Dynamic Network Energy Management via Proximal Message Passing

Matt Kraning
Stanford University
mkraning@stanford.edu

Eric Chu
Stanford University
echu508@stanford.edu

Javad Lavaei
Columbia University
lavaei@ee.columbia.edu

Stephen Boyd
Stanford University
boyd@stanford.edu

the essence of knowledge
Boston — Delft

Foundations and Trends® in Optimization

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

M. Kraning, E. Chu, J. Lavaei, and S. Boyd. *Dynamic Network Energy Management via Proximal Message Passing.* Foundations and Trends® in Optimization, vol. 1, no. 2, pp. 73–126, 2014.

This Foundations and Trends® issue was typeset in LaTeX using a class file designed by Neal Parikh. Printed on acid-free paper.

ISBN: 978-1-60198-715-0

Foundations and Trends® in Optimization
Volume 1, Issue 2, 2014

Editorial Scope

Topics

Foundations and Trends® in Optimization publishes survey and tutorial articles on methods for and applications of mathematical optimization, including the following topics:

- Algorithm design, analysis, and implementation (especially on modern computing platforms)
- Models and modeling systems
- New optimization formulations for practical problems
- Applications of optimization in:
  - Machine learning
  - Statistics
  - Data analysis
  - Signal and image processing
  - Computational economics and finance
  - Engineering design
  - Scheduling and resource allocation
  - and other areas

Information for Librarians

Foundations and Trends® in Optimization, 2013, Volume XX, 4 issues. ISSN paper version 2167-3888. ISSN online version 2167-3918. Also available as a combined paper and online subscription.

Foundations and Trends® in Optimization
Vol. 1, No. 2 (2014) 73–126

DOI: 10.1561/2400000002

Dynamic Network Energy Management via Proximal Message Passing

Matt Kraning
Stanford University
mkraning@stanford.edu

Eric Chu
Stanford University
echu508@stanford.edu

Javad Lavaei
Columbia University
lavaei@ee.columbia.edu

Stephen Boyd
Stanford University
boyd@stanford.edu

Contents

Abstract

We consider a network of devices, such as generators, fixed loads, deferrable loads, and storage devices, each with its own dynamic constraints and objective, connected by AC and DC lines. The problem is to minimize the total network objective subject to the device and line constraints over a time horizon. This is a large optimization problem with variables for consumption or generation for each device, power flow for each line, and voltage phase angles at AC buses in each period.

We develop a decentralized method for solving this problem called *proximal message passing*. The method is iterative: At each step, every device exchanges simple messages with its neighbors in the network and then solves its own optimization problem, minimizing its own objective augmented by a term determined by the messages it has received. We show that this message passing method converges to a solution when the device objective and constraints are convex. The method is completely decentralized and needs no global coordination other than iteration synchronization; the problems to be solved by each device can typically be solved extremely efficiently and in parallel.

The proximal message passing method is fast enough that even a serial implementation can solve substantial problems in reasonable time frames. We report results for several numerical experiments, demonstrating the method's speed and scaling, including the solution of a problem instance with over 30 million variables in 5 minutes for a serial implementation; with decentralized computing, the solve time would be less than one second.

M. Kraning, E. Chu, J. Lavaei, and S. Boyd. *Dynamic Network Energy Management via Proximal Message Passing*. Foundations and Trends® in Optimization, vol. 1, no. 2, pp. 73–126, 2014.

DOI: 10.1561/2400000002.

1

Introduction

1.1 Overview

A traditional power grid is operated by solving a number of optimization problems. At the transmission level, these problems include unit commitment, economic dispatch, optimal power flow (OPF), and security-constrained OPF (SC-OPF). At the distribution level, these problems include loss minimization and reactive power compensation. With the exception of the SC-OPF, these optimization problems are static with a modest number of variables (often less than 10000) and are solved on time scales of 5 minutes or more.

However, the operation of next generation electric grids (*i.e.*, smart grids) will rely critically on solving large-scale, dynamic optimization problems involving hundreds of thousands of devices jointly optimizing tens to hundreds of millions of variables, on the order of seconds rather than minutes [16, 41]. More precisely, the distribution level of a smart grid will include various types of active dynamic devices, such as distributed generators based on solar and wind, batteries, deferrable loads, curtailable loads, and electric vehicles, whose control and scheduling amount to a very complex power management problem [59, 9].

In this paper, we consider a general problem, which we call the *dy-*

namic optimal power flow problem (D-OPF), in which dynamic devices are connected by both AC and DC lines, and the goal is to jointly minimize a network objective subject to local constraints on the devices and lines. The network objective is the sum of the device objective functions. These objective functions extend over a given time horizon and encode operating costs such as fuel consumption and constraints such as limits on power generation or consumption. In addition, the objective functions encode dynamic objectives and constraints such as limits on ramp rates for generators or charging and capacity limits for storage devices. The variables for each device consist of its consumption or generation in each time period and can also include local (private) variables which represent internal states of the device over time, such as a storage device's state of charge.

When all device objective functions and line constraints are convex, D-OPF is a convex optimization problem, which in principle can be solved efficiently [7]. If not all device objective functions are convex, we can solve a relaxed form of the D-OPF which can be used to find good, local solutions to the D-OPF. The optimal value of the relaxed D-OPF also gives a lower bound for the optimal value of the D-OPF which can be used to evaluate the suboptimality of a local solution, or, when the local solution has the same value, as a certificate of global optimality.

For any network, the corresponding D-OPF contains at least as many variables as the number of devices and lines multiplied by the length of the time horizon. For large networks with hundreds of thousands of devices and a time horizon with tens or hundreds of time periods, the extremely large number of variables present in the corresponding D-OPF makes solving it in a centralized fashion computationally impractical, even when all device objective functions are convex.

We propose a decentralized optimization method which efficiently solves the D-OPF by distributing computation across every device in the network. This method, which we call *proximal message passing*, is iterative: At each iteration, every device passes simple messages to its neighbors and then solves an optimization problem that minimizes the sum of its own objective function and a simple regularization term that only depends on the messages it received from its neighbors in the previous iteration. As a result, the only non-local coordination needed between devices for proximal message passing is synchronizing iterations. When all device objective functions are convex, we show that proximal message passing converges to a solution of the D-OPF.

Our algorithm can be implemented in several ways. A traditional implementation is to collect all the device constraints and objectives on a single computer or cluster and solve the problem. Our implementation takes this approach and runs on a single 32-core computer with hyperthreading (64 independent threads). A more interesting implementation is a peer-to-peer architecture, in which each device contains its own processor, which carries out the required local dynamic optimization and exchanges messages with its neighbors on the network. In this setting, the devices do not need to divulge their objectives or constraints; they only need to support a simple protocol for interacting with their neighbors. Our algorithm ensures that the network power flows and AC bus phase angles will converge to their optimal values, even though each device has very little information about the rest of the network, and only exchanges limited messages with its immediate neighbors.

Due to recent advances in convex optimization [61, 46, 47], in many cases the optimization problems that each device solves in each iteration of proximal message passing can be executed at millisecond or even microsecond time-scales on inexpensive, embedded processors. Since this execution can happen in parallel across all devices, the entire network can execute proximal message passing at kilohertz rates. We present a series of numerical examples to illustrate this fact by using proximal message passing to solve instances of the D-OPF with over 30 million variables serially in 5 minutes. Using decentralized computing, the solve time would be essentially independent of the size of the network and require just a fraction of a second.

We note that although a primary application for proximal message passing is power management, it can easily be adapted to more general resource allocation and graph-structured optimization problems [51, 2].

1.2 Related work

The use of optimization in power systems dates back to the 1920s and has traditionally concerned the optimal dispatch problem [22], which aims to find the lowest-cost method for generating and delivering power to consumers, subject to physical generator constraints. With the advent of computer and communication networks, many different ways to numerically solve this problem have been proposed [62], and more sophisticated variants of optimal dispatch have been introduced, such as OPF, economic dispatch, and dynamic dispatch [12], which extend optimal dispatch to include various reliability and dynamic constraints. For reviews of optimal and economic dispatch as well as general power systems, see [4] and the book and review papers cited above.

When modeling AC power flow, the D-OPF is a dynamic version of the OPF [8], extending the latter to include many more types of devices such as storage units. Recent smart grid research has focused on the ability of storage devices to cut costs and catalyze the consumption of variable and intermittent renewables in the future energy market [23, 44, 13, 48]. With D-OPF, these storage concerns are directly addressed and modeled in the problem formulation with the introduction of a time horizon and coupling constraints between variables across periods.

Distributed optimization methods are naturally applied to power networks given the graph-structured nature of the transmission and distribution networks. There is an extensive literature on distributed optimization methods, dating back to the early 1960s. The prototypical example is dual decomposition [14, 17], which is based on solving the dual problem by a gradient method. In each iteration, all devices optimize their local (primal) variables based on current prices (dual variables). Then the dual variables are updated to account for imbalances in supply and demand, with the goal being to determine prices for which supply equals demand.

Examples of distributed algorithms in the power systems literature include two phase procedures that resemble a single iteration of dual decomposition. In the first phase, dynamic prices are set over a given time horizon (usually hourly over the following 24 hours) by some mechanism (*e.g.*, centrally by an ISO [28, 29], or through information aggregation in a market [57]). In the second phase, these prices allow individual devices to jointly optimize their power flows with minimal (if any) additional coordination over the time horizon. More recently, building on the work of [39], a distributed algorithm was proposed [38] to solve the dual OPF using a standard dual decomposition on subsystems that are maximal cliques of the power network.

Dual decomposition methods are not robust, requiring many technical conditions, such as strict convexity and finiteness of all local cost functions, for both theoretical and practical convergence to optimality. One way to loosen the technical conditions is to use an augmented Lagrangian [25, 49, 5], resulting in the method of multipliers. This subtle change allows the method of multipliers to converge under mild technical conditions, even when the local (convex) cost functions are not strictly convex or necessarily finite. However, this method has the disadvantage of no longer being separable across subsystems. To achieve both separability and robustness for distributed optimization, we can instead use the *alternating direction method of multipliers* (ADMM) [21, 20, 15, 6]. ADMM is very closely related to many other algorithms, and is identical to Douglas-Rachford operator splitting; see, *e.g.*, the discussion in [6, §3.5].

Augmented Lagrangian methods (including ADMM) have previously been applied to the study of power systems with static, single period objective functions on a small number of distributed subsystems, each representing regional power generation and consumption [35]. For an overview of related decomposition methods applied to power flow problems, we direct the reader to [36, 1] and the references therein.

Our decentralized proximal message passing method is similar in spirit to flow control on a communication network, where each source modulates its sending rate based only on information about the number of un-acknowledged packets; if the network state remains constant, the flows converge to levels that satisfy the constraints and maximize a total utility function [33, 42]. In Internet flow control, this is called *end-point control*, since flows are controlled (mostly) by devices on the edges of the network. The proximal message passing method is closer to local control, since decision making is based only on interaction with neighbors on the network. Another difference is that our method uses *virtual*, proposed energy flows in its messages and not actual energy flows. (Once converged, of course, they can become actual energy flows.)

1.3 Outline

The rest of this paper is organized as follows. In chapter 2 we give the formal definition of our network model. In chapter 3 we give examples of how to model specific devices such as generators, deferrable loads and energy storage systems in our formal framework. In Chapter 4, we describe the role that convexity plays in the D-OPF and introduce the idea of convex relaxations as a tool to find solutions to the D-OPF in the presence of non-convex device objective functions. In Chapter 5 we derive the proximal message passing equations. In Chapter 6 we present a series of numerical examples, and in Chapter 7 we discuss how our framework can be extended to include use cases we do not explicitly cover in this paper.

References

[1] R. Baldick. *Applied Optimization: Formulation and Algorithms for Engineering Systems.* Cambridge University Press, 2006.

[2] S. Barman, X. Liu, S. Draper, and B. Recht. Decomposition methods for large scale LP decoding. *Submitted, IEEE Transactions on Information Theory*, 2012.

[3] A. Bemporad. Model predictive control design: New trends and tools. In *Proceedings of 45th IEEE Conference on Decision and Control*, pages 6678–6683, 2006.

[4] A. Bergen and V. Vittal. *Power Systems Analysis.* Prentice Hall, 1999.

[5] D. P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods.* Academic Press, 1982.

[6] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3:1–122, 2011.

[7] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[8] J. Carpentier. Contribution to the economic dispatch problem. *Bull. Soc. Francaise Elect.*, 3(8):431–447, 1962.

[9] E. A. Chakrabortty and M. D. Ilic. *Control & Optimization Methods for Electric Smart Grids.* Springer, 2012.

[10] M. Chiang, S. Low, A. Calderbank, and J. Doyle. Layering as optimization decomposition: A mathematical theory of network architectures. *Proc. of the IEEE*, 95(1):255–312, 2007.

[11] E. G. Cho, K. A. Thoney, T. J. Hodgson, and R. E. King. Supply chain planning: Rolling horizon scheduling of multi-factory supply chains. In *Proceedings of the 35th conference on Winter simulation: driving innovation*, pages 1409–1416, 2003.

[12] B. H. Chowdhury and S. Rahman. A review of recent advances in economic dispatch. *IEEE Transactions on Power Systems*, 5(4):1248–1259, 1990.

[13] A. O. Converse. Seasonal energy storage in a renewable energy system. *Proceedings of the IEEE*, 100(2):401–409, Feb 2012.

[14] G. B. Dantzig and P. Wolfe. Decomposition principle for linear programs. *Operations Research*, 8:101–111, 1960.

[15] J. Eckstein. Parallel alternating direction multiplier decomposition of convex programs. *Journal of Optimization Theory and Applications*, 80(1):39–62, 1994.

[16] J. H. Eto and R. J. Thomas. Computational needs for the next generation electric grid. In *Department of Energy*, 2011. http://certs.lbl.gov/pdf/lbnl-5105e.pdf.

[17] H. Everett. Generalized Lagrange multiplier method for solving problems of optimum allocation of resources. *Operations Research*, 11(3):399–417, 1963.

[18] A. Eydeland and K. Wolyniec. *Energy and Power Risk Management: New Developments in Modeling, Pricing and Hedging*. Wiley, 2002.

[19] G.D. Forney. Codes of graphs: Normal realizations. *IEEE Transactions on Information Theory*, 47(2):520–548, 2001.

[20] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximations. *Computers and Mathematics with Applications*, 2:17–40, 1976.

[21] R. Glowinski and A. Marrocco. Sur l'approximation, par elements finis d'ordre un, et la resolution, par penalisation-dualité, d'une classe de problems de Dirichlet non-lineares. *Revue Française d'Automatique, Informatique, et Recherche Opérationelle*, 9:41–76, 1975.

[22] H. H. Happ. Optimal power dispatch — a comprehensive survey. *IEEE Transactions on Power Apparatus and Systems*, 96(3):841–854, 1977.

[23] E. K. Hart, E. D. Stoutenburg, and M. Z. Jacobson. The potential of intermittent renewables to meet electric power demand: current methods and emerging analytical techniques. *Proceedings of the IEEE*, 100(2):322–334, Feb 2012.

[24] F. Herzog. *Strategic Portfolio Management for Long-Term Investments: An Optimal Control Approach*. PhD thesis, ETH, Zurich, 2005.

[25] M. R. Hestenes. Multiplier and gradient methods. *Journal of Optimization Theory and Applications*, 4:302–320, 1969.

[26] K. Heussen, S. Koch, A. Ulbig, and G. Andersson. Energy storage in power system operation: The power nodes modeling framework. In *Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 2010 IEEE PES*, pages 1–8, Oct 2010.

[27] T. Hovgaard, L. Larsen, J. Jorgensen, and S. Boyd. Nonconvex model predictive control for commercial refrigeration. *International Journal of Control*, 86(8):1349–1366, 2013.

[28] M. Ilic, L. Xie, and J.-Y. Joo. Efficient coordination of wind power and price-responsive demand — Part I: Theoretical foundations. *IEEE Transactions on Power Systems*, 26(4):1875–1884, Nov 2011.

[29] M. Ilic, L. Xie, and J.-Y. Joo. Efficient coordination of wind power and price-responsive demand—part ii: Case studies. *IEEE Transactions on Power Systems*, 26(4):1885–1893, Nov 2011.

[30] J. L. Jerez, P. J. Goulart, S. Richter, G. A. Constantinides, E. C. Kerrigan, and M. Morari. Embedded online optimization for model predictive control at megahertz rates. *Submitted, IEEE Transactions on Automatic Control*, 2013.

[31] V. Kekatos and G. Giannakis. Joint power system state estimation and breaker status identification. In *Proceedings of the 44th North American Power Symposium*, 2012.

[32] V. Kekatos and G. Giannakis. Distributed robust power system state estimation. *IEEE Transactions on Power Systems*, 2013.

[33] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan. Rate control in communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49:237–252, 1998.

[34] A. Kiana and A. Annaswamy. Wholesale energy market in a smart grid: A discrete-time model and the impact of delays. In A. Chakrabortty and M. Ilic, editors, *Control and Optimization Methods for Electric Smart Grids*, volume 3 of *Power Electronics and Power Systems*, pages 87–110. Springer US, 2012.

[35] B. H. Kim and R. Baldick. Coarse-grained distributed optimal power flow. *IEEE Transactions on Power Systems*, 12(2):932–939, 1997.

[36] B. H. Kim and R. Baldick. A comparison of distributed optimal power flow algorithms. *IEEE Transactions on Power Systems*, 15(2):599–604, 2000.

[37] M. Kraning, Y. Wang, E. Akuiyibo, and S. Boyd. Operation and configuration of a storage portfolio via convex optimization. In *Proceedings of the 18th IFAC World Congress*, pages 10487–10492, 2011.

[38] A. Lam, B. Zhang, and D. Tse. Distributed algorithms for optimal power flow problem. `http://arxiv.org/abs/1109.5229`, 2011.

[39] J. Lavaei and S. Low. Zero duality gap in optimal power flow problem. *IEEE Transactions on Power Systems*, 27(1):92–107, 2012.

[40] J. Lavaei, D. Tse, and B. Zhang. Geometry of power flows in tree networks. *IEEE Power & Energy Society General Meeting*, 2012.

[41] J. Liang, G. K. Venayagamoorthy, and R. G. Harley. Wide-area measurement based dynamic stochastic optimal power flow control for smart grids with high variability and uncertainty. *IEEE Transactions on Smart Grid*, 3:59–69, 2012.

[42] S. H. Low, L. Peterson, and L. Wang. Understanding tcp vegas: a duality model. In *Proceedings of the 2001 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, SIGMETRICS '01, pages 226–235, New York, NY, USA, 2001. ACM.

[43] J. Maciejowski. *Predictive Control with Constraints*. Prentice Hall, 2002.

[44] S. H. Madaeni, R. Sioshansi, and P. Denholm. How thermal energy storage enhances the economic viability of concentrating solar power. *Proceedings of the IEEE*, 100(2):335–347, Feb 2012.

[45] G. Malewicz, M. H. Austern, A. J. C. Bik, J. C. Dehnert, I. Horn, N. Leiser, and G. Czajkowski. Pregel: A system for large-scale graph processing. In *Proceedings of the 2010 International Conference on Management of Data*, pages 135–146, 2010.

[46] J. Mattingley and S. Boyd. CVXGEN: Automatic convex optimization code generation. `http://cvxgen.com/`, 2012.

[47] J. Mattingley, Y. Wang, and S. Boyd. Receding horizon control: Automatic generation of high-speed solvers. *IEEE Control Systems Magazine*, 31:52–65, 2011.

[48] W. F. Pickard. The history, present state, and future prospects of underground pumped hydro for massive energy storage. *Proceedings of the IEEE*, 100(2):473–483, Feb 2012.

[49] M. J. D. Powell. A method for nonlinear constraints in minimization problems. In R. Fletcher, editor, *Optimization*. Academic Press, 1969.

[50] S. J. Qin and T. A. Badgwell. A survey of industrial model predictive control technology. *Control Engineering Practice*, 11(7):733–764, 2003.

[51] P. Ravikumar, A. Agarwal, and M. J. Wainwright. Message-passing for graph-structured linear programs: Proximal methods and rounding schemes. *Journal of Machine Learning Research*, 11:1043–1080, 2010.

[52] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[53] D. Shah. Gossip algorithms. *Foundations and Trends in Networking*, 3(2):1–125, 2008.

[54] S. Sojoudi and J. Lavaei. Physics of power networks makes hard problems easy to solve. *To appear, IEEE Power & Energy Society General Meeting*, 2012.

[55] S. Sojoudi and J. Lavaei. Convexification of generalized network flow problem with application to power systems. Preprint available at `http://www.ee.columbia.edu/~lavaei/Generalized_Net_Flow.pdf`, 2013.

[56] S. Sojoudi and J. Lavaei. Semidefinite relaxation for nonlinear optimization over graphs with application to power systems. Preprint available at `http://www.ee.columbia.edu/~lavaei/Opt_Over_Graph.pdf`, 2013.

[57] N. Taheri, R. Entriken, and Y. Ye. A dynamic algorithm for facilitated charging of plug-in electric vehicles. *arxiv:1112:0697*, 2011.

[58] K. T. Talluri and G. J. Van Ryzin. *The Theory and Practice of Revenue Management*. Springer, 2004.

[59] K. Turitsyn, P. Sulc, S. Backhaus, and M. Chertkov. Options for control of reactive power by distributed photovoltaic generators. *Proceedings of the IEEE*, 99:1063–1073, 2011.

[60] L. G. Valiant. A bridging model for parallel computation. *Communications of the ACM*, 33(8):111, 1990.

[61] Y. Wang and S. Boyd. Fast model predictive control using online optimization. *IEEE Transactions on Control Systems Technology*, 18:267–278, 2010.

[62] J. Zhu. *Optimization of Power System Operation*. Wiley-IEEE Press, 2009.