Full text available at: http://dx.doi.org/10.1561/2400000027

Information Relaxations and Duality in Stochastic Dynamic Programs: A Review and Tutorial

Other titles in Foundations and Trends® in Optimization

Atomic Decomposition via Polar Alignment: The Geometry of Structured Optimization
Zhenan Fan, Halyun Jeong, Yifan Sun and Michael P. Friedlander
ISBN: 978-1-68083-742-1

Optimization Methods for Financial Index Tracking: From Theory to Practice
Konstantinos Benidis, Yiyong Feng and Daniel P. Palomar
ISBN: 978-1-68083-464-2

The Many Faces of Degeneracy in Conic Optimization
Dmitriy Drusvyatskiy and Henry Wolkowicz
ISBN: 978-1-68083-390-4

Information Relaxations and Duality in Stochastic Dynamic Programs: A Review and Tutorial

David B. Brown
Fuqua School of Business
Duke University
USA
dbbrown@duke.edu

James E. Smith
Tuck School of Business
Dartmouth College
USA
jim.smith@dartmouth.edu

now
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

D. B. Brown and J. E. Smith. *Information Relaxations and Duality in Stochastic Dynamic Programs: A Review and Tutorial*. Foundations and Trends® in Optimization, vol. 5, no. 3, pp. 246–339, 2022.

ISBN: 978-1-68083-963-0

Foundations and Trends® in Optimization
Volume 5, Issue 3, 2022

Editorial Scope

Topics

Foundations and Trends® in Optimization publishes survey and tutorial articles in 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, and 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, 2022, Volume 5, 4 issues. ISSN paper version 2167-3888. ISSN online version 2167-3918. Also available as a combined paper and online subscription.

Contents

Information Relaxations and Duality in Stochastic Dynamic Programs: A Review and Tutorial

David B. Brown[1] and James E. Smith[2]

[1] *Fuqua School of Business, Duke University, USA; dbbrown@duke.edu*
[2] *Tuck School of Business, Dartmouth College, USA; jim.smith@dartmouth.edu*

ABSTRACT

In this monograph, we provide an overview of the information relaxation approach for calculating performance bounds in stochastic dynamic programs (DPs). The technique involves (1) relaxing the temporal feasibility (or nonanticipativity) constraints so the decision-maker (DM) has additional information before making decisions, and (2) incorporating a penalty that punishes the DM for violating the temporal feasibility constraints. The goal of this monograph is to provide a self-contained overview of the key theoretical results of the information relaxation approach as well as a review of research that has successfully used these techniques in a broad range of applications. We illustrate the information relaxation approach on applications in inventory management, assortment planning, and portfolio optimization.

Keywords: stochastic dynamic programs; information relaxations; approximate dynamic programming

David B. Brown and James E. Smith (2022), "Information Relaxations and Duality in Stochastic Dynamic Programs: A Review and Tutorial", Foundations and Trends® in Optimization: Vol. 5, No. 3, pp 246–339. DOI: 10.1561/2400000027.

1

Introduction

In principle, dynamic programming (DP) provides a powerful framework for modeling complex decision problems where uncertainty is resolved and decisions are made over time. However, in practice, the "curse of dimensionality" – the fact that the size of the state space typically grows exponentially in the number of state variables considered – severely limits the complexity of problems that can be solved using DP methods. In contrast, Monte Carlo simulation methods typically scale well with the number of state variables considered and, given a control policy, it is not difficult to simulate a complex dynamic system with many uncertainties. Simulating with a feasible policy provides a lower bound on the expected reward (or upper bound on the expected cost) with an optimal policy, but Monte Carlo simulation typically does not provide a good way to identify an optimal policy or provide a *performance bound*, i.e., an upper bound on the expected reward (or lower bound on expected cost) with an optimal policy. Consequently, researchers and practitioners using heuristic control policies often wonder how good a policy is and whether it is "good enough" to use in practice.

In this monograph, we review the information relaxation approach for calculating performance bounds in stochastic DPs, following Brown, Smith, and Sun (2010) (hereafter BSS (2010)) and related works. The information relaxation approach consists of two elements: (1) we relax the temporal feasibility (or nonanticipativity) constraints that require decisions to depend only on the information available at the time a decision is made and (2) we impose a penalty that punishes violations of these relaxed constraints. Relaxing the temporal feasibility constraint allows the decision-maker (DM) to make decisions using more information than is truly available and thus leads to an upper bound on value. Without any penalty for using this additional information, the resulting performance bound is often quite weak. Informally, we say a penalty is dual feasible if it does not penalize temporally feasible policies. Though there exists a dual feasible penalty that provides a bound that is equal to the optimal value for the primal DP (i.e., strong duality holds), these ideal penalties are based on the optimal value function, which is typically not available in the applications of interest – if the value function were available, we would not need performance bounds. In practice, we typically use penalties based on approximate value functions to generate performance bounds.

By relaxing the temporal feasibility constraints, we can often greatly simplify the problem by reducing a complex stochastic DP to a series of scenario-specific deterministic optimization problems solved within a Monte Carlo simulation. To illustrate this idea, we will consider a dynamic assortment problem, where a retailer decides which products to offer for sale ("display") when facing uncertain demand, drawn from a distribution with unknown parameters. Here a perfect information relaxation assumes the DM knows all demands and distribution parameters before deciding which products to display. With this information, the problem of choosing products to display is a deterministic optimization problem. The information relaxation performance bound can be estimated using Monte Carlo simulation by repeatedly drawing random demands and distributions and averaging the results. We can also consider imperfect information relaxations where, for example, the DM knows the demand distribution but not the realized demands.

1.1 Outline of the Monograph

The goal of this monograph is to provide a summary of the key ideas of information relaxation methods for stochastic DPs and demonstrate their use in several examples. The idea is to provide a "one-stop-shop" (or at least a "first stop") for researchers seeking to learn the key ideas and tools for using information relaxation methods.

Following a brief history and literature review in Section 1.2, in Sections 2–4, we describe the theory associated with the information relaxation approach. Section 2 establishes the basic framework and Section 3 presents the key theoretical results, both following BSS (2010). In Section 4, we study DPs with a convex structure and show how the use of "gradient" penalties leads to inner problems that are easy to solve; this section draws on Brown and Smith (2014b). Before considering specific examples in detail, in Section 5 we provide a summary of the information relaxation approach and advice on how to proceed in applications.

In Sections 6–8, we consider illustrative applications. Section 6 illustrates the basic results and methods in a simple inventory management example with and without uncertainty about the state of the world; this problem is simple enough that it can be solved to optimality, allowing us to compare the information relaxation performance bounds to the optimal value. In Section 7, we consider a more complex example based on the dynamic assortment problem studied in Caro and Gallien (2007); our discussion draws on Brown and Smith (2020). In Section 8, we illustrate the use of gradient penalties (introduced in Section 4) on dynamic portfolio optimization problems with transaction costs, building on the model and results of Brown and Smith (2011).

A reader eager to see examples could read Section 6 describing the inventory example and perhaps Section 7 on the dynamic assortment example in parallel with Sections 2–3 describing the general framework and main results. Similarly, one could read Section 8 describing the portfolio optimization example in parallel with Section 4 describing the theory for convex DPs.

In Sections 9 and 10, we briefly review other work that has advanced information relaxation methodology and successfully applied the information relaxation approach. Section 11 offers a few concluding remarks and suggestions for future research.

1.2 History and Literature Review

Our interest in information relaxation methods for DPs began with BSS (2010). As discussed in BSS (2010), we were motivated by the need to evaluate the quality of heuristic policies in applications. As an example of one such application, Lai *et al.* (2010) consider the problem of managing natural gas storage over time in the presence of stochastic price dynamics. In the model, the merchant may inject or withdraw natural gas in each period. This problem is naturally formulated as a stochastic DP but is challenging because the natural gas forward curve involves a high-dimensional model that leads to a very large state space for the stochastic DP. Lai *et al.* (2010) develop some policies based on approximations of the value function. Naturally, one might wonder how good these policies are: could one do better with other – perhaps more complex – policies or is the current one "good enough?" Such questions are common when studying complex dynamic models.

The information relaxation approach to calculating performance bounds for DPs in BSS (2010) was inspired by Haugh and Kogan (2004)'s "duality approach" for placing bounds on the value of an American option; Rogers (2002) independently proposed a similar approach, also applied to option pricing. Both Haugh and Kogan (2004) and Rogers (2002) consider the use of what we call perfect information relaxations and establish their main results using martingale arguments. Haugh and Kogan (2004) propose a particular method for generating penalties or, in their terminology, "dual martingales" based on approximate value functions and demonstrate the use of this method in high-dimensional option pricing problems. Andersen and Broadie (2004) propose an alternative method for generating dual martingales based on approximate policies. Glasserman (2003) provides a nice overview of this work. Subsequent work (e.g., Meinshausen and Hambly, 2004; Schoenmakers, 2012) in financial engineering extended these dual methods to multiple stopping problems, for example, derivatives with several exercise rights such as "swing options" in electricity markets or "chooser caps" in interest rate markets.

BSS (2010) generalizes Haugh and Kogan (2004), Rogers (2002), and Andersen and Broadie (2004) in several ways. First, rather than focusing exclusively on option pricing problems, it considers general stochastic DPs. Second, rather than focusing exclusively on perfect information relaxations, it considers general information relaxations. BSS (2010) also presents a general method for constructing good penalties that includes and extends the methods proposed by Haugh and Kogan (2004) and Andersen and Broadie (2004).

The idea of relaxing temporal feasibility (or nonanticipativity) constraints has also been studied in the stochastic programming literature (see, for example, Rockafellar and Wets, 1976; Shapiro *et al.*, 2009). The stochastic programming formulation typically requires the reward functions and set of feasible actions to be convex and the penalties to be linear functions of the actions; they consider only perfect information relaxations. In contrast, the information relaxation approach described here allows general reward functions and action spaces, allows general penalty functions, and considers imperfect as well as perfect information relaxations. The connection between the stochastic programming formulation and the information relaxation approach is discussed in more detail in Appendix B of BSS (2010). That appendix also discusses connections between the information relaxation results and standard Lagrangian duality results for linear programs (LPs). In the LP formulation of the information relaxation problem, the decision variables are mixing weights on policies and the objectives and constraints (including the temporal feasibility constraints) are linear functions of these decision variables. In this LP formulation, the penalties of the information relaxation approach correspond to the Lagrange multipliers associated with the temporal feasibility constraints. However, as shown in Section 3, we can also use simple, direct arguments to establish the key information relaxation duality results without considering mixed policies or LP duality results.

We view this information relaxation approach as a complement to the use of simulation methods and approximate dynamic programming methods for studying DPs (see, for example, Bertsekas and Tsitsiklis, 1996; de Farias and Van Roy, 2003; Powell, 2007; Adelman and Mersereau, 2008). As mentioned earlier, given a candidate policy (perhaps identified using a heuristic reasoning or using approximate DP techniques), we can use standard simulation techniques to estimate the expected value with this policy and thereby generate a lower bound on the expected reward with an optimal policy. The information relaxation performance bound can often be estimated with little additional effort in the same simulation and, as discussed, can help determine whether the proposed policy is "good enough" or if we should continue searching for a better policy, perhaps using more complex ADP techniques.

"Hindsight bounds" – perfect information bounds with no penalties – are popular in the theoretical computer science literature (see, for example, Feldman *et al.*, 2010). These bounds are used to establish theoretical guarantees, for example showing that an algorithm is guaranteed to produce a solution that is within, say, 50% of the optimal solution. As we will see in our numerical examples, perfect information bounds with no penalty are often quite weak. Balseiro and Brown (2019) show how one can incorporate penalties in such theoretical studies and improve the theoretical guarantees to show, for example, that an algorithm or policy is asymptotically optimal in a given setting (see Section 9 for more).

References

Adelman, D. and A. J. Mersereau (2008). “Relaxations of weakly coupled stochastic dynamic programs”. *Operations Research*. 56(3): 712–727.

Ahuja, R. K., T. L. Magnanti, and J. B. Orlin (1988). *Network Flows*. Cambridge, Mass.: Alfred P. Sloan School of Management.

Andersen, L. M. and M. Broadie (2004). “Primal-dual simulation algorithm for pricing multidimensional American options”. *Management Science*. 50(9): 1222–1234.

Balseiro, S. R. and D. B. Brown (2019). “Approximations to stochastic dynamic programs via information relaxation duality”. *Operations Research*. 67(2): 577–597.

Balseiro, S. R., D. B. Brown, and C. Chen (2018). “Static routing in stochastic scheduling: Performance guarantees and asymptotic optimality”. *Operations Research*. 66(6): 1641–1660.

Bender, C., C. Gärtner, and N. Schweizer (2018). “Pathwise dynamic programming”. *Mathematics of Operations Research*. 43(3): 965–995.

Bernstein, F., Y. Li, and K. Shang (2016). “A simple heuristic for joint inventory and pricing models with lead time and backorders”. *Management Science*. 62(8): 2358–2373.

Bertsekas, D. P. (2017). *Dynamic Programming and Optimal Control*. 4th edn. Vol. 1. Athena Scientific.

Bertsekas, D. P. and J. N. Tsitsiklis (1996). *Neuro-Dynamic Programming*. Athena Scientific.

Bertsimas, D. and J. N. Tsitsiklis (1997). *Introduction to Linear Optimization*. Athena Scientific Series in Optimization and Neural Computation, 6. Athena Scientific.

Broadie, M. and W. Shen (2016). "High-dimensional portfolio optimization with transaction costs". *International Journal of Theoretical and Applied Finance*. 19(4).

Broadie, M. and W. Shen (2017). "Numerical solutions to dynamic portfolio problems with upper bounds". *Computational Management Science*. 14(2): 215–227.

Brown, D. B. and M. B. Haugh (2017). "Information relaxation bounds for infinite horizon Markov decision processes". *Operations Research*. 65(5): 1355–1379.

Brown, D. B. and J. E. Smith (2011). "Dynamic portfolio optimization with transaction costs: Heuristics and dual bounds". *Management Science*. 57(10): 1752–1770.

Brown, D. B. and J. E. Smith (2013). "Optimal sequential exploration: Bandits, clairvoyants, and wildcats". *Operations Research*. 61(3): 644–665.

Brown, D. B. and J. E. Smith (2014a). "Dynamic portfolio optimization with transaction costs: Heuristics and dual bounds (addendum on gradient penalties)".

Brown, D. B. and J. E. Smith (2014b). "Information relaxations, duality, and convex stochastic dynamic programs". *Operations Research*. 62(6): 1394–1415.

Brown, D. B. and J. E. Smith (2020). "Index policies and performance bounds for dynamic selection problems". *Management Science*. 66(7): 3029–3050.

Brown, D. B., J. E. Smith, and P. Sun (2010). "Information relaxations and duality in stochastic dynamic programs". *Operations Research*. 58(4-part-1): 785–801.

Caro, F. and J. Gallien (2007). "Dynamic assortment with demand learning for seasonal consumer goods". *Management Science*. 53(2): 276–292.

Casella, G. and R. L. Berger (2002). *Statistical Inference*. 2nd edition. *Duxbury Advanced Series*. New Delhi: Wadsworth.

Chandramouli, S. S. and M. B. Haugh (2012). "A unified approach to multiple stopping and duality". *Operations Research Letters*. 40(4): 258–264.

Chen, N., X. Ma, Y. Liu, and W. Yu (2020). "Information relaxation and a duality-driven algorithm for stochastic dynamic programs". *arXiv preprint arXiv:2007.14295*.

de Farias, D. P. and B. Van Roy (2003). "The linear programming approach to approximate dynamic programming". *Operations Research*. 51(6): 850–865.

Dean, B. C., M. X. Goemans, and J. Vondrák (2008). "Approximating the stochastic Knapsack problem: The benefit of adaptivity". *Mathematics of Operations Research*. 33(4): 945–964.

Desai, V., V. Farias, and C. Moallemi (2011). "Bounds for Markov decision processes". *Reinforcement Learning and Approximate Dynamic Programming for Feedback Control*: 452–473. Ed. by F. L. Lewis and D. Liu.

Desai, V., V. F. Farias, and C. C. Moallemi (2012). "Pathwise optimization for optimal stopping problems". *Management Science*. 58(12): 2292–2308.

Devalkar, S., R. Anupindi, and A. Sinha (2011). "Integrated optimization of procurement, processing, and trade of commodities". *Operations Research*. 59(6): 1369–1381.

El Shar, I. and D. Jiang (2020). "Lookahead-bounded Q-learning". In: *International Conference on Machine Learning*. PMLR. 8665–8675.

Farahani, M. H., M. Dawande, and G. Janakiraman (2020). "Order now, pickup in 30 minutes: Managing queues with static delivery guarantees". Forthcoming in *Operations Research*, URL: https://pubsonline.informs.org/doi/10.1287/opre.2021.2203.

Federgruen, A., D. Guetta, and G. Iyengar (2015). "Information relaxation-based lower bounds for the stochastic lot sizing problem with advanced demand information". Technical report, Working paper, Columbia University.

Feldman, J., M. Henzinger, N. Korula, V. S. Mirrokni, and C. Stein (2010). "Online stochastic packing applied to display ad allocation". In: *Proceedings of the 18th Annual European Conference on Algorithms: Part I. ESA'10*. Springer-Verlag. 182–194.

Florian, M., J. K. Lenstra, and A. Rinnooy Kan (1980). "Deterministic production planning: Algorithms and complexity". *Management Science*. 26(7): 669–679.

Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*. Vol. 53. Springer Science & Business Media.

Goodson, J. C., J. W. Ohlmann, and B. W. Thomas (2013). "Rollout policies for dynamic solutions to the multivehicle routing problem with stochastic demand and duration limits". *Operations Research*. 61(1): 138–154.

Haugh, M. B. and L. Kogan (2004). "Pricing American options: A duality approach". *Operations Research*. 52(2): 258–270.

Haugh, M. B. and O. R. Lacedelli (2019). "Information relaxation bounds for partially observed Markov decision processes". *IEEE Transactions on Automatic Control*. 65(8): 3256–3271.

Haugh, M. B. and A. E. Lim (2012). "Linear-quadratic control and information relaxations". *Operations Research Letters*. 40(6): 521–528.

Haugh, M., G. Iyengar, and C. Wang (2016). "Tax-aware dynamic asset allocation". *Operations Research*. 64(4): 849–866.

Haugh, M. and C. Wang (2014a). "Dynamic portfolio execution and information relaxations". *SIAM Journal on Financial Mathematics*. 5(1): 316–359.

Haugh, M. and C. Wang (2014b). "Information relaxations and dynamic zero-sum games". *arXiv preprint arXiv:1405.4347*.

Hawkins, J. T. (2003). "A Langrangian decomposition approach to weakly coupled dynamic optimization problems and its applications". *PhD thesis*. Massachusetts Institute of Technology.

Henderson, S. G. and P. W. Glynn (2002). "Approximating martingales for variance reduction in Markov Process Simulation". *Mathematics of Operations Research*. 27(2): 253–271.

Hinz, J. and J. Yee (2017). "Stochastic switching for partially observable dynamics and optimal asset allocation". *International Journal of Control*. 90(3): 553–565.

Hinz, J. and J. Yee (2018). "Optimal forward trading and battery control under renewable electricity generation". *Journal of Banking & Finance*. 95: 244–254.

Jiang, D. R., L. Al-Kanj, and W. B. Powell (2020). "Optimistic Monte Carlo tree search with sampled information relaxation dual bounds". *Operations Research*. 68(6): 1678–1697.

Kogan, L. and I. Mitra (2019). "Near-rational equilibria in heterogeneous-agent models: A verification method". Working paper. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3465120.

Kullman, N. D., J. C. Goodson, and J. E. Mendoza (2021). "Electric vehicle routing with public charging stations". *Transportation Science*. 55(3): 637–659.

Lai, G., F. Margot, and N. Secomandi (2010). "An approximate dynamic programming approach to benchmark practice-based heuristics for natural gas storage valuation". *Operations Research*. 58(3): 564–582.

Lin, Q., S. Nadarajah, and N. Soheili (2020). "Revisiting approximate linear programming: Constraint-violation learning with applications to inventory control and energy storage". *Management Science*. 66(4): 1544–1562.

Lu, T., C.-Y. Lee, and L.-H. Lee (2020). "Coordinating pricing and empty container repositioning in two-depot shipping systems". *Transportation Science*. 54(6): 1697–1713.

Luenberger, D. G. and Y. Ye (2016). *Linear and Nonlinear Programming: Fourth Edition*. Springer.

Marla, L. and A. Bassamboo (2020). "Information relaxation bounds for evaluating ambulance dispatch and allocation policies". Working paper. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3519306.

Mei, X. and F. J. Nogales (2018). "Portfolio selection with proportional transaction costs and predictability". *Journal of Banking & Finance*. 94: 131–151.

Meinshausen, N. and B. M. Hambly (2004). "Monte Carlo methods for the valuation of multiple-exercise options". *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*. 14(4): 557–583.

Min, S., C. Maglaras, and C. C. Moallemi (2019). "Thompson sampling with information relaxation penalties". In: *Advances in Neural Information Processing Systems*. 3549–3558.

Nadarajah, S., F. Margot, and N. Secomandi (2015). "Relaxations of approximate linear programs for the real option management of commodity storage". *Management Science*. 61(12): 3054–3076.

Nadarajah, S., F. Margot, and N. Secomandi (2017). "Comparison of least squares Monte Carlo methods with applications to energy real options". *European Journal of Operational Research*. 256(1): 196–204.

Nadarajah, S. and N. Secomandi (2018). "Merchant energy trading in a network". *Operations Research*. 66(5): 1304–1320.

Powell, W. B. (2007). *Approximate Dynamic Programming: Solving the curses of dimensionality*. Vol. 703. John Wiley & Sons.

Puterman, M. L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. 1st edn. USA: John Wiley & Sons, Inc.

Rockafellar, R. T. and R.-B. Wets (1976). "Nonanticipativity and L 1-martingales in stochastic optimization problems". In: *Stochastic Systems: Modeling, Identification and Optimization, II*. Springer. 170–187.

Rogers, L. (2002). "Monte Carlo valuation of American options". *Mathematical Finance*. 12: 271–286.

Rogers, L. (2007). "Pathwise stochastic optimal control". *SIAM Journal on Control and Optimization*. 46: 1116–1132.

Schoenmakers, J. (2012). "A pure martingale dual for multiple stopping". *Finance and Stochastics*. 16(2): 319–334.

Secomandi, N. (2015). "Merchant commodity storage practice revisited". *Operations Research*. 63(5): 1131–1143.

Shapiro, A., D. Dentcheva, and A. Ruszczyński (2009). *Lectures on Stochastic Programming: Modeling and Theory*. MOS-SIAM Series on Optimization 9. Society for Industrial and Applied Mathematics (SIAM).

Shumsky, R. A., J. E. Smith, A. G. Hoen, and M. Gilbert (2021). "Allocating COVID-19 vaccines: Save one for the second dose?" Working paper, Dartmouth College.

Song, J.-S. and P. Zipkin (1993). "Inventory control in a fluctuating demand environment". *Operations Research*. 41(2): 351–370.

Talluri, K. and G. van Ryzin (1998). "An analysis of bid-price controls for network revenue management". *Management Science*. 44(11): 1577–1593.

Thompson, W. R. (1933). "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples". *Biometrika*. 25(3/4): 285–294.

Treharne, J. T. and C. R. Sox (2002). "Adaptive inventory control for nonstationary demand and partial information". *Management Science*. 48(5): 607–624.

Trivella, A., D. Mohseni Taheri, and S. Nadarajah (2018). "Meeting corporate renewable power targets". Forthcoming in *Management Science*, URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3294724.

Trivella, A., S. Nadarajah, S.-E. Fleten, D. Mazieres, and D. Pisinger (2021). "Managing shutdown decisions in merchant commodity and energy production: A social commerce perspective". *Manufacturing and Service Operations Management*. 23(2): 311–330.

Wang, Z. and A. J. Mersereau (2017). "Bayesian inventory management with potential change-points in demand". *Production and Operations Management*. 26(2): 341–359.

Whittle, P. (1980). "Multi-armed bandits and the Gittins index". *Journal of the Royal Statistical Society: Series B (Methodological)*. 42(2): 143–149.

Yang, B., S. Nadarajah, and N. Secomandi (2020). "Pathwise optimization for merchant energy production". Working paper, URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3510676.

Ye, F. and E. Zhou (2015). "Information relaxation and dual formulation of controlled Markov diffusions". *IEEE Transactions on Automatic Control*. 60(10): 2676–2691.

Zipkin, P. (2008). "On the structure of lost-sales inventory models". *Operations Research*. 56(4): 937–944.

Zipkin, P. H. (2000). *Foundations of Inventory Management*. Boston: McGraw Hill.