimise the matrix equations [7], [8], but these methods only consider uncertainty heuristically and superficially without incorporating their probabilistic nature within the algorithmic framework.

As a probabilistic algorithm MP [1] offers a principled, distributed optimisation method which passes messages locally in order to obtain a global optimum; this allows the computational complexity to increase as $\mathcal{O}(N)$ where N is the number of vertices in the grid. Its probabilistic properties allow it to consider generator uncertainties inherently within the calculations.

In Section II we describe the MP method and explain how it can accommodate uncertainties; in Section III we show

results validating the use of these methods on synthetically generated systems. In Section IV we discusses the advantages, disadvantages and possible extensions of the MP algorithm.

II. Methodology

A. Message Passing

Resource allocation is the distribution of a resource in a network according to some objective function and constraints. An efficient resource allocation method based on statistical physics methodology was presented in [1], both macroscopically and microscopically. The framework is based on assigning a cost to each state of the system with the aim of minimising it to obtain the *ground state* or the state of the lowest cost for a given system. This is carried out by minimising the free energy F that accounts for both the related cost of a state and the multiplicity of states of the same cost. Links between the macroscopic state of a system and algorithms aimed at finding the state of each variable microscopically for a given system realisation have been explained in the literature [9], [10] and utilised in a number of hard computational tasks, from error-correcting codes [11] to graph colouring [12] and routing [13].

In a resource allocation system with N nodes, each connected to c other nodes and with a capacity Λ_j, representing either power generation (positive values) or demand (negative values), we want to satisfy the demands, i.e. render the final capacities of all nodes to positive or zero, while minimising an objective function. To do this we dispatch resources from node j to i along an edge, $y_{ij} = -y_{ji}$. Satisfying the constraint requires that every node's capacity plus any resource dispatched from neighbouring nodes to be greater than or equal to zero. This is written mathematically as:

$$\sum_{i=1}^{N} \mathcal{A}_{ij} y_{ij} + \Lambda_j \geq 0 \tag{1}$$

where $\mathcal{A}_{ij}$ is an adjacency matrix; if the node i and j are connected $\mathcal{A}_{ij} = 1$, or $\mathcal{A}_{ij} = 0$ otherwise. The summation is over all nodes in the network.

As well as enforcing the capacity of all nodes to be non-negative, the algorithm minimises the transportation costs, which in this paper we will consider to be quadratic; $\phi(y_{ij}) = \frac{y_{ij}^2}{2}$ but can assume other functional forms as well. We can write the energy function of the system as $E = \sum_{(ij)} \mathcal{A}_{ij}\phi(y_{ij})$, where the summation over (ij) is all pairs of nodes in the network.

The partition function Z is a normalization constant, subject to the constraint, and can be written as:

$$\begin{aligned} Z \;=\; & \left(\prod_{(ij)} \int \mathrm{d}y_{ij}\right) \prod_j \Theta\Big(\sum_i \mathcal{A}_{ij} y_{ij} + \Lambda_j\Big) \\ & \times e^{-\frac{1}{T}\sum_{(ij)} \mathcal{A}_{ij}\phi(y_{ij})} \end{aligned} \tag{2}$$

where T is the temperature of the system, a parameter that determines how strictly the optimisation is carried out, and $\Theta(*)$ is the step function; it returns 0 if $*$ is negative, and 1 otherwise. The free energy to be minimised is written as $F = -T \ln Z$. However, it is generally difficult to minimise the free energy due to the computational cost involved and one resorts to approximations, some more principled than others.

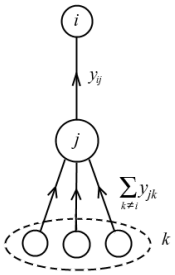

Fig. 1. Tree of node j with ancestor i and descendants k.

The Bethe approximation is based on the assumption that the network is sparsely connected, i.e., that the number of connections per node is much smaller than the systems size N. As the probability of finding loops is small, the network can be approximated by considering just one node and its neighbours at a time, assuming a locally tree-like topology and ignoring long-range correlations. This gives rise to *local* passing of messages, conditional probabilities, since once the node to be studied is removed (hence the term *cavity* method in the physics literature) neighbouring nodes are weakly correlated with one another and are considered probabilistically independent. Considering a node j, we set aside one of its neighbours to be an ancestor, i, and the remaining $c-1$ neighbours to be descendants, k. Using this we can now write the free energy of node j according to the free energy of its descendants, $F(y_{jk}|\mathbf{T}_k)$, this is sent to node i. $\mathbf{T}_k$ represents the tree of node k and its descendants.

$$\begin{aligned} F(y|\mathbf{T}) = -T \ln \Bigg\{ & \prod_{k=1}^{c-1} \left(\int \mathrm{d}y_{jk} \right) \Theta\left(\sum_{k=1}^{c-1} y_{jk} - y_{ij} + \Lambda_j \right) \\ & \times \exp\left[-\frac{1}{T} \sum_{k=1}^{c-1} \big(F(y_{jk}|\mathbf{T}_k) + \phi(y_{jk}) \big) \right] \Bigg\} \end{aligned} \tag{3}$$

The free energy can be split into two parts, $F(y|\mathbf{T}) = N_{\mathbf{T}} F_{av} + F_V(y|T)$ where $N_{\mathbf{T}}$ is the number of nodes in tree $\mathbf{T}$, F_{av} is the average free energy per node and $F_V(y|\mathbf{T})$ is the Vertex Free Energy (VFE). When a node is added to a tree this changes its free energy. The addition of a node means that the VFE is found when the average free energy of the remainder of the tree is subtracted from the change in free energy. Rearranging this gives us the recursion relation in terms of VFE's (for the derivation see [1]):

$$\begin{aligned} F_V(y|\mathbf{T}) = -T \ln \Bigg\{ & \prod_{k=1}^{c-1} \big(\int \mathrm{d}y_{jk} \big) \Theta\left(\sum_{k=1}^{c-1} y_{jk} - y_{ij} + \Lambda_j \right) \\ & \times \exp\left[-\frac{1}{T} \sum_{k=1}^{c-1} \big(F_V(y_{jk}|\mathbf{T}_k) + \phi(y_{jk}) \big) \right] \Bigg\} \\ & - F_{av}. \end{aligned} \tag{4}$$

In the case of strict optimisation, or when temperature is set to zero:

$$F_V(y_{ij}|\mathbf{T}) = \min\left[\sum_{k=1}^{c-1}\left(F_V(y_{jk}|\mathbf{T}_k)+\phi(y_{jk})\right)\right] - F_{av}$$
$$\text{s.t.} \sum_{k=1}^{c-1} y_{jk} - y_{ij} + \Lambda_j \geq 0. \tag{5}$$

This is the message to be passed between nodes in order to find the global optimal. However, in the case of continuous variables this is a continuous function and it is difficult to send it using MP. The alternative is to represent the message-passing function by a small number of parameters; in [1] it was assumed that the function can be accurately represented by its first and second derivatives. Instead of sending the full function the messages consisted of $[A_{jk}, B_{jk}] = \left[\frac{\partial F_V}{\partial y_{jk}}, \frac{\partial^2 F_V}{\partial y_{jk}^2}\right]$. We calculate the A and B messages using the Taylor expansion of the VFE (4) around some adjustment of y_{jk}, ε_{jk} and with a Lagrange multiplier μ_{ij} to consider the constraint:

$$F_{ij} = \sum_{k\neq i}\left[(A_{jk}+\phi'_{jk})\varepsilon_{jk} + \frac{1}{2}(B_{jk}+\phi''_{jk})\varepsilon_{jk}^2\right]$$
$$+\mu_{ij}\Big(\sum_{k\neq i}(y_{jk}+\varepsilon_{jk}) - y_{ij} + \Lambda_j\Big). \tag{6}$$

Optimising this function gives rise to the messages:

$$A_{ij} \leftarrow \max\left[0, \frac{y_{ij} - \sum_{k\neq i}(y_{jk} - \frac{A_{jk}+\phi'_{jk}}{B_{jk}+\phi''_{jk}}) - \Lambda_j}{\sum_{k\neq i}\frac{1}{B_{jk}+\phi''_{jk}}}\right], \tag{7}$$

$$B_{ij} \leftarrow \frac{\Theta(A_{ij}-\epsilon)}{\sum_{k\neq i}\frac{1}{B_{jk}+\phi''_{jk}}}, \tag{8}$$

where ϵ represents a small positive value, and a message to node i indicating its new current according to the adjustments ε_{ij} is written as:

$$y_{ij} \leftarrow \frac{B_{ij}y_{ij} - A_{ij} - B_{ji}y_{ji} + A_{ji} - \phi'_{ij} + \phi''_{ij}y_{ij}}{B_{ij}+B_{ji}+\phi''_{ij}}. \tag{9}$$

The algorithm chooses a node j and ancestor i at random, it uses the messages from descendant nodes k, and according to these sends the values of A_{ij}, B_{ij} and y_{ij} to node i. This is done iteratively for all nodes in the network until convergence, by which time the results of the edges y_{ij} give an approximate optimal solution. Figure 2 shows an example of this resource allocation on a small network.

B. Message passing for optimal dispatch in a power grid

In this paper we assume a simple network. Power stations, sub-stations and consumers (this paper concentrates on the transmission and not the distribution stage of a power grid so by a consumer we mean the sub-station leading to a group of consumers) are represented as nodes with a starting value

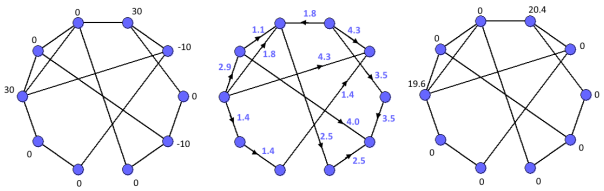

Fig. 2. An example of the algorithm applied to a 10-Node network. Left: The original capacities Λ at each node. Centre: the calculated values of the resource passed along each edge y. Right: The nodes' capacities after receiving or giving resource, all non-negative.

Λ_j indicating the power at the node (the value of Λ_j is positive for generators and negative for consumers). Power lines are represented by edges and any two nodes connected by a power line are described as neighbours. Excess at an uncertain node indicates the remaining reserve power at a node (we assume that satisfying all nodes is more important than unused capacity [14]).

Here, we extend the MP algorithm [1] to include volatile nodes by minimising the *expected* VFE considering the probability distribution that represents the volatility of each node. Nodes with a definite power generation or consumption will have a dirac delta distribution at $\bar{\Lambda}_j$, and any uncertain nodes will be assumed to have a Gaussian distribution with mean, $\bar{\Lambda}_j$, and variance, σ_j^2. Carrying out averages with respect to the distributions is non trivial and we will consider two cases: annealed and quenched averaging.

1) Annealed approximation: Calculating the expected free energy is difficult as $\langle F\rangle = -T\langle \ln Z\rangle$ (angled brackets $\langle\rangle$ denote an average over the capacity distribution) and averages over a logarithm are non-trivial. However, averaging over the partition function first, $\langle Z\rangle$ and then taking the logarithm is easier. This is called annealed approximation, $\langle F\rangle_{annealed} = -T\ln\langle Z\rangle$ and it is a good approximation in many cases.

As the uncertain nodes are now drawn from a Gaussian distribution which has an infinite tail, the hard constraint that each node must be non-negative is unsatisfiable, since very low capacity values are bound to exist in the large system limit. Instead, we require that the probability of the total resource at a node being smaller than zero, after the program converges, should be less than a predetermined value p. This gives a constraint of:

$$\left\langle \Theta\left[\sum_{k\neq i}\mathcal{A}_{jk}(y_{jk}+\varepsilon_{jk}) - y_{ij} + \Lambda_j\right]\right\rangle_{\Lambda} =$$
$$\frac{1}{2}\text{erfc}\left(\frac{-\sum_{k\neq i}\mathcal{A}_{jk}(y_{jk}+\varepsilon_{jk}) + y_{ij} - \bar{\Lambda}_j}{\sqrt{2\sigma_j^2}}\right) > 1-p, \tag{10}$$

As the only dependence on Λ_j in (6) is found in the constraint, this is the only place averaging needs to be considered. The difference these calculations give is in the A messages:

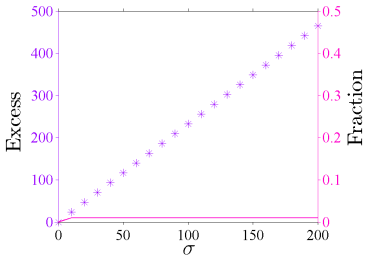

Fig. 3. The excess capacity at a node as a function of the standard deviation, σ using the annealed approximation (purple), presented jointly with the probability of the node being unsatisfied (pink). In a 20-node network with one uncertain consumer with mean $\bar{\Lambda} = -200$ and fraction of unsatisfied nodes $p = 0.01$.

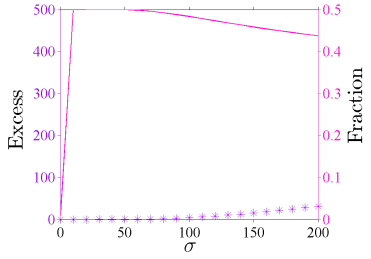

Fig. 4. The amount of excess at a node as standard deviation σ increases using quenched averaging (purple), with the probability of the node being unsatisfied (pink). On a 20-node network with one uncertain consumer of mean $\bar{\Lambda} = -200$.

$$A_{ij} \leftarrow \min\left[0, \frac{y_{ij} - \sum_{k\neq i}(y_{jk} - \frac{A_{jk}+\phi'_{jk}}{B_{jk}+\phi''_{jk}}) - \zeta - \bar{\Lambda}_j}{\sum_{k\neq i}\frac{1}{B_{jk}+\phi''_{jk}}}\right], \quad (11)$$

where $\zeta = \sqrt{2\sigma_j^2}\text{erf}^{-1}(2p-1)$.

Considering a scenario where demand and supply fluctuate, the MP method will satisfy the problem with probability, $1-p$. It does this by leaving volatile node with extra power in case it produces less than expected (generators) or uses more than expected (consumers).

This approximation allows us to determine a confidence level. It may be helpful for calculating what level of reserve to set at controllable power stations.

To find the z-value of a Gaussian distribution we use the equation:

$$z = \bar{\Lambda}_j + \sigma_j\sqrt{2}\text{erf}^{-1}(2p-1) \quad (12)$$

This gives the value on the x-axis that satisfies the distribution $(1-p)\%$ of the time and proves that the annealed averaging works as intended, and is effectively equivalent to the changing the constraint.

Figure 3 shows how the excess (the extra power an uncertain node receives to protect against fluctuations) given to an uncertain node increases as the standard deviation of the fluctuations at that node increases.

2) Quenched averaging: Although annealed averaging is a good approximation it is not exact. Small fluctuations between samples are amplified due to emerging correlations between higher terms in the partition function which may lead to very rare samples dominating the partition function. To carry out quenched averaging one has to calculate $\langle F\rangle = -T\langle \ln Z\rangle$. The same framework as before, e.g. Eq. (4) but consider the expected value of the vertex free energy with respect to the capacity fluctuations. We expand the *expected VFE* with respect to ε_{jk}, using the notation $[A_{jk}^{<>}, B_{jk}^{<>}] = [\frac{\partial\langle F_V\rangle}{\partial y_{jk}}, \frac{\partial^2\langle F_V\rangle}{\partial y_{jk}^2}]$ for the corresponding derivatives; note that this has the implicit assumption that these messages will be sufficient to approximate the new expected messages. The corresponding optimal F_{ij}^* becomes:

$$\langle F_{ij}^*\rangle_\Lambda = \frac{1}{2}\sum_{k\neq i}\mathcal{A}_{jk}\frac{\langle\mu_{ij}^2\rangle_\Lambda - (A_{jk}^{<>}+\phi'_{jk})^2}{B_{jk}^{<>}+\phi''_{jk}}, \quad (13)$$

when μ_{ij} is the Lagrange multiplier. To evaluate F_{ij}^* we need to find $\langle\mu_{ij}^2\rangle$. The optimal μ_{ij} is found, squared, and the average is taken, resulting in the messages:

$$A_{ij}^{<>} \longleftarrow \frac{1}{2}\frac{\Big(\text{erf}(\frac{x-\bar{\Lambda}_j}{\sqrt{2}\sigma_j})+1\Big)(x-\bar{\Lambda}_j) + \frac{2\sigma^2 e^{-(\frac{x-\bar{\Lambda}_j}{\sqrt{2}\sigma})_j^2}}{\sqrt{2\pi\sigma^2}}}{\sum_{k\neq i}\mathcal{A}_{jk}\Big(\frac{1}{B_{jk}^{<>}+\phi''_{jk}}\Big)} \quad (14)$$

where $x \equiv y_{ij} - \sum_{k\neq i}\mathcal{A}_{jk}(y_{jk} + \frac{A_{jk}^{<>}+\phi'_{jk}}{B_{jk}^{<>}+\phi''_{jk}})$.

$$B_{ij}^{<>} \longleftarrow \frac{1}{2\sum_{k\neq i}\mathcal{A}_{jk}\frac{1}{B_{jk}^{<>}+\phi''_{jk}}}\left[\text{erf}\Big(\frac{x-\bar{\Lambda}_j}{\sqrt{2}\sigma}\Big)+1\right], \quad (15)$$

and with a forwards message as before. At $\sigma \to 0$ the equations of $A_{ij}^{<>}$ and $B_{ij}^{<>}$ return to those found in [1] and Eqs. (7) and (8), validating the results.

Figure 4 shows that after convergence with the quenched averaging based algorithm, the excess given to the volatile node is almost zero. This means that the probability of the node being *unsatisfied* is very high, which is unacceptable for any functioning power distribution mechanism.

We speculate that the reasoning behind such a small excess is that the function we minimise is the expected VFE; it includes a hard constraint which outputs 0 if the node is unsatisfied, and 1 if it is. It therefore may be biased towards a small range of values that dominate the expected VFE due to their probability of occurrence coupled with a high probabilistic weight, possibly due to high cumulative transportation costs, at the expense of unsatisfied nodes with a lower probabilistic weight.

3) Quenched with soft constraint: To mitigate the weaknesses of the quenched averaging-based algorithm we introduced soft constraints into the quenched averaging, similar to the constraint used in annealed averaging. The reasoning behind this is to consider uncertainty inherently, but with a constraint that enforces a higher probability of satisfied nodes. Therefore we will continue to use the equations from quenched averaging,

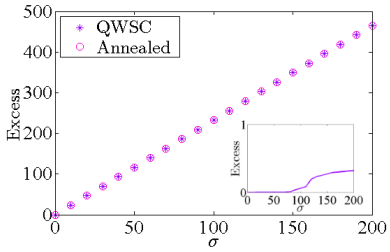

Fig. 5. The excess of quenched with soft constraints (purple *) and the annealed approximation (pink o). On a 20-node network with one uncertain consumer with mean, $\Lambda_j = -200$ and allowed fraction of unsatisfied nodes $p = 0.01$. The inset shows quenched averaging with soft constraints minus annealed results.

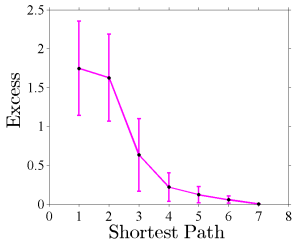

Fig. 6. Excess provided to a volatile consumer node as the generators move further from the node, increasing the transportation length.

but adjust Λ_j to $\Lambda_j + \sqrt{2\sigma_j^2}\text{erf}^{-1}(2p-1)$.

Figure 5 shows that this new method of averaging gives similar results to the annealed. The inset looks at Quenched With Soft Constraints (QWSC) excess values minus those from annealed averaging in order to understand the difference between the two methods. This shows that as the standard deviation σ increases the quenched averaging adds extra capacity to the node. This is much better as it protects against larger uncertainties, arguably due to weighting against transportation cost again.

III. Results

From the results presented it can be seen that with smaller uncertainties the annealed averaging technique will do, but as the uncertainties grow a quenched average based technique ensures that the network has extra security.

Figure 6 shows that the excess given to a volatile node decreases as the shortest path between generators and consumers increases. This demonstrates that the quenched averaging excess is effected largely by transportation cost and reduces the costs incurred by moving resource from far off locations at the expense of providing additional capacity. Future work will examine how to weight the two objectives more appropriately. Using the QWSC algorithm we calculate the effect of uncertainties on the quadratic transportation cost. Figure 7 considers 5 cases of different fractions of volatile nodes in the grid, the *overall* σ value is distributed between. It shows that the distribution of uncertainty makes a huge difference to the transportation cost and the importance of having smaller and more spread out renewable sources throughout the grid

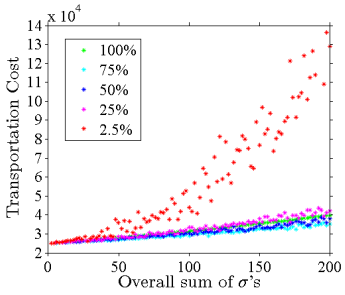

Fig. 7. Transportation cost as the standard deviation σ increases using QWSC averaging, for when the overall standard deviation σ is distributed between 100%, 75%, 50%, 25% or 2.5% of nodes (in a 40-node network).

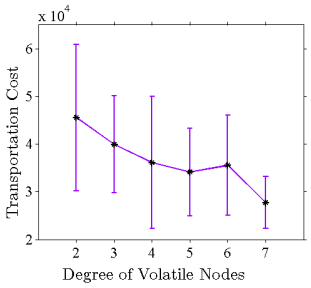

Fig. 8. The transportation cost when uncertainties are put on 4 nodes each with a degree 2,3,4,5,6 and 7, in an 80-node network.

instead of many concentrated in one place. It is known that a decentralised grid increases its robustness against structural perturbations [15]. The graph also shows that if all nodes share the given volatility, this does not give the lowest costs either, of the cases available, 75% shows the lowest cost. Figure 8 shows that a higher connectivity of volatile nodes within a power grid reduces transportation costs too. This may be because a volatile node is more likely to be connected to a node with the power it needs, and so the power is sent along less edges, reducing the cost.

IV. Conclusion

Using simple synthetic examples we demonstrate the potential of MP algorithms to inherently consider volatility within a grid, and may facilitate higher network robustness.

The MP algorithms presented could be improved by making the constraint softer in order to quantify the levels of importance between transportation cost and satisfaction at a given node, this may make a huge difference in the excess given to volatile nodes.

Distributive, probabilistic algorithms of this type have a lot of scope for future work and one could consider issues specific to power grids such as including resistance or changing from passing power over edges to adjusting voltages or phase angles at each node. Also there are other objectives of power grids to be considered such as:

- Bandwidth - There is a maximum power that power lines can handle without overheating and tripping. This would

be a constraint that prevents edges from exceeding their limit [16] jointly with fluctuations.

- Minimising generation costs - The transportation cost in this paper represents minimising power loss through heat in the grid. However, electricity producers are more concerned with the cost of power generation (economic or environmental). A similar algorithm can be developed to minimise generation costs for given demand, in a distributive and computationally efficient manner.
- Minimising load shedding - In the scenario where the amount of power being produced is insufficient to satisfy all consumers, we can assign weights to the consumers reflecting the importance of keeping their power supply intact. Algorithms of the type used here can be employed to minimise the damage caused by insufficient power.

We are currently working on developing appropriate algorithms to address these issues.

ACKNOWLEDGMENTS

We would like to thank Alstom and the EPSRC (Industrial CASE Studentship 12330048) for supporting this research. This work is also partially supported by the Research Grants Council of Hong Kong (grant numbers 604512, 605813). Special thanks also to Sajjad Fekriasl for his interest and enthusiasm and Bill Yeung for help coding and sharing his understanding and insight.

REFERENCES

[1] K. Y. M. Wong and D. Saad, "Inference and optimization of real edges on sparse graphs: A statistical physics perspective," Phys. Rev. E. (76) 011115, 2007.

[2] L. H. Lam *et al.*, "Optimal Location of Power Plant in Power Market," in *2014 International Conference on Intelligent Green Building and Smart Grids*, Taipei, 2014, pp 1-4.

[3] J. Carpentier, "Optimal Power Flows," Int. Journal of Elect. Power & Energy Systems. 1(1), pp 3-15, 1978.

[4] *Report from the Commission to the European Parliment, the Council, the European Economic and Social Committee and the Committee of the Regions,* European Commission, 2015.

[5] D. Bienstock *et al.*, "Chance constrained optimal power flow: Risk-aware network control under uncertainty," SIAM Rev. 56(3) pp 461-495, 2013.

[6] R. Bent *et al*., "Synchronization-Aware and Algorithm-Efficient Chance Constrained Optimal Power Flow," *2013 IREP Symposium-Bulk Power System Dynamics and Control -IX*. Rethymno, 2013.

[7] Q. Jiang *et al.*, "An Efficient Implementation of Automatic Differentiation in Interior Point Optimal Power Flow," IEEE Trans. on Power Systems. 25(1) pp 147-155, 2010.

[8] W-M. Lin *et al.*, "A Hybrid Current-Power Optimal Power Flow Technique," IEEE Trans. on Power Systems. 23(1) pp 1-9, 2008.

[9] M. Mézard and A. Montanari, "The random energy model," in *Information, Physics, and Computation*, Oxford University Press Inc., New York. 2009, pp 101-103.

[10] H. Nishimori, "Mean Field Theory of Spin Glasses" in *Statistical Physics of Spin Glasses and Information Processing*, Oxford University Press Inc., New York. 1999, pp 11-13.

[11] Y. Kabashima *et al.*, "Typical Performance of Gallager-type Error-Correcting Codes," Phys. Rev. Lett. 84(6), pp 1355-1358, 2000.

[12] K. Y. M. Wong and D. Saad, "Minimising Unsatisfaction in Colourful Neighbourhoods," Jour. Phys. A. 41(32) pp 324023, 2008.

[13] C. H. Yeung *et al.*, "From the Physics of Interacting Polymers to Optimizing Routes on the London Underground," Proc, of the Nat. Academy of Sciences. 110(34), pp 13717-13722, 2013.

[14] K. Methaprayoon *et al.*, "An Integration of ANN Wind Power Estimation in UC Considering the Forecasting Uncertainty," in *Industrial and Commercial Power Systems Tech. Conf.*, New York, 2005, pp 116-124.

[15] M. Rohden *et al.*, "Self-organised sychronisation in decentralized power grids," Phys. Rev. Lett. 109(064101), pp 1-5, 2012.

[16] C. H. Yeung and K. Y. M. Wong, "Optimal resource allocation in random networks with transportation bandwidths," Jour. Stat. Mech. P03029, 2009.