University Library



A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Lang, A;Cantoni, M;Farokhi, F;Shames, I

Title:
Rigid-Profile Input Scheduling under Constrained Dynamics with a Water Network Application

Date:
2021-11-01

Citation:
Lang, A., Cantoni, M., Farokhi, F. & Shames, I. (2021). Rigid-Profile Input Scheduling under Constrained Dynamics with a Water Network Application. IEEE Transactions on Control Systems Technology, 29 (6), pp.2457-2472. https://doi.org/10.1109/TCST.2020.3043443.

Persistent Link:
https://hdl.handle.net/11343/268324

Rigid-profile input scheduling under constrained dynamics with a water network application

Adair Lang, Michael Cantoni, Farhad Farokhi, Iman Shames

arXiv:2012.01626v1 [math.OC] 3 Dec 2020

***Abstract*—The motivation for this work stems from the problem of scheduling requests for flow at supply points along an automated network of open-water channels. The off-take flows are rigid-profile inputs to the system dynamics. In particular, the channel operator can only shift orders in time to satisfy constraints on the automatic response to changes in the load. This leads to a non-convex semi-infinite programming problem, with sum-separable cost that encodes the collective sensitivity of end users to scheduling delays. The constraints encode the linear time-invariant continuous-time dynamics and limits on the state across a continuous scheduling horizon. Discretization is used to arrive at a more manageable approximation of the semi-infinite program. A method for parsimoniously refining the discretization is applied to ensure continuous-time feasibility for solutions of the approximate problem. It is then shown how to improve cost without loss of feasibility. Supporting analysis is provided, along with simulation results for a realistic irrigation channel setup to illustrate the approach.**

***Index Terms*—Continuous-time dynamics, load input scheduling, semi-infinite programming.**

I. Introduction

The problem considered in this paper is motivated by issues that arise in the management of automated open-channel networks for rural water distribution [1]. Specifically, the scheduling problem of interest pertains to the timing of rigid-profile off-take flows at supply points, given constraints on the transient channel water-level and flow responses to load changes, while minimizing the collective cost of scheduling delays to end users. Problems of this kind may also be relevant in other domains, e.g., energy systems and process control.

Some of the most widely studied scheduling problems, such as job and machine allocation, typically involve only static relationships between the variables of interest [2]–[4]. Similarly, within the literature on irrigation networks, most scheduling studies are limited to static (or steady-state) capacity constraint satisfaction, with no regard for the transient behaviour [5], [6].

Scheduling problems that involve constraints on the transient behaviour of a dynamical system are considered in [7]–[11]. In these works, the dynamics correspond to a discrete-time model, or a uniformly sampled continuous-time model, with constraint satisfaction only enforced on a uniformly sampled subset of the scheduling horizon. The issue of continuous-time constraint satisfaction is not addressed. Specific consideration of rigid-profile input scheduling subject to constraints on discrete-time dynamics is studied in [1], [12].

Irrigation channels are complex physical systems with continuous-time dynamics. For channels operating under closed-loop control, linear time-invariant models are suitable when there is ample in-channel storage [13], [14]. This motivates the approach pursued below, where attention is focused on optimizing scheduling decisions subject to constraints, over a continuous interval, on the evolution of a continuous-time model. In particular, a non-convex semi-infinite programming formulation of the scheduling problem is considered for linear time-invariant dynamics. Scheduling problems subject to constrained continuous-time dynamics are also considered in [15]–[18]. However, these works do not explicitly consider important aspects of computing feasible solutions for the semi-infinite program. This motivates the focus here on ensuring constraint satisfaction over the continuous scheduling horizon. Particular emphasis is given to the construction of direct discretizations with parsimonious non-uniform sampling of time and constraints such that solutions of the corresponding approximate problem are continuous-time feasible. This is the aim of the proposed first stage of the approach. In the second and final stage, the cost of the first-stage feasible schedule is improved by a sequential quadratic programming approximation of the original problem. While it is possible to extend the approach to more general dynamics, the linear time-invariant context of this paper enables explicit characterization of the state and derivatives with respect to the scheduling variable, at any specified time for a given initial condition.

In the rest the introduction, the semi-infinite program is formulated. Aspects of the discretization-based approximations are then introduced. To conclude, the contributions are summarized and an outline is given for the paper.

A. Problem formulation

Consider a linear time-invariant system with dynamics

$$\dot{x}(t) = Ax(t) + Bu(t) + \sum_{j=1}^{m} E_j v_j(t - \tau_j),\ x(0) = x_0. \quad (1)$$

In this model, $x(t) \in \mathbb{R}^{n_x}$ is the state, $u(t) \in \mathbb{R}^{n_u}$ is the control input, and each $v_j(t - \tau_j) \in \mathbb{R}$ is a shifted load input at time $t \in \mathbb{R}_{\geq 0}$. Let $\mathcal{T} := [0, T]$ be the scheduling horizon

This work was supported by Rubicon Water Pty. Ltd., the Australian Research Council (LP130100605, LP160100666), and a McKenzie Fellowship.

A. Lang is with Dept. of Electrical and Electronic Engineering, The University of Melbourne, VIC 3010, Australia, and Rubicon Water, 1 Cato St, Hawthorn East, VIC 3123, Australia. aclang@student.unimelb.edu.au

M. Cantoni is with Dept. of Electrical and Electronic Engineering, The University of Melbourne, VIC 3010, Australia. cantoni@unimelb.edu.au

F. Farokhi is with the Dept. of Electrical and Electronic Engineering, The University of Melbourne, VIC 3010, Australia, ffarokhi@unimelb.edu.au

I. Shames is with Dept. of Electrical and Electronic Engineering, The University of Melbourne, VIC 3010, Australia. ishames@unimelb.edu.au

and let $\mathbb{N}_{[a,b]}$ denote the integers between $a > 0$ and $b \geq a$ inclusive. Given the admissible shift interval $\mathcal{D}_j := [\underline{\tau}_j, \overline{\tau}_j]$ for each piecewise continuous load input request

$$v_j = (t \in [-\overline{\tau}_j, T - \underline{\tau}_j] \mapsto v_j(t) \in \mathbb{R}), \quad j \in \mathbb{N}_{[1,m]},$$

the scheduling problem is to select the piecewise continuous control input $u : \mathcal{T} \to \mathbb{R}^m$ and load input schedule shift variables $\tau_j \in \mathcal{D}_j$ to satisfy linear constraints on the state

$$x(t; u, (v_j, \tau_j)_{j=1}^m) = \bar{x}_0^u(t) + \sum_{j=1}^m \bar{x}_j^v(t - \tau_j), \quad t \in \mathcal{T}, \quad (2a)$$

where

$$\bar{x}_0^u(t) := \exp(At)x_0 + \int_0^t \exp(A(t-s))Bu(s)\mathrm{d}s, \quad (2b)$$

$$\bar{x}_j^v(t) := \int_0^t \exp(A(t-s))E_j v_j(s)\mathrm{d}s. \quad (2c)$$

The best choice, among the feasible possibilities, minimizes $f(\tau_1, \ldots, \tau_m) := \sum_{j=1}^m h_j(\tau_j)$. The differentiable $h_j : \mathcal{D}_j \to \mathbb{R}_{\geq 0}$ represents the sensitivity of user j to shifting its request (e.g., production cost impacts of delayed delivery).

PROBLEM 1 (FIXED-PROFILE LOAD SCHEDULING): Given set $\mathcal{U}$ of admissible control inputs, and state constraint parameters $C \in \mathbb{R}^{n_c \times n_x}$ and $c \in \mathbb{R}^{n_c}$, solve

$$f^* := \min_{u, (\tau_j)_{j=1}^m} f(\tau_1, \ldots, \tau_m) \quad (3a)$$

$$\text{s.t.} \quad Cx(t; u, (v_j, \tau_j)_{j=1}^m) \leq c, \ t \in \mathcal{T}, \quad (3b)$$

$$\tau_j \in \mathcal{D}_j, \ j \in \mathbb{N}_{[1,m]}, \quad (3c)$$

$$u \in \mathcal{U}. \quad (3d)$$

Several factors make it challenging to solve (3) in general:

1) non-convexity of the constraint set with respect to the decision variables $(\tau_j)_{j=1}^m$;
2) infinitely many constraint in (3b); and
3) infinite dimensionality of the control input u.

The first factor relates to the influence of $(\tau_j)_{j=1}^m$ on the state $x(t; u, (v_j, \tau_j)_{j=1}^m)$ in the inequality constraints (3b); see [17]. Non-convexity can make it difficult to distinguish between potentially multiple local minima and the global minimum. With the second factor, the problem can thus be classified as a non-convex semi-infinite program (SIP); see [19] for a comprehensive overview of such optimization problems. The third factor relates to the complexity of solving constrained continuous-time optimal control problems. Given the linear dependence of (2a) on u, this can be largely overcome by restricting the set of admissible controls to be finite-dimensional, which is a commonly employed technique [20]. Doing so yields a problem in which the main difficulty rests with how the state dynamics are affected by the decision variables $(\tau_j)_{j=1}^m$; i.e., factors 1) and 2). As such, the rest of the paper is focused on these difficulties, while simplifying the developments by assuming that the control input is constant; i.e., $\mathcal{U} = \{u = (t \in \mathcal{T} \mapsto u_0 \in \mathbb{R}^{n_u})\}$. Indeed, u is subsequently dropped as a decision variable.

ASSUMPTION 1.1 (STRICT FEASIBILITY): There exists at least one schedule of load request shifts $(\tau_j)_{j=1}^m$, with $\tau_j \in \mathcal{D}_j$, such that $Cx(t; u_0, (v_j, \tau_j)_{j=1}^m) - c < 0$ for all $t \in \mathcal{T}$.

Problem 1 could be reformulated in terms of a switched hybrid model of the dynamics, in which the switching instances correspond to the scheduling parameters τ_j in (2a). However, it appears that doing so leads to even more challenging problems. Firstly, the set of switch states grows exponentially in the number of users, since the evolution of the channel dynamics depends on the combination in which requested load profiles are applied. Secondly, the natural reformulation does not match better studied forms of switched dynamics optimization problems [21]–[23]. Specifically, the switching sequence would not be pre-determined as in [21], this reformulation does not decompose in the way considered in [22], and it does not seem possible to comply with the rigid profile requirements within the framework of [23], where switching sequence and switching instants are decision variables together. As such, switched system models are not considered further. Instead, direct discretization approximations are pursued for obtaining a first-stage feasible solution, used to initialize a continuous variable local approximation method for the original SIP.

B. Direct discretization-based approximations

A common approach to SIPs is to relax the problem by sampling the constraints across the time horizon [19]. Even so, Problem 1 remains difficult. Hence related work has also considered discretization of the decision spaces to yield an integer linear program [1], [12]. Whilst still difficult, standard methods and solvers exist for such problems; e.g., see [24]. On the other hand, shortcomings of this existing work relate to the use of uniform discretization, which can lead to unnecessarily large problems in terms of the number of decision variables and number of constraints, and the lack of guarantees on continuous-time feasibility of the solutions obtained. Aspects of the discretization approach are now presented as a precursor to a summary of the main contributions made in this paper.

Replacing $\mathcal{D}_j$ with the finite discretized decision space

$$\hat{\mathcal{D}}_j := \{\tau_j^{(1)}, \ldots, \tau_j^{(N_j)}\} \subset \mathcal{D}_j \quad (4)$$

for $j \in \mathbb{N}_{[1,m]}$ yields the problem

$$\hat{f}^* := \min_{(\tau_j)_{j=1}^m} f(\tau_1, \ldots, \tau_m) \quad (5a)$$

$$\text{s.t.} \quad Cx(t; (\tau_j)_{j=1}^m) \leq c, \ t \in \mathcal{T}, \quad (5b)$$

$$\tau_j \in \hat{\mathcal{D}}_j \text{ for } j \in \mathbb{N}_{[1,m]}, \quad (5c)$$

where explicit dependence of x on the fixed u_0 and given requests v_j has been dropped for convenience. In principle, problem (5) could be solved by exhaustive search as the decision space is finite. A more sophisticated approach is presented in Section II. The approach is based on iteratively solving the following related finite-dimensional optimization problems.

The optimization problem (5) is a restriction of (3), and thus, $\hat{f}^* \geq f^*$. In fact, it is possible that (5) is infeasible for instances of the finite decision sets $(\hat{\mathcal{D}}_j)_{j=1}^m$. Under Assumption 1.1, this can be overcome by adding elements to these sets. Parsimonious augmentation is desirable as the complexity of solving (5) grows with the cardinality N_j of $\hat{\mathcal{D}}_j$, $j \in \mathbb{N}_{[1,m]}$. A corresponding method is proposed in Section II-B.

Replacing $\mathcal{T}$ in (5) with a finite subset of sample times at which the state constraints must hold yields a relaxation. With

$$\hat{\mathcal{T}}_i := \{t_i^{(1)}, \ldots, t_i^{(T_i)}\} \subset \mathcal{T}, \quad i \in \mathbb{N}_{[1,n_c]}, \tag{6}$$

the problem becomes

$$\hat{f}^{*,L} := \min_{(\tau_j)_{j=1}^m} f(\tau_1, \ldots, \tau_m) \tag{7a}$$

$$\text{s.t.} \quad C_i x(t; (\tau_j)_{j=1}^m) \leq c_i, \ t \in \hat{\mathcal{T}}_i, \ i \in \mathbb{N}_{[1,n_c]}, \tag{7b}$$

$$\tau_j \in \hat{\mathcal{D}}_j, \ j \in \mathbb{N}_{[1,m]}, \tag{7c}$$

where C_i and c_i denote the i-th row of C and i-th entry of c, respectively. The cardinality of each constraint discretization set $\hat{\mathcal{T}}_i$ is denoted by T_i, which may be zero. Further, $\hat{f}^{*,L} \leq \hat{f}^*$, by definition.

Problem (7) is tractable in the sense that it can be transformed into an integer linear program. This is shown in Section II-C. However, it is possible for solutions of (7) to violate (5b), and thus, be infeasible for problem (3). In view of Assumption 1.1, and continuity of the state with respect to $t \in \mathbb{R}_{\geq 0}$, this issue can be overcome by adding sample times to the sets $\hat{\mathcal{T}}_i \subset \mathcal{T}$ and tightening the constraint (7b). Specifically, the tightened problem takes the form

$$\hat{f}^{*,U} := \min_{(\tau_j)_{j=1}^m} f(\tau_1, \ldots, \tau_m) \tag{8a}$$

$$\text{s.t.} \quad C_i x(t; (\tau_j)_{j=1}^m) \leq c_i - \epsilon^g, \ t \in \hat{\mathcal{T}}_i, \ i \in \mathbb{N}_{[1,n_c]}, \tag{8b}$$

$$\tau_j \in \hat{\mathcal{D}}_j, \ j \in \mathbb{N}_{[1,m]}, \tag{8c}$$

with $\epsilon^g > 0$. This is a restriction of (7), whereby $\hat{f}^{*,U} \geq \hat{f}^{*,L}$.

Problem (8) is neither a restriction nor a relaxation of (3). However, as suggested above, suitable constructions of $\hat{\mathcal{T}}_i$ and ϵ^g can be made so that solutions of (8) satisfy (5b), and thus, (3b). In this case, $\hat{f}^{*,U} \geq \hat{f}^* \geq \hat{f}^{*,L}$. To this end, a recent algorithm for general SIPs from [25] is adapted to (5), as detailed in Section II.

C. Contribution

The main contribution is a two-stage approach to Problem 1; see (3). The first stage involves iterative construction of the discrete decision spaces $\hat{\mathcal{D}}_j$, constraint discretization sets $\hat{\mathcal{T}}_i$, and constraint restriction parameter ϵ^g in (8), from initial values. The resulting finite-dimensional problem yields a good schedule of shifts for which (3b) holds; i.e., a schedule that is close in objective value to the optimal value in (5), where only the decision space is discretized, and feasible for the original problem. The aim of the second stage is to improve this schedule while maintaining feasibility. To this end, the original continuous decision space is re-instated, and a sequential quadratic programming (SQP) method is used to approximately solve (3), initialized from the feasible schedule.

The proposed two-stage approach goes beyond prior work in ensuring continuous-time feasibility via parsimoniously constructed discretizations. In particular, a method for general SIPs from [25] is adapted to the scheduling problem and extended to accommodate combined discretization of the decision space and the constraints. The resulting discretizations are not necessarily uniform, unlike the purely discrete-time formulations considered in [1], [12]. This can lead to smaller integer linear programming problems in the first-stage. In the second-stage, an SQP method from [26], [27] is used to improve the cost without loss of feasibility. As in [17], which develops a continuous-variable penalty-function based method that does not necessarily lead to improvement of the cost, analytic formulations of the derivatives required to implement the SQP method are possible by virtue of the linear time-invariance of the underlying dynamics. The proposed algorithm is ultimately demonstrated on a non-trivial simulation example that is based on models of an Australian irrigation channel that are used operationally in the field and historically realistic demand profiles.

D. Outline

The rest of the paper is organized as follows. Section II includes discussion of the SIP discretization procedure from [25] and the modifications made to ensure finite termination when it is applied to (5) from an initially infeasible discretization of the decision spaces. The continuous-variable SQP approach to improving the first-stage feasible schedule is developed in Section III. A formal characterization of the combined two-stage algorithm is presented in Section IV. Supporting analytical results are provided throughout. Numerical results based on non-trival irrigation channel scenarios are discussed in Section V. The paper is concluded with discussion of future research directions in Section VI.

II. Stage 1 - Discretization with feasibility

The Hybrid-SIP algorithm (HSIPA) from [25], which builds on [28], [29], provides a mechanism for generating constraint discretization sets $\hat{\mathcal{T}}_i$, and a constraint restriction parameter ϵ^g, such that an optimizing schedule for (8) is feasible for (5) with $\hat{f}^{*,U} - \hat{f}^*$ less than a specified tolerance $\epsilon^f > 0$.

The HSIPA requires the following input parameters:

- initial constraint discretization sets $\hat{\mathcal{T}}_i \subset \mathcal{T}$, $i \in \mathbb{N}_{[1,n_c]}$;
- initial constraint restriction $\epsilon^g \in \mathbb{R}_{>0}$ in (8); and
- desired optimality tolerance $\epsilon^f \in \mathbb{R}_{>0}$ for problem (5).

Given these inputs the HSIPA iteratively determines upper and lower bounds for $\hat{f}^*$. These bounds converge to within the specified tolerance ϵ^f of $\hat{f}^*$. The bounds are generated by successively solving iterations of the following three subproblems: i) lower-bound problem (7); ii) upper-bound problem (8); and iii) refinement problem:

$$-\eta^* := \min_{(\tau_j)_{j=1}^m, \eta} -\eta \tag{9a}$$

$$\text{s.t.} \quad \sum_{j=1}^m h_j(\tau_j) - f^R \leq 0 \tag{9b}$$

$$C_i x(t; (\tau_j)_{j=1}^m) \leq c_i - \eta, \ t \in \hat{\mathcal{T}}_i, \ i \in \mathbb{N}_{[1,n_c]}, \tag{9c}$$

$$\tau_j \in \hat{\mathcal{D}}_j, \ j \in \mathbb{N}_{[1,m]}, \text{ and } \eta \in \mathbb{R}, \tag{9d}$$

for given target objective value $f^R > 0$. The results are used to update the constraint discretization sets $(\hat{\mathcal{T}}_i)_{i=1}^{n_c}$, the restriction parameter ϵ^g, and target objective f^R, as described subsequently. Iterations proceed, until the upper and lower bounds obtained lie within ϵ^f of each other. Each solve of one of the three subproblems results in a candidate scheduling solution $(\tau_j)_{j=1}^m$. For given candidate solution $(\tau_j)_{j=1}^m$, the maximum level of constraint violation for each constraint $i \in \mathbb{N}_{[1,n_c]}$ is given by:

$$g_i^*((\tau_j)_{j=1}^m)) := \max_{t \in \mathcal{T}} C_i x(t; (\tau_j)_{j=1}^m) - c_i. \tag{10}$$

If $g_i^*((\tau_j)_{j=1}^m)) > 0$, then the candidate schedule is not feasible for (5). In this case, a point in time that corresponds to this maximum level of constraint violation is added to the constraint discretization set. Specifically, given such a candidate scheduling solution $(\tau_j)_{j=1}^m$, the constraint discretization set is updated as follows:

$$\hat{\mathcal{T}}_i \leftarrow \begin{cases} \hat{\mathcal{T}}_i \cup \{t_i^*((\tau_j)_{j=1}^m)\} & \text{if } g_i^*((\tau_j)_{j=1}^m)) > 0 \\ \hat{\mathcal{T}}_i & \text{otherwise} \end{cases}, \tag{11}$$

where

$$t_i^*((\tau_j)_{j=1}^m) \in \mathcal{T}_i^*((\tau_j)_{j=1}^m) := \arg\max_{t \in \mathcal{T}} \left(C_i x(t; (\tau_j)_{j=1}^m) - c_i \right) \tag{12}$$

is a corresponding maximizer of the maximum level of constraint violation. Feasibility of a schedule with respect to the infinite constraints (3b) corresponds to

$$g_i^*((\tau_j)_{j=1}^m)) \leq 0, \quad i \in \mathbb{N}_{[1,n_c]}. \tag{13}$$

The update mechanisms of the constraint restriction parameter ϵ^g and target objective f^R are discussed within the next section. The optimality tolerance ϵ^f should be chosen to reflect the desired optimality gap for solving (5), which likely depends on the given setup.

Algorithm 2 in [25] is extended here in two ways. The first extension provides a way to deal with the finite sets $\hat{\mathcal{D}}_j$, and thus, a potentially infeasible problem (5) for the given algorithm initialization. The introduction of a new parameter enables interlacing of the HSIPA with an update procedure for the sets $\hat{\mathcal{D}}_j$ that achieves eventual feasibility of (5) from an initially infeasible discretization. The corresponding update procedure is presented in Section II-B. Secondly, the development as presented here elucidates the handling of a finitely indexed set of infinite constraints (i.e., $n_c > 1$).

A. HSIPA for the scheduling problem

Next, the notation in this paper is mapped to that used in [25], and the key algorithm parameters are identified. Aspects of the procedures associated with the aforementioned lower-bound, upper-bound and refinement problems are described in relation to generating the required constraint discretization sets $\hat{\mathcal{T}}_i$ and constraint restriction parameter ϵ^g. Generation of the required discrete decision spaces $\hat{\mathcal{D}}_j$ is the topic of Section II-B. Modifications made here in adapting the HSIPA from [25] to the scheduling problem are highlighted below.

[25]	This paper	Description
(SIP)	(5)	SIP to be solved
$f(\boldsymbol{x})$	$f(\tau_1, \ldots, \tau_m)$	Objective function
f^*	$\hat{f}^*$	Optimal objective value
$\boldsymbol{x}$	$(\tau_j)_{j=1}^m$	Decision variable(s)
$\mathcal{X}$	$(\hat{\mathcal{D}}_j)_{j=1}^m$	Decision space
$g(\boldsymbol{x}, \boldsymbol{y})$	$Cx(t; (\tau_j)_{j=1}^m) - c$	Inequality constraint(s)
$\mathcal{Y}$	$\mathcal{T}$	Index set of infinite constraints
$\mathcal{Y}^D$	$\hat{\mathcal{T}}_i,\ i \in \mathbb{N}_{[1,n_c]}$	Constraint discretization set(s)
(LBD)	(7)	Lower-bound sub-problem
(UBD)	(8)	Upper-bound sub-problem
(RES)	(9)	Refinement sub-problem
(LLP)	(10)	Lower level program

TABLE I
A MAP OF NOTATION/LABELS IN [25] TO THE NOTATION IN THIS PAPER

1) HSIPA notation and parameters: Table I relates the labels and notation used in this paper to those used in [25]. Notice, in particular, the multiplicity of constraint discretization sets here, one for each row of C in (5b), compared to one set in the formulation of [25]. The HSIPA involves many parameters. The parameters listed at start of Section II are important within the specific context of the subproblems (7) and (8). The following parameter is introduced here to enable a decision space update procedure to be activated when (5) is infeasible:

- $k_r^{max} \in \mathbb{N}$, the maximum number of consecutive constraint restriction updates allowed in the upper-bound procedure.

2) Lower-bound: The lower-bound procedure of the HSIPA corresponds to Lines 2–11 of [25, Algorithm 2]. It pertains to solving the relaxed problem (7) to determine a lower-bound $\hat{f}^{*,L}$ for $\hat{f}^*$. Given a feasible solution of (7), the constraint discretization sets $\hat{\mathcal{T}}_i$ are updated as per (11). This update, and those made in subsequent procedures of the HSIPA, ensure that successive runs of the lower-bound procedure result in a bound that converges to within a specified optimality tolerance of $\hat{f}^*$. To ensure finite termination of the HSIPA for an initially infeasible problem (5), the lower-bound procedure is modified here to initially check if (7) is feasible. When it is not, the procedure terminates without updating the sets $\hat{\mathcal{T}}_i$, after setting a flag that is used in the extension of Section II-B to trigger an update the discrete decision space $\hat{\mathcal{D}}_j$.

3) Upper-bound: Lines 12–24 of [25, Algorithm 2] constitute the upper-bound procedure of the HSIPA. This procedure relates to solving (8), which restricts the discretized constraints by ϵ^g. If (8) is infeasible, then the restriction is gradually reduced, in steps $\epsilon^g \leftarrow \epsilon^g / r^g$ for specified $r_g > 1$, until it becomes feasible. The modification made here, relative to [25], is to limit the number of such steps to k_r^{max}, after which the upper-bound procedure terminates and the HSIPA returns to the lower-bound procedure, which could itself terminate as infeasible, triggering augmentation of the decision spaces $\hat{\mathcal{D}}_j$. When (8) is feasible, and the resulting schedule satisfies (13), the corresponding $\hat{f}^{*,U}$ is an upper-bound for $\hat{f}^*$; i.e., the infinite constraints (5b) are feasible. The upper-bound procedure then terminates, after a further step of ϵ^g reduction, and the HSIPA proceeds to the refinement procedure. Otherwise, when (13) does not hold, the constraint

discretization sets $\hat{\mathcal{T}}_i$ are updated, as per (11), and the upper-bound procedure is repeated, including such updates, until an upper bound is found for $\hat{f}^*$. These updates, and those of successive subprocedures of the HSIPA, eventually lead to an upper-bound that lies within a specified tolerance of $\hat{f}^*$, as established in [25, Lemma 4].

4) Refinement procedure: The role of the refinement procedure is to improve both the upper- and lower-bounds whilst avoiding over-population of the sets $\hat{\mathcal{T}}_i$. The approach is adapted from [30] in the HSIPA developed in [25]. The subprocedure corresponds to Lines 26–42 in [25, Algorithm 2]. A target objective value, f^R, is selected as the mid-point between the current upper and lower bounds. The corresponding refinement problem (9) is then solved and updates are made on the basis of the outcome. The following lemmas reveal how solving (9) may be used to improve the upper- or lower-bounds, $\hat{f}^{*,U}$ and $\hat{f}^{*,L}$, respectively, as detailed in the subsequent remarks.

LEMMA 2.1: If η^* in (9) satisfies $\eta^* < 0$, then $f^R < \hat{f}^*$.

Proof: See Appendix A-A. ∎

LEMMA 2.2: If an optimal schedule $(\tau_j^*)_{j=1}^m$ for (9) satisfies (13), then $f^R \geq \hat{f}^*$.

Proof: Since the resulting schedule satisfies (13), it is feasible for (5), and hence, $\hat{f}^* \leq f(\tau_1^*, \ldots, \tau_m^*)$. By the constraint (9b), $f(\tau_1^*, \ldots, \tau_m^*) \leq f^R$. As such, $\hat{f}^* \leq f^R$. ∎

REMARK 2.1: The refinement procedure is repeated whilst the conditions of Lemma 2.1 are satisfied, updating the lower-bound to f^R accordingly for each such run. This can improve the overall computational time of the HSIPA as each run halves the difference between the upper and lower bound.

REMARK 2.2: Satisfaction of (13) in Lemma 2.2 implies $\eta^* \in \mathbb{R}_{\geq 0}$; which is thus an upper-bound for the smallest constraint restriction that retains feasibility. Updating ϵ^g accordingly yields improvement of $\hat{f}^{*,U}$ in subsequent runs of the upper-bound procedure, as described in [25].

REMARK 2.3: It is possible that neither Lemma 2.1 nor Lemma 2.2 apply, leading to Line 36 of [25, Algorithm 2]. In this case, the constraint discretization is updated, as per (11), and (9) is re-solved. To moderate growth in the cardinality of the constraint discretization sets $\hat{\mathcal{T}}_i$, this process terminates after a specified $l_{max} \in \mathbb{N}$ consecutive solves of (9), and the HSIPA returns to the lower-bound procedure.

In practice, the optimization problem (9) can only be solved to a specified tolerance, say $\epsilon^{RES} > 0$. So as outlined in [25, Proposition 1] it is possible to get an inconclusive result. In this case the HSIPA returns to the lower-bound procedure.

5) HSIPA Illustration: Fig. 1 shows an example trajectory of a constraint across a continuous scheduling horizon in order to illustrate the key differences between problems (7), (8) and (9). The original problem (3) requires this trajectory to remain under the blue line (c_i) for the entire time-horizon. The lower bound problem (7) is feasible if the constraint trajectory lies below the blue dots at the specified sampled time-points, which holds true in this example. For the upper-bound problem (8) the constraints are still enforced only at specified sample times, but the trajectory needs to be under the red-dashed line $c_i - \epsilon^g$. In this example, the constraint at $t_i^{(2)}$ is violated. When solving the refinement problem (9), assuming this example trajectory has an objective value less than f^R, the maximum constraint restriction possible is η shown as distance from c_i to the open blue dot at $t_i^{(2)}$. In all cases, the problem is not SIP-feasible since g_i^* is greater than 0. As such, the time sample t_i^* would be added to the sampled time-horizon for this constraint for the next iterate.

B. Extended HSIPA (EHSIPA)

Theorem 2 in [25] establishes conditions for finite termination of the HSIPA at an ϵ^f-optimal solution of the SIP at hand. For the result to apply here, the corresponding problem (5) must be feasible. This may not be the case for certain initializations of the decision space discretizations. The modifications made to the HSIPA presented in the previous section ensure finite termination, without a solution, when (5) is infeasible. In this case, the HSIPA sets a flag used to trigger augmentation of the decision space dicretizations. This continues until feasibility of (5) is achieved . The subsequent run of the HISPA then augments the constraint discretization sets until the sub-problems (8) yields an ϵ^f-optimal solution of (5), as described in the preceding subsection.

Given decision space discretizations $\hat{\mathcal{D}}_j = \{\tau_j^{(1)}, \ldots, \tau_j^{(N_j)}\}$, N_j finite, $j \in \mathbb{N}_{[1,m]}$, the update mechanism is defined by the following:

$$\hat{\mathcal{D}}_j \leftarrow \hat{\mathcal{D}}_j \cup \left\{ \frac{\underline{\tau}_j + \tau_j^{(1)}}{2} \right\} \cup \left\{ \frac{\tau_j^{(N_j)} + \overline{\tau}_j}{2} \right\} \cup \left\{ \frac{\tau_j^{(l-1)} + \tau_j^{(l)}}{2} \ : \ l \in \mathbb{N}_{[2,N_j]} \right\}. \quad (14)$$

LEMMA 2.3: Under Assumption 1.1, a finite number of decision space discretization updates (14) leads to a feasibile problem (5).

Proof: See Appendix A-B. ∎

REMARK 2.4: Each update (14) leads to a doubling of the cardinality of every $\hat{\mathcal{D}}_j$, and consequently larger integer linear programs to solve in the HSIPA. It is of interest to devise a more parsimonious mechanism, which is the topic of future work. Dependence of run time on the discretised decision space cardinalities is explored empirically in Section V.

ASSUMPTION 2.1 (SOLUTION TOLERANCES): The feasibility of problems (7) and (8) is assessed without tolerance to infeasibility. When found to be feasible, these problems are solved to within specified global optimality tolerances $\epsilon^{LBP}, \epsilon^{UBP} \in \mathbb{R}_{>0}$, respectively. Further, it is possible to solve problem (10) to arbitrary optimality tolerance.

REMARK 2.5: In principle, (7) and (8) can be solved exactly by exhaustive search. However, this would be impractical for any reasonably sized problem. As shown in the next subsection, the discretization of decision spaces enables transcription to standard linear integer programs. While still difficult to solve, existing algorithms for such problems offer much greater efficiency than exhaustive search; see [24], [31]. Problem (10) can be solved in principle by simulation of the linear time-invariant dynamics with sufficient accuracy.

With Lemma 2.3, the conditions required to apply [25, Theorem 2] are met. Its application yields the following result.

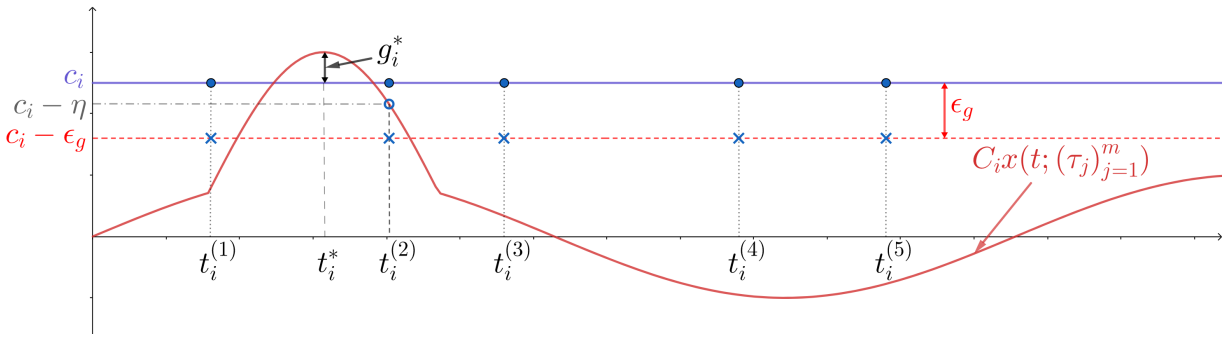


Fig. 1. An example constraint trajectory to highlight the difference between the subproblems within the HSIPA.

PROPOSITION 2.4: Given desired optimality tolerance ϵ^f for problem (5), under Assumption 2.1, if there exists $\tilde{\epsilon}^f \in \mathbb{R}_{>0}$ such that $\tilde{\epsilon}^f \geq f(\tau_1^f, \ldots, \tau_m^f) - \hat{f}^*$ for a schedule $(\tau_j^f)_{j=1}^m$ with $\max_{i \in \mathbb{N}_{[1,n_c]}} g_i^*((\tau_j^f)_{j=1}^m)) < 0$ (i.e., strictly feasible) and

$$\epsilon^f > \tilde{\epsilon}^f + \epsilon^{LBP} + \epsilon^{UBP}, \tag{15}$$

then the EHSIPA terminates after finite steps at a schedule $(\tau_j^*)_{j=1}^m$ that satisfies (3b) and $f(\tau_1^*, \ldots, \tau_m^*) \leq \hat{f}^* + \epsilon^f$.

Proof: First note that $\mathcal{T}$ and $\hat{\mathcal{D}}_j$ are compact sets; the former is a bounded real interval and the latter a finite set [32]. Therefore, [25, Assumption 1] is satisfied. Now from (2a), note that the constraints are all continuous in the schedule variables, as is the objective function, since the h_j are differentiable. Hence for the case of a single constraint, $n_c = 1$, [25, Assumption 2] is satisfied. In the case of $n_c > 1$, the multiple constraints can be replaced with a single constraint given by:

$$\max_{i \in \mathbb{N}_{[1,n_c]}} (C_i x(t;(\tau_j)_{j=1}^m) - c_i) \leq 0, \ t \in [0,T], \tag{16}$$

which is continuous since the maximum of a set of continuous functions is continuous. Hence, under Assumption 2.1, all requirements for the proof in [25] to follow are satisfied. ■

REMARK 2.6: Proposition 2.4 only guarantees the number of steps required to reach optimality is finite. It does not give any guarantees on the size of this number. The chosen tolerance ϵ^f can have a significant effect on the overall number of steps required. Instead of letting the algorithm run until desired optimality guarantees are achieved, the algorithm can be stopped after at least one valid upper bound has been found. The schedule associated with the lowest upper-bound can then be returned, since this schedule is feasible, i.e., satisfies (3b). The sub-optimality gap achieved with such an approach would be highly problem and configuration dependent.

C. Sub-problems as Integer Linear Programs

For the linear dynamics considered here, the HSIPA sub-problems (7), (8), and (9), can be equivalently reformulated as binary linear programs. To this end, define the following:

$$\Psi_{ji} = \begin{bmatrix} \bar{x}_j^{v(1,1)} & \bar{x}_j^{v(1,2)} & \cdots & \bar{x}_j^{v(1,N_j)} \\ \bar{x}_j^{v(2,1)} & \bar{x}_j^{v(2,2)} & \cdots & \bar{x}_j^{v(2,N_j)} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{x}_j^{v(T_i,1)} & \bar{x}_j^{v(T_i,2)} & \cdots & \bar{x}_j^{v(T_i,N_j)} \end{bmatrix},$$

$$\bar{\Psi}_j = \begin{bmatrix} (I_{T_1} \otimes C_1)\Psi_{j1} \\ (I_{T_2} \otimes C_2)\Psi_{j2} \\ \vdots \\ (I_{T_{n_c}} \otimes C_{n_c})\Psi_{jn_c} \end{bmatrix}, \qquad b_i = \begin{bmatrix} \bar{x}_0^u(t_i^{(1)}) \\ \bar{x}_0^u(t_i^{(2)}) \\ \vdots \\ \bar{x}_0^u(t_i^{(T_i)}) \end{bmatrix},$$

$$\bar{c} = \begin{bmatrix} (\mathbb{1}_{T_1} \otimes c_1) - (I_{T_1} \otimes C_1)b_1 \\ (\mathbb{1}_{T_2} \otimes c_2) - (I_{T_2} \otimes C_2)b_2 \\ \vdots \\ (\mathbb{1}_{T_{n_c}} \otimes c_{n_c}) - (I_{T_{n_c}} \otimes C_{n_c})b_{n_c} \end{bmatrix},$$

$$f_j = \begin{bmatrix} h_j(\tau_j^{(1)}) & h_j(\tau_j^{(2)}) & \cdots & h_j(\tau_j^{(N_j)}) \end{bmatrix}^\top,$$

where $\bar{x}_j^{v(k,l)} = \bar{x}_j^v(t_i^{(k)} - \tau_j^{(l)})$ with $\tau_j^{(l)}$ and $t_i^{(k)}$ as defined in (4) and (6), respectively.

Now the lower bound problem (7) can be rewritten as an equivalent binary linear program given by

$$\min_{(z_j)_{j=1}^m} \sum_{j=1}^m f_j^\top z_j, \tag{17a}$$

$$\text{s.t.} \quad \sum_{j=1}^m \bar{\Psi}_j z_j - \bar{c} \leq 0, \tag{17b}$$

$$z_j \in \{0,1\}^{N_j}, \quad \mathbb{1}^\top z_j = 1, \quad j \in \mathbb{N}_{[1,m]}. \tag{17c}$$

Problem (8) can be reformulated similarly, except $\bar{c}$ is formed with $c_i - \epsilon^g$ in place of c_i. Similarly, reformulation of (9) results in

$$-\eta^* = \min_{(z_j)_{j=1}^m, \eta} -\eta \tag{18a}$$

$$\text{s.t} \quad \sum_{j=1}^m f_j^\top z_j - f^R \leq 0 \tag{18b}$$

$$\sum_{j=1}^m \bar{\Psi}_j z_j - \bar{c} \leq -\bar{\eta}, \tag{18c}$$

$$z_j \in \{0,1\}^{N_j}, \quad \mathbb{1}^\top z_j = 1, \quad j \in \mathbb{N}_{[1,m]}, \tag{18d}$$

where $\bar{\eta} = \begin{bmatrix} (\mathbb{1}_{T_1} \otimes \eta)^\top & (\mathbb{1}_{T_2} \otimes \eta)^\top & \ldots & (\mathbb{1}_{T_{n_c}} \otimes \eta)^\top \end{bmatrix}^\top$.

The reformulations above generalize the approach in [1], [12], where purely discrete-time dynamics were considered. The generalization allows for the non-uniform sampling and different sampling sets for each constraint generated by the EHSIPA. Each reformulation involves a change of variables

from τ_j to z_j for each user j. The corresponding relationship between z_j and τ_j is given by

$$\begin{bmatrix}\tau_j^{(1)} & \dots & \tau_j^{(N_j)}\end{bmatrix} z_j = \tau_j. \tag{19}$$

This change of variable depends on $N_j, T_i < \infty$, $j \in \mathbb{N}_{[1,m]}$, $i \in \mathbb{N}_{[1,n_c]}$, which is not the case for the original formulation (3) where the corresponding sets are continuous intervals.

The costs and constraints in (17) and (18) are linear in the decision variables $(z_j)_{j=1}^m$, except for the binary requirement on the entries of z_j. Dedicated linear-integer programming solvers are readily available for such problems; e.g., see [24]. The difficulty of solving discrete problems in non-standard forms, such as (7), (8) and (9), is highlighted in [31].

The reformulations (17) and (18) can be relaxed to linear programs by replacing the binary constraint with $z_j^{(q)} \in (0,1)$, $q \in \mathbb{N}_{[1,N_j]}$ and $j \in \mathbb{N}_{[1,m]}$. This relaxation leads to much nicer problems to solve, but at the cost of losing the rigidity constraint, as discussed in [1]. So it is only applicable where the load-profile is not constrained to a fixed shape.

III. Stage 2 - Localized cost improvement

The EHSIPA described above involves discretization to arrive at a schedule $(\tau_j^*)_{j=1}^m$ that is feasible for (3). This schedule is ϵ^f-optimal with respect to the restriction (5) associated with the final discretized decision spaces $\hat{\mathcal{D}}_j$. Since $\hat{\mathcal{D}}_j \subset \mathcal{D}_j$ by definition, it may be possible to improve the sum of user sensitivities to scheduling delay, $f(\tau_1^*, \dots, \tau_m^*)$, by returning to the original continuous interval decision spaces. To this end, an SQP algorithm is used in a second stage of the proposed approach to Problem 1. Starting with the feasible schedule generated by the EHSIPA, the approach leads to localised improvement of the cost whilst maintaining feasibility with respect to the original constraints in (3).

A. SQP approximation

SQP methods involve solving quadratic programs formulated to approximate the original problem around each iterate [33]. Adaptations of such methods to SIPs, like problem (3), can be found in [19], [34]. The usual approach is to discretize the infinite constraints. However, such relaxation may result in solutions that are infeasible for the original problem. This could be overcome by refining the constraint discretization, in a similar fashion to the decision space update (14) in the EHSIPA, for example. However, this can lead to a large number of constraints. The approach here is to sample constraint $i \in \mathbb{N}_{[1,n_c]}$ at the maximizers $\mathcal{T}_i^*((\tau_j)_{j=1}^m)$ defined by (12) for the given feasible schedule iterate $(\tau_j)_{j=1}^m$. When the number of maximizers is finite, as subsequently assumed, a tractable SQP algorithm ensues [26].

To construct the approximate quadratic program at each step, differentiability of the constraints is required. For the problem (3) at hand, these derivatives can be formulated explicitly.

Lemma 3.1: The left-hand side of constraint $C_i x(t;(\tau_j)_{j=1}^m) - c_i \le 0$, $i \in \mathbb{N}_{[1,n_c]}$, in (3) is differentiable in the schedule $(\tau_j)_{j=1}^m$. Moreover,

$$\frac{\partial}{\partial \tau_\ell} C_i x(t;(\tau_j)_{j=1}^m) = -C_i E_\ell v_\ell(t-\tau_\ell) - C_i A \bar{x}_\ell^v(t-\tau_\ell)$$

Proof: See Appendix A-C. ∎

For given feasible schedule $(\tau_j)_{j=1}^m$, the quadratic program to solve at each step of the SQP algorithm takes the form

$$\tilde{f}^* := \min_{p \in \mathbb{R}^m} \ \frac{1}{2} p^\top B p + \sum_{j=1}^m p_j \frac{d}{d\tau_j} h_j(\tau_j) + \sum_{j=1}^m h_j(\tau_j) \tag{20a}$$

$$\text{s.t.} \quad a_i(t_i;(\tau_j)_{j=1}^m)^\top p + b_i(t_i;(\tau_j)_{j=1}^m) \le 0, \quad t_i \in \mathcal{A}_i, \ i \in \mathbb{N}_{[1,n_c]}, \tag{20b}$$

$$p_j \in [\underline{\tau}_j - \tau_j, \overline{\tau}_j - \tau_j], \quad j \in \mathbb{N}_{[1,m]}, \tag{20c}$$

$$\|p\|_\infty \le \Gamma. \tag{20d}$$

The decision variable $p = \begin{bmatrix} p_1 & p_2 & \cdots & p_m \end{bmatrix}^\top$ is the update direction from the current iterate $(\tau_j)_{j=1}^m$ to a new schedule,

$$a_i(t_i;(\tau_j)_{j=1}^m) = \nabla_{(\tau_j)_{j=1}^m} \left(C_i x(t_i;(\tau_j)_{j=1}^m) \right) \in \mathbb{R}^m$$

and

$$b_i(t_i;(\tau_j)_{j=1}^m) = C_i x(t_i;(\tau_j)_{j=1}^m) - c_i \in \mathbb{R}$$

are constants, for the given feasible schedule and the finite cardinality sets $\mathcal{A}_i$ contain the maximizers $\mathcal{T}_i^*((\tau_j)_{j=1}^m)$ defined in (12) for $i \in \mathbb{N}_{[1,n_c]}$. The positive scalar Γ is a trust region for the current iterate. The matrix B is a symmetric positive definite approximation of the Hessian of the Lagrangian, with multipliers $((\lambda_{il})_{l=1}^{|\mathcal{A}_i|})_{i=1}^{n_c}$, given by

$$L\left((\tau_j)_{j=1}^m; ((\lambda_{il})_{l=1}^{|\mathcal{A}_i|})_{i=1}^{n_c}\right) \tag{21}$$
$$= \sum_{j=1}^m h_j(\tau_j) + \sum_{i=1}^{n_c} \sum_{l=1}^{|\mathcal{A}_i|} \lambda_{il} \left(C_i x(t_i^{(l)};(\tau_j)_{j=1}^m) - c_i \right),$$

where $|\mathcal{A}_i|$ denotes the cardinality of $\mathcal{A}_i$, $i \in \mathbb{N}_{[1,n_c]}$.

Remark 3.1: Note (20) is feasible provided the schedule around which the problem is approximated is feasible. This can be seen by setting $p = 0$.

The quadratic program (20) is a local approximation of the original problem (3) relative to the schedule $(\tau_j)_{j=1}^m$. The SQP algorithm (SQPA) proceeds by updating this schedule to the schedule $(\tau_j + \gamma p_j^*)_{j=1}^m$, where p^* solves (20) and a line search is performed to find the largest $\gamma \in [0,1)$ such that

$$g_i^*((\tau_j + \gamma p_j^*)_{j=1}^m) \le 0, \ \forall i \in \mathbb{N}_{[1,n_c]}, \tag{22}$$

and

$$\sum_{j=1}^m h_j(\tau_j + \gamma p_j^*) \le \sum_{j=1}^m h_j(\tau_j) + \eta\gamma \sum_{j=1}^m p_j^* \frac{d}{d\tau_j} h(\tau_j), \tag{23}$$

where $\eta \in (0,1)$ is a constant algorithm parameter. Corresponding updates of a_i, b_i, and $\mathcal{A}_i$, $i \in \mathbb{N}_{[1,n_c]}$, h_j, $\frac{d}{dt} h_j$, $j \in \mathbb{N}_{[1,m]}$, B, and Γ are then made to become consistent with the new iterate of the schedule. These steps repeat until a stopping criterion is met. Here this criterion relates to the step size becoming smaller than a tolerance $\epsilon^s > 0$; i.e.,

$$\gamma \|p^*\|_\infty < \epsilon^s. \tag{24}$$

To guarantee a unique solution to (20), the matrix B must remain positive definite. Since the Lagrangian (21) is non-convex in the schedule a standard Broyden-Fletcher-Goldfarb-Shanno (BFGS) update formula may lead to non-positive definite matrices. Here a damped BFGS update formula from [33, p.537] is used. The damping ensures a curvature condition is satisfied and that B remains positive definite if the initialization is positive definite. The update formula from [33, p.537] requires an estimate of the Lagrangian multipliers at current iterate which can be obtained in the process of solving (20).

The trust region Γ is used to reflect confidence in the previous approximation. This can be dynamically updated to allow for larger steps when the approximation is deemed more accurate, and refined as the approximation becomes coarse, in order to reduce the number of iterations needed to perform the line search. Here, [33, Algorithm 4.1] is used to update the trust region.

THEOREM 3.2: Given $\epsilon^s > 0$ and $\eta \in (0, 1)$, suppose the SQPA is initialized with $\Gamma > 0$, B positive definite, and a schedule that is feasible for (3) with cost f^0. Then

i) every schedule iterate is feasible for (3), and
ii) the algorithm terminates in a finite number of steps, such that when the initial solution of (20) is non-zero,

$$\sum_{j=1}^{m} h_j(\tau_j) < f^0 \tag{25}$$

at termination.

Proof: See Appendix A-D ■

REMARK 3.2: Either the SQPA terminates immediately, with the initial schedule, or a new SIP feasible schedule with strictly improved cost is found in finite steps. By contrast, strict improvement cannot be guaranteed by the penalty methods proposed in [17].

REMARK 3.3: If the requested load inputs v_j are restricted to continuous functions, and certain regularity assumptions hold, and the sets $\mathcal{A}_i$ at each step are modified to include points that can be *extended* (in an appropriate sense) to each of the global maximizers of another schedule that is in the neighborhood of the iterate $(\tau_j)_{j=1}^m$, then it is possible to show that the SQPA, with sufficiently small $\epsilon^s > 0$, converges to a point that satisfies the first order optimality conditions of (3); i.e., to a local-stationary point. This result is established in [26]. However, improvement in the cost is the primary aim here, and since it is difficult to show the regularity assumption holds and to ensure $\mathcal{A}_i$ contains the necessary points, the conditions are relaxed to those in Theorem 3.2, at the expense of not ensuring local optimality.

REMARK 3.4: The algorithm presented in [26] also allows for infeasible initial schedules via an exact penalty method. However, it is observed in practice that this approach is very sensitive to initialization. Within the context of the numerical examples presented in Section V, initialization of the approach at realistic original flow requests failed to yield a schedule that satisfies (13). This is not unexpected since stationary points for the exact penalty reformulation of the problem are not necessarily feasible. This motivated development of the EHSIPA.

IV. CHARACTERISTICS OF THE TWO-STAGE APPROACH

The two-stage approach to Problem 1 is as follows. First, the EHSIPA described in Section II is applied to generate a schedule that is feasible for (3) and ϵ^f-optimal for the restriction associated with a discretization of the decision spaces. In the second stage, this SIP feasible schedule is then improved in cost by application of the SQPA described in Section III.

THEOREM 4.1: Under Assumptions 1.1 and 2.1, suppose $\epsilon^f, \epsilon^{LBD}$, and ϵ^{UBP} satisfy (15). Then the two-stage approach terminates finitely with a schedule $(\tilde{\tau}_j^*)_{j=1}^m$ such that $i)$ $\tilde{\tau}_j^* \in \mathcal{D}_j$, $j \in \mathbb{N}_{[1,m]}$, $ii)$ (3b) holds, and $iii)$

$$\underline{f} \leq f^* \leq f(\tilde{\tau}_1^*, \ldots, \tilde{\tau}_m^*) \leq \hat{f}^* + \epsilon^f, \tag{26}$$

where f^* and $\hat{f}^*$ are defined by (3) and (5), respectively, and

$$\underline{f} := \min_{(\tau_j)_{j=1}^m} f(\tau_1, \ldots, \tau_m), \text{ s.t } \tau_j \in \mathcal{D}_j, \ j \in \mathbb{N}_{[1,m]}. \tag{27}$$

Proof: Proposition 2.4 gives the right hand side of (26). Theorem 3.2 ensures that the SQPA maintains or improves the cost whilst maintaining SIP feasibility. Hence, the bound and feasibility still hold after stage 2. The left-hand side of (26) follows because the cost of any feasible schedule is an upper bound for the optimal cost and that the optimization problem defining $\underline{f}$ is a relaxation of (3); i.e., there are no constraints on the dynamics. ■

REMARK 4.1: Theorem 4.1 holds for any second stage algorithm that takes an initial schedule that is feasible for (3) and terminates after a finite number of steps with a schedule satisfying (25) and (13).

REMARK 4.2: The inequality (26) bounds the obtained optimality gap. Specifically, $f(\tau_1^*, \ldots, \tau_m^*) - f^* \leq f(\tau_1^*, \ldots, \tau_m^*) - \underline{f} \leq \hat{f}^* + \epsilon^f - \underline{f}$. If an upper bound is known for $\hat{f}^*$ then the last expression can be calculated *a priori*. However, in practice these bounds are conservative since $\underline{f} \leq f^*$. This bound relies on calculation of (27) which should be easier than (3) since there are no constraints from the dynamics. In particular if h_j are convex then (27) is a convex problem which could be solved using standard convex optimization methods such as those in [35].

REMARK 4.3: The main result relies on Assumption 2.1. If there exist algorithms that satisfied Assumption 2.1 with the original continuous decision spaces, then the HSIPA algorithm for constraint discretization would find the optimal solution, negating the need for the second localized SQPA based refinement stage. The decision spaces are discretized to arrive at tractable problems that satisfy Assumption 2.1. The localized refinement stage can then lead to improvement.

Although Theorem 4.1 guarantees finite termination it does not say anything about the convergence rate or computational complexity and specifically how to choose the initial discretizations of the decision spaces and constraints. The subsequent discussion relates to why a good choice of initial discretizations is not obvious.

A. Effect of Initial Discretizations

The two-stage approach based on EHSIPA and SQPA is guaranteed to converge for any choice of initial discretiza-

tions of the decision spaces and constraints in (3). However, this choice can have significant impact on both the overall computational time and the achieved optimality gap.

1) Discretization and Computation Time: The integer programs in the EHSIPA have dimensions linked explicitly to N_j and T_i, the cardinalities of the discretized decision spaces and constraint discretization sets. Hence, it may seem intuitive to start with these being small in order to reduce the solve time of these sub-problems. However, sparse initial sets may lead to slower overall computation time.

To see this, first consider sparse initial constraint discretization sets $\hat{\mathcal{T}}_i$. Using these may lead to an initial lower-bound within the EHSIPA that is far from the optimal, e.g., if the constraint set were empty, then the initial lower bound would be $\underline{f}$. Further, the first call to the upper-bound procedure may require many iterations to add sufficient points to the constraint discretization sets to find a feasible schedule.

Second, consider sparse initial discretized decision spaces $\hat{\mathcal{D}}_j$. This could yield an infeasible problem (5) and several iterations of the EHSIPA could be required before (5) becomes feasible. Additionally, a sparse decision space may increase the likelihood that the resulting schedule from the first stage is far from a stationary point of the original problem (3). This can increase the convergence time of the SQPA.

These issues are discussed further within the context of a particular example in the numerical results Section V.

2) Discretization and Optimality Gap: The SQPA is guaranteed to improve the initial schedule if possible. However, due to non-convexity of (3), the iterates tend towards local stationary points of (3). In general, there is no measure of how far these local stationary points are from the global optimum. Hence, the sub-optimality gap achieved for the two-stage approach is strongly linked to the sub-optimality of the first stage solution; i.e., the distance between $\hat{f}^*$ and f^*.

REMARK 4.4: Nothing can be said about the achieved optimality gap if the SQPA is initialized at an arbitrary feasible schedule, rather than an optimizer for (5).

The difference $\hat{f}^* - f^*$ depends on the discretization of the decision spaces. This difference can be made arbitrarily small by adding sufficiently enough points to the discretzized decision spaces $(\hat{\mathcal{D}}_j)_{j=1}^m$. However, this approach is impractical since this may lead to prohibitively large integer programs to be solved as part of the HSIPA. This is highlighted further in Section V.

As mentioned in Remark 2.4, it is of interest to investigate the use of additional information to inform the discretization(s) of smaller decision spaces. Note that it is possible for a smaller discretization to result in a smaller optimality gap. For instance, consider the case where the discretization for each user is chosen with cardinality equal to one, containing only the shift from an optimal schedule for (3); i.e., $\hat{\mathcal{D}}_j = \{\tau_j^*\}$ for $j \in \mathbb{N}_{[1,m]}$. In this case, the smallest optimality gap is achieved even though the discretizations are very sparse.

A possibility in this direction is to re-run the EHSIPA after the first feasible schedule is found, with a finer discretization centered at the solution of the previous iterate such that the cardinality of the decision spaces remain the same. This preliminary idea, along with the effects of the other aforementioned complexities regarding the initial discretizations are demonstrated for a practical example from automated irrigation networks in the subsquent section.

V. NUMERICAL RESULTS

In this section, a realistic scheduling setup is investigated numerically based on operational data for an irrigation channel in south-eastern Australia. Results obtained from applying the first stage of the approach developed above are presented for nine instances of the initial discreizations of the decision space and constraints. These initializations range from fine uniform discretizations of both the decision space and the constraints to coarse discretizations of both. The outcomes are compared in terms of feasibility, optimality and computational burden. The results reveal that a dense initialization does not necessarily yield the best outcome. For the example considered it is observed that coarse initialization can yield a feasible schedule that is as good as one obtained from a much denser starting point, without the optimization sub-problems involved becoming overly large. The effectiveness of the second stage is also explored, including comparisons with the penalty based gradient methods from [17], in terms of improvement in cost and computational burden.

A. Application Gravity Fed Irrigation Networks

An irrigation channel is a cascade of pools. Following [13], [36], each pool is modeled in the Laplace domain as

$$y_i(s) = \frac{c_{\mathrm{in},i}}{s} e^{-t_{\mathrm{d},i}s} q_i(s) - \frac{c_{\mathrm{out},i}}{s} q_{i+1}(s) - \frac{c_{\mathrm{out},i}}{s} o_i(s),$$

where $c_{\mathrm{in},i}$ and $c_{\mathrm{out},i}$ (in $1/(\min\sqrt{\mathrm{m}})$) are discharge rates determined by the physical characteristics of the gates used to set the flow between neighbouring pools, and $t_{\mathrm{d},i}$ (in min) is the delay associated with the transport of water along the pool. Here, $o_i(s)$ denotes the overall off-take load on pool i, that is, the sum of all the water supplied to the farms connected to this pool. Moreover, $q_i(s)$ represents the head over the upstream gate of pool i raised to the power of $3/2$ which is proportional to the flow (in $\mathrm{m}^3/\mathrm{min}$) of water from pool $i-1$ to pool i and $y_i(s)$ denotes the water level (in m) in pool i. For the purpose of this example, the delays are replaced with a first-order Padé approximation[1]. Each pool is controlled, locally, by

$$q_i(s) = \frac{\kappa_i(\phi_i s + 1)}{s(\rho_i s + 1)}(u_i(s) - y_i(s)) + \gamma_i q_{i+1}(s),$$

where κ_i, ϕ_i, ρ_i and γ_i are appropriately selected control parameters. Furthermore, $u_i(s)$ (in m) denotes the water-level reference signal of pool i. The model for a series of pools using the aforementioned model, with first-order Padé approximation, can be represented in the form of (1) where there are four states per pool, two for the controller and two for the level dynamics.

[1]Note that the choice of a first-order Padé approximation is justifiable as the pool delays are all parts of closed-loops (with local controllers), with loop-gain cross-overs that are sufficiently small to make the overall closed-loop behavior insensitive to the approximation error [14].

TABLE II
PARAMETERS, $c_{in,i}$, $c_{out,i}$ (IN $1/(\min\sqrt{\text{m}})$), $t_{d,i}$ (IN min), AND UNITLESS CONTROL PARAMETERS κ_i, ϕ_i AND ρ_i FOR $i \in \mathbb{N}_{[1,10]}$ FOR THE DYNAMICAL SYSTEM USED IN THE SIMULATION.

Pool No	$c_{\text{in},i}$	$c_{\text{out},i}$	$t_{\text{d},i}$	κ_i	ϕ_i	ρ_i
$i=1$	0.1079	0.1080	1	0.0156	46.648	3.452
$i=2$	0.0777	0.0777	1.67	0.0091	72.403	5.213
$i=3$	0.0586	0.0586	2.33	0.0065	99.274	7.084
$i=4$	0.1269	0.1269	1.67	0.0084	60.305	3.972
$i=5$	0.0313	0.0313	1.83	0.0092	110.85	8.878
$i=6$	0.0456	0.0507	4	0.0036	152.73	10.36
$i=7$	0.0725	0.0725	1.33	0.0119	64.978	4.885
$i=8$	0.0216	0.0216	3.67	0.0080	147.65	10.28
$i=9$	0.0366	0.0366	1.67	0.0100	98.231	7.816
$i=10$	0.2062	0.2331	2	0.0100	48.156	2.101

B. Example Parameters and Setup

1) Problem Data: Table II shows the parameters for the 10 pool system used in this example, which are from validated models of a real channel provided by Rubicon Water Pty Ltd. A fixed $\gamma_i = 0.7$ is used for all pools $i \in \mathbb{N}_{[1,10]}$. The reference input u is a set-point to a lower-level feedback controller. In this context it is common that the reference be kept constant as it is desired to maintain a constant level of supply to the offtakes. For this setup it is fixed to $1m$ for all pools, i.e., $u_0 = \mathbb{1}_{10}$.

The state constraints are envelope constraints on the water level such that $y_i(t) \in [\underline{y}_i, \overline{y}_i]$ for all $t \in [0, T]$ where $\underline{y}_i = 0.9$ (m) and $\overline{y}_i = 1.075$(m) $\quad \forall i \in \mathbb{N}_{[1,4]} \cup \{9, 10\}$ and $\underline{y}_i = 0.88$ (m) and $\overline{y}_i = 1.10 \quad \forall i \in \mathbb{N}_{[5,8]}$.

In this example, there are two users for each pool. Each user j, in each pool i, has a requested off-take profile defined by a start time, s_{ij}, duration d_{ij} (both in min), and magnitude m_{ij}, which is proportional, via a discharge coefficient, to flow (in m^3/min). This gives typical pulse shape for off-take flows in irrigation networks, i.e., $v_{ij}(t) = m_{ij}(H(t - s_{ij}) - H(t - s_{ij} - d_{ij}))$ where $H : \mathbb{R} \to \{0, 1\}$ denotes the continuous time Heaviside step function. Table III shows the particular off-take load parameters used in this example. The top of Fig. 2. displays the requested orders, which follow a realistic demand pattern based on historic orders. It is of note the initial requests cause significant violation of the constraints; see top of Fig. 3.

The bounds on the allowable shifts are set to $\underline{\tau}_{ij} = -180$ and $\overline{\tau}_{ij} = 180$ for all $i \in \mathbb{N}_{[1,10]}, j \in \{1, 2\}$, i.e., the orders can be scheduled by shifting the requested load forward or backwards by up to three hours. The end-user sensitivity is measured with a quadratic function $h_j(\tau) = 0.01\tau^2$ for all users, i.e., each user experiences a greater cost the further away from the requested start-time. The ideal schedule for each user is to not have their order shifted at all. The planning horizon is set to $T = 1440$min (1 day), for all simulations. In open-channel irrigation networks, it is desirable to reduce the required lead-time for users to request desired off-take profiles. In practice the lead time is typically in the order of days, but there are network operators striving for lead times of around two hours. Hence, a computational time in the order of an hour for solving the scheduling problem is tolerable. The aforementioned setup gives all the problem data needed to formulate (3).

TABLE III
PARAMETERS FOR OFF-TAKE LOADS USED IN THE SIMULATION. s_{ij}, d_{ij} (IN min) AND m_{ij} MULTIPLIED BY A DISCHARGE COEFFICIENT HAS UNITS OF m^3/min

Pool No	s_{i1}	d_{i1}	m_{i1}	s_{i2}	d_{i2}	m_{i2}
$i=1$	200	360	0.0645	500	360	0.0322
$i=2$	200	180	0.0322	500	180	0.0322
$i=3$	200	360	0.0322	500	360	0.0645
$i=4$	200	180	0.0645	500	180	0.0322
$i=5$	200	360	0.0322	500	360	0.0322
$i=6$	200	180	0.0290	500	180	0.0580
$i=7$	200	360	0.0580	500	360	0.0290
$i=8$	200	180	0.0322	500	180	0.0322
$i=9$	200	360	0.0322	500	360	0.0645
$i=10$	800	360	0.0285	500	180	0.0285

2) Algorithm parameters for the EHSIPA and SQPA stages: In addition to the parameters mentioned throughout Section II, the HSIPA requires specification of the tolerance for problem (9), $\epsilon^{RES} \in \mathbb{R}_{>0}$, an initial tolerance for (10), $\epsilon^{LLP} \in \mathbb{R}_{>0}$, which is refined throughout the HSIPA by steps $\epsilon^{LLP} \leftarrow \epsilon^{LLP}/r^{LLP}$ for specified $r^{LLP} > 1$, as detailed in [25, Algorithm 2]. Each parameter is independently varied to determine a set of parameters that give reasonable and reliable computational performance. This results in the set of parameters chosen as $(\epsilon^f, \epsilon^g, r^g, r^{LLP}, \epsilon^{LBP}, \epsilon^{UBP}, \epsilon^{RES}, \epsilon^{LLP}, l_{max}, k_r^{max}) = (5, 10^{-3}, 1.5, 1.2, 0.5, 0.5, 10^{-4}, 0.01, 10, 10)$. The focus of the simulation study is directed to the effect of the initial discretization of both the decision spaces and constraints.

The parameters for the SQPA are chosen similarly, resulting in parameters $(\Gamma, \eta, B, \epsilon_{step}) = (5, 0.33, I_m, 0.5)$.

3) EHSIPA Initializations: The discretizations $\hat{\mathcal{D}}_{ij}$ are chosen to be a uniformly sampled version of the continuous sets $\mathcal{D}_{ij} = [-180, 180]$ with sample period Δ_τ, i.e., $\hat{\mathcal{D}}_{ij} = \{\underline{\tau}_{ij}, \underline{\tau}_{ij} + \Delta_\tau, \ldots, \overline{\tau}_{ij}\}$ for all $i \in \mathbb{N}_{[1,10]}, j \in \{1, 2\}$, and the initial discrete constraint sets $\hat{\mathcal{T}}_i$ are chosen to be uniformly sampled time points with a period Δ. Nine different simulation study configurations are considered as outlined below:

1) $(\Delta_\tau, \Delta) = (60, 15)$;
2) $(\Delta_\tau, \Delta) = (30, 15)$;
3) $(\Delta_\tau, \Delta) = (15, 15)$;
4) $(\Delta_\tau, \Delta) = (5, 15)$;
5) $(\Delta_\tau, \Delta) = (15, \infty)$ (i.e., initial constraint set is empty);
6) $(\Delta_\tau, \Delta) = (15, 30)$;
7) $(\Delta_\tau, \Delta) = (15, 1)$;
8) First the EHSIPA is run with $(\Delta_\tau, \Delta) = (30, 15)$ then the EHSIPA is rerun with discretization centered around the optimal solution with $(\Delta_\tau, \Delta) = (2.5, 15)$. Specifically, for user j on pool i denote τ_{ij}^* as the optimal shift from configuration 2. The restricted set is chosen as $\hat{\mathcal{D}}_{ij} = \{\tau_{ij}^* - 30, \tau_{ij}^* - 27.5, \ldots, \tau_{ij}^*, \ldots, \tau_{ij}^* + 30\}$;
9) $(\Delta_\tau, \Delta) = (15, 0.25)$ with a single iteration of (7) only.

The first four configurations are used to explore the effect of the decision space discretizations, while configurations 3 and 5-7 explore the effect of the initial constraint discretization. Configuration 8 examines the potential for an improved method of decision space updates as discussed in Section IV-A2. Configuration 9, which does not involve update of the initial discretizations, is equivalent to the discrete-time system methods in [1]. It is included for comparison with the

approach proposed here.

C. Implementation of proposed two-stage approach

1) EHSIPA implementation: Numerical solver `ode45` in Matlab is used to evaluate the integrals in the components $\bar{x}^v_{ij}(t)$, $\bar{x}_0(t)$, $i \in \mathbb{N}_{[1,10]}, j \in \mathbb{N}_{[1,2]}$ required to compute (10). By exploiting linearity and time-invariance of (1), each constraint k is calculated for a given schedule by summing $\bar{x}_0(t)$ with shifted versions of $\bar{x}^v_j(t)$ and multiplying the result by the corresponding row C_k. The integrals are evaluated on a uniform discretized set with an initial sample period chosen as $\delta_s = 0.1$. If, during the HSIPA, the maximum difference between any two consecutive samples of the constraint is greater than the current required tolerance ϵ^{LLP}, the sample period δ_s is reduced accordingly and the intergrals are re-evaluated on discretization set with the smaller sample period.

The sub-problems in the HSIPA are formed as the linear (mixed-)integer programs described in Section II-C and solved using Gurobi [37], interfaced with Matlab.

2) Localized refinement stage: For the second localized refinement stage the following three algorithms are compared:

i) the proposed SQPA, described in Section III;
ii) the projected gradient algorithm in [17] with log-penalty and $\epsilon = 1500$;
iii) the projected gradient algorithm in [17] with exponential-penalty and $\vartheta = 1500$.

The line search in the projected gradient algorithm in [17] used for ii) and iii) is modified slightly to include a feasibility condition similar to (22). This allows the comparison to focus on the final objective value and computation time. The derivatives required to implement the SQPA can be constructed via Lemma 3.1 from simulations of the linear system dynamics, as used to evaluate (10) in the EHSIPA. By contrast, the derivatives of the augmented penalty terms needed for the penalty based algorithms are computed using a first order backward finite difference method; i.e., the partial derivative of given function $l(\cdot)$ with respect to shift τ_ℓ at the current schedule $(\tau_j)_{j=1}^m$ is approximated as

$$\frac{\partial l}{\partial \tau_\ell} \approx \frac{l((\tau_j)_{j=1}^m) - l((\tau_j)_{j\neq\ell}; \tau_\ell - \delta_s)}{\delta_s}, \qquad (28)$$

with δ_s set to be smaller than the final sample period from all simulations required for the EHSIPA (for this example $\delta_s = 0.01$). This approximation is computationally much more efficient and numerically robust, compared to calculating the derivatives of the penalty functions on the basis of explicit formulae, as these do not directly correspond to a simulation of the linear dynamics.

All simulations were carried out with Matlab 2018b on a Windows PC with 16GB RAM and Intel Core i7-4790K CPU @4.00GHz processor.

D. Results

The top of Fig 2 shows the requested (unscheduled) off-take profiles and the bottom shows the off-take profiles shifted with a feasible sub-optimal schedule obtained from the proposed

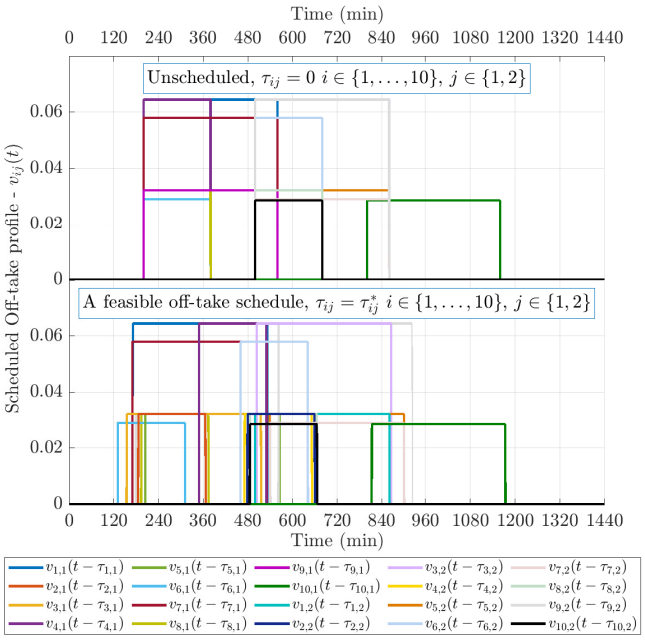


Fig. 2. Top: Unscheduled 20 off-take profiles. Bottom: Off-take profiles under a feasible schedule using proposed approach with configuration 4 followed by SQPA. This feasible sub-optimal schedule is close to original, as per users desire.

algorithm initialized according to configuration 4. The corresponding water level trajectories are displayed in Fig. 3. The shifted off-takes remain "close" to the requested ones under the sub-optimal schedule, and the levels remain within the constraints. Achieving constraint satisfaction relies on the first stage, EHSIPA, to find a feasible schedule. From Fig. 4 it can be seen that the purely discrete-time (uniform sampling) approach from [1] is not able to meet the constraints even with fine discretization (i.e., $\Delta = 0.25$, configuration 9). By contrast, the EHSIPA is able to yield a feasible solution with only 95 constraints (configuration 5); see Fig. 5. In addition, in configurations 3, 5 and 6, which have the same initial decision spaces, the proposed algorithm is almost a whole order of magnitude faster than in configuration 9; see Fig. 6.

The runtime of the EHSIPA increases in proportion to the size of the decision spaces as highlighted by configurations 1-4 in Fig. 6. Configurations 3 and 5-7, have $N_j = 24$ possible choices for each of the 20 off-takes, which corresponds to 24^{20} possible combinations. As such, it is intractable to find the optimal solution via a brute-force (exhaustive search) approach. This highlights the importance of formulating the sub-problems as (mixed-) linear-integer programs as such problems are amendable to widely available dedicated solvers. Configurations with the same initial decision spaces, i.e., 3 and 5-7, are used to explore the effect of the initial constraint discretization. Of these, configuration 6, with $\Delta = 30$, resulted in the fastest run time of the first stage (EHSIPA). Configuration 5 corresponds to the coarsest initial and final discretization, but more constraints are added during the algorithm execution. As such, it requires many more calls to the lower-level problem (10) and iterations of the sub-problems. Note in particular, that the initial lower-bound is much worse for configurations 5 and

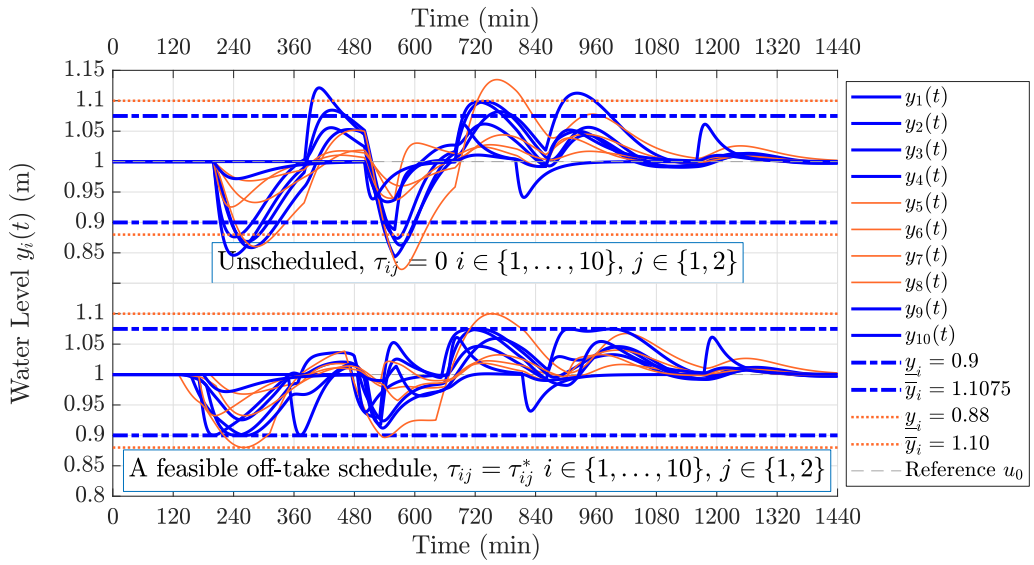


Fig. 3. Levels $y_i(t)$ (in m) in response to nominal off-takes, where constraints are clearly violated on both the upper and lower bounds (top figure); and feasible schedule from configuration 4, where the levels are within the constraint limits (bottom figure).

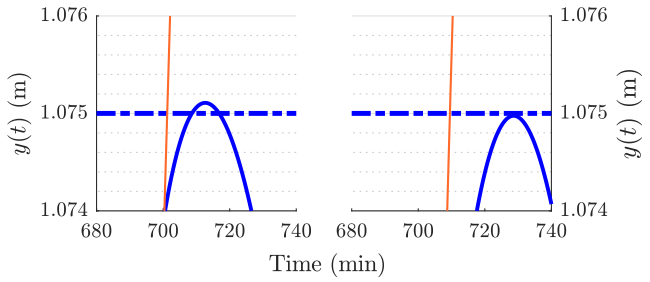


Fig. 4. Levels $y_i(t)$ in response to the off-take schedule using the uniform discretization method from [1], i.e., configuration 9, (left), which does not meet constraints, and a off-tale schedule from configuration 3 combined with SQPA, where the levels are within the constraint limits.

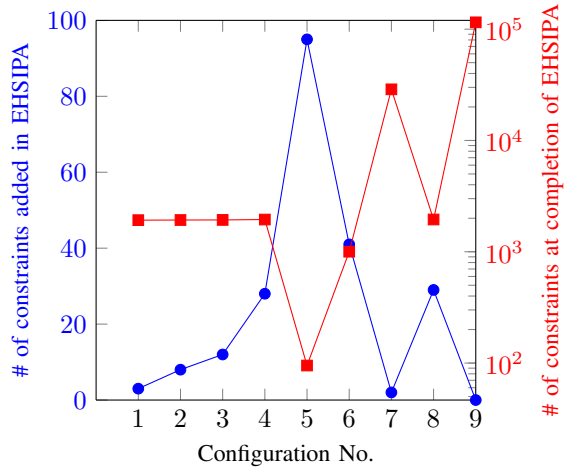


Fig. 5. Number of constraints added throughout the EHSIPA algorithm and the total number of constraints for the 9 different configurations.

6, as shown in Fig. 7. This accentuates the points discussed in Section IV-A1.

The configurations with finer discretization resulted in smaller objective $\hat{f}^*$ obtained from EHSIPA; see Fig. 9. Configuration 8 results in a final objective that is 0.67% lower than configuration 4, however it achieves this in 8.4% of the time. This highlights the potential for further improvements to decision space update methods.

The arrows in Fig. 9 show the impact of the second stage on the overall objective and computation time. Both a longer run time and larger improvement in objective is made with stage 2 for configuration 1. The logarithmic penalty method was the fastest for all configurations but only resulted in marginal improvement of the objective in each case. The proposed method resulted in improvement in all configurations and had the greatest improvement for configurations 3, 4 and 8. Interestingly, for configuration 4 and 8, the exponential penalty method resulted in an increase in the overall objective, shown by the positive gradient of the dotted arrow. That is, the penalty method cannot guarantee a strict improvement of the schedule, unlike the SQPA method; see (25).

Finally, to highlight the applicability of the proposed method in practice, the water levels for the unscheduled and a feasible sub-optimal scheduled off-take loads are simulated using high fidelity models which include, low-level control system actuator saturation and the St Venant PDE model for the open-water dynamics [38]. Fig. 8. shows the optimal schedule only slightly violates the constraints for only two of the pools.

VI. Conclusions and Future Work

A two-stage approach to finding a sub-optimal feasible solution to a rigid-profile input scheduling problem for a continuous-time linear time-invariant dynamical system is proposed. The first stage pertains to the discretization of both the decision spaces and constraints. Through iteratively refining the discretizations and solving tractable (mixed-)integer linear programs as sub-problems, a schedule that is feasible for the original continuous-time problem is generated. This schedule is then improved by the second stage in a way that guarantees improvement in the objective and feasibility with respect to the original problem. A sequential quadratic programming method is applied for this stage and shown to

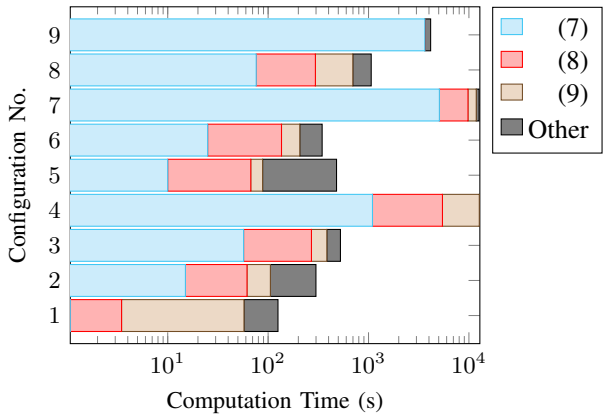


Fig. 6. The breakdown of the total runtime for the EHSIPA for each configuration.

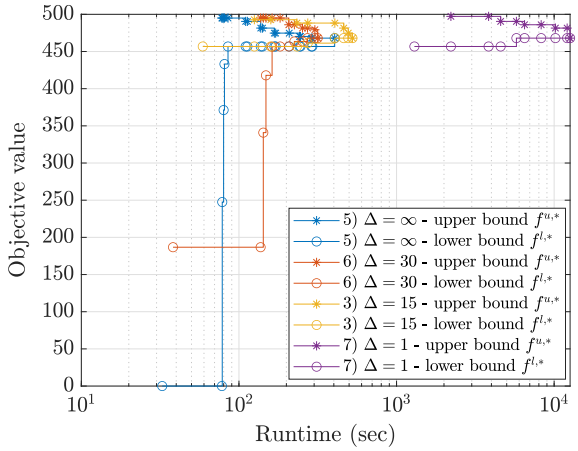


Fig. 7. Figure showing convergence of upper and lower bounds of EHSIPA, for different initial discretizations of $\tilde{\mathcal{T}}_i$. A less dense discretization, $\Delta = \infty$, results in more a conservative initial lower bound. Whereas, a dense discretization, $\Delta = 1$ takes much longer to calculate initial lower bound.

meet both of these requirements. Simulation results from a realistic example from automated irrigation networks illustrate the advantages and some of the properties, trade-offs for the proposed algorithm(s). Future work is to explore how prior information or multiple iterations of the first stage of the algorithm can be used to inform more suitable choices of restricted decision spaces. Other work can be done on characterizing the convergence for the various algorithms, and development of feasible methods for computation of useful optimality gaps to detect acceptable solutions. Finally, further work could explore how to update the schedule, perhaps on a receding horizon basis, to accommodate for model mismatches and to also account for changes to requested orders.

References

[1] J. Alende, Y. Li, and M. Cantoni, "A {0, 1} linear program for fixed-profile load scheduling and demand management in automated irrigation channels," in *Proceedings of the 48th IEEE Conference on Decision and Control*, pp. 597–602, 2009.

[2] M. L. Pinedo, *Scheduling: theory, algorithms, and systems*. Springer, 2016.

[3] J. Blazewicz, M. Drabowski, and J. Weglarz, "Scheduling multiprocessor tasks to minimize schedule length," *IEEE Transactions on Computers*, vol. 100, no. 5, pp. 389–393, 1986.

[4] A. Allahverdi, "The third comprehensive survey on scheduling problems with setup times/costs," *European Journal of Operational Research*, vol. 246, no. 2, pp. 345–378, 2015.

[5] S. Hong, P.-O. Malaterre, G. Belaud, and C. Dejean, "Optimization of irrigation scheduling for complex water distribution using mixed integer quadratic programming ({MIQP})," in *Proceedings of the 10th International Conference on Hydroinformatics (HIC 2012)*, 2012.

[6] J. M. Reddy, B. Wilamowski, and F. Cassel-Sharmasarkar, "Optimal scheduling of irrigation for lateral canals," *ICID Journal on Irrigation and Drainage*, vol. 48, no. 3, pp. 1–12, 1999.

[7] M. Zachar and P. Daoutidis, "Scheduling and supervisory control for cost effective load shaping of microgrids with flexible demands," *Journal of Process Control*, vol. 74, pp. 202 – 214, 2019. Efficient energy management.

[8] L. S. Dias, R. C. Pattison, C. Tsay, M. Baldea, and M. G. Ierapetritou, "A simulation-based optimization framework for integrating scheduling and model predictive control, and its application to air separation units," *Computers & Chemical Engineering*, vol. 113, pp. 139 – 151, 2018.

[9] A. Allman, M. J. Palys, and P. Daoutidis, "Scheduling-informed optimal design of systems with time-varying operation: A wind-powered ammonia case study," *AIChE Journal*, vol. 65, no. 7, p. e16434, 2019.

[10] J. I. Otashu and M. Baldea, "Scheduling chemical processes for frequency regulation," *Applied Energy*, vol. 260, p. 114125, 2020.

[11] B. Burnak, N. A. Diangelakis, J. Katz, and E. N. Pistikopoulos, "Integrated process design, scheduling, and control using multiparametric programming," *Computers & Chemical Engineering*, vol. 125, pp. 164 – 184, 2019.

[12] F. Farokhi, M. Cantoni, and I. Shames, "A game-theoretic approach to distributed scheduling of rigid demands on dynamical dystems," in *Proceedings of the Australian Control Conference*, pp. 147–152, 2016.

[13] E. Weyer, "System identification of an open water channel," *Control Engineering Practice*, vol. 9, no. 12, pp. 1289 – 1299, 2001.

[14] M. Cantoni, E. Weyer, Y. Li, S. K. Ooi, I. Mareels, and M. Ryan, "Control of Large-Scale Irrigation Networks," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 75–91, 2007.

[15] A. Flores-Tlacuahuac and I. E. Grossmann, "Simultaneous cyclic scheduling and control of a multiproduct cstr," *Industrial & Engineering Chemistry Research*, vol. 45, no. 20, pp. 6698–6712, 2006.

[16] J. Zhuge and M. G. Ierapetritou, "Integration of scheduling and control with closed loop implementation," *Industrial & Engineering Chemistry Research*, vol. 51, no. 25, pp. 8550–8565, 2012.

[17] F. Farokhi, M. Cantoni, and I. Shames, "Scheduling rigid demands on continuous-time linear shift-invariant systems," in *Proceedings of the IEEE Conference on Decision and Control*, pp. 5358–5363, 2015.

[18] C. Tsay, A. Kumar, J. Flores-Cerrillo, and M. Baldea, "Optimal demand response scheduling of an industrial air separation unit using data-driven dynamic models," *Computers & Chemical Engineering*, vol. 126, pp. 22 – 34, 2019.

[19] R. Hettich and K. O. Kortanek, "Semi-infinite programming: theory, methods, and applications," *SIAM review*, vol. 35, no. 3, pp. 380–429, 1993.

[20] J. T. Betts, *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming*. Society for Industrial and Applied Mathematics, second ed., 2010.

[21] X. Xu and P. J. Antsaklis, "Optimal control of switched systems based on parameterization of the switching instants," *IEEE Transactions on Automatic Control*, vol. 49, pp. 2–16, jan 2004.

[22] F. Zhu and P. J. Antsaklis, "Optimal control of hybrid switched systems: A brief survey," *Discrete Event Dynamic Systems*, vol. 25, no. 3, pp. 345–364, 2015.

[23] T. Das and R. Mukherjee, "Optimally switched linear systems," *Automatica*, vol. 44, no. 5, pp. 1437–1441, 2008.

[24] L. A. Wolsey, *Integer Programming*. Wiley Series in Discrete Mathematics and Optimization, Wiley, 1998.

[25] H. Djelassi and A. Mitsos, "A hybrid discretization algorithm with guaranteed feasibility for the global solution of semi-infinite programs," *Journal of Global Optimization*, vol. 68, no. 2, pp. 227–253, 2017.

[26] C. J. Price and I. D. Coope, "An exact penalty function algorithm for semi-infinite programmes.," *BIT. Numerical Mathematics*, vol. 30, p. 723, jan 1990.

[27] C. J. Price, "Non-linear semi-infinite programming," 1992.

[28] A. Mitsos, "Global optimization of semi-infinite programs via restriction of the right-hand side," *Optimization*, vol. 60, no. 10-11, pp. 1291–1308, 2011.

[29] A. Mitsos and A. Tsoukalas, "Global optimization of generalized semi-infinite programs via restriction of the right hand side," *Journal of Global Optimization*, vol. 61, no. 1, 2015.

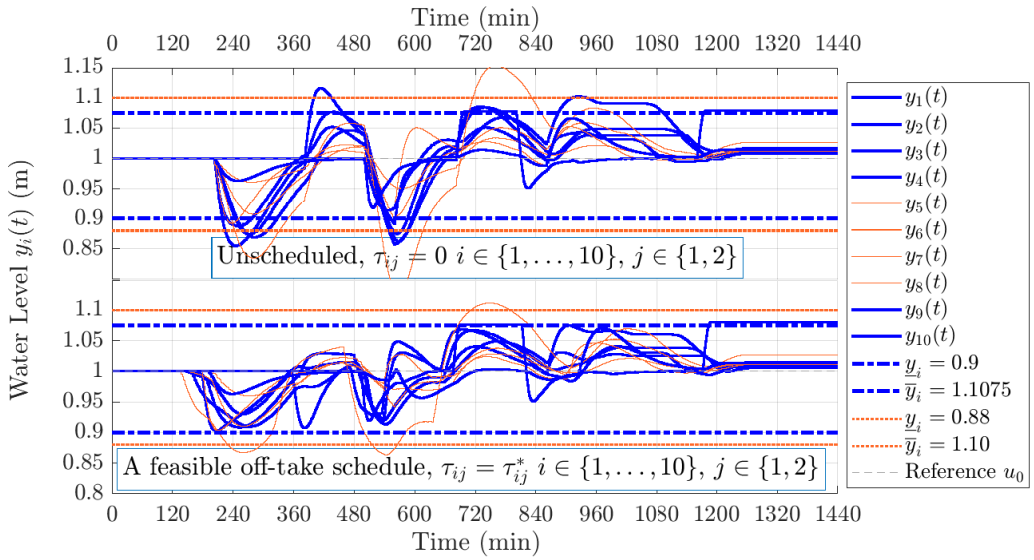


Fig. 8. High fidelity St Venant PDE models based simulation of levels $y_i(t)$ in response to nominal off-takes, where constraints are clearly violated on both the upper and lower bounds (top figure); and a feasible off-take schedule from configuration 4, where the levels are almost all within the constraint limits (bottom figure).

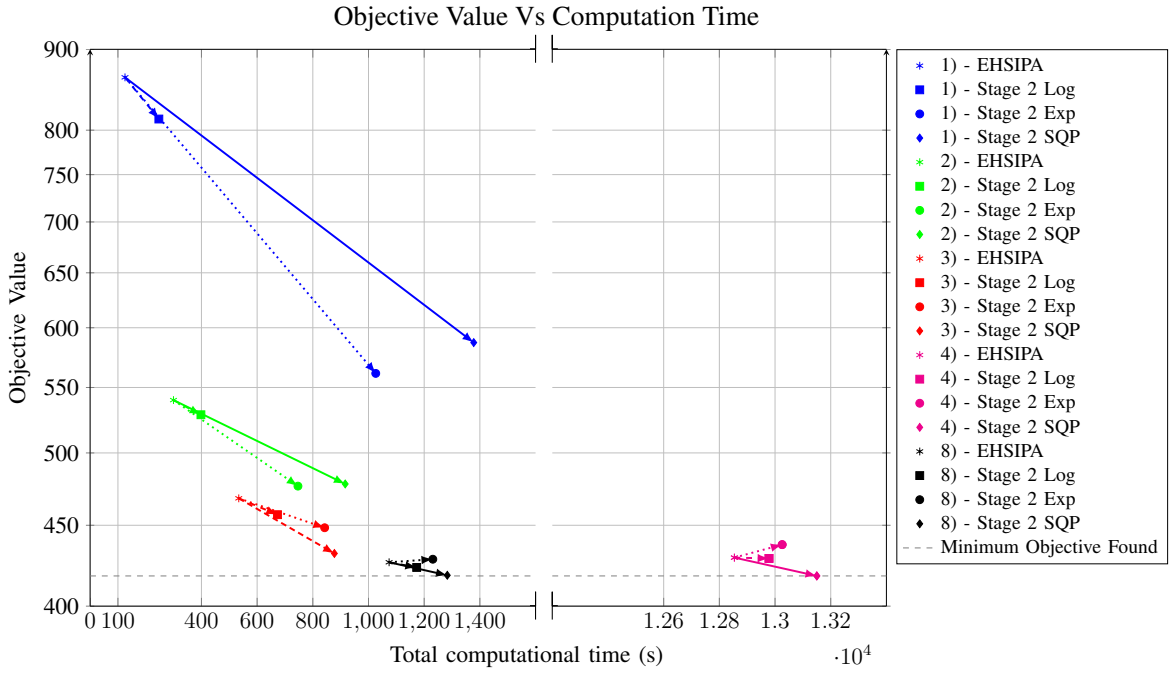


Fig. 9. Runtime vs objective for the different choices of $\hat{\mathcal{D}}_{ij}$ and different algorithms for stage 2 of proposed approach

[30] A. Tsoukalas and B. Rustem, "A feasible point adaptation of the Blankenship and Falk algorithm for semi-infinite programming," *Optimization Letters*, vol. 5, no. 4, pp. 705–716, 2011.

[31] P. Belotti, C. Kirches, S. Leyffer, J. Linderoth, J. Luedtke, and A. Mahajan, "Mixed-integer nonlinear optimization," *Acta Numerica*, vol. 22, pp. 1–131, 2013.

[32] W. Rudin, *Functional Analysis*. International series in pure and applied mathematics, McGraw-Hill, 1991.

[33] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering, New York, NY : Springer New York, 2006., 2006.

[34] R. Reemtsen and J.-J. Rückmann, *Semi-Infinite Programming. [electronic resource]*. Nonconvex Optimization and Its Applications: 25, Springer US, 1998.

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

[36] I. Mareels, E. Weyer, S. K. Ooi, M. Cantoni, Y. Li, and G. Nair, "Systems engineering for irrigation systems: Successes and challenges," *Annual Reviews in Control*, vol. 29, no. 2, pp. 191–204, 2005.

[37] I. Gurobi Optimization, "Gurobi Optimizer Reference Manual," 2016.

[38] M. H. Chaudhry, *Open-channel flow*. Springer Science & Business Media, 2007.

APPENDIX A
PROOFS OF LEMMAS AND THEOREMS

A. *Proof of Lemma 2.1*

Assume $\eta^* < 0$. Since, the objective of (9) is $-\eta$ then optimality of $\eta^* < 0$ implies that for $\eta = 0$ there are no feasible schedules $(\tau_j)_{j=1}^m$ for (9). Let $\mathcal{R} := \{(\tau_j)_{j=1}^m : \tau_j \in \hat{\mathcal{D}}_j, \forall j \in \mathbb{N}_{[1,m]}, \sum_{j=1}^m h_j(\tau_j) \leq f^R\}$, which is non-empty since f^R is chosen to be greater than a lower-bound $\hat{f}^{*,L}$ as

per [25, Algorithm 2]. Then, since $\eta = 0$ is not feasible for (9), for all $(\tau_j)_{j=1}^m \in \mathcal{R}$ there exists a constraint $i \in \mathbb{N}_{[1,n_c]}$ and $t \in \hat{\mathcal{T}}_i$ such that

$$C_i x(t;(\tau_j)_{j=1}^m) - c_i > 0. \tag{29}$$

Therefore, since $\hat{\mathcal{T}}_i \subset \mathcal{T}_i$, there is no schedule in $\mathcal{R}$ that satisfies (5b). Hence $f^R < \hat{f}^*$ which completes the proof.

B. Proof of Lemma 2.3

Let $(\tau_j^c)_{j=1}^m$ be a schedule satisfying Assumption 1.1; i.e., a strictly feasible schedule. Then there exists $\gamma > 0$ such that

$$g_i^*((\tau_j^c)_{j=1}^m) \leq -\gamma. \tag{30}$$

In view (2b) and (2c), the dependence of the left-hand side of the constraint (3b) is continuous in $(\tau_j)_{j=1}^m$, whereby g_i^* is continuous as the max of continuous functions. Therefore, for every $\epsilon > 0$ there exists $\delta(\epsilon) > 0$ such that

$$\max_j \|\tau_j - \tau_j^c\| < \delta(\epsilon) \tag{31a}$$

$$\Rightarrow g_i^*((\tau_j)_{j=1}^m) < g_i^*((\tau_j^c)_{j=1}^m) + \epsilon \leq -\gamma + \epsilon. \tag{31b}$$

In particular, with $\epsilon = \gamma$, if given schedule $(\tau_j)_{j=1}^m$ satisfies the corresponding condition (31a), then it is feasible for the original problem by (31b).

Define the distance between τ_j^c and given $\hat{\mathcal{D}}_j$ as $d_j = \min_{\tau_j \in \hat{\mathcal{D}}_j} |\tau_j - \tau_j^c|$, and let $\bar{d} = \max_j d_j$. It follows by construction of the update (14) that this distance is halved with each update. To achieve feasibility the distance must be made smaller than $\delta(\gamma)$, which from an initial value of $\bar{d}_0$ takes $\lceil \log_2 \frac{\bar{d}_0}{\delta(\gamma)} \rceil$ updates.

C. Proof of Lemma 3.1

Note that

$$\begin{aligned}
&\frac{\partial}{\partial \tau_\ell}\left(C_i \bar{x}_0(t) + C_i \sum_{j=1}^m \bar{x}_j^v(t-\tau_j)\right) \\
&= \frac{\partial}{\partial \tau_\ell} C_i \int_0^{t-\tau_\ell} \exp(A(t-\tau_\ell-\beta)) E_\ell v_\ell(\beta) d\beta \\
&= -C_i E_\ell v_\ell(t-\tau_\ell) \\
&\quad + C_i \int_0^{t-\tau_\ell} \frac{\partial}{\partial \tau_\ell} \exp(A(t-\tau_\ell-\beta)) E_\ell v_\ell(\beta) d\beta \\
&= -C_i E_\ell v_\ell(t-\tau_\ell) \\
&\quad - C_i \int_0^{t-\tau_\ell} A \exp(A(t-\tau_\ell-\beta)) E_\ell v_\ell(\beta) d\beta
\end{aligned}$$

Proof follows from piecewise continuity of $v_j(t-\tau)$.

D. Proof of Theorem 3.2

To facilitate the development, relevant variables are indexed by iteration k in what follows. The initial iteration state is $k = 0$. Each iteration of the SQPA increments the index.

i) Assume that $(\tau_j[k])_{j=1}^m$ satisfies (13) and $\tau_j[k] \in \mathcal{D}_j$ for all $j \in \mathbb{N}_{[1,m]}$. The updated schedule at $k+1$ is

$$\tau_j[k+1] = \tau_j[k] + \gamma p_j[k], \quad \forall j \in \mathbb{N}_{[1,m]}.$$

Note that $\tau_j[k+1] \in \mathcal{D}_j$ since (20c) must hold and through the line search condition (22) then

$$g_i^*((\tau_j[k] + \gamma p_j[k])_{j=1}^m) = g_i^*((\tau_j[k+1])_{j=1}^m) \leq 0.$$

Hence $(\tau_j[k+1])_{j=1}^m$ satisfies (13). As $(\tau_j[0])_{j=1}^m$ is assumed to be feasible, statement one follows by induction.

ii) Since $B[0]$ is postive definite, $B[k]$ is positive definite for all $k \in \mathbb{N}_0$; see [33, Chapter 18]. Let

$$m_k(p) := \frac{1}{2} p^\top B[k] p + \sum_{j=1}^m p_j \frac{d}{d\tau_j} h_j(\tau_j[k]) + \sum_{j=1}^m h_j(\tau_j[k]).$$

Since $B[k]$ is positive definite, for $p \neq 0$,

$$m_k(p) > \sum_{j=1}^m p_j \frac{d}{d\tau_j} h_j(\tau_j[k]) + \sum_{j=1}^m h_j(\tau_j[k]). \tag{32}$$

The optimal objective value for (20) is given by

$$\tilde{f}^*[k] = m_k(p^*[k]). \tag{33}$$

An upper bound for $\tilde{f}^*[k]$ is $m_k(0)$, because $p = 0$ is a feasible solution for (20). Combining this with (33) and (32) gives

$$\begin{aligned}
\sum_{j=1}^m p_j^*[k] \frac{d}{d\tau_j} h_j(\tau_j[k]) + \sum_{j=1}^m h_j(\tau_j[k]) &< \tilde{f}^*[k] \\
&\leq \sum_{j=1}^m h_j(\tau_j[k]).
\end{aligned}$$

Thus, $\sum_{j=1}^m p_j^*[k] \frac{d}{d\tau_j} h_j(\tau_j[k]) < 0$ for $p^*[k] \neq 0$.

The last inequality, the line-search condition (23), and the fact that γ and η are positive, combine to give

$$\sum_{j=1}^m h_j(\tau_j[k] + \gamma p_j^*[k]) = \sum_{j=1}^m h_j(\tau_j[k+1]) < \sum_{j=1}^m h_j(\tau_j[k]),$$

provided $p^*[k] \neq 0$. This is equivalent to

$$m_{k+1}(0) < m_k(0) \tag{34}$$

when $p^*[k] \neq 0$. Consider first that $p^*[k] = 0$ for some $k \neq 0$. Then the termination condition (24) is satisfied and statement two holds by (34) and the fact $k > 0$. Otherwise, $p^*[k] \neq 0$ for all $k \in \mathbb{N} \cup \{0\}$ by hypothesis. In this case, the sequence $(m_k(0))_{k=0}^\infty$ is monotonically decreasing, and it is bounded since h_j are continuously differentiable and each delay must lie in a closed bounded set $\mathcal{D}_j$. As a result, this sequence converges to a value denoted $\underline{m}$ i.e., $\lim_{k\to\infty} m_k(0) = \underline{m} < m_0(0)$. This implies that for any $\epsilon > 0$ there exists a positive integer K such that

$$m_{k-1}(0) - m_k(0) < \epsilon, \quad k \geq K. \tag{35}$$

Let $\epsilon < -\frac{\epsilon_{step}}{\Gamma[k]} \eta \sum_{j=1}^m p_j^*[k-1] \frac{d}{d\tau_j} h_j(\tau_j[k-1])$. Then combining the line-search condition (23) and (35) leads to

$$\gamma < \frac{\epsilon^s}{\Gamma[k]}, \quad k \geq K.$$

Consequently, $\gamma \|p^*[k-1]\|_\infty < \epsilon^s$, $k \geq K$, as $\|p^*[k-1]\|_\infty \leq \Gamma[k]$. Hence, the SQPA terminates after $K \geq 1$ steps and $m_K(0) < m_0(0)$.