

Two-Stage Dual Dynamic Programming with Application to Nonlinear Hydro Scheduling

Journal Article**Author(s):**

Flamm, Benjamin; Eichler, Annika ; Warrington, Joseph; Lygeros, John 

Publication date:

2021-01

Permanent link:

<https://doi.org/10.3929/ethz-b-000388081>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

Originally published in:

IEEE Transactions on Control Systems Technology 29(1), <https://doi.org/10.1109/TCST.2019.2961645>

Funding acknowledgement:

787845 - Optimal control at large (EC)

Two-Stage Dual Dynamic Programming with Application to Nonlinear Hydro Scheduling

Benjamin Flamm, Annika Eichler, *Member, IEEE*, Joseph Warrington, *Member, IEEE*,
and John Lygeros, *Fellow, IEEE*

Abstract—We present an approximate method for solving nonlinear control problems over long time horizons, in which the full nonlinear model is preserved over an initial part of the horizon, while the remainder of the horizon is modeled using a linear relaxation. As this approximate problem may still be too large to solve directly, we present a Benders decomposition-based solution algorithm that iterates between solving the nonlinear and linear parts of the horizon. This extends the Dual Dynamic Programming approach commonly employed for optimization of linearized hydro power systems. We prove that the proposed algorithm converges after a finite number of iterations, even when the nonlinear initial stage problems are solved inexactly. We also bound the suboptimality of the split-horizon method with respect to the original nonlinear problem, in terms of the properties of a map between the linear and nonlinear state-input trajectories. We then apply this method to a case study concerning a multiple reservoir hydro system, approximating the nonlinear head effects in the second stage using McCormick envelopes. We demonstrate that near-optimal solutions can be obtained in a shrinking horizon setting when the full nonlinear model is used for only a short initial section of the horizon. For this example, the approach is shown to be more practical than both conventional dynamic programming and a multi-cell McCormick envelope approximation from literature.

Index Terms—dual dynamic programming, optimal control, nonlinear model predictive control, hydro optimization.

I. INTRODUCTION

THE control of energy storage devices over a long planning horizon is gaining in importance, as the number of renewable and variable energy sources on the electric power grid increases. U.S. installed wind and solar generation capacity increased by 11% and 52% respectively in 2016 [1]. Energy storage is particularly well-suited to complement the resulting unpredictable power generation. Long-term energy storage exists in different forms, with varying technical maturity and efficiency [2]. Examples include natural gas for heating, compressed air electrical storage, ground thermal energy storage, and hydro reservoir systems; the latter have been extensively utilized due to their large storage capacity, technological maturity, and proliferation. Key challenges for the integration of seasonal storage into next generation energy systems are the long horizons and nonlinear system dynamics, which can render long-term control of the storage computationally difficult.

We typically wish to operate long-term storage devices in response to underlying energy demand or supply patterns that

are cyclical over a period of months or years. Examples of such patterns include yearly snow melt, as well as seasonal heating and cooling demands.

Several methods have been proposed to tackle nonlinear seasonal storage problems. Many involve approximating the nonlinear dynamics through modeling simplifications and heuristic methods such as timescale separation [3]. Specifically for hydro optimization, a fundamental difficulty is the presence of nonlinear head effects when converting between stored water and electrical energy. Many papers ignore nonlinear head effects and represent energy conversion as a constant efficiency [4], [5]. Methods that do account for head effects usually generate convex hulls of the power production function, either by fitting a set of piecewise linear constraints to the true model [6], [7], or by rewriting nonlinear terms using convex approximations such as McCormick envelopes [8].

Other methods take advantage of improvements in computing power, which allows increasingly large nonlinear control problems to be solved exactly. Approaches that use nonlinear models usually consider only short horizons [9]. Techniques such as spatial branch-and-bound and dynamic programming allow for the solution of nonlinear problems to high precision, but scale poorly with the state space size.

Dual dynamic programming (DDP), also known as multistage Benders decomposition, was introduced as a method for seasonal hydro storage scheduling in [10], but has since been used primarily for solving linear approximations of these problems. DDP was extended to convex problems using generalized Benders decomposition in [11]. Roughly speaking, DDP bounds the value function of a convex problem by a piecewise affine function of the initial state. Since DDP cannot solve problems with nonconvex value functions, convex approximations of the model are required when using DDP for nonlinear problems. Recent extensions of this approach have considered integer programs [12], locally-valid Benders cuts [13], and polynomial-based moment relaxation [14].

In [15], a method was introduced to treat general receding horizon nonlinear optimal control problems. The full problem horizon was split into short- and long-term parts, with high model accuracy in the short term, and reduced model accuracy in the long term. Here we expand this approach by introducing several convergence and optimality results related to solving this approximation of the underlying nonlinear problem. Unlike [15], in Theorem 1 we provide a guarantee on convergence of the proposed DDP algorithm, even when the nonlinear first-stage problem is solved suboptimally. Additionally, inspired by [16], where linear subproblems were solved to a known

The authors are with the Automatic Control Laboratory of ETH Zürich, Physikstrasse 30, 8092 Zürich, Switzerland. {flammb, eichlean, warrington, lygeros}@control.ee.ethz.ch

tolerance in DDP, in Theorems 2 and 3 we provide bounds on the suboptimality of the solution of the two-stage approximation. These bounds are a function of the first-stage nonlinear solver tolerance as well as the approximation error between linear and nonlinear models in the second stage. We also demonstrate the generality of the approach by applying it to a different kind of nonlinearity than in [15], involving bilinear as opposed to integer terms. Finally, we quantify how the control horizon (the decision length used each time the model predictive control (MPC) problem is solved) and modeling accuracy affect optimality in simulation.

The underlying stochasticity of real-time optimal control problems is an additional source of computational complexity, and lends itself to the use of DDP. This can be treated by introducing scenarios to capture different potential realizations of the stochastic processes [17]. With additional variables introduced for different scenarios, long-term problems with even simple models can push the limits of tractability. While much of the hydro optimization literature incorporates stochasticity, we consider a deterministic setting here to focus on the underlying computational method and treatment of nonlinearity.

This paper is structured as follows. Section II introduces the general case of the split-horizon approximate problem, as well as a DDP-based algorithm to solve it. The optimality of the solution produced by the algorithm is derived as a function of the optimality of the first-stage solution. In Section III, a bound is found for the error of the two-stage approximation relative to the exact problem. Section IV formulates the nonlinear hydro optimal control problem, providing specific models and approximations of the underlying hydro system. In Section V, we analytically determine the error bound on McCormick envelope approximations of bilinear terms occurring in the problem. Section VI presents simulation results for a representative hydro system, illustrating the computational benefits of the multistage method. Section VII concludes with analysis of possible improvements and extensions to the proposed method.

II. SPLIT-HORIZON PROBLEM AND DDP ALGORITHM

A. General Problem Formulation

We wish to solve a generic discrete-time optimal control problem for a system with dynamics f_t , input and state constraints defined by the set \mathcal{Z}_t , and cost function g_t over a horizon of length T . The state trajectory is denoted by (x_0, \dots, x_T) , while the input trajectory is (u_0, \dots, u_{T-1}) , where $x_t \in \mathbb{R}^{n_t}$ and $u_t \in \mathbb{R}^{m_t}$. The optimal inputs are determined by solving the nonlinear program (NLP)

$$\min_{\substack{x_1, \dots, x_T \\ u_0, \dots, u_{T-1}}} \sum_{t=0}^{T-1} g_t(x_t, u_t) + g_T(x_T) \quad (1a)$$

$$\text{s.t. } x_{t+1} = f_t(x_t, u_t), \quad t = 0, \dots, T-1 \quad (1b)$$

$$(x_t, u_t) \in \mathcal{Z}_t, \quad t = 0, \dots, T-1 \quad (1c)$$

$$x_0 \text{ given}, \quad (1d)$$

where $f_t(x_t, u_t)$, $g_t(x_t, u_t)$, and $g_T(x_T)$ are general nonlinear functions on the set \mathcal{Z}_t . The dynamics, costs, and constraints can all be time-varying. This formulation is general in terms of

variable type. For example, integer variables can be considered by adding constraints to \mathcal{Z}_t .

Note that this is a quite general nonlinear programming problem whose complexity grows rapidly in the dimension of the state and input, as well as the horizon length.

B. Two-Stage Approximation

In [15], to solve (1) over a long horizon, the authors proposed solving an approximate problem consisting of a short initial stage of length T_1 where the exact NLP holds, and a subsequent longer stage of length $T - T_1$ where the problem is approximated by a linear program (LP):

$$\min_{\substack{x_1, \dots, x_{T_1} \\ u_0, \dots, u_{T_1-1} \\ \tilde{x}_{T_1+1}, \dots, \tilde{x}_T \\ \tilde{u}_{T_1}, \dots, \tilde{u}_{T-1}}} \sum_{t=0}^{T_1-1} g_t(x_t, u_t) + \sum_{t=T_1}^{T-1} (c_t^\top \tilde{x}_t + d_t^\top \tilde{u}_t) + c_T^\top \tilde{x}_T \quad (2a)$$

$$\text{s.t. } \left. \begin{aligned} x_{t+1} &= f_t(x_t, u_t), \\ (x_t, u_t) &\in \mathcal{Z}_t, \end{aligned} \right\} t = 0, \dots, T_1 - 1 \quad (2b)$$

$$\left. \begin{aligned} \tilde{x}_{t+1} &= A_t \tilde{x}_t + B_t \tilde{u}_t, \\ E_t \tilde{x}_t + F_t \tilde{u}_t &\leq h_t, \end{aligned} \right\} t = T_1, \dots, T - 1 \quad (2c)$$

$$x_{T_1} = \tilde{x}_{T_1}, \quad (2d)$$

$$x_0 \text{ given}. \quad (2e)$$

Here, the first stage of the problem, from $t = 0$ to $T_1 - 1$, has the exact nonlinear cost function, dynamics, and constraints. In the second stage of the problem, these are replaced with linear costs, dynamics, and constraints.

Note that the state of the approximate linear part at time t is denoted by \tilde{x}_t (instead of x_t), to stress that the second-stage approximation is different from the original problem, potentially incorporating different state variables at each timestep (for example, a relaxation of integer variables in the original problem to real-valued variables in the linear part). However, by (2d), the states x_t and \tilde{x}_t coincide at $t = T_1$.

We now recast the nonlinear approximate problem (2) as an equivalent two-stage problem. The nonlinear first stage can be written as

$$\min_{\substack{x_1, \dots, x_{T_1} \\ u_0, \dots, u_{T_1-1}}} \sum_{t=0}^{T_1-1} g_t(x_t, u_t) + \tilde{G}_{T_1}(x_{T_1}) \quad (3a)$$

$$\text{s.t. } x_{t+1} = f_t(x_t, u_t), \quad t = 0, \dots, T_1 - 1 \quad (3b)$$

$$(x_t, u_t) \in \mathcal{Z}_t, \quad t = 0, \dots, T_1 - 1 \quad (3c)$$

$$x_0 \text{ given}. \quad (3d)$$

In (3a), $\tilde{G}_{T_1}(x_{T_1})$ is a value function that represents the cost of the linear second stage of (2) as a function of the state at the end of the first stage, with

$$\tilde{G}_{T_1}(x_{T_1}) = \min_{\substack{\tilde{x}_{T_1+1}, \dots, \tilde{x}_T \\ \tilde{u}_{T_1}, \dots, \tilde{u}_{T-1}}} \sum_{t=T_1}^{T-1} (c_t^\top \tilde{x}_t + d_t^\top \tilde{u}_t) + c_T^\top \tilde{x}_T \quad (4a)$$

$$\text{s.t. } \tilde{x}_{t+1} = A_t \tilde{x}_t + B_t \tilde{u}_t, \quad (4b)$$

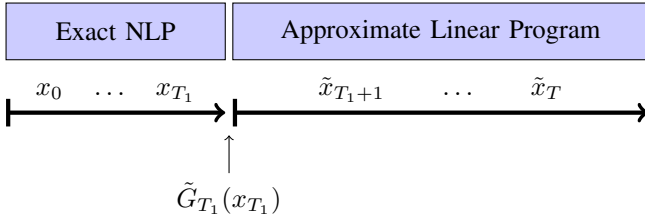


Fig. 1. Diagram of split-horizon method, with exact short-term nonlinear problem (NLP) and intermediate value function representing the approximate linear long-term problem.

$$E_t \tilde{x}_t + F_t \tilde{u}_t \leq h_t, \quad (4c)$$

$$t = T_1, \dots, T-1,$$

$$\tilde{x}_{T_1} = x_{T_1} \text{ given.} \quad (4d)$$

The division of the horizon into the two stages and intermediate value function is depicted in Figure 1.

To ensure that the subsequent theory and analysis is meaningful, we assume that both (1) and (2) are feasible. We additionally make the following assumption, which facilitates the iterative solution of the two-stage problem.

Assumption 1. *The linear second-stage problem (4) is feasible for all x_{T_1} that are feasible for the first-stage problem (3).*

This assumption, known in the mathematical programming literature as *complete recourse*, is not strictly necessary, as the theory related to Benders decomposition used here also works for problems where the second stage is not feasible for a particular x_{T_1} . In that case, Benders feasibility cuts on x_{T_1} can be iteratively added to the first-stage problem (3), as is done in [11]. However, for simplicity we omit this case here.

C. Split-Horizon Approximate DDP Algorithm

In [15], we presented an algorithm that converged in a finite number of iterations to an optimum of the approximate two-stage problem (2), provided the first-stage NLP (3) was solved to global optimality. In certain cases, it may not be desirable or possible to solve (3) to global optimality, due to an algorithmic choice of solution tolerance [16], [18], or constraints on available solution time.

Algorithm 1 presents an adaptation of this setting, where the first-stage NLP (3) is not solved to global optimality. Before analyzing the convergence of this algorithm, as described in Theorem 1, we first explain the value function approximation used therein. We wish to solve a two-stage problem (2) using DDP. To do this, we iteratively solve the two stages of the problem, using DDP to progressively construct a set of lower-bounding hyperplanes for the second-stage value function $\tilde{G}_{T_1}(x_{T_1})$. The lower-bounding hyperplanes are found using the following lemma.

Lemma 1. (Lemma 1 in [15]) *Given a solution to (4) for some x_{T_1} , let λ and ν be the dual variables corresponding to the constraints (4b) and (4c) at timestep $t = T_1$. Then, $(A_{T_1}^\top \lambda + E_{T_1}^\top \nu + c_{T_1})^\top x_{T_1} - \nu^\top h_{T_1}$ is a lower bound on $\tilde{G}_{T_1}(x_{T_1})$ for all x_{T_1} .*

Algorithm 1 Split-horizon DDP with approximate solution to nonlinear first stage

Define: T_1 : length of nonlinear first stage;

x_0 : starting system state

Initialize: $\mathcal{H} = \emptyset$;

$UB = +\infty, LB = -\infty$;

1. Solve first-stage NLP (3), with $\tilde{G}_{T_1}(x_{T_1}) = 0$:
 $((u_0, \dots, u_{T_1-1}), (x_1, \dots, x_{T_1})) \leftarrow$ feasible solution to (3);

while (true) do

2. Solve second-stage linear program (4) with argument x_{T_1} as found in previous step. Returned arguments are $(\tilde{u}_{T_1}, \dots, \tilde{u}_{T-1})$ and $(\tilde{x}_{T_1+1}, \dots, \tilde{x}_T)$. Let dual variables λ and ν correspond to (4b) and (4c) at $t = T_1$;

3. Set $UB = \sum_{t=T_1}^{T-1} (c_t^\top \tilde{x}_t + d_t^\top \tilde{u}_t) + c_T^\top \tilde{x}_T$;

4. **if** $UB = LB$ **then**

 | **break**;

end

5. Let $a = A_{T_1}^\top \lambda + E_{T_1}^\top \nu + c_{T_1}$ and $b = -\nu^\top h_{T_1}$, and update collection of hyperplanes bounding $\tilde{G}_{T_1}(x_{T_1})$ as $\mathcal{H} = \mathcal{H} \cup \{(a, b)\}$;

6. Solve first-stage NLP (3), replacing $\tilde{G}_{T_1}(x_{T_1})$ by $\underline{G}_{T_1}(x_{T_1}) = \max_{(a,b) \in \mathcal{H}} (a^\top x_{T_1} + b)$:
 $((u_0, \dots, u_{T_1-1}), (x_1, \dots, x_{T_1})) \leftarrow$ feasible solution to (3);

7. Set $LB = \underline{G}_{T_1}(x_{T_1})$;

end

Output: $u_{\text{full}} = (u_0, \dots, u_{T_1-1}, \tilde{u}_{T_1}, \dots, \tilde{u}_{T-1})$;

$x_{\text{full}} = (x_1, \dots, x_{T_1}, \tilde{x}_{T_1+1}, \dots, \tilde{x}_T)$;

Lemma 1 allows one to build a collection of lower-bounding hyperplanes for $\tilde{G}_{T_1}(x_{T_1})$. Let $a = A_{T_1}^\top \lambda + E_{T_1}^\top \nu + c_{T_1}$ and $b = -\nu^\top h_{T_1}$ specify the parameters of the lower bound in Lemma 1, found for a particular x_{T_1} . Each iteration of Algorithm 1 solves (4) for a different x_{T_1} , leading to a new hyperplane (a, b) . We can combine the hyperplanes into the set $\mathcal{H} = \{(a, b)\}$, and construct the approximate value function

$$\underline{G}_{T_1}(x_{T_1}) = \max_{(a,b) \in \mathcal{H}} (a^\top x_{T_1} + b), \quad (5)$$

which is itself a lower bound on $\tilde{G}_{T_1}(x_{T_1})$. Thus, each iteration of Algorithm 1 adds a hyperplane constraint to $\underline{G}_{T_1}(x_{T_1})$.

In words, Algorithm 1 starts by solving the first-stage NLP with no information about the second stage. This provides a guess x_{T_1} for the initial state of the second-stage LP. In step 2, the LP is solved using this x_{T_1} . The resulting objective, computed in step 3, is the exact second-stage value function at x_{T_1} . In step 4, if this objective equals the previously-found value function approximation evaluated at x_{T_1} , i.e. the current value function approximation is tight at the new x_{T_1} , then the problem terminates. Otherwise, in the key step 5 of the algorithm, Lemma 1 is used to generate a new lower-bounding hyperplane for the second-stage value function $\tilde{G}_{T_1}(x_{T_1})$. This hyperplane is then incorporated into the second-stage value

function approximation $\underline{G}_{T_1}(x_{T_1})$, which is a lower bound on $\tilde{G}_{T_1}(x_{T_1})$. In step 6, we again solve the first-stage NLP, this time with the updated $\underline{G}_{T_1}(x_{T_1})$. This iterative solution of the first and second stages is repeated until the previously-found value function approximation is tight at the x_{T_1} found when solving the NLP in step 6.

Theorem 1. *Provided Assumption 1 holds, steps 1 and 6 return feasible solutions of (3), and step 2 solves (4) to optimality, then Algorithm 1 terminates in a finite number of iterations, returning a feasible solution of the split-horizon problem. In addition,*

- (a) *If the nonconvex first stage is solved in step 6 to local optimality, then the returned solution is locally optimal for the two-stage problem (2).*
- (b) *If the nonconvex first stage is solved in step 6 to global optimality, then the returned solution is globally optimal for the two-stage problem (2).*

Proof. We first show that Algorithm 1 terminates in a finite number of iterations.

Suppose step 6 (or step 1 for the first iteration) has previously returned a particular feasible solution x_{T_1} of (3). Due to Assumption 1, solving (4) at x_{T_1} returns feasible $(\tilde{u}_{T_1}, \dots, \tilde{u}_{T-1})$ and $(\tilde{x}_{T_1+1}, \dots, \tilde{x}_T)$. Since the second-stage LP (4) is solved to optimality,

$$\tilde{G}_{T_1}(x_{T_1}) = \sum_{t=T_1}^{T-1} (c_t^\top \tilde{x}_t + d_t^\top \tilde{u}_t) + c_T^\top \tilde{x}_T \quad (6)$$

is by definition the exact second-stage value function evaluated at the particular x_{T_1} .

Step 5 adds the dual-feasible vertex (λ, ν) to \mathcal{H} . By the strong duality of (4), after $\underline{G}_{T_1}(x_{T_1})$ is updated in step 6,

$$\underline{G}_{T_1}(x_{T_1}) = \tilde{G}_{T_1}(x_{T_1}). \quad (7)$$

at the particular x_{T_1} .

Since (4) is a linear program, there are a finite number of distinct dual-feasible vertices (λ, ν) . If solving the LP (4) in step 2 yields a dual-feasible vertex that has already been found in a previous iteration of the algorithm, then $\underline{G}_{T_1}(x_{T_1})$ is unchanged. If solving the NLP in step 6 returns the same x_{T_1} as in the previous iteration, (7) already holds. Thus, the value function approximation is tight at x_{T_1} , $UB = LB$, and Algorithm 1 terminates in step 4.

Otherwise, we add a hyperplane to \mathcal{H} parameterized by the new dual-feasible vertex. Since the number of such vertices is finite, one either achieves a complete characterization of $\tilde{G}_{T_1}(x_{T_1})$, or else sets $UB = LB$ before this point.

The resulting first- and second-stage arguments are each feasible, and can be concatenated (as in the algorithm) to form feasible solutions u_{full} and x_{full} of the split-horizon problem.

Proof of (a). We use Algorithm 1 to solve (2) to convergence. This returns $(u_0, \dots, u_{T_1-1}, \tilde{u}_{T_1}, \dots, \tilde{u}_{T-1})$, $(x_1, \dots, x_{T_1}, \tilde{x}_{T_1+1}, \dots, \tilde{x}_T)$, and the approximate value function $\underline{G}_{T_1}(x_{T_1})$.

Each algorithm iteration, we solve an approximation of the first-stage NLP (3), replacing $\tilde{G}_{T_1}(x_{T_1})$ with $\underline{G}_{T_1}(x_{T_1})$, and solving to local optimality. This means that when the algorithm

has converged, there exist $\delta_x > 0$ and $\delta_u > 0$ such that, for all $\hat{x}_t : \|\hat{x}_t - x_t\| \leq \delta_x$ for $t = 1, \dots, T_1$, $\hat{u}_t : \|\hat{u}_t - u_t\| \leq \delta_u$ for $t = 0, \dots, T_1 - 1$,

$$\sum_{t=0}^{T_1-1} g_t(x_t, u_t) + \underline{G}_{T_1}(x_{T_1}) \leq \sum_{t=0}^{T_1-1} g_t(\hat{x}_t, \hat{u}_t) + \underline{G}_{T_1}(\hat{x}_{T_1}). \quad (8)$$

By Lemma 1, since the second stage is an LP,

$$\underline{G}_{T_1}(x_{T_1}) \leq \tilde{G}_{T_1}(x_{T_1}), \quad \forall x_{T_1}. \quad (9)$$

Considering the case where the algorithm has converged to a particular x_{T_1} , and combining (7), (8), and (9),

$$\begin{aligned} \sum_{t=0}^{T_1-1} g_t(x_t, u_t) + \tilde{G}_{T_1}(x_{T_1}) &= \sum_{t=0}^{T_1-1} g_t(x_t, u_t) + \underline{G}_{T_1}(x_{T_1}) \\ &\leq \sum_{t=0}^{T_1-1} g_t(\hat{x}_t, \hat{u}_t) + \underline{G}_{T_1}(\hat{x}_{T_1}) \\ &\leq \sum_{t=0}^{T_1-1} g_t(\hat{x}_t, \hat{u}_t) + \tilde{G}_{T_1}(\hat{x}_{T_1}). \end{aligned}$$

This result holds for all $\hat{x}_t : \|\hat{x}_t - x_t\| \leq \delta_x$ for $t = 1, \dots, T_1$, $\hat{u}_t : \|\hat{u}_t - u_t\| \leq \delta_u$ for $t = 0, \dots, T_1 - 1$. Thus, (x_1, \dots, x_{T_1}) and (u_0, \dots, u_{T_1-1}) are locally optimal solutions to the exact (3).

To complete the proof of (a), note that the linear second-stage problem (4) is solved to optimality for the given first-stage arguments.

Proof of (b) follows from Theorem 1 of [15]. \square

We point out that the result holds regardless of whether the first stage is solved with a deterministic solver. Note that when solving the first stage NLP (3) with a given $\underline{G}_{T_1}(x_{T_1})$, a nondeterministic solver might not return the same feasible solution in subsequent iterations. Nevertheless, the solver will return an x_{T_1} , where either the existing hyperplanes provide a tight bound to the value function (in which case the UB and LB agree in Algorithm 1), or a new hyperplane will be added to $\underline{G}_{T_1}(x_{T_1})$. As there are a limited number of hyperplanes to add, the algorithm will converge in a finite number of iterations.

III. ERROR BOUND ON TWO-STAGE APPROXIMATION

The results in the previous section establish the properties of solutions for the two-stage problem (2) returned by Algorithm 1. Since the original intention was to approximate the nonlinear program (1), one would also like to know how far the optimal solutions of the two-stage approximate problem (2) are from those of the true nonlinear problem (1). Related results are provided in [16], which gives approximation bounds for a setting where each stage of a multi-stage linear program is solved to a certain tolerance. Here, we extend this approach to address the case of (2), where instead of solving a linear program to a known tolerance, we approximate the second stage of the nonlinear problem (1) with a linear program, and solve that linear program exactly.

For a given timestep t , let $\tilde{\mathcal{Z}}_t$ denote the set of arguments of the second-stage LP (4) that satisfy (4c). Recall that \mathcal{Z}_t denotes the set of arguments satisfying (1c) at time t .

Assumption 2. For all $t = T_1, \dots, T-1$, there exists a map $M_t : \tilde{\mathcal{Z}}_t \rightarrow \mathcal{Z}_t$ such that

- (a) There exists a $\delta_t \geq 0$, such that for all $(\tilde{x}_t, \tilde{u}_t) \in \tilde{\mathcal{Z}}_t$,
- $$c_t^\top \tilde{x}_t + d_t^\top \tilde{u}_t \leq g_t(M_t(\tilde{x}_t, \tilde{u}_t)) \leq c_t^\top \tilde{x}_t + d_t^\top \tilde{u}_t + \delta_t.$$
- (b) M_t is surjective, i.e. for all $(x_t, u_t) \in \mathcal{Z}_t$, there exists $(\tilde{u}_t, \tilde{x}_t) \in \tilde{\mathcal{Z}}_t$ such that $M_t(\tilde{x}_t, \tilde{u}_t) = (x_t, u_t)$.

In other words, at each timestep, there is a transformation M_t from feasible arguments of the LP to feasible arguments of the NLP, such that the LP objective is an underestimate of the NLP objective, with a maximum underestimate of δ_t . Additionally, all feasible solutions of the NLP can be found from a feasible solution of the LP through the transformation.

Note that by (2d), at $t = T_1$, $M_{T_1}(\tilde{x}_{T_1}, \tilde{u}_{T_1})$ leaves \tilde{x}_{T_1} unchanged, and maps the input \tilde{u}_{T_1} to a feasible input of the NLP.

We first consolidate the terminology for the exact NLP (1) and two-stage approximation (2):

- $G_0(x_0)$ is the optimal value of the exact NLP (1), as a function of the initial state x_0 .
- $G_{T_1}(x_{T_1})$ is the value of the second part of the exact NLP (1) ($t = T_1, \dots, T-1$), as a function of the intermediate state x_{T_1} . Note that

$$G_0(x_0) = \min_{\substack{u_0, \dots, u_{T_1-1} \\ x_1, \dots, x_{T_1}}} \sum_{t=0}^{T_1-1} g_t(x_t, u_t) + G_{T_1}(x_{T_1}) \quad (10)$$

s.t. $x_{t+1} = f_t(x_t, u_t)$, $t = 0, \dots, T_1 - 1$
 $(x_t, u_t) \in \mathcal{Z}_t$, $t = 0, \dots, T_1 - 1$.

- $\tilde{G}_{T_1}(x_{T_1})$ is the value of the second-stage LP (4), as a function of the intermediate state x_{T_1} .
- $\underline{G}_{T_1}(x_{T_1})$ is the approximate value function for the second-stage LP (4) in Algorithm 1. It consists of a finite number of hyperplanes bounding $\tilde{G}_{T_1}(x_{T_1})$ from below. Hyperplanes are iteratively added to this set via DDP.

Now, we consider the objective for the split-horizon problem (2) after Algorithm 1 has converged, and define it as

$$\underline{G}_0(x_0) = \min_{\substack{u_0, \dots, u_{T_1-1} \\ x_1, \dots, x_{T_1}}} \sum_{t=0}^{T_1-1} g_t(x_t, u_t) + \underline{G}_{T_1}(x_{T_1}) \quad (11)$$

s.t. $x_{t+1} = f_t(x_t, u_t)$, $t = 0, \dots, T_1 - 1$
 $(x_t, u_t) \in \mathcal{Z}_t$, $t = 0, \dots, T_1 - 1$.

Our aim is to bound the difference between $G_0(x_0)$ and $\underline{G}_0(x_0)$.

Theorem 2. If Assumptions 1 and 2 hold, and step 6 is solved to ϵ -optimality in each iteration, then

$$\underline{G}_0(x_0) \leq G_0(x_0) \leq \underline{G}_0(x_0) + \epsilon + \sum_{t=T_1}^{T-1} \delta_t.$$

Proof. By the definitions of (1) and (2), the NLP and split-horizon problems are equivalent for the first part of the horizon, and differ only in the second stage.

By Lemma 1, the Benders decomposition underestimates the objective of the second-stage LP, and $\underline{G}_{T_1}(x_{T_1}) \leq \tilde{G}_{T_1}(x_{T_1})$. By Assumption 2, for any solution to the true NLP, a corresponding solution to the LP approximation exists with a lower objective. Combining these two observations, for any x_{T_1} ,

$$\underline{G}_{T_1}(x_{T_1}) \leq \tilde{G}_{T_1}(x_{T_1}) \leq G_{T_1}(x_{T_1}). \quad (12)$$

If we solve (11) to ϵ -optimality in step 6, returning (u_0, \dots, u_{T_1-1}) and (x_1, \dots, x_{T_1}) , then

$$\underline{G}_0(x_0) \leq \sum_{t=0}^{T_1-1} g_t(x_t, u_t) + \underline{G}_{T_1}(x_{T_1}) \leq \underline{G}_0(x_0) + \epsilon. \quad (13)$$

By (12), since (u_0, \dots, u_{T_1-1}) and (x_1, \dots, x_{T_1}) meet the constraints of (10) (and (11)), the objective of (11) is less than or equal to the objective of (10), and thus

$$0 \leq G_0(x_0) - \underline{G}_0(x_0) \quad (14)$$

$$\leq G_0(x_0) - \sum_{t=0}^{T_1-1} g_t(x_t, u_t) - \underline{G}_{T_1}(x_{T_1}) + \epsilon \quad (15)$$

$$\leq G_{T_1}(x_{T_1}) - \underline{G}_{T_1}(x_{T_1}) + \epsilon. \quad (16)$$

The second inequality (15) comes from substituting in the right inequality of (13) into (14). The final inequality (16) is because $G_0(x_0) \leq \sum_{t=0}^{T_1-1} g_t(x_t, u_t) + G_{T_1}(x_{T_1})$ for any feasible solution, in particular the found ϵ -suboptimal solution.

Assumption 2 states that at each second-stage timestep, all feasible arguments of $G_{T_1}(\cdot)$ are the result of the mapping M_t from feasible arguments of $\tilde{G}_{T_1}(\cdot)$, such that the objective of $\tilde{G}_{T_1}(\cdot)$ is an underestimate of the objective of $G_{T_1}(\cdot)$ by at most δ_t . Taking the worst case for each timestep from T_1 to $T-1$, for any x_{T_1} ,

$$G_{T_1}(x_{T_1}) - \tilde{G}_{T_1}(x_{T_1}) \leq \sum_{t=T_1}^{T-1} \delta_t. \quad (17)$$

When Algorithm 1 has converged, by (7), $\tilde{G}_{T_1}(x_{T_1}) = \underline{G}_{T_1}(x_{T_1})$, and thus

$$G_{T_1}(x_{T_1}) - \underline{G}_{T_1}(x_{T_1}) \leq \sum_{t=T_1}^{T-1} \delta_t. \quad (18)$$

Combining (14), (16), and (18),

$$\underline{G}_0(x_0) \leq G_0(x_0) \leq \underline{G}_0(x_0) + \epsilon + \sum_{t=T_1}^{T-1} \delta_t. \quad (19)$$

□

We now consider the difference between the optimum of the exact problem, and the objective of the exact problem evaluated using the arguments found in the approximate problem.

Theorem 3. If Assumptions 1 and 2 hold, step 6 is solved to ϵ -optimality, and step 6 returns (u_0, \dots, u_{T_1-1}) and (x_1, \dots, x_{T_1}) when Algorithm 1 converges, then

$$G_0(x_0) \leq \sum_{t=0}^{T_1-1} g_t(x_t, u_t) + G_{T_1}(x_{T_1}) \leq G_0(x_0) + \epsilon + \sum_{t=T_1}^{T-1} \delta_t.$$

Proof. The first inequality is straightforward: $G_0(x_0)$ is the optimal objective of the NLP (1). Evaluating the objective of the NLP at any other feasible solution cannot lead to a lower objective.

Now, for the sake of notational ease, denote the resulting arguments when step 6 is solved to ϵ -optimality as $u = (u_0, \dots, u_{T_1-1})$ and $x = (x_1, \dots, x_{T_1})$. Furthermore, denote the first-stage cost as $C_0(x, u) = \sum_{t=0}^{T_1-1} g_t(x_t, u_t)$.

Due to the ϵ -optimality of (x, u) , for all feasible (\hat{x}, \hat{u}) ,

$$C_0(x, u) + \underline{G}_{T_1}(x_{T_1}) \leq C_0(\hat{x}, \hat{u}) + \underline{G}_{T_1}(\hat{x}_{T_1}) + \epsilon. \quad (20)$$

To prove the second inequality in the theorem, consider the difference between the NLP objective evaluated using (x, u) , and the optimal NLP objective $G_0(x_0)$, with corresponding arguments (x^*, u^*) :

$$\begin{aligned} & C_0(x, u) + G_{T_1}(x_{T_1}) - G_0(x_0) \\ & \leq C_0(x, u) + \underline{G}_{T_1}(x_{T_1}) + \sum_{t=T_1}^{T-1} \delta_t - G_0(x_0) \end{aligned} \quad (21)$$

$$\leq C_0(x^*, u^*) + \underline{G}_{T_1}(x_{T_1}^*) + \epsilon + \sum_{t=T_1}^{T-1} \delta_t - G_0(x_0) \quad (22)$$

$$\begin{aligned} & = C_0(x^*, u^*) + \underline{G}_{T_1}(x_{T_1}^*) + \epsilon + \sum_{t=T_1}^{T-1} \delta_t \\ & \quad - C_0(x^*, u^*) - G_{T_1}(x_{T_1}^*) \end{aligned} \quad (23)$$

$$\begin{aligned} & = \epsilon + \sum_{t=T_1}^{T-1} \delta_t + \underline{G}_{T_1}(x_{T_1}^*) - G_{T_1}(x_{T_1}^*) \\ & \leq \epsilon + \sum_{t=T_1}^{T-1} \delta_t. \end{aligned} \quad (24)$$

The inequality (21) comes from substituting (18) for $G_{T_1}(x_{T_1})$. The inequality (22) is due to choosing (x^*, u^*) in (20). The equality (23) comes from expanding $G_0(x_0)$ in terms of its optimal arguments. Finally, (24) holds because $\underline{G}_{T_1}(x_{T_1}^*) \leq G_{T_1}(x_{T_1}^*)$, as in (12). \square

By solving the first stage to a higher level of optimality and using a tighter linearization, the difference between the exact and approximate problem can be reduced. Nevertheless, the numerical results presented below suggest that very good performance can be obtained for a relatively coarse approximation (see Section V-D for a comparison with a state-of-the-art method).

IV. HYDRO RESERVOIR SYSTEM MODEL AND APPROXIMATION

We apply the above split-horizon approximation method to a system of N interconnected reservoirs, with a specific topology of pumps and turbines that allow the reservoirs to exchange water. Denote the set of reservoirs to which reservoir i can charge or discharge as $\mathcal{N}^{i \rightarrow}$ and the set of reservoirs which can charge or discharge to reservoir i as $\mathcal{N}^{\rightarrow i}$. We can transfer water volume $V^{i \rightarrow j}$ from each reservoir i to any $j \in \mathcal{N}^{i \rightarrow}$. The power associated with this action,

$P^{i \rightarrow j}$, is positive when pumping water to a higher elevation reservoir or negative when releasing water through a turbine to a lower elevation reservoir. We buy and sell power on the electricity spot market at price p . This convention on device power and electricity price means that a negative objective value corresponds to a profit, while a positive objective value represents a loss. The aim is to maximize profit by taking advantage of spot price fluctuations. We assume the volume stored in reservoir i is linearly proportional to the reservoir level ℓ^i , with a proportionality constant γ^i reflecting the surface area of the reservoir. That is, if $V^{i \rightarrow j}$ is transferred from reservoir i to j , ℓ^i will decrease by $V^{i \rightarrow j}/\gamma^i$ (and ℓ^j will increase by $V^{i \rightarrow j}/\gamma^j$). For simplicity, we assume that there are no inflows or evaporative losses in this system.

Intuitively, the problem can be cast in the finite horizon optimal control framework considered here by setting $x = (\ell_0, \dots, \ell_T)$, with $\ell_t = (\ell_t^1, \dots, \ell_t^N)$ for all t and $u = (V_0, \dots, V_{T-1})$, with $V_t = (\{V_t^{i \rightarrow j}\}_{j \in \mathcal{N}^{i \rightarrow}})_{i=1, \dots, N}$. We wish to consider a horizon that is sufficiently long to capture seasonal price fluctuations.

A. Nonlinear Exact Model

We assume that the energy associated with transferring a unit volume of water from reservoir i to j depends on the net head between reservoirs i and j [8]:

$$E_t^{i \rightarrow j} = \alpha^{i \rightarrow j} + \beta^{i \rightarrow j} (\ell_t^i - \ell_t^j).$$

Letting $P_t^{i \rightarrow j} = V_t^{i \rightarrow j} E_t^{i \rightarrow j}$, we write the optimal control problem as

$$\min_{P, \ell, V} \sum_{t=0}^{T-1} \sum_{i=1}^N \sum_{j \in \mathcal{N}^{i \rightarrow}} p_t P_t^{i \rightarrow j} + c_T^\top \ell_T \quad (25a)$$

$$\begin{aligned} \text{s.t. } & P_t^{i \rightarrow j} = V_t^{i \rightarrow j} \left(\alpha^{i \rightarrow j} + \beta^{i \rightarrow j} (\ell_t^i - \ell_t^j) \right) \\ & \forall j \in \mathcal{N}^{i \rightarrow} \end{aligned} \quad (25b)$$

$$\ell_{t+1}^i = \ell_t^i + \frac{1}{\gamma_i} \left(\sum_{k \in \mathcal{N}^{\rightarrow i}} V_t^{k \rightarrow i} - \sum_{j \in \mathcal{N}^{i \rightarrow}} V_t^{i \rightarrow j} \right) \quad (25c)$$

$$\underline{\ell}^i \leq \ell_t^i \leq \bar{\ell}^i, \quad \ell_0^i \text{ given} \quad (25d)$$

$$\underline{V}^{i \rightarrow j} \leq V_t^{i \rightarrow j} \leq \bar{V}^{i \rightarrow j} \quad \forall j \in \mathcal{N}^{i \rightarrow} \quad (25e)$$

$$i = 1, \dots, N; \quad t = 0, \dots, T-1.$$

By absorbing equation (25b) into the objective function, we can move the bilinear terms from the constraints into the objective, and eliminate the variables $P_t^{i \rightarrow j}$. The objective (25a) then becomes

$$\sum_{t=0}^{T-1} \sum_{i=1}^N \sum_{j \in \mathcal{N}^{i \rightarrow}} p_t V_t^{i \rightarrow j} \alpha^{i \rightarrow j} + c_T^\top \ell_T + \text{bilinear term} \quad (26)$$

with the bilinear term equal to

$$\begin{aligned} & \sum_{t=0}^{T-1} \sum_{i=1}^N \sum_{j \in \mathcal{N}^{i \rightarrow}} p_t V_t^{i \rightarrow j} \beta^{i \rightarrow j} (\ell_t^i - \ell_t^j) \\ & = \sum_{t=0}^{T-1} \sum_{i=1}^N \sum_{j \in \mathcal{N}^{i \rightarrow}} p_t \beta^{i \rightarrow j} \left(V_t^{i \rightarrow j} \ell_t^i - V_t^{i \rightarrow j} \ell_t^j \right). \end{aligned} \quad (27)$$

The reformulated problem thus has linear constraints (since (25b) has been removed), and an objective with linear and bilinear terms. For sufficiently small horizons, we can find the global optimum in a reasonable length of time using YALMIP's built-in *BMIBNB* spatial branch-and-bound solver [19]. We have observed that Matlab's *fmincon* also tends to return the global optimum for small problem instances in the simulations presented in Section V.

B. Linear Approximate Model

A common way of solving problem (25) is to solve a convex outer approximation. One way to do this is to replace the constraint (25b) with

$$P_t^{i \rightarrow j} = V_t^{i \rightarrow j} \alpha^{i \rightarrow j} + \beta^{i \rightarrow j} (\chi_t^{(i \rightarrow j, i)} - \chi_t^{(i \rightarrow j, j)}). \quad (28)$$

Here, we have introduced the additional variables $\chi_t^{(i \rightarrow j, i)}$ and $\chi_t^{(i \rightarrow j, j)}$ to represent the bilinear terms $V_t^{i \rightarrow j} \ell_t^i$ and $V_t^{i \rightarrow j} \ell_t^j$ respectively. Incorporating this linear constraint into the objective, the bilinear term (27), is thus approximated by

$$\sum_{t=0}^{T-1} \sum_{i=1}^N \sum_{j \in \mathcal{N}^{i \rightarrow}} p_t \beta^{i \rightarrow j} (\chi_t^{(i \rightarrow j, i)} - \chi_t^{(i \rightarrow j, j)}). \quad (29)$$

We then linearize using McCormick envelopes [8], adding to the problem for each $\chi_t^{(i \rightarrow j, i)}$ the inequalities

$$\chi_t^{(i \rightarrow j, i)} \geq V_t^{i \rightarrow j} \bar{\ell}_t^i + \bar{V}_t^{i \rightarrow j} \underline{\ell}_t^i - \bar{V}_t^{i \rightarrow j} \bar{\ell}_t^i, \quad (30a)$$

$$\chi_t^{(i \rightarrow j, i)} \geq V_t^{i \rightarrow j} \underline{\ell}_t^i + \underline{V}_t^{i \rightarrow j} \bar{\ell}_t^i - \underline{V}_t^{i \rightarrow j} \bar{\ell}_t^i, \quad (30b)$$

$$\chi_t^{(i \rightarrow j, i)} \leq V_t^{i \rightarrow j} \bar{\ell}_t^i + \underline{V}_t^{i \rightarrow j} \bar{\ell}_t^i - \underline{V}_t^{i \rightarrow j} \bar{\ell}_t^i, \quad (30c)$$

$$\chi_t^{(i \rightarrow j, i)} \leq V_t^{i \rightarrow j} \underline{\ell}_t^i + \bar{V}_t^{i \rightarrow j} \bar{\ell}_t^i - \bar{V}_t^{i \rightarrow j} \bar{\ell}_t^i. \quad (30d)$$

Here, the upper and lower McCormick bounds, $[\underline{V}_t^{i \rightarrow j}, \bar{V}_t^{i \rightarrow j}]$ for $V_t^{i \rightarrow j}$, and $[\underline{\ell}_t^i, \bar{\ell}_t^i]$ for ℓ_t^i , can depend on the timestep. This results in a linear program, with objective (26) (with the bilinear term replaced by (29)), and constraints (25c)-(25e), (30), and a similar set of McCormick bounds for each $\chi_t^{(i \rightarrow j, j)}$.

C. Error Bound on McCormick Envelope Approximation

We now calculate the maximum error from approximating the bilinear relation $\chi_t^{(i \rightarrow j, i)} = V_t^{i \rightarrow j} \ell_t^i$ by its McCormick envelope in (30). We drop the time, source, and destination indices in the subsequent analysis for cleaner notation.

For each bounding hyperplane in (30), the points where the first derivative of the approximation error with respect to either V or ℓ is zero have zero approximation error. For example, for the first hyperplane (30a), the approximation error is $(V \bar{\ell} + \bar{V} \ell - \bar{V} \bar{\ell}) - V \ell$. The first derivatives are zero where $\ell = \bar{\ell}$ or $V = \bar{V}$, where the error is also zero.

Thus, the largest errors must occur at the intersection of the bounding hyperplanes. Considering Figure 2, the largest overestimate must occur at the intersection of the two upper-bounding hyperplanes (30c) and (30d), and the largest underestimate at the intersection of the two lower-bounding hyperplanes (30a) and (30b) (the intersection of upper- and lower-bounding hyperplanes, e.g., (30a) and (30c), is an exact representation of the underlying function).

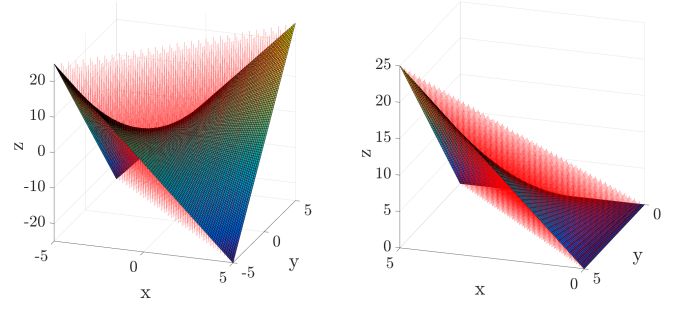


Fig. 2. McCormick envelope convex approximation of $z = xy$, over various bounds.

We determine the largest overestimate analytically. At the intersection of the two upper-bounding hyperplanes,

$$V \bar{\ell} + \underline{V} \ell - \underline{V} \bar{\ell} = V \underline{\ell} + \bar{V} \ell - \bar{V} \underline{\ell}.$$

Solving for V in terms of ℓ ,

$$V_{cd}(\ell) = \frac{\ell(\bar{V} - \underline{V}) + \underline{V} \bar{\ell} - \bar{V} \underline{\ell}}{\bar{\ell} - \underline{\ell}}$$

The approximation error at the intersection of the two upper-bounding hyperplanes is

$$\begin{aligned} \bar{e}(\ell) &= (V_{cd}(\ell) \bar{\ell} + \underline{V} \ell - \underline{V} \bar{\ell}) - V_{cd}(\ell) \ell \\ &= \left(\frac{\ell(\bar{V} - \underline{V}) + \underline{V} \bar{\ell} - \bar{V} \underline{\ell}}{\bar{\ell} - \underline{\ell}} \right) (\bar{\ell} - \ell) + \underline{V} \ell - \underline{V} \bar{\ell}. \end{aligned}$$

Taking the derivative with respect to ℓ , we have $\frac{d\bar{e}(\ell)}{d\ell} = -2\ell + \bar{\ell} + \underline{\ell}$. The derivative of the approximation error is zero at $\ell^* = (\bar{\ell} + \underline{\ell})/2$ with the corresponding $V^* = (\bar{V} + \underline{V})/2$. This leads to a maximum overestimate of

$$\bar{e}^* = (\bar{V} - \underline{V})(\bar{\ell} - \underline{\ell})/4. \quad (31)$$

By similar calculations, or by symmetry, the maximum underestimate \underline{e}^* occurs at the same V^* and ℓ^* , with a negative sign in the error.

D. Error Bound on Two-Stage Approximation

We solve the two-stage approximate problem using Algorithm 1. Since the two-stage approximation differs from the exact problem only in the second stage, a bound on the two-stage approximation error can be found by summing the worst-case approximation error (31) for each second-stage bilinear term in (27). In the context of Assumption 2, at each second-stage timestep t , the worst-case bound between the exact and approximate objectives is

$$\delta_t = \sum_{i=1}^N \sum_{j \in \mathcal{N}^{i \rightarrow}} \frac{|p_t \beta^{i \rightarrow j}|}{4} (\bar{V}_t^{i \rightarrow j} - \underline{V}_t^{i \rightarrow j}) (\bar{\ell}_t^i - \underline{\ell}_t^i + \bar{\ell}_t^j - \underline{\ell}_t^j). \quad (32)$$

Solutions to the two-stage approximation are also feasible for the exact problem, so the map M_t between solutions of the two problem formulations is simply identity.

We assume that the first stage is solved exactly, i.e., $\epsilon = 0$. Denoting the approximate solution as $\underline{G}_0(x_0)$ and optimal

solution as $G_0(x_0)$, and applying Theorem 2, $\underline{G}_0(x_0)$ achieves a maximum underestimate of

$$\sum_{t=T_1}^{T-1} \sum_{i=1}^N \sum_{j \in \mathcal{N}^{i \rightarrow}} \frac{|p_t \beta^{i \rightarrow j}|}{4} \left(\bar{V}_t^{i \rightarrow j} - \underline{V}_t^{i \rightarrow j} \right) \left(\bar{\ell}_t^i - \underline{\ell}_t^i + \bar{\ell}_t^j - \underline{\ell}_t^j \right). \quad (33)$$

The absolute limits on the reservoir levels (25d) and volume flows (25e) can be used in the McCormick envelopes, but these crude estimates can be improved. Tighter bounds on the volume flows are not known a priori, but the bounds on the reservoir levels can be tightened by incorporating the system dynamics (25c), and assuming maximum inflows and outflows at each timestep:

$$\bar{\ell}_t^i = \min \left(\ell_0^i + \frac{1}{\gamma^i} \sum_{s=0}^{t-1} \sum_{j \in \mathcal{N}^{\rightarrow i}} \bar{V}_s^{j \rightarrow i}, \bar{\ell}^i \right) \quad (34)$$

$$\underline{\ell}_t^i = \max \left(\ell_0^i - \frac{1}{\gamma^i} \sum_{s=0}^{t-1} \sum_{j \in \mathcal{N}^{i \rightarrow}} \bar{V}_s^{i \rightarrow j}, \underline{\ell}^i \right). \quad (35)$$

For the split-horizon problem, the bounds must be found relative to the starting level for the entire problem, rather than the starting level for the linear part. Otherwise, the approximation of the bilinear terms changes at each iteration, leading to a different linear program each iteration and voiding the convergence guarantee of Theorem 1.

V. NUMERICAL RESULTS

A. Experimental System

We apply the split-horizon approximation and solution method of Algorithm 1 to the system of two reservoirs and an infinite-capacity basin depicted in Figure 3. This example system is inspired by [7], which considers a single reservoir of the same volume and height used here. The reservoirs have identical capacities of $33 \times 10^6 \text{ m}^3$, and their bottoms are separated by a height of $h_0 = 200 \text{ m}$. The water volume in each reservoir is proportional to the water level relative to the reservoir bottom, with maximum levels $\bar{\ell}^a = 85 \text{ m}$ and $\bar{\ell}^b = 100 \text{ m}$ for the upper and lower reservoir, respectively. The reservoir bottoms are connected by a reversible 100 MW pump/turbine which operates with $\mu = 90\%$ one-way conversion efficiency. The lower reservoir is connected to an infinite basin ($\bar{\ell}^{bas} = \underline{\ell}^{bas} = 0 \text{ m}$) situated 300 m below its bottom, via another identical reversible 100 MW pump/turbine. The reservoir levels are constrained between empty ($\underline{\ell}^a = \underline{\ell}^b = 0 \text{ m}$) and their maximums. The reservoirs are initially half-full.

The energy conversion parameters for (25b) can be calculated from first principles, since the energy required to raise a given volume a height h is $E = \rho g h$, with $\rho = 1000 \text{ kg/m}^3$ and $g = 9.81 \text{ m/s}^2$. Taking into account the one-way conversion efficiency μ , the energy conversion parameters for (25b) are $\alpha^{a \rightarrow b} = -k_{H_2O} h_0 \mu$, $\beta^{a \rightarrow b} = -k_{H_2O} \mu$, $\alpha^{b \rightarrow a} = k_{H_2O} h_0 / \mu$, and $\beta^{b \rightarrow a} = k_{H_2O} / \mu$, where $k_{H_2O} = 0.002725 \text{ kWh}$. The maximum flow rates in (25e) are calculated as the flows that produce or consume 100 MW at the initial reservoir levels. Reservoir a takes 156 hours to empty

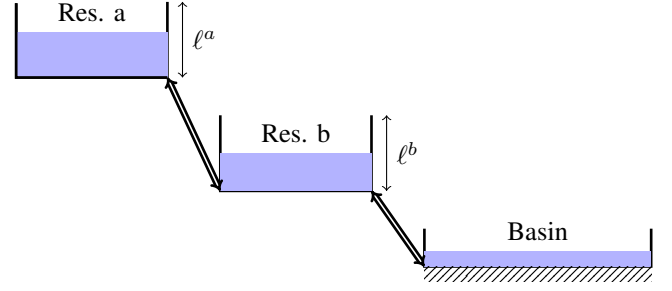


Fig. 3. Test system with two reservoirs and reversible turbines/pumps, and an infinite basin.

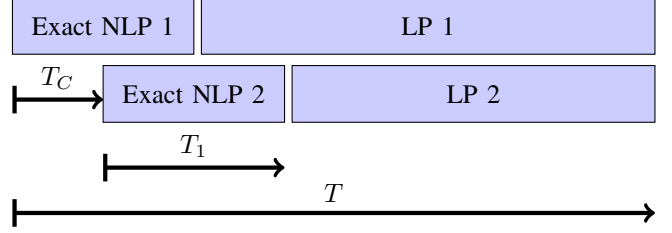


Fig. 4. Depiction of shrinking-horizon simulation over two simulation steps. Simulation is run over total horizon T , with nonlinear first stage of length T_1 , and control horizon of length T_C where current solution is applied.

into Reservoir b at maximum release flow, while Reservoir b takes 283 hours to empty into the lower basin.

The terminal cost c_T of water stored in reservoirs is chosen as the energy embodied in a unit volume of water at the beginning of the time horizon, multiplied by the average price throughout the time horizon. This energy includes conversion efficiency losses, and is relative to the zero potential level of the system. Thus, the value of water stored in the upper reservoir takes into account its potential passage through the lower reservoir as well. Price data are taken from the Swiss day-ahead EPEX spot market, starting on January 1, 2017 [20], and are assumed to be perfectly known over the entire horizon.

B. Solution of Nonlinear and Split-Horizon Problems

We seek to minimize the exact NLP (25), with the total horizon $T = 20$ days. As a reference estimate of the global optimal value, we grid both the state and input space into 32 equally-spaced points and solve the approximate dynamic program using the DPM toolbox [21]. We compare this to the solution of the split-horizon problem (consisting of (25) with modifications (28), (29), and (30)) over a shrinking horizon, as depicted in Figure 4. Here, we solve the split-horizon problem and apply the first T_C hours of the solution. We then move forward in time by T_C , compute a solution to the subsequent, shortened split-horizon problem, and so forth. Note that for the problem considered here, solutions to the LP are also feasible for the NLP, and thus the relative length of T_C versus T_1 presents no issue for feasibility.

When solving the split-horizon problem, the nonlinear first stage is also solved using the DPM toolbox. The linear second stage is solved using CPLEX 12.7.0. The simulations are run in MATLAB, formulated by YALMIP, and solved on a six core Intel Core i7-5820K with 16 GB of RAM.

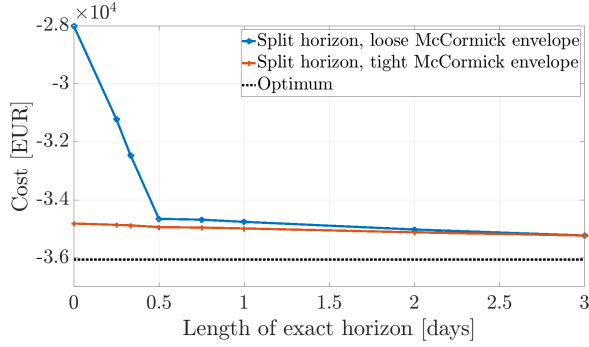


Fig. 5. Optimality of split-horizon DDP method over 20 days with varying length of exactly-modeled first stage, and control horizon of 12 hours. Split-horizon method uses McCormick envelopes in linear stage, with linearization limits either set to the given reservoir level limits (loose) or as in (34) and (35) (tight). Method applied to two-reservoir system of Figure 3.

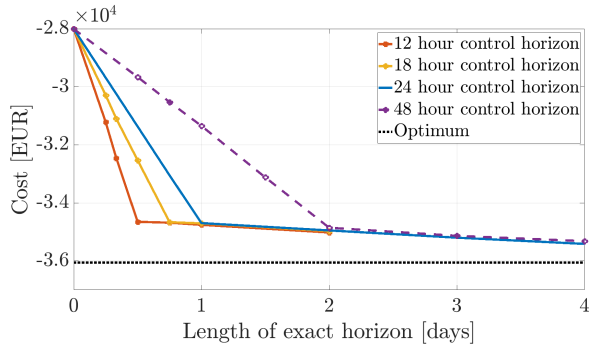


Fig. 6. Optimality of split-horizon DDP method over 20 days with varying length of exactly-modeled period. Results are for two-reservoir system. Control horizon of MPC decisions is varied.

In Figure 5, the two-stage problem is solved over various exact horizon lengths T_1 . The figure shows that as T_1 increases (for a fixed total horizon), the cost decreases. When the McCormick bounds are set to the reservoir level limits, using the split-horizon method with $T_1 = 12$ hours results in a problem objective which is 23.7% less than that from the linearized problem (where $T_1 = 0$ in the figure), and is within 3.9% of the optimum. Since the suboptimality decreases with the exact horizon length, we would choose the longest T_1 that allows Algorithm 1 to converge in the allotted time. For the hydro application here, the allotted time would typically be on the order of one hour, as in [3].

When using the tightened McCormick bounds (34) and (35), there is a 0.4% improvement due to modeling 12 hours exactly. Thus, the utility of the split-horizon method depends on the level of accuracy in the linearization. Note that here, the improvement in linearization due to the tightened McCormick bounds is reduced as the length of the exact horizon increases.

The theoretical bounds from (33) state that for the problem considered in Figure 5, if the two-stage problem with $T_1 = 12$ hours is solved to global optimality, the objective found for the first timestep is within 61.5% of the optimum for the case of the loose McCormick bounds, and within 62.6% of the optimum for the case of the tightened bounds. These bounds are conservative compared to the suboptimality achieved in

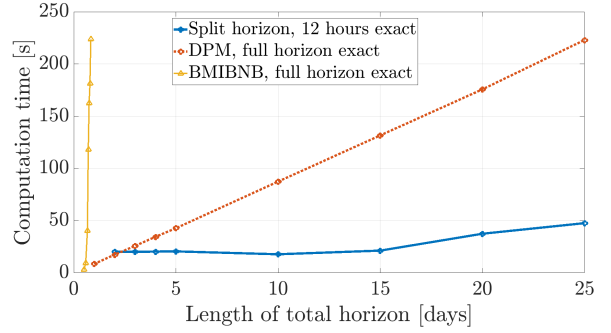


Fig. 7. Computation time for two-reservoir system of split-horizon DDP method with 12 hours exactly modeled, compared to solving full bilinear problem approximately using DPM dynamic programming toolbox and exactly using BMIBNB global solver.

simulation for two reasons. First, the McCormick envelope approximation considers the worst case combination of inputs, which rarely occurs in practice. This conservatism accumulates at each timestep, rendering bounds at more distant timesteps even more conservative. Second, the bounds are given for a single problem instance, and do not take into account that only part of each solution is used in the receding horizon setting.

In Figure 6, we vary the control horizon T_C over which the computed action is applied. As T_1 is also varied, a “knee” appears in the graph, coinciding with T_C . Since T_C is usually a fixed problem parameter, this result suggests we should choose $T_1 \geq T_C$. The objective continues to improve as T_1 is increased, but not as significantly as before the “knee.” Also, as T_C gets shorter, the objective improves. This is expected from the receding horizon context, as measurements from the true system are incorporated at a more frequent rate.

In Figure 7, we compare the computation time for the split-horizon approximation (2) versus solving the exact bilinear problem (25) using the DPM toolbox and BMIBNB solver. The split-horizon problem computation time is for solving the first instance of the problem i.e., we compute the solution for the full horizon once, and do not account for solving the problem repeatedly over the shrinking horizon. This reflects the envisioned receding horizon setting: in practice we would solve the problem once, use the solution for the first timestep, and then re-solve a new problem that incorporates updated problem data for the next timestep. We see that the split-horizon method with $T_1 = 12$ hours scales well with respect to the total horizon length. In practice, the increase in solution time depends mainly on that of the second stage LP, as the number of Benders cuts for the second stage tends to increase sublinearly as the second stage length grows. For example, the split horizon method in Figure 7 results in two cuts for total horizons less than five days, between three and four cuts for horizons from 5-19 days, and between four and five cuts for horizons up to 25 days. As expected theoretically, the computation time for solving the exact bilinear problem scales exponentially in horizon length when using the global BMIBNB solver, and scales linearly when solving approximately using the DPM toolbox.

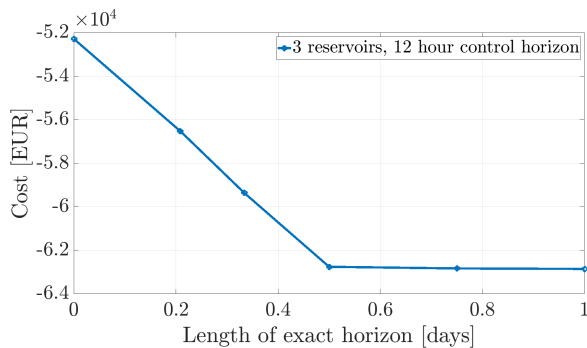


Fig. 8. Optimality of split-horizon DDP method over 20 days with varying length of exactly-modeled period. 12 hour control horizon used throughout. Results are for three reservoir system, where dimension of system is too large for DPM toolbox to solve.

C. Solution of Problems with Higher State Dimension

For the two-reservoir system considered above, the DPM toolbox can solve the exact problem to within the tolerance of the given grid. However, the chief disadvantage of dynamic programming is that the computation time scales exponentially in the problem dimension. When we modify the system of Figure 3 to include an additional reservoir, increasing the state and input dimensions each to three, the DPM toolbox fails to return a solution due to reaching computer-specific RAM usage limits. It is possible to make the DP discretization coarser, but this comes at a cost to optimality.

In contrast to dynamic programming, where one must solve over the entire state space, local solution methods can be used to solve the first-stage nonlinear problem of the split-horizon approximation. The length of the first stage can be chosen short enough so that it is computationally tractable to solve. The experimental results presented here suggest that using even short first stages for the split-horizon approximation leads to near-optimal solutions in a receding horizon setting. In Figure 8, we display the results of using the split-horizon method on the three reservoir system mentioned in the previous paragraph. Using the BMIBNB global solver (solved to a relative tolerance of 1×10^{-6}) with a control horizon of 12 hours, we found that the split-horizon method with an exact horizon of 12 hours improved the objective by 20%, relative to the solution of the linearized problem. Note that solving the 20 day exact NLP to optimality is computationally intractable.

D. Comparison to Multi-cell Approximation

For a comparison with other methods which consider local approximations of the bilinearity ([6]-[8]), we also implement the multi-cell McCormick envelope approximation method of [8]. The method is used to approximate the bilinearity (25b) over the entire horizon. Here, we choose to split each volume flow and reservoir level variable into two equal intervals, and generate a McCormick envelope for each of the four resulting cells. For the two-reservoir model of Figure 3, this introduces one binary variable per timestep for each of the four devices and two reservoirs. Considering the 20 day horizon with hourly timesteps, this results in a mixed-integer linear program (MILP) with 2880 binary variables. We solve the

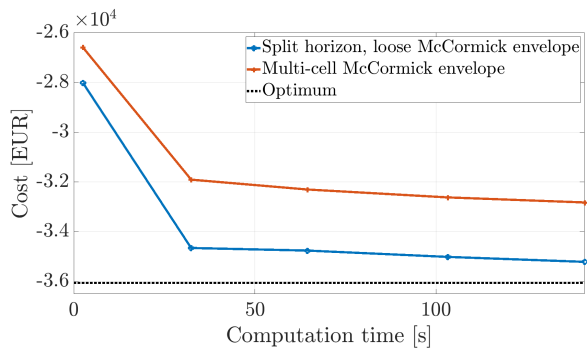


Fig. 9. Pareto curves for split-horizon DDP method versus multi-cell McCormick method [8], with volume flow and reservoir level variables each split into two intervals. Data points result from varying length of exactly-modeled first stage in DDP method, taking computation time, and setting this as maximum computation time for multi-cell method. Displayed is performance over 20 days with 12 hour control horizon, for two-reservoir system of Fig. 3.

MILP in the receding horizon setting of Figure 4, allotting a solution time equivalent to the time used by the split-horizon method for a given exact horizon length.

We produce Figure 9 by varying the exact horizon length of the split-horizon method. This can be viewed as exploration of a Pareto front between computation time and solution optimality. We see that the multi-cell McCormick method performs worse than the split-horizon method. For a 20 day horizon, when both methods use a computation time of 32.4 seconds (the solution time of the split-horizon method with an exact horizon of 0.5 days), the multi-cell method achieves an objective which is within 11.5% of the optimum, while the split-horizon method is within 3.9%. Moreover, we expect the computation time of the split-horizon method to scale better with increasing problem size, compared to the exponential complexity of solving a larger MILP.

VI. CONCLUSION

We have shown that the DDP-based Algorithm 1, when applied to a split-horizon approximation of a bilinear hydro optimization problem, achieves accurate and computationally efficient results compared to solving either the full exact nonlinear problem or its linearization. When the optimization is conducted in a receding horizon manner with exact horizons on the order of the control horizon, the performance is nearly optimal, with significant computational savings when modeling higher-dimensional systems.

The experimental results presented here demonstrate that accurately modeling the first part of the horizon can significantly improve the result relative to a linearized model. This improvement is due to the mismatch between the linearized and true model. Although better linearizations reduce the need for the split-horizon method, it might still be advantageous to use DDP with local linearizations like [8] in the second stage. However, additional research along the lines of [12] is needed to use DDP with the resulting integer variables. We note that the success of the split-horizon method depends on the type of nonlinearity considered, and how well applicable solution methods scale as the first stage length increases.

We also derived bounds for the suboptimality of the split-horizon method. These bounds were found to be quite conservative for the problem here. Error bounds can be tighter for other problem cases and classes. For example, the suboptimality relative to an integer relaxation of an MILP can also be bounded, with a tightness that depends on the problem parameters. However, our case study here shows that while the bounds on a McCormick envelope approximation of a bilinearity are conservative, the split-horizon method used in a receding horizon setting can still achieve good results. Future research could explore how these bounds differ for a receding horizon setting, compared to the single-period optimization for which they were derived.

The split-horizon method is well suited to be extended to stochastic problems e.g., for uncertain prices here. As noted above, stochasticity, perhaps incorporated using a scenario tree approach as in [17], further increases the complexity of solving the problem exactly over long horizons.

ACKNOWLEDGMENT

The authors would like to thank Roy Smith and members of the Building Control Group at the ETH Automatic Control Laboratory for their input into this work. This work is supported by the SCCER FEED&D project and the European Research Council under the project OCAL, ERC-2017-ADG-787845.

REFERENCES

- [1] U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, "2016 Renewable Energy Data Book," <https://www.nrel.gov/docs/fy18osti/70231.pdf>. Accessed Dec. 1, 2017.
- [2] H. S. Chen, T. N. Cong, W. Yang, C. Q. Tan, Y. L. Li, and Y.L. Ding, "Progress in electrical energy storage system: A critical review," *Progress in Natural Science*, vol. 19, no. 3, pp. 291–312, 2009.
- [3] H. Abgottsson and G. Andersson, "Multi-horizon Modeling in Hydro Power Planning," *Energy Procedia*, vol. 87, pp. 2–10, 2016.
- [4] G. Pritchard, A. B. Philpott, and P.J. Neame, "Hydroelectric reservoir optimization in a pool market," *Mathematical Programming*, vol. 103, no. 3, pp. 445–461, 2005.
- [5] T. A. Rotting and A. Gjelsvik, "Stochastic dual dynamic programming for seasonal scheduling in the Norwegian power system," *IEEE Transactions on Power Systems*, vol. 7, no. 1, pp. 273–279, Feb. 1992.
- [6] A. Diniz and M. Maceira, "A Four-Dimensional Model of Hydro Generation for the Short-Term Hydrothermal Dispatch Problem Considering Head and Spillage Effects," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1298–1308, 2008.
- [7] A. Borghetti, C. D'Ambrosio, A. Lodi, and S. Martello, "An MILP Approach for Short-Term Hydro Scheduling and Unit Commitment With Head-Dependent Reservoir" *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1115–1124, Aug. 2008.
- [8] S. Cerisola, J.M. Latorre, and A. Ramos, "Stochastic dual dynamic programming applied to nonconvex hydrothermal models," *European Journal of Operational Research*, vol. 218, no. 3, pp. 687–697, 2012.
- [9] J. P. S. Catalao, S. J. P. S. Mariano, V. M. F. Mendes, and L. A. F. M. Ferreira, "Scheduling of Head-Sensitive Cascaded Hydro Systems: A Nonlinear Approach," *IEEE Transactions on Power Systems*, vol. 24, no. 1, pp. 337–346, Feb. 2009.
- [10] M. V. F. Pereira and L. M. V. G. Pinto, "Multi-stage stochastic optimization applied to energy planning," *Mathematical Programming*, vol. 52, pp. 359–375, 1991.
- [11] N.V. Sahinidis and I.E. Grossmann, "Convergence Properties of Generalized Benders Decomposition," *Computers and Chemical Engineering*, vol. 15, no. 7, pp. 481–491, 1991.
- [12] J. Zou, S. Ahmed, and X. A. Sun, "Stochastic dual dynamic integer programming," *Mathematical Programming*, vol. 175, no. 1–2, pp. 461–502, 2019.

- [13] H. Abgottsson, "Hydro power planning: Multi-horizon modeling and its applications," ETH Zürich PhD Dissertation No. 22729, 2015.
- [14] M. Hohmann, J. Warrington, and J. Lygeros, "A Moment and Sum-of-Squares Extension of Dual Dynamic Programming with Application to Nonlinear Energy Storage Problems," *European Journal of Operational Research*, 2019.
- [15] B. Flamm, A. Eichler, J. Warrington, and J. Lygeros, "Dual Dynamic Programming for Nonlinear Control Problems over Long Horizons," *2018 European Control Conference (ECC)*, Limassol, 2018, pp. 471–476.
- [16] V. Guigues, "Inexact Cuts in Deterministic and Stochastic Dual Dynamic Programming Applied to Linear Optimization Problems," 2018. <https://arxiv.org/abs/1801.04243>.
- [17] S. Rebennack, "Combining sampling-based and scenario-based nested Benders decomposition methods: application to stochastic dual dynamic programming," *Mathematical Programming*, vol. 156, no. 1–2, pp. 343–389, 2016.
- [18] R. Rahmani, T. G. Crainic, M. Gendreau, and W. Rei, "The Benders decomposition algorithm: A literature review," *European Journal of Operational Research*, vol. 259, no. 3, pp. 801–817, 2017.
- [19] J. Löfberg, "YALMIP: a toolbox for modeling and optimization in MATLAB," *2004 IEEE International Conference on Robotics and Automation*, pp. 284–289, 2004.
- [20] European Power Exchange Swiss Spot Market Price. www.epexspot.com/en. Accessed Oct. 25, 2017.
- [21] O. Sundstrom and L. Guzzella, "A generic dynamic programming Matlab function," In *Proceedings of the 18th IEEE International Conference on Control Applications*, pp. 1625–1630, Saint Petersburg, Russia, 2009.



Benjamin Flamm received B.S. and M.S. degrees in electrical engineering from Stanford University in 2009 and 2011 respectively. From 2011–2015 he worked as a robotics controls engineer at Intuitive Surgical in Sunnyvale, California. He is currently pursuing his Ph.D. in the Automatic Control Lab (IfA) at ETH Zurich. His research interests include large-scale stochastic optimization and modeling of energy storage devices, with applications to the optimal control of energy storage.



Annika Eichler received her B.E in General Engineering Sciences and Dipl.-Ing. in Mechatronics from the Hamburg University of Technology in 2008 and 2011 respectively. She finished her Ph.D on distributed systems and control in May 2015 at the Institute of Control Systems at the Hamburg University of Technology, Germany. From 2015–2019, she was a postdoctoral and senior researcher at the Automatic Control Lab at ETH Zurich. Currently she is a senior researcher at DESY in Hamburg, Germany. Her research interests are in distributed, data-driven, robust and stochastic control with the focus on energy systems and particle accelerators, where tools from model predictive control, dynamic programming and robust and stochastic optimization are applied.



Joseph Warrington is a Senior Scientist in the Automatic Control Lab (IfA) at ETH Zurich, Switzerland. His Ph.D. is from ETH Zurich (2013), and his B.A. and M.Eng. degrees in Mechanical Engineering are from the University of Cambridge (2008). From 2014-2016 he worked as an energy consultant at Baringa Partners LLP, London, UK, and he has also worked as a control systems engineer at Wind Technologies Ltd., Cambridge, UK, and privately as an operations research consultant. He is the recipient of the 2015 ABB Research Prize for an outstanding

PhD thesis in automation and control, and a Simons-Berkeley Fellowship for the period January-May 2018. His research interests include dynamic programming, large-scale optimization, and predictive control, with applications including power systems and transportation networks.



John Lygeros completed a B.Eng. degree in electrical engineering in 1990 and an M.Sc. degree in Systems Control in 1991, both at Imperial College of Science Technology and Medicine, London, UK. In 1996 he obtained a Ph.D. degree from the Electrical Engineering and Computer Sciences Department, University of California, Berkeley. During the period 1996–2000 he held a series of post-doctoral researcher appointments at the Laboratory for Computer Science, M.I.T., and the Electrical Engineering and Computer Sciences Department at

U.C. Berkeley. Between 2000 and 2003 he was a University Lecturer at the Department of Engineering, University of Cambridge, UK, and a Fellow of Churchill College. Between 2003 and 2006 he was an Assistant Professor at the Department of Electrical and Computer Engineering, University of Patras, Greece. In July 2006 he joined the Automatic Control Laboratory at ETH Zurich, where he is currently serving as the Head of the Automatic Control Laboratory and the Head of the Department of Information Technology and Electrical Engineering. His research interests include modelling, analysis, and control of hierarchical, hybrid, and stochastic systems, with applications to biochemical networks, automated highway systems, air traffic management, power grids and camera networks. John Lygeros is a Fellow of the IEEE, and a member of the IET and the Technical Chamber of Greece; since 2013 he has served as the Treasurer of the International Federation of Automatic Control.