

Optimal distribution in smart grids with volatile renewable sources using a message passing algorithm

Elizabeth Harrison^a, David Saad^a, K. Y. Michael Wong^b

^a School of Engineering and Applied Sciences, Aston University, Birmingham, B4 7ET, UK

^b Department of Physics, The Hong Kong University of Science and Technology, Hong Kong, China

Abstract

The design of future electricity grids will allow for renewable energy generators to be effectively incorporated into the network. Current methods of economic dispatch were not designed to accommodate the level of volatility and uncertain nature of sources such as wind and solar; here we demonstrate how an optimisation algorithm called message passing, which is based on principled statistical physics methodologies and is inherently probabilistic, could be an alternative way of considering source volatility efficiently and reliably. The algorithm iteratively passes probabilistic messages in order to find an approximate global optimal solution with moderate computational complexity and inherently consider source volatility. We demonstrate the capabilities of message passing as a distribution algorithm in the presence of uncertainty on synthetic benchmark IEEE networks and show how the volatility increase effects distribution costs.

Keywords: Message-Passing, optimisation, networks, distribution, power flow, uncertainty, renewable energy, electricity, smart grid

1. Introduction

The development of smart electricity grids will facilitate a steady increase in the penetrations of renewable energy sources allowing for harmful and unsustainable sources such as fossil fuels and nuclear to be gradually phased out. Wind is among the most mature forms of renewable 'clean' energy sources and is of significant potential [1]; however, sources such as wind and solar are intermittent and volatile, and can only partially be controlled. The current method of power distribution at the economic dispatch stage was not designed to consider fluctuating sources at that level and a new model is sought to facilitate the distribution under inherent uncertainties. Power generation by renewable sources can be represented as probability distributions; we will use a principled probabilistic algorithm to find an approximate global optimal solution for power distribution in this regime.

The current method for power distribution, Optimal Power Flow (OPF) [2] has been in use for a long time and has seen limited changes over the years. It uses a two-step approach: Starting from some initial voltages, the first step uses Newton Raphson or a fast-decoupled technique to adjust voltages to satisfy Kirchoff's law. The second step uses steepest descent to search for a better solution. The algorithm repeats the two steps until an optimum is found. This method has been used effectively for over 50 years but relies on the fact that changes are small and linear search methods are assumed to be effective. The main disadvantage is that the method is capable of considering only small scale fluctuations and was not designed to consider the scale of volatility that renewable energy brings. An effective power distribution should adhere to the physical constraints while considering objective functions, such as: costs and power loss. It also needs to be robust against deliberate and random attacks as well as run smoothly and reliably for all consumers.

Various economic dispatch methods have been suggested. A leading approach is Chance Constrained OPF [3]-[5] which adds probabilities to the hard constraints of the OPF as a way of considering uncertainties such as the ones induced by renewable sources. Alternatively, the interior point method uses matrices to satisfy the network constraints and a predictor-corrector strategy for minimising solution error [6], [7]. Other optimisation methods have been suggested such as particle swarm [8], cuckoo search algorithm [9] and the shuffled frog leaping algorithm [10]. These methods are mostly heuristic and do not consider high levels of volatility. Being probabilistic, message passing algorithms inherently consider the probabilistic nature of sources/demand; they search for the optimal solution by passing probabilistic messages locally, resulting in an efficient algorithm whose complexity increases only linearly with the system size [11].

We briefly explain message passing in real-valued variables [11] and show how this optimisation method can be adjusted to consider uncertainties in Section 2. Section 3 shows how the algorithm has been successfully applied to synthetic benchmark IEEE networks and the effects of uncertainties on the network. Finally, Section 4 will discuss the advantages of the suggested method and future research directions.

2. Methodology

2.1. Message passing

The tertiary stage of power distribution (economic dispatch) considers the next 15-60 minutes time-window. Distribution algorithms consider predictions of future expected demand in this time window and manage production by setting controllable generators accordingly. Message passing utilizes the *expected* generation prediction and its variance to optimise the distribution of power.

Message passing for resource allocation minimises an objective function, such as power loss, while considering constraints such as the minimum and maximum power a generator can produce. Such an algorithm has been applied microscopically [11] and results give rise to macroscopic properties. Within the statistical physics terminology, this is analogous to minimising the free energy in order to find the *ground state* of a system. The free energy relates to the probability of being in a given state considering the related cost and multiplicity of such solutions. The link between finding the microscopic state of a system and its macroscopic properties is explained in literature [12], [13] and has been studied in many hard computational problems such as graph colouring [14], routing [15] and error-correcting codes [16].

The distribution of power in a system with N nodes, each with connectivity c , is done by assigning a capacity Λ_j to each node j , which denotes generation (positive) or consumption (negative). Power is passed between nodes over an edge, $y_{ij} = -y_{ji}$. The aim of the distribution algorithm is to make all capacities positive or zero (demand satisfied within production limits) by passing power over edges, while minimising the objective functions and considering constraints. The constraint that all nodes should be non-negative can be written mathematically for each node j as

$$\sum_{i=1}^N A_{ij}y_{ij} + \Lambda_j \geq 0 \quad (1)$$

where A_{ij} is an adjacency matrix; $A_{ij} = 1$ if nodes i and j are connected by an edge and zero otherwise.

The objective function considered in this paper is to minimise power loss, $\phi = y_{ij}^2$, for instance if one considers y_{ij} to represent current from i to j ; the algorithm is capable of considering other objective functions as required. With this objective function we can write the general energy function

$$E = \sum_{(ij)} A_{ij} \phi(y_{ij}) \quad (2)$$

where the sum of (ij) is over every pair of nodes in the network. The free energy is defined as $F = -T \ln Z$, where Z is the normalisation constant. Including both costs and constraints it assumes the form:

$$Z = \left(\prod_{(ij)} \int dy_{ij} \right) \prod_j \theta \left(\sum_i A_{ij} y_{ij} + \Lambda_j \right) e^{-\frac{1}{T} \sum_{(ij)} A_{ij} \phi(y_{ij})} \quad (3)$$

where T is the temperature of the system determining how strictly the optimisation is carried out. $\theta(*)$ is the step function returning 1 if the argument is positive and 0 otherwise. It is difficult to minimise the free energy as the computational complexity increases exponentially with the system size, so a principled approximation is necessary. Wong and Saad [11] use the Bethe approximation which exploits the fact that the network is sparsely connected ($N \gg c$). Due to the small number of connections, the probability of loops is negligible; the network can be assumed to be locally tree-like and long range correlations can be ignored. *Local* messages of conditional probabilities can be passed using the assumption that if a node is removed its neighbours would be weakly correlated and can be considered probabilistically independent.

The algorithm works by randomly choosing a node j and one of its neighbours i , the ancestor. The rest of the neighbours are termed descendants and are labelled by the index k (Fig. 1a). This allows for the free energy to be written as a conditional probability $F(y_{ij} | \mathbf{T}_j)$ depending on the free energy of its descendants, where \mathbf{T}_j represents the tree of node j .

$$F(y_{ij} | \mathbf{T}_j) = -T \ln \left\{ \prod_{k=1}^{c-1} \left(\int dy_{jk} \right) \theta \left(\sum_{k=1}^{c-1} y_{jk} - y_{ij} + \Lambda_j \right) \exp \left[-\frac{1}{T} \sum_{k=1}^{c-1} \left(F(y_{jk} | \mathbf{T}_k) + \phi(y_{jk}) \right) \right] \right\} \quad (4)$$

If a node was added to a network of $N - 1$ nodes, the free energy of the system would change. The new free energy of the system minus the average free energy of the tree before the additional node F_{av} , gives the vertex free energy (VFE). This value represents the individual free energy contribution of the added node. The VFEs constitute a recursion relation for each node in the network (see [11] for full derivations):

$$F_V(y_{ij} | \mathbf{T}_j) = -T \ln \left\{ \prod_{k=1}^{c-1} \left(\int dy_{jk} \right) \theta \left(\sum_{k=1}^{c-1} y_{jk} - y_{ij} + \Lambda_j \right) \exp \left[-\frac{1}{T} \sum_{k=1}^{c-1} \left(F_V(y_{jk} | \mathbf{T}_k) + \phi(y_{jk}) \right) \right] \right\} - F_{av} \quad (5)$$

Which becomes

$$F_V(y_{ij} | \mathbf{T}_j) = \min_{\sum_{k=1}^{c-1} y_{jk} - y_{ij} + \Lambda_j \geq 0} \left[\sum_{k=1}^{c-1} \left(F_V(y_{jk} | \mathbf{T}_k) + \phi(y_{jk}) \right) \right] - F_{av} \quad (6)$$

When the temperature, $T \rightarrow 0$. The VFE is a probabilistic local message which can be iteratively passed in order to find the global optimal solution. As it is continuous function message passing is difficult, so

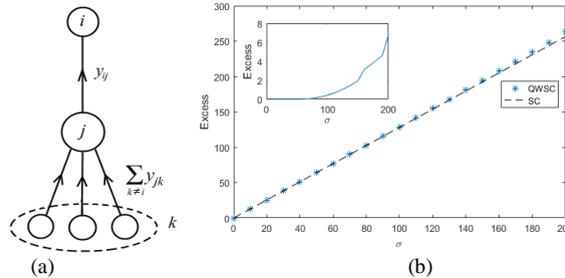


Fig. 1. (a) The tree of node j , with ancestor i and descendants k . (b) The excess given to a volatile node when optimised using message passing with a soft constraint (--) and when it considers fluctuations (*); inset shows the difference between the two methods. Quenched averaging gives more resource to volatile nodes than soft constraints.

alternatively one can pass the first and second derivative of the VFE [11], $[A_{jk}, B_{jk}] = \left[\frac{dF_V}{dy_{jk}}, \frac{d^2F_V}{dy_{jk}^2} \right]$, using the assumption that the function can be accurately represented by the first two terms of a Taylor expansion about the values y_{jk} :

$$F_{ij} = \sum_{k \neq i} \mathcal{A}_{jk} \left((A_{jk} + \phi'_{jk}) \varepsilon_{jk} + \frac{1}{2} (B_{jk} + \phi''_{jk}) \varepsilon_{jk}^2 \right) + \mu_{ij} (\sum_{k \neq i} (y_{jk} + \varepsilon_{jk}) - y_{ij} + \Lambda_j) \quad (7)$$

where ϕ'_{jk} and ϕ''_{jk} represent the first and second derivatives of the objective function ϕ with respect to y_{jk} and the second term represents the constraint on the local positivity of currents/power at each node, enforced by a Lagrange multiplier μ_{ij} . Optimising this function gives the messages (derivatives)

$$A_{ij} \leftarrow \max \left[0, \frac{y_{ij} - \sum_{k \neq i} \mathcal{A}_{jk} \left(y_{jk} \frac{A_{jk} + \phi'_{jk}}{B_{jk} + \phi''_{jk}} \right) - \Lambda_j}{\sum_{k \neq i} \mathcal{A}_{jk} \frac{1}{B_{jk} + \phi''_{jk}}} \right] \quad (8)$$

$$B_{ij} \leftarrow \frac{\Theta(-\mu_{ij}, -\epsilon)}{\sum_{k \neq i} \mathcal{A}_{jk} \frac{1}{B_{jk} + \phi''_{jk}}} \quad (9)$$

where ϵ represents a small positive value; a message to descendants indicating their new current according to the adjustments ε_{jk} is written as:

$$y_{jk} \leftarrow y_{jk} + \varepsilon_{jk} = y_{jk} - \frac{A_{jk} + \phi'_{jk} + \mu_{ij}}{B_{jk} + \phi''_{jk}} \quad (10)$$

The algorithm works by iteratively and randomly choosing a node j and a neighbour i ; passing messages (5) and (6) to node i and adjusting edges to descendants k (7) until the algorithm converges. After convergence, pseudoposterior estimates of current marginal can be calculated determining the solution.

2.2. Message passing as a method of optimising power distribution

We assume a simplified version of a power network with power generators, consumers and substations represented as nodes with appropriate capacities. Only the transmission stage will be considered here. Two stations connected by a power line are represented by two neighbouring nodes connected by an edge. This algorithm does not currently enforce the requirement that all power (generated/consumed) must sum to zero and after convergence some nodes will have positive power, representing reserve or unused capability; satisfying all consumer demands is considered more important than full utilisation of capacity [17]. This derivation extends the message passing algorithm [11] to the case of power grids with volatile renewable generators; it aims to optimise networks with fluctuating nodes by minimising the *expected* VFE, where each node generates a capacity value drawn from Gaussian probability distribution of mean $\bar{\Lambda}_j$ and variance σ_j .

2.3. Quenched averaging

The average of the free energy with respect to random demands/generation can be found by averaging it with respect to specific instances, termed quenched average $\langle F \rangle = -T \langle \ln Z \rangle$, where $\langle * \rangle$ indicates the average over a probability distribution of the demand/generation. This is a non-trivial calculation and as the capacities of each node are now drawn from a Gaussian distribution with an infinite tail, the hard constraint that each node must always be non-negative is now unsatisfiable. Therefore, we soften the constraint by requiring it to be satisfied within a certain probability, p :

$$\langle \Theta \left[\sum_{k \neq i} A_{jk} (y_{jk} + \varepsilon_{jk}) - y_{ij} + \Lambda_j \right] \rangle_{\Lambda} = \frac{1}{2} \operatorname{erfc} \left(\frac{-\sum_{k \neq i} A_{jk} (y_{jk} + \varepsilon_{jk}) + y_{ij} - \Lambda_j}{\sqrt{2\sigma_j^2}} \right) > 1-p \quad (11)$$

where p is the predetermined probability of the constraint being unsatisfied. As before, we calculate the zero temperature vertex free energy but now consider the expected value with respect to the capacity fluctuations. Here the values $[A_{jk}^{\langle \rangle}, B_{jk}^{\langle \rangle}] = \left[\frac{d\langle F_V \rangle}{dy_{jk}}, \frac{d^2\langle F_V \rangle}{dy_{jk}^2} \right]$ are considered, where angled brackets indicate that they refer to the expansion of the *expected vertex free energies* of nodes. We assume that these will be sufficient to describe the expected VFE. Optimising with respect to the first two terms the expected Taylor expansion of the VFE becomes

$$\langle F_{ij}^* \rangle_{\Lambda} = \frac{1}{2} \sum_{k \neq i} A_{jk} \frac{\langle \mu_{ij}^2 \rangle - (A_{jk}^{\langle \rangle} + \phi'_{jk})^2}{B_{jk}^{\langle \rangle} + \phi''_{jk}} \quad (12)$$

where μ_{ij} is the Lagrange multiplier. Evaluating $\langle \mu_{ij}^2 \rangle$ results in the modified messages:

$$A_{ij}^{\langle \rangle} \leftarrow \frac{1}{2} \frac{\left(\operatorname{erf}\left(\frac{x - \hat{\Lambda}_j}{\sqrt{2}\sigma_j}\right) + 1 \right) (x - \hat{\Lambda}_j) + \frac{2\sigma_j^2}{\sqrt{2\pi\sigma_j^2}} e^{-\left(\frac{x - \hat{\Lambda}_j}{\sqrt{2}\sigma_j}\right)^2}}{\sum_{k \neq i} A_{jk} \frac{1}{B_{jk}^{\langle \rangle} + \phi''_{jk}}} \quad (13)$$

$$B_{ij}^{\langle \rangle} \leftarrow \frac{1}{2} \frac{\left(\operatorname{erf}\left(\frac{x - \hat{\Lambda}_j}{\sqrt{2}\sigma_j}\right) + 1 \right)}{\sum_{k \neq i} A_{jk} \frac{1}{B_{jk}^{\langle \rangle} + \phi''_{jk}}} \quad (14)$$

where $x = y_{ij} - \sum_{k \neq i} A_{jk} (y_{jk} - \frac{A_{jk}^{\langle \rangle} + \phi'_{jk}}{B_{jk}^{\langle \rangle} + \phi''_{jk}})$, $\hat{\Lambda}_j = \bar{\Lambda}_j + \sqrt{2}\sigma_j \operatorname{erf}^{-1}(2p - 1)$ and a backwards message

$y_{jk} \leftarrow y_{jk} - \frac{A_{jk}^{\langle \rangle} + \phi''_{jk} + (\mu_{ij})}{B_{jk}^{\langle \rangle} + \phi''_{jk}}$. This is consistent with (8) and (9) for non-fluctuating capacities with $\sigma_j = 0$.

Messages from leaf nodes, connected to just one other node, should be considered separately.

When both demand and generation fluctuate the algorithm provides extra resource to fluctuating nodes in case they use more than or generate less than the expected values. Fig. 1(b) shows how the extra resource changes as fluctuations increase. Here we can see the difference between inherently incorporating uncertainties in the algorithm in comparison to superficially adjusting constraints by modifying the threshold; when fluctuations are higher the algorithm gives nodes some extra resource. The softer constraint determined by the value of p , allows for the confidence level to be predetermined; a small p results in a more reliable system, in addition to the inherent consideration of fluctuations [18].

3. Results

Fig. 2 shows an example of the IEEE 118-Bus [19] with normally generated capacities (about $\mathcal{N}(1,1)$) and fixed standard deviation of 0.5. Three algorithms have been applied to optimise the power flow: based on message passing without considering fluctuations (8)-(9); with soft constraints (11) where $p=0.3$; and with soft-constraints and message passing that incorporates fluctuations (13)-(14) where $p=0.3$. The

red dots signify the overall capacity deficit for each node over a 60 minute time window when every second the *actual* capacity of each node was randomly generated from it's probability distribution $\mathcal{N}(\bar{\Lambda}_j, 0.5)$. Incorporating soft-constraints with message passing that considers fluctuations gave the most reliable results. The soft-constraints were set to satisfy the network with 70% confidence for demonstration only; real networks would require a much higher confidence level. The figure highlights the need for incorporating fluctuations in the optimisation algorithm. Its related computational cost was negligible.

Fig. 3(a) shows the level of extra resource given to each node of some variance. The black dashed line represents the extra resource a node receives according to the predetermined confidence level, p (this is the same dashed line as in Fig. 1(b)). Blue 'x' symbols show network solution values when abundance of resource is available. Orange '•' symbols show that as the overall volatility increases the generators' excess decreases since more resource is required to satisfy a given threshold. Yellow 'o' symbols represent a case of the network not having enough overall power to satisfy the soft constraint at each node, the algorithm nevertheless aims to satisfy each node appropriately according to its volatility. This graph shows how the algorithm addresses fluctuative generation/demand and even when soft constraints are unsatisfiable it appears to provide an appropriate optimal solution. The inset shows the excess given to

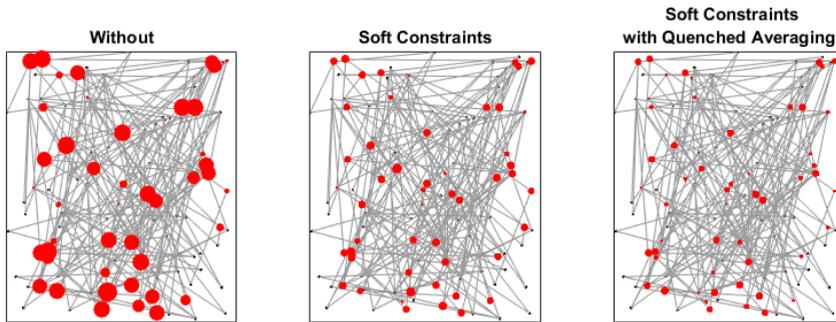


Fig. 2. Red dots indicate constraint violations (the sum of the capacity deficit of each node when capacities were chosen from the random variable using $\mathcal{N}(\bar{\Lambda}_j, 0.25)$ every second over a 60 minute time window) when network capacities are fluctuating (capacities were chosen from a normal distribution $\mathcal{N}(1,1)$ and each node had a standard deviation of 0.5). Power flow was calculated without considering fluctuations (left); when fluctuations are considered in soft constraints with $p=0.3$ (centre); when we consider fluctuations within derivations with $p=0.3$ (right).

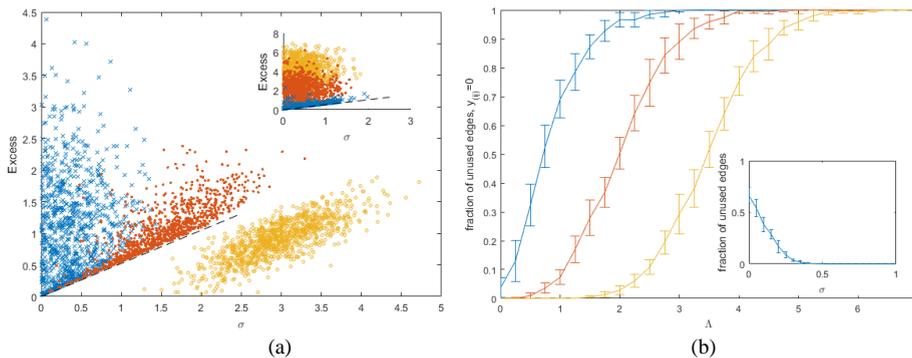


Fig. 3.(a) - Excess given to each node against its standard deviation σ , drawn from a truncated (the positive part of a) Gaussian of standard deviation 0.5 and means 0 (blue x), 1.5 (orange •), 3 (yellow o); for randomly generated 100-node networks of connectivity 5, randomly connected, with Gaussian capacities ($\mathcal{N}(1,1)$). Inset shows the case where average capacities are randomly generated from a Gaussian of mean 0 (blue x), 2 (orange•), 4 (yellow o), and variance 1. (b) - Shows how increasing average Λ in a randomly generated 100-node network effects the fraction of unused edges; when standard deviation at every node is 0 (blue, left), 0.25 (orange, centre) and 0.5 (yellow, right), with average capacities randomly drawn from $\mathcal{N}(\bar{\Lambda}, 1)$. Inset shows how increasing standard deviation (with average capacities 1) effects the fraction of unused edges.

Each node according to its standard deviation, considering different average capacity values, showing that the overall excess in the network increases with the average capacity.

Fig. 3.(b) indicates that as the average capacity of a network decreases, the number of edges used increases until all are being used. We can also see that for higher volatility edges become fully utilised for lower mean values due to the need for additional resource across the network. The inset demonstrates that the more volatile a network (higher standard deviation), the more edges are used.

4. Conclusion

We propose message passing as a probabilistic, computationally efficient and principled tool for optimising power distribution in a power grid that incorporates volatile renewable sources. The derivation of the algorithm has been outlined in two main scenarios, where fluctuations translate to softer constraints and where they are incorporated in the message passing derivation itself. The algorithm has been tested on synthetic network problems to demonstrate its efficacy. In particular we demonstrate its performance on the benchmark IEEE 118 Bus system. The results show how the method increases reliability and robustness of a network with volatile sources. Incorporated within existing control mechanisms the method will enhance the performance of current distribution methods and facilitate the further penetration of volatile renewable sources.

Future work to be considered within the frameworks of message passing includes load shedding minimisation, the inclusion of bandwidth [20], minimising generation costs, Unit Commitment on/off decisions and adjusting voltages and phase angles as the network variables directly. Work is currently underway to address these challenges.

Acknowledgement

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 helping with the code and sharing his understanding and insight.

References

- [1] Jaber S. Environmental impacts of wind energy. *Journal of Clean Energy Technologies*, 2013; 1(3).
- [2] Carpentier J. Contribution to the economic dispatch problem. *Bulletin de la Societe Francoise des Electriciens*, 1962; 3(8):431-447.
- [3] Bent R, Beinstick D, Chertkov M. Synchronization-Aware and algorithm-efficient chance constrained optimal power flow. In: *Proc. of IREP Symposium-Bulk Power System Dynamics and Control-IX*. Rethymno, 2013.
- [4] Bienstock D, Chertkov M, Harnett S. Chance constrained optimal power flow: risk-aware network control under uncertainty. *SIAM Review*, 2013; 56(3):461-495.
- [5] Zhang H. Chance constrained programming for optimal power flow under uncertainty. *IEEE Transactions on Power Systems*, 2011; 26(4):2417-2424.
- [6] Lin WM, Huang CH, Zhan TS. A hybrid current-power optimal power flow technique. *IEEE Trans. on Power Systems*, 2008; 23(1):1-9.
- [7] Jiang Q, Geng G, Guo C, Cao Y. An efficient implementation of automatic differentiation in interior point optimal power flow. *IEEE Transactions on Power Systems*, 2010; 25(1):147-155.
- [8] Yoshida H, Kawata K, Fukuyama Y, Takayama S, Nakanishi Y. A particle swarm optimization for reactive power and voltage control considering voltage security assessment. *IEEE Trans on Power Systems*, 2000; 15(4):1232-1239.
- [9] Govindaraj T, Udayakumar S. Optimal reactive power planning and real power loss minimization using cuckoo search algorithm. *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering*, 2014; 2(2):1-5.
- [10] Rahiminejad A, Alimardani A, Vahidi B, Hosseinian SH. Shuffled frog leaping algorithm optimization for ac-dc optimal power flow dispatch. *Turkish Journal of Electrical Engineering & Computer Sciences*, 2014; 22:874-892.
- [11] Wong KYM, Saad D. Inference and optimization of real edges on sparse graphs: a statistical physics perspective. *Physical*

- Review E*, 2007; (76):011115.
- [12] Mézard M, Montanari A. The random energy model. *Information, Physics, and Computation*, Oxford University Press Inc., New York; 2009:101-103.
- [13] Nishimori H. Mean field theory of spin glasses. In: *Statistical Physics of Spin Glasses and Information Processing*, Oxford University Press Inc., New York; 1999:11-13.
- [14] Wong KYM, Saad D. Minimising unsatisfaction in colourful neighbourhoods. *Journal Physical A*, 2008; 41(32):324023.
- [15] Yeung CH, Saad D, Wong KYM. From the physics of interacting polymers to optimizing routes on the London Underground. *Proceedings of the National Academy of Sciences*, 2013; 110(34):13717-13722.
- [16] Kabashima Y, Murayama T, Saad D. Typical performance of gallager-type error-correcting codes. *Physical Review Letters*, 2000; 84(6):1355-1358.
- [17] Methaprayoon K, Yingvivanapong C, Lee WJ, Liao JR. An integration of ANN wind power estimation in UC considering the forecasting uncertainty. In: *Proc. of Industrial and Commercial Power Systems Technical Conference*, New York; 2005:116-124.
- [18] Harrison E, Saad D, Wong KYM. Message passing for distributed optimisation of power allocation with renewable resources. In: *Proc. of International Conference on Intelligent Green Building & Smart Grid*, 2016. In press
- [19] Christie R. IEEE 118 Bus Test Case. *University of Washington*, 2009.
- [20] Yeung CH, Wong KYM. Optimal resource allocation in random networks with transportation bandwidths. *Journal of Statistical Mechanics*, 2009:P03029.