Operator Splitting Methods in Control

Operator Splitting Methods in Control

Giorgos Stathopoulos
École Polytechnique Fédérale de Lausanne (EPFL)

Harsh Shukla
École Polytechnique Fédérale de Lausanne (EPFL)

Alexander Szűcs
Slovak University of Technology in Bratislava

Ye Pu
École Polytechnique Fédérale de Lausanne (EPFL)

Colin N. Jones
École Polytechnique Fédérale de Lausanne (EPFL)

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Systems and Control

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

G. Stathopoulos, H. Shukla, A. Szűcs, Y. Pu and C. N. Jones. *Operator Splitting Methods in Control*. Foundations and Trends® in Systems and Control, vol. 3, no. 3, pp. 249–362, 2016.

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

ISBN: 978-1-68083-174-0

Foundations and Trends® in Systems and Control

Volume 3, Issue 3, 2016

Editorial Board

Editorial Scope

Topics

Foundations and Trends® in Systems and Control publishes survey and tutorial articles in the following topics:

- Control of:
  - Hybrid and discrete event systems
  - Nonlinear systems
  - Network systems
  - Stochastic systems
  - Multi-agent systems
  - Distributed parameter systems
  - Delay systems
- Systems
  - Energy storage
  - Grid integration
  - Conversion technologies
  - Underpinning materials developments
- Filtering, estimation, and identification
- Optimal control
- Systems theory
- Control applications

Information for Librarians

Foundations and Trends® in Systems and Control, 2016, Volume 3, 4 issues. ISSN paper version 2325-6818. ISSN online version 2325-6826. Also available as a combined paper and online subscription.

Foundations and Trends® in Systems and Control
Vol. 3, No. 3 (2016) 249–362

DOI: 10.1561/2600000008

now
the essence of knowledge

Operator Splitting Methods in Control

Giorgos Stathopoulos
École Polytechnique Fédérale de Lausanne (EPFL)

Harsh Shukla
École Polytechnique Fédérale de Lausanne (EPFL)

Alexander Szűcs
Slovak University of Technology in Bratislava

Ye Pu
École Polytechnique Fédérale de Lausanne (EPFL)

Colin N. Jones
École Polytechnique Fédérale de Lausanne (EPFL)

Contents

Abstract

The significant progress that has been made in recent years both in hardware implementations and in numerical computing has rendered real-time optimization-based control a viable option when it comes to advanced industrial applications. More recently, the need for control of a process in the presence of a limited amout of hardware resources has triggered research in the direction of embedded optimization-based control. At the same time, and standing at the other side of the spectrum, the field of big data has emerged, seeking for solutions to problems that classical optimization algorithms are incapable to provide. This triggered some interest to revisit the family of first order methods commonly known as *decomposition schemes* or *operator splitting methods*. Although it is established that splitting methods are quite beneficial when applied to large-scale problems, their potential in solving small to medium scale embedded optimization problems has not been studied so extensively. Our purpose is to study the behavior of such algorithms as solvers of control-related problems of that scale. Our effort focuses on identifying special characteristics of these problems and how they can be exploited by some popular splitting methods.

G. Stathopoulos, H. Shukla, A. Szűcs, Y. Pu and C. N. Jones. *Operator Splitting Methods in Control*. Foundations and Trends® in Systems and Control, vol. 3, no. 3, pp. 249–362, 2016.

DOI: 10.1561/2600000008.

1

Introduction

The significant progress that has been made in recent years both in hardware implementations and in numerical computing has rendered real-time optimization-based control a viable option when it comes to advanced industrial applications. More recently, the need for control of a process in the presence of a limited amout of hardware resources has triggered research in the direction of embedded optimization-based control. Many efficient high-speed solvers have been developed for both linear and nonlinear control, based on either *first order methods* (FiOrdOs [133], QPgen [57],[59], DuQuad [88]), *interior point (IP) methods* (FORCES [43], CVXGEN [84]) and *active set methods* (QPOASES [50]).

In this work we focus on systems with linear dynamics, giving rise to convex control problems. The purpose of the survey is to explore a family of first order methods known as *decomposition schemes* or *operator splitting methods*. The abstract form of the problem at hand is the minimization of the sum of two convex functions subject to linear equality constraints, and can be written as

$$\text{minimize} \quad f(z) + g(Lz) \ , \tag{1.1}$$

with variables $z \in \mathbb{R}^n$, where f and g are closed, proper convex functions and $A : \mathbb{R}^n \to \mathbb{R}^p$ is a linear map. A splitting method can be applied to the above problem after rewriting it as

$$\begin{array}{ll} \text{minimize} & f(z) + g(y) \\ \text{subject to} & Lz = y \ , \end{array} \tag{1.2}$$

by alternatingly (or simultaneously) minimizing over y and z. Clearly, the solutions of problems (1.2) and (1.1) are identical. Inequality constraints that might appear are already embedded in one of the two functions in the form of indicator functions, *i.e.*, a membership function for a set $\mathcal{C}$

$$\delta_{\mathcal{C}}(z) = \begin{cases} 0 & z \in \mathcal{C} \\ \infty & \text{otherwise,} \end{cases} \tag{1.3}$$

which is the reason why both f and g are considered to be *extended-real-valued functions* (see [18, § 3.1.2]). Formulations similar to the above have been studied extensively and we can look for their roots in the method of multipliers [75], [110], the Arrow-Hurwicz method [3], Douglas-Rachford splitting [44] and ADMM [60], [55]. Decomposition of the original problem into simpler ones is beneficial when distributed computation tools are available. This potential is already suggested in the classical references [15] and [45]. It was not until recently, though, that decomposition algorithms were indeed applied in modern engineering problems (signal and image processing, big data analysis, machine learning, [17] and [27]), in cases where off-the-shelf interior point solvers simply fail due to the large dimensions involved. The thesis [47] provides a comprehensive description of the connection of several splitting algorithms under a common framework. Finally, the book [7] provides a mathematically rigorous introduction to operator splitting methods in general Hilbert spaces.

The plethora of different approaches for solving problem (1.2) is partly a consequence of the problem-dependent behavior of first order methods. This behavior has both its pros and cons; on one hand, sensitivity to the problem's structure and data requires pre-processing and tuning of several parameters, a procedure that can be cumbersome. However, it is exactly this procedure that gives the flexibility to customize the solver to the problem at hand, and, in many cases,

outperform by several orders of magnitude general purpose solvers. Consequently, there are numerous approaches, each of which can be less or more pertinent for the specific problem. Mentioning some of the most important categorizations, we can solve either the *primal* problem, the *dual* problem, or a *primal-dual* formulation. Regarding primal approaches, the most popular one is the primal decomposition method [15], [19], where the original problem is decomposed into a master problem and two subproblems. The two subproblems have both local and shared (complicating) variables, while the master subproblem manipulates only the complicating variables. Primal decomposition works well when the complicating variables for the two subproblems are few.

Dualization plays a crucial role in more complicated problems. It can be performed by means of *Lagrangian relaxations* (dual decomposition [35], [49], [123], [14]), *augmented Lagrangian relaxations* [13], [117], [116], *alternating minimization (Gauss-Seidel) augmented Lagrangian schemes* (ADMM), mixture of Lagrangian with augmented Lagrangian schemes (AMA [131]), *linearized augmented Lagrangians* or *approximate minimization* schemes ([23], [4]) and, finally, *mixtures of alternating minimization with partial linearization* (PDHG [139], [48], [22], [30] and several similar primal-dual schemes [28], [134], [16]).

Although it is well-established that splitting methods are quite beneficial when applied to large-scale problems, their potential in solving small to medium scale embedded optimization problems has not been studied so extensively. It was not until very recently that the first works attempting to apply decomposition methods in control problems started making their appearance [102], [56], [57], [59], [105]. Our purpose is to study the behavior of such algorithms as solvers of control-related convex problems of that scale, *i.e.*, from tens to a few hundreds of variables. Our effort focuses on identifying special characteristics of these problems and how they can be exploited by some popular splitting methods. Some of the questions that we attempt to answer are:

1. It is very common in practice that optimal control problems come with a quadratic objective, since in this way stability can be proven for regulation or tracking purposes. What is the best way

to exploit this smooth term, along with the special structure of the dynamics equation?

2. Given that a control problem has to be solved repeatedly (*e.g.*, MPC), how does warm-starting of the solution affect the speed?

3. Given the structure of the problem at hand, which algorithms will converge more quickly?

4. Are there ways to precondition the problem in order to reduce the solve time?

In what follows we present three well-understood splitting algorithms, the *alternating direction method of multipliers (ADMM)*, the *alternating minimization algorithm (AMA)* and a *primal-dual algorithm (PDA)*, the most popular representative of several primal-dual schemes that have been recently developed. These three methods come from different sides of the spectrum described above, but also hold very strong similarities. Our choice is motivated from the fact that the methods are analyzed and extended from several communities, and hence their properties are well-understood.

The paper is organized as follows: In Chapter 2 we formulate the problem we want to solve and look at it from three different perspectives, resulting in the three algorithms we use. Subsequently we introduce the algorithms under a unified scheme and report their properties. In Chapters 3 and 4 we build on the basic variants of the methods presented before, introduce several enhanced versions and focus on their applicability for solving optimization problems. More specifically, in Chapter 3 we review how one can exploit the structure of the problem to accelerate the theoretical convergence rates. In Chapter 4 we extend the discussion on acceleration to more practical schemes, *i.e.*, stepsize selection and preconditioning. We provide a comprehensive literature review of existing methods and we present generic preconditioned versions of the three algorithms. In Chapter 5 we discuss the computational aspects; we identify the bottlenecks in each method and propose ways to speed up the computation. In Chapter 6 we summarize the

observations that we have made and attempt to construct a guideline about how to choose a splitting scheme given a problem. Finally, the algorithms are illustrated with three examples in Chapter 6.

1.1 Notation and Definitions

Let $\mathcal{Z} = (\mathbb{R}^n, \langle\cdot,\cdot\rangle)$ be a Euclidean space equipped with the inner product $\langle z, x\rangle = z^\top x$ and the corresponding norm $\|z\| = \sqrt{\langle z, z\rangle}$. Symmetric n-dimensional matrices are denoted with $\mathbb{S}^n$, while positive (semi)definite matrices are denoted with $(\mathbb{S}^n_+)\mathbb{S}^n_{++}$. We also consider the scaled norm $\|z\|_P = \sqrt{\langle z, Pz\rangle}$, with $P \in \mathbb{S}_+$. The matrix norm of the linear operator $M \in \mathbb{R}^{m\times n}$ is defined as $\|M\| = \sup_{z\neq 0} \frac{\|Mz\|}{\|z\|}$. The minimum and maximum eigenvalue of a matrix $Q \in \mathbb{R}^{n\times n}$ are denoted by $\lambda_{\min}(Q)$ and $\lambda_{\max}(Q)$, respectively.

The domain of the extended-real-valued function f is defined as $\mathbf{dom}\, f = \{z \in \mathcal{Z} : f(z) < +\infty\}$ and f is proper if $\mathbf{dom}\, f \neq \emptyset$ and $f > -\infty$. The function f is closed if its epigraph $\mathbf{epi}\, f = \{(z,t) \in \mathbb{R}^n \times \mathbb{R} : f(z) \leq t\}$ is a closed nonempty convex set. The range of extended-real-valued functions is denoted with $\mathbb{R} \cup \{+\infty\} = \overline{\mathbb{R}}$. We denote the conjugate of a convex function with $f^\star$, while a minimizer is denoted by an asterisk, *i.e.*, $f(z^*) \leq f(z)\ \forall z \in \mathcal{Z}$. Finally, for succinctness in the notation, we denote the class of all proper, closed, convex functions from $\mathcal{Z}$ to $\overline{\mathbb{R}}$ with $\Gamma_0(\mathcal{Z})$.

The indicator function of a convex set $\mathcal{C}$ is denoted with $\delta_{\mathcal{C}}(\cdot)$. For the common norm balls the notation changes to $\delta_i(z,\alpha),\ i = 1, 2, \infty$, which denotes the constraint $\|z\|_i \leq \alpha$. Similar notation to the 2-norm ball is used for the second-order cone constraint, with the difference that the second argument is a scalar affine function itself, *i.e.*, $\delta_2(Ax+b, c^\top z+d)$ denotes the constraint $\|Az + b\|_2 \leq c^\top z + d$. The most common pairs of indicator functions with their conjugate representation are given in Table A.1.

References

[1] A. B. Açikmese and L. Blackmore. Lossless convexification of a class of optimal control problems with non-convex control constraints. *Automatica*, 2011.

[2] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 3rd edition, 1999.

[3] K. J. Arrow, L. Hurwicz, and H. Uzawa. *Studies in linear and non-linear programming*. Stanford University Press, 1958.

[4] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran. Alternating proximal algorithms for weakly coupled convex minimization problems. Applications to dynamical games and PDE's. *Journal of Convex Analysis*, 15(3):485, 2008.

[5] H. Attouch and M. Soueycatt. Augmented Lagrangian and Proximal Alternating Direction Methods of Multipliers in Hilbert spaces. Applications to Games, PDE's and Control. *Pacific Journal of Optimization 5*, 2008.

[6] D. Axehill and M. Morari. An alternative use of the Riccati recursion for efficient optimization. *Systems & Control Letters*, 61(1):37–40, 2012.

[7] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. Springer Science+ Business Media, 2011.

[8] A. Beck and M. Teboulle. A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems. *SIAM Journal of Imaging Sciences*, 2(1):183–202, 2009.

[9] A. Beck and M. Teboulle. Smoothing and first order methods: A unified framework. *SIAM Journal on Optimization*, 22(2):557–580, 2012.

[10] A. Beck and M. Teboulle. A fast dual proximal gradient algorithm for convex minimization and applications. *Operations Research Letters*, 2014.

[11] S. Becker and J. Fadili. A quasi-Newton proximal splitting method. In *Advances in Neural Information Processing Systems 25*, pages 2618–2626. 2012.

[12] M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 5 2005.

[13] D. P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods (Optimization and Neural Computation Series)*. Athena Scientific, 1996.

[14] D. P Bertsekas. *Nonlinear Programming*. Athena Scientific, 1999.

[15] D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Prentice-Hall, Inc., 1989.

[16] R. I. Bot, E. R. Csetnek, and A. Heinrich. On the convergence rate improvement of a primal-dual splitting algorithm for solving monotone inclusion problems. *arXiv preprint arXiv:1303.2875*, 2013.

[17] 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):1–122, 2011.

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

[19] S. Boyd, L. Xiao, A. Mutapcic, and J. Mattingley. Notes on decomposition methods. *Notes for EE364B, Stanford University*, 2007.

[20] A. M. Bradley. Algorithms for the Equilibration of Matrices and Their Application to Limited-Memory Quasi-Newton Methods,. PhD thesis, Stanford ICME, 2010.

[21] A. Chambolle and C. Dossal. On the convergence of the iterates of "FISTA". 2014.

[22] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 2011.

[23] G. Chen and M. Teboulle. A proximal-based decomposition method for convex minimization problems. *Mathematical Programming*, 1994.

[24] E. Chu, B. O'Donoghue, N. Parikh, and S. Boyd. A Primal-Dual Operator Splitting Method for Conic Optimization. Technical report, Stanford Internal Report, 2013.

[25] G. Cohen. Auxiliary problem principle and decomposition of optimization problems. *Journal of Optimization Theory and Applications*, 32(3):277–305, 1980.

[26] P. L. Combettes, L. Condat, J. C. Pesquet, and B. C. Vũ. A forward-backward view of some primal-dual optimization methods in image recovery. In *The IEEE International Conference on Image Processing*, pages 4141–4145, 2014.

[27] P. L. Combettes and J. C. Pesquet. Proximal splitting methods in signal processing. In *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pages 185–212. Springer New York, 2011.

[28] P. L. Combettes and J. C. Pesquet. Primal-dual splitting algorithm for solving inclusions with mixtures of composite, lipschitzian, and parallel-sum type monotone operators. *Set-Valued and Variational Analysis*, 20(2):307–330, 2012.

[29] P. L. Combettes and B. C. Vũ. Variable metric forward–backward splitting with applications to monotone inclusions in duality. *Optimization*, 63(9):1289–1318, 2014.

[30] L. Condat. A Primal-Dual Splitting Method for Convex Optimization Involving Lipschitzian, Proximable and Linear Composite Terms. *Journal of Optimization Theory and Applications*, 158(2):460–479, 2013.

[31] A. R. Conn, N. I. M. Gould, and P. L. Toint. A globally convergent augmented lagrangian algorithm for optimization with general constraints and simple bounds. *SIAM Journal on Numerical Analysis*, 28:545–572, 1991.

[32] C. Conte, T. Summers, M. N. Zeilinger, M. Morari, and C. N. Jones. Computational aspects of distributed optimization in model predictive control. In *The 51st IEEE Annual Conference on Decision and Control*, pages 6819–6824, 2012.

[33] D. B. Crawley, L. K. Lawrie, F. C. Winkelmann, W. F. Buhl, Y. J. Huang, C. O. Pedersen, R. K. Strand, R. J. Liesen, D. E. Fisher, M. J. Witte, and J. Glazer. EnergyPlus: creating a new-generation building energy simulation program. *Energy and Buildings*, 33(4):319–331, 2001.

[34] E. Cuthill and J. McKee. Reducing the Bandwidth of Sparse Symmetric Matrices. In *Proceedings of the ACM 24th National Conference*, pages 157–172, New York, NY, USA, 1969.

[35] G. B. Dantzig and P. Wolfe. Decomposition Principle for Linear Programs. *Operations Research*, 1960.

[36] D. Davis. Convergence rate analysis of primal-dual splitting schemes. *SIAM Journal on Optimization*, 25(3):1912–1943, 2015.

[37] D. Davis and W. Yin. Convergence rate analysis of several splitting schemes. *arXiv preprint arXiv:1406.4834*, 2014.

[38] D. Davis and W. Yin. A Three-Operator Splitting Scheme and its Optimization Applications. Technical Report CAM 15-13, University of California, Los Angeles, 2015.

[39] T. Davis. *Direct Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, 2006.

[40] W. Deng and W. Yin. On the Global and Linear Convergence of the Generalized Alternating Direction Method of Multipliers. *Technical Report Rice CAAM TR12-14*, 2012.

[41] O. Devolder, F. Glineur, and Y. Nesterov. Double smoothing technique for large-scale linearly constrained convex optimization. *SIAM Journal on Optimization*, 22(2):702–727, 2012.

[42] O. Devolder, F. Glineur, and Y. Nesterov. First-order methods of smooth convex optimization with inexact oracle. *Mathematical Programming*, 146(1-2):37–75, 2014.

[43] A. Domahidi, A. Zgraggen, M. N. Zeilinger, M. Morari, and C. N. Jones. Efficient Interior Point Methods for Multistage Problems Arising in Receding Horizon Control. In *The IEEE Conference on Decision and Control*, pages 668 – 674, Maui, HI, USA, 2012.

[44] J. Douglas and H. H. Rachford. On the Numerical Solution of Heat Conduction Problems in Two and Three Space Variables. *Transaction of the American Mathematical Society*, 82:421–489, 1956.

[45] J. Eckstein and D. P. Bertsekas. On the Douglas-Rachford Splitting Method and the Proximal Point Algorithm for Maximal Monotone Operators. *Mathematical Programming*, 1992.

[46] J. Eckstein and P. J. S. Silva. A practical relative error criterion for augmented lagrangians. *Mathematical Programming*, 141:319–348, 2013.

[47] E. Esser. Primal Dual Algorithms for Convex Models and Applications to Image Restoration, Registration and Nonlocal Inpainting. PhD thesis, UCLA, 2010.

[48] E. Esser, X. Zhang, and T. F. Chan. A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science. *SIAM Journal of Imaging Sciences*, 2010.

[49] H. Everett. Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources. *Operations Research*, 1963.

[50] H. J. Ferreau, H. G. Bock, and M. Diehl. An online active set strategy to overcome the limitations of explicit MPC. *International Journal of Robust and Nonlinear Control*, 2008.

[51] C. Fougner and S. Boyd. Parameter selection and pre-conditioning for a graph form solver. *arXiv preprint arXiv:1503.08366*, 2015.

[52] G. Frison. Numerical methods for model predictive control. 2012.

[53] G. Frison and J. B. Jørgensen. Efficient implementation of the Riccati recursion for solving linear-quadratic control problems. In *The IEEE International Conference on Control Applications*, pages 1117–1122, 2013.

[54] G. Frison, H. B. Sørensen, B. Dammann, and J. B. Jørgensen. High-performance small-scale solvers for linear Model Predictive Control. In *The IEEE European Control Conference*, pages 128–133, 2014.

[55] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite-element approximations. *Computers & Mathematics with Applications*, 1976.

[56] E. Ghadimi, A. Teixeira, I. Shames, and M. Johansson. Optimal parameter selection for the alternating direction method of multipliers (admm): quadratic problems. *IEEE Transactions on Automatic Control*, 60(3):644–658, 2015.

[57] P. Giselsson and S. Boyd. Linear Convergence and Metric Selection in Douglas Rachford Splitting and ADMM. *To appear in IEEE Transactions on Automatic Control*.

[58] P. Giselsson and S. Boyd. Monotonicity and restart in fast gradient methods. In *The 53rd IEEE Annual Conference on Decision and Control*, pages 5058–5063, 2014.

[59] P. Giselsson and S. Boyd. Metric Selection in Fast Dual Forward Backward Splitting. *Automatica*, 2015.

[60] Marroco A. Glowinski, R. Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires. *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique*, 9(R2):41–76, 1975.

[61] R. Glowinski and P. Le Tallec. *Augmented Lagrangian And Operator-splitting Methods In Nonlinear Mechanics*. Society for Industrial and Applied Mathematics, 1989.

[62] T. Goldstein, E. Esser, and R. Baraniuk. Adaptive primal-dual hybrid gradient methods for saddle-point problems. *arXiv preprint arXiv:1305.0546*, 2013.

[63] T. Goldstein, B. O'Donoghue, S. Setzer, and R. Baraniuk. Fast alternating direction optimization methods. *SIAM Journal on Imaging Sciences*, 7(3):1588–1623, 2014.

[64] T. Goldstein and S. Osher. The Split Bregman Method for l_1-Regularized Problems. *SIAM Journal of Imaging Sciences*, 2(2):323–343, 2009.

[65] T. Goldstein, C. Studer, and R. Baraniuk. A Field Guide to Forward-Backward Splitting with a FASTA Implementation. *arXiv preprint arXiv:1411.3406*, 2015.

[66] G. G. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, 3rd edition, 1996.

[67] T. T. Gorecki, F. A. Qureshi, and C. Jones. Openbuild : An integrated simulation environment for building control. In *The Multi-Conference on Systems and Control*, 2015.

[68] O. Güler. On the Convergence of the Proximal Point Algorithm for Convex Minimization. *SIAM Journal on Control and Optimization*, 29(2):403–419, 1991.

[69] O. Güler. New proximal point algorithms for convex minimization. *SIAM Journal on Optimization*, 2(4):649–664, 1992.

[70] A. Hamdi and S. K. Mishra. *Decomposition Methods Based on Augmented Lagrangians: A Survey*, pages 175–203. 2011.

[71] E. N. Hartley, J. L. Jerez, A. Suardi, J. M. Maciejowski, E. C. Kerrigan, and G. Constantinides. Predictive Control using an FPGA with Application to Aircraft Control. *IEEE Transactions on Control Systems Technology*, 2013.

[72] B. He, H. Yang, and S.L. Wang. Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities. *Journal of Optimization Theory and Applications*, 2000.

[73] B. He and X. Yuan. Convergence Analysis of Primal-Dual Algorithms for a Saddle-Point Problem: From Contraction Perspective. *SIAM Journal on Imaging Sciences*, 5(1):119–149, 2012.

[74] B. He and X. Yuan. On the $O(1/n)$ Convergence Rate of the Douglas-Rachford Alternating Direction Method. *SIAM Journal on Numerical Analysis*, 2012.

[75] R. M. Hestenes. Multiplier and gradient methods. *Journal of Optimization Theory and Applications*, 1969.

[76] J. B. Hiriart-Urruty and C. Lemarechal. *Convex Analysis and Minimization Algorithms: Part 2: Advanced Theory and Bundle Methods*. Springer, 2010.

[77] J. H. Hours and C. N. Jones. An alternating trust region algorithm for distributed linearly constrained nonlinear programs, application to the AC optimal power flow. EPFL-REPORT-205056, 2015.

[78] J. H. Hours and C. N. Jones. A parametric nonconvex decomposition algorithm for real-time and distributed nmpc. *IEEE Transactions on Automatic Control*, 2016.

[79] P. Latafat and P. Patrinos. Asymmetric Forward-Backward-Adjoint Splitting for Solving Monotone Inclusions Involving Three Operators. *arXiv.org*, 2016.

[80] Z. Liu, A. Hansson, and L. Vandenberghe. Nuclear norm system identification with missing inputs and outputs. *Systems & Control Letters*, 62(8):605–612, 2013.

[81] Z-Q. Luo M. Hong. On the linear convergence of the alternating direction method of multipliers, 2012.

[82] B. Martinet. Régularisation d'inéquations variationnelles par approximations successives. *Revue Française de Informatique et Recherche Opérationelle*, 4(R3):154–158, 1970.

[83] B. Martinet. Détermination approchée d'un point fixe d'une application pseudo-contractante. *C.R. Acad. Sci. Paris*, 274A:163–165, 1972.

[84] J. Mattingley and S. Boyd. CVXGEN: a code generator for embedded convex optimization. *Optimization and Engineering*, 2012.

[85] A. Moradifam and A. Nachman. Convergence of the alternating split Bregman algorithm in infinite-dimensional Hilbert spaces. *arXiv:1112.1960*, page 113, 2011.

[86] J. J. Moreau. Fonctions convexes duales et points proximaux dans un espace hilbertien. *Comptes Rendus de l'Académie des Sciences (Paris), Série A*, 255, 1962.

[87] J. J. Moreau. Proximité et dualité dans un espace Hilbertien. *Bulletin de la Société Mathématique de France*, 93(2):273–299, 1965.

[88] I. Necoara and A. Patrascu. DuQuad: an inexact (augmented) dual first order algorithm for quadratic programming. *arXiv preprint arXiv:1504.05708*, 2015.

[89] I. Necoara, A. Patrascu, and A. Nedić. *Complexity Certifications of First-Order Inexact Lagrangian Methods for General Convex Programming: Application to Real-Time MPC*, pages 3–26. Springer International Publishing, 2015.

[90] I. Necoara and J. A. K. Suykens. Application of a smoothing technique to decomposition in convex optimization. *IEEE Transactions on Automatic Control*, 53(11):2674–2679, 2008.

[91] A. Nemirovski and D. B. Yudin. *Problem complexity and method efficiency in optimization*. Wiley-Interscience series in discrete mathematics. Wiley-Interscience, Chichester, New York, 1983.

[92] Y. Nesterov. A method for solving a convex programming problem with rate of convergence $O(1/k^2)$. *Doklady Mathematics*, 269(3):543–547, 1983.

[93] Y. Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course*. Springer, 2004.

[94] Y. Nesterov. Smooth minimization of non-smooth functions. *Mathematical Programming*, 2005.

[95] Y. Nesterov. Gradient methods for minimizing composite objective function. Technical report, CORE, Catholic University of Louvain, 2007.

[96] Y. Nesterov. How to advance in structural convex optimization. *OPTIMA: Mathematical Programming Society Newsletter*, 78:2–5, 2008.

[97] I. Nielsen, D. Ankelhed, and D. Axehill. Low-rank modifications of riccati factorizations with applications to model predictive control. In *The 52nd IEEE Annual Conference on Decision and Control*, pages 3684–3690, 2013.

[98] R. Nishihara, L. Lessard, B. Recht, A. Packard, and M. I. Jordan. A general analysis of the convergence of ADMM. In *In 32nd International Conference on Machine Learning*, 2015.

[99] J. Nocedal and S. J. Wright. *Numerical Optimization 2nd edition.* Springer, 2006.

[100] B. O'Donoghue and E.J. Candes. Adaptive Restart for Accelerated Gradient Schemes. *arXiv.org*, 2012.

[101] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd. Operator splitting for conic optimization via homogeneous self-dual embedding. *arXiv preprint arXiv:1312.3039*, 2013.

[102] B. O'Donoghue, G. Stathopoulos, and S. Boyd. A splitting method for optimal control. *IEEE Transactions on Control Systems Technology*, 2012.

[103] N. Parikh. Proximal operators. https://github.com/cvxgrp/proximal.

[104] N. Parikh and S. Boyd. Proximal algorithms. *Foundations and Trends® in Optimization*, 1(3):127–239, 2014.

[105] P. Patrinos and A. Bemporad. An accelerated dual gradient-projection algorithm for embedded linear model predictive control. *IEEE Transactions on Automatic Control*, 59(1):18–33, 2014.

[106] P. Patrinos, L. Stella, and A. Bemporad. Douglas-Rachford splitting: Complexity estimates and accelerated variants. In *The 53rd IEEE Annual Conference on Decision and Control*, pages 4234–4239, 2014.

[107] P. Patrinos, L. Stella, and A. Bemporad. Forward-backward truncated Newton methods for convex composite optimization. *arXiv:1402.6655*, 2014.

[108] T. Pock and A. Chambolle. Diagonal preconditioning for first order primal-dual algorithms in convex optimization. In *The IEEE International Conference on Computer Vision*, pages 1762–1769. IEEE, 2011.

[109] B. T. Polyak. Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics* , 4(5):1–17, 1964.

[110] M. J. D. Powell. A method for nonlinear constraints in minimization problems. In R. Fletcher, editor, *Optimization*, pages 283–298. Academic Press, 1969.

[111] Y. Pu, M. N. Zeilinger, and C. N. Jones. Inexact Fast Alternating Minimization Algorithm for Distributed Model Predictive Control. In *The 53rd IEEE Annual Conference on Decision and Control*, pages 5915–5921, 2014.

[112] A. U. Raghunathan and S. Di Cairano. Infeasibility detection in alternating direction method of multipliers for convex quadratic programs. In *The 53rd IEEE Annual Conference on Decision and Control*, pages 5819–5824, 2014.

[113] A. U. Raghunathan and S. Di Cairano. Optimal step-size selection in alternating direction method of multipliers for convex quadratic programs and model predictive control. In *International Symposium on Mathematical Theory of Networks and Systems*, pages 807–814, 2014.

[114] H. Raguet, J. Fadili, and G. Peyré. A Generalized Forward-Backward Splitting. *SIAM Journal on Imaging Sciences*, 6(3):1199–1226, 2013.

[115] S. Richter. Computational complexity certification of gradient methods for real-time model predictive control. PhD thesis, ETH Zurich, 2012.

[116] R. T. Rockafellar. Augmented Lagrangians and Applications of the Proximal Point Algorithm in Convex Programming. *Mathematics of Operations Research*, 1976.

[117] R. T. Rockafellar. Monotone Operators and the Proximal Point Algorithm. *SIAM Joiurnal on Control and Optimization*, 1976.

[118] R. T. Rockafellar and R. Rockajellm. *Monotone Operators Associated with Saddle Functions and Minimax Problems*, pages 241–250. American Mathematical Society, 1970.

[119] E. K. Ruy and S. Boyd. A Primer on Monotone Operator Methods. *Applied and Computational Mathematics*, 15(1), 2016.

[120] R. Scattolini. Architectures for distributed and hierarchical model predictive control - a review. *Journal of Process Control*, 19(5):723–731, 2009.

[121] M. Schmidt, N. L. Roux, and F. Bach. Convergence rates of inexact proximal-gradient methods for convex optimization. In *The 25th Annual Conference on Neural Information Processing Systems*, pages 6819–6824, 2011.

[122] R. Shefi and M. Teboulle. Rate of Convergence Analysis of Decomposition Methods Based on the Proximal Method of Multipliers for Convex Minimization. *SIAM Journal on Optimization*, 2014.

[123] N. Z. Shor, K. C. Kiwiel, and A. Ruszcayǹski. *Minimization Methods for Non-differentiable Functions*. Springer-Verlag New York, Inc., 1985.

[124] R. Sinkhorn and P. Knopp. Concerning nonnegative matrices and doubly stochastic matrices. *Pacific Journal of Mathematics*, 21(2):343–348, 1967.

[125] G. Stathopoulos, M. Korda, and C. N. Jones. Solving the infinite-horizon constrained LQR problem using splitting techniques. In *The 19th IFAC World Congress*, 2014.

[126] G. Stathopoulos, M. Korda, and C. N. Jones. Solving the Infinite-Horizon Constrained LQR Problem using Accelerated Dual Proximal Methods. *To appear in IEEE Transactions on Automatic Control*, 2016.

[127] T. H. Summers and J. Lygeros. Distributed model predictive consensus via the alternating direction method of multipliers. In *The 50th Annual Allerton Conference on Communication, Control and Computing*, pages 79–84, 2012.

[128] B. F. Svaiter. On weak convergence of the Douglas-Rachford method. *SIAM Journal on Control and Optimization*, 2011.

[129] Q. Tran-Dinh, I. Necoara, and M. Diehl. A dual decomposition algorithm for separable nonconvex optimization using the penalty framework. In *The 52nd IEEE Annual Conference on Decision and Control*, 2013.

[130] Q. Tran-Dinh, I. Necoara, and M. Diehl. Fast inexact decomposition algorithms for large-scale separable convex optimization. *Optimization*, pages 1–32, 2015.

[131] P. Tseng. Applications of splitting algorithm to decomposition in convex programming and variational inequalities. *SIAM Journal on Control and Optimization*, 1991.

[132] P. Tseng. On accelerated proximal gradient methods for convex-concave optimization. *submitted to SIAM Journal on Optimization*, 2008.

[133] E. Ullmann. A Matlab toolbox for C-code generation for first order methods. Master's thesis, ETH Zurich, 2011.

[134] B. C. Vũ. A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Advances in Computational Mathematics*, 38(3):667–681, 2013.

[135] L. Vandenberghe. Optimization methods for large-scale systems. UCLA EE 236C lecture notes, 2010.

[136] D. S. Watkins. *Fundamentals of matrix computations*. Pure and applied mathematics. Wiley-Interscience, New York, 2010.

[137] X. Xu, P. F. Hung, and Y. Ye. A simplified homogeneous and self-dual linear programming algorithm and its implementation. *Annals of Operations Research*, 62(1):151–171, 1996.

[138] Y. Ye, M. J. Todd, and S. Mizuno. An $O(\sqrt{nL})$-iteration homogeneous and self-dual linear programming algorithm. *Mathematics of Operations Research*, 19:53–67, 1994.

[139] M. Zhu and T. Chan. An efficient primal-dual hybrid gradient algorithm for total variation image restoration. UCLA CAM Report, 2008.