A distributed optimization approach for complete vehicle energy management
Romijn, T.C.J.; Donkers, M.C.F.; Kessels, J.T.B.A.; Weiland, S.

Published in:
IEEE Transactions on Control Systems Technology

DOI:
10.1109/TCST.2018.2789464

Accepted/In press: 01/01/2018

Document Version
Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:
- A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 21. Nov. 2018
A Distributed Optimization Approach for Complete Vehicle Energy Management

T.C.J. Romijn, M.C.F. Donkers, J.T.B.A. Kessels, and S. Weiland

Abstract—In this paper, a distributed optimization approach is proposed to solve the complete vehicle energy management problem of a hybrid truck with several controllable auxiliaries. The first part of the approach is a dual decomposition, which allows the underlying optimal control problem to be solved for every subsystem separately. For the second part of the approach, the optimal control problem for every subsystem is further decomposed by splitting the control horizon into several smaller horizons. Two methods for splitting the control horizon are used; the first method uses Alternating Direction Method of Multipliers and divides the horizon a priori, while the second method divides the horizon iteratively by solving unconstrained optimization problems analytically. We demonstrate the approach by solving the complete vehicle energy management problem of a hybrid truck with a refrigerated semi-trailer, an air supply system, an alternator, a DCDC converter, a low-voltage battery, and a climate control system. Simulation results show that the fuel consumption can be reduced up to 0.52% by including smart auxiliaries in the energy management problem. More interestingly, the computation time is reduced by a factor of 64 up to 1825, compared with solving a centralized convex optimization problem.

Index Terms—Energy Management, Optimal Control, Distributed Optimization, Hybrid Vehicles, Smart Auxiliaries

I. INTRODUCTION

REDUCING fuel consumption has been one of the top priorities of the automotive industry in the last decades. A low fuel consumption is important for a sustainable society and is also one of the largest competitive factors for sales. Over the last decades, several new technologies have been introduced in vehicles to improve fuel efficiency. A major improvement is obtained with the introduction of hybrid technology. By adding an electric motor with a high-voltage battery to the powertrain, braking energy can be recuperated and the internal combustion engine can be controlled at a more efficient operating point. The energy management strategy in this case is restricted to determining the (optimal) powersplit between the electric motor and the internal combustion engine. The energy consumption is, however, not only related to the driving functionality of the vehicle, especially for heavy-duty vehicles, as part of the energy is used for auxiliaries, e.g., an air supply system, a refrigerated semi-trailer and a climate control system. These auxiliaries have a potential energy buffer that can be utilized by the energy management system to schedule energy flows and thereby further improving the vehicle energy efficiency, i.e., Complete Vehicle Energy Management (CVEM, see [1]). At the same time, heavy-duty vehicles are augmented with many different type of auxiliaries, which requires the energy management system to be flexible and scalable to reduce development time and costs. Moreover, as heavy-duty vehicles typically drive long distances, it is needed to solve the CVEM problem over large horizons to demonstrate and benchmark the benefit of CVEM.

Many different solution strategies for solving the energy management problem have been proposed over the past decades. The proposed solution strategies can be divided into so-called online and offline solution strategies [2]–[4]. Online solution strategies are real-time implementable and can be further divided in rule-based strategies (see, e.g., [5]), equivalent consumption minimization strategies (ECMS, see, e.g., [6]–[9]) or model predictive control (MPC, see, e.g., [10], [11]). The online solution strategies rely on feedback and/or predictions and can therefore not guarantee optimality of the solution. To verify the performance of the online solution strategies and to analyze different powertrain topologies or component sizes, so-called offline solution strategies have been developed based on, e.g., dynamic programming (DP, see, e.g., [12]–[15]), Pontryagin’s minimum principle (PMP, see, e.g., [16]–[18]) or static convex optimization (see, e.g., [19]–[21]). The offline solution strategies require all disturbances to be known (e.g., the driving cycle) so that the (close-to) optimal solution can be computed and can therefore not be implemented in real-time. It should be noted that any offline solution method uses approximations (e.g., discretisation of the state space in DP, finite accuracy of the shooting methods used to solve the two-point boundary value problem in PMP, and the models approximation used in the approaches based on static convex optimization). Still, the approximation errors are generally small and the offline methods do provide a benchmark for online solution methods, thereby making them valuable tools. We focus on offline solution strategies in this paper.

While online optimization methods based on ECMS might be able to handle the complexity of the CVEM problem, the aforementioned offline optimal control methods cannot. It should be noted that multi-state energy management problems, e.g., including battery state-of-health [22], battery aging [23], [24], thermal management [25], [26] and the control of a waste heat recovery system [27] are all based on ECMS, meaning that tuning is needed and the quality of the solution
needs to be benchmarked with an offline solution method. For the offline optimization methods, scalability is poor as the computational complexity of DP increases exponentially with the number of states and solving the two-point boundary value problem resulting from PMP is difficult, particularly when state constraints are present, see, e.g., [25] in the context of thermal dynamics. Finally, a convex approximation of the energy management problem can handle high-order systems (unlike DP) and incorporate state constraints (unlike PMP), but still requires a large-scale optimization problem to be solved.

For this reason, distributed solutions for energy management start to appear. In [28], [29], an online implementable game-theoretic approach to CVEM is shown. In [30], scalability is obtained by using the Alternating Direction Method of Multipliers (ADMM) while ideas based on ECMS are used to calculate the equivalent costs at a supervisory level. Still, these distributed solutions are all online solution methods for which a (close-to) optimal solution is not guaranteed.

In this paper, we propose to use methods from distributed optimization to solve the convex approximation of the CVEM problem and obtain the optimal solution (for this convex optimization problem) which can be used to verify the performance of online solution methods and to analyze the fuel-reduction potential of different auxiliaries. Moreover, the method presented in this paper serves as a basis for the distributed MPC strategy presented in [11]. We use the dual decomposition approach for CVEM that we first introduced in [31] in combination with efficient algorithms to solve the dual functions that we first introduced in [32]. Dual decomposition has been used in large-scale optimization since the early 1960s [33] and have since then been applied to problems with large-scale dynamical systems (see, e.g., [34]–[36]).

With the dual decomposition approach [31], the optimal control problem is decomposed into smaller dual problems. Each of the dual problems is then either solved explicitly, with an ADMM algorithm or solved with a Lagrangian Method [32]. Both papers showed good performance on a simplified CVEM problem which considered only an internal combustion engine, an electric motor, a high-voltage battery and a refrigerated semi-trailer while real scalability of the method with the number of components has not been demonstrated. Contrary to [31] and [32], we present the general CVEM problem as a quadratically constrained linear program (QCLP), where we apply a relaxation to ensure convexity of the CVEM optimal control problem. This allows the method to be applied to various vehicle configurations. Moreover, the dual decomposition approach is extended with a Newton dual update algorithm to improve convergence speed. The ADMM algorithm is extended for multi-state dynamical systems and the Lagrangian Method is extended to handle systems with saturated states on the lower or upper bound. Finally, to fully demonstrate the performance of the distributed optimization approach, the optimal control of a hybrid truck with an internal combustion engine, an electric motor, a high-voltage battery, a refrigerated semi-trailer, an air supply system, an alternator, a DCDC converter, a low-voltage battery and a climate control system is solved. Moreover, the proposed algorithm is compared with the state-of-the-art solver CPLEX [37].

The remainder of this paper is organized as follows. The general optimal control problem and the application of the dual decomposition is given in Section II. In Section III, solution methods are presented to solve the dual functions that resulted from the dual decomposition. The CVEM problem is casted as an optimal control problem in Section IV and, finally, simulation results are presented in Section V.

II. DISTRIBUTED OPTIMIZATION OF POWER NETS

In this section, we consider the optimal control of the energy flows in a power net, which is illustrated in Fig. 1. The power net consist of energy storage devices, e.g., a high-voltage battery, and energy converters, e.g., an electric machine. The storage devices are connected to the converters, while the outputs $y_{m,k}$ and inputs $u_{m,k}$ of the converters are connected to each other via nodes according to a specific topology, i.e., energy can be exchanged directly between converters but not directly between storage devices. At each node $n \in \{1, \ldots, N\}$, there is also a known exogenous load signal $v_{n,k}$ given for each time instant $k$. Subsystems are composed of a combination of a converter, possibly with an energy storage device. The goal of the power net is to minimize the cumulative energy losses of all subsystems, while meeting constraints on the inputs, outputs and states in each subsystem. In this section, we will introduce the optimal control problem for this power net and we will give a dual decomposition approach to solve the optimal control problem. In Section IV we will show that the Complete Vehicle Energy Management (CVEM) problem can be represented as a power net, where minimizing the energy losses is equivalent to minimizing the fuel consumption.

A. Optimal Control Problem

The optimal control problem for the power net is given by
\begin{equation}
\min_{\{u_{m,k}, y_{m,k}\}} \sum_{m \in M} \sum_{k \in K} c_m u_{m,k} - d_m y_{m,k},
\end{equation}

where $u_{m,k} \in \mathbb{R}$ and $y_{m,k} \in \mathbb{R}$ are the (scalar) inputs and outputs of the converter in subsystem $m \in M = \{1, \ldots, M\}$ with $M$ the number of subsystems and at time instant $k \in K = \{0, 1, \ldots, K - 1\}$, with $K$ the horizon length. In (1a), $c_m \in \mathbb{R}$ and $d_m \in \mathbb{R}_+$ are coefficients that...
define the energy losses in converter \( m \). Moreover, we use the notation \( \{u_{m,k}, y_{m,k}\} \) to indicate \( \{u_{m,k}, y_{m,k}\} \in \mathcal{M}, k \in \mathcal{K} \). This notation will be used throughout the paper for minimizing over a set. The optimization problem (1a) is to be solved subject to (s.t.) a quadratic equality constraint describing the input-output behavior of each converter, i.e.,

\[
\frac{1}{2} q_{m,k} u_{m,k}^2 + f_{m,k} u_{m,k} + e_{m,k} + y_{m,k} = 0, \tag{1b}
\]

with \( q_{m,k} \in \mathbb{R}_+ \), \( f_{m,k} \in \mathbb{R} \) and \( e_{m,k} \in \mathbb{R} \) being efficiency coefficients of the converter \( m \in \mathcal{M} \) at time instant \( k \in \mathcal{K} \), and subject to linear system dynamics of the storage device in subsystem \( m \in \mathcal{M} \), i.e.,

\[
x_{m,k+1} = A_m x_{m,k} + B_{m,w} w_{m,k} + B_{m,u} u_{m,k}, \tag{1c}
\]

for all \( k \in \mathcal{K} \), where the initial state \( x_{m,0} \) and final state \( x_{m,K} \) of the storage device are assumed to be given, \( u_{m,k} \) is a known load signal at every time instant \( k \) and the input \( u_{m,k} \in \mathbb{R} \) is subject to linear inequality constraints, i.e.,

\[
\underline{u}_{m,k} \leq u_{m,k} \leq \overline{u}_{m,k}, \tag{1d}
\]

for all \( k \in \mathcal{K} \), \( m \in \mathcal{M} \), and the state \( x_{m,k} \) is subject to linear inequality constraints, i.e.,

\[
\underline{x}_{m,k} \leq x_{m,k} \leq \overline{x}_{m,k}, \tag{1e}
\]

for all \( k \in \mathcal{K} \) and \( m \in \mathcal{M} \). Finally, the optimization problem is solved subject to a linear equality constraint describing the interconnection of the subsystems, i.e.,

\[
\sum_{m \in \mathcal{M}} \Gamma_{m,u} u_{m,k} + \Theta_{m,y} y_{m,k} + \frac{1}{M} v_k = 0, \tag{1f}
\]

for all \( k \in \mathcal{K} \), where \( \Gamma_{m} \in \mathbb{R}^{N} \) and \( \Theta_{m} \in \mathbb{R}^{N} \) are vectors with the \( n \)-th element being \(-1\) if the power flow to node \( n \) is positive, \( 0 \) if there is no power flow to node \( n \) and \( 1 \) if the power flow to node \( n \) is negative. Here, \( N \) is the number of nodes in the topology where power is aggregated. Furthermore, the load signal \( v_k = [v_{1,k}, \ldots, v_{N,k}]^T \in \mathbb{R}^N \) is assumed to be known at each time instant \( k \in \mathcal{K} \). We define the primal optimal solution as the solution \( \{u_{m,k}, y_{m,k}\} \) that satisfies (1) and \( p^* \) as the primal optimal value of (1). Note that quadratic behavior (1b) is a good assumption for a wide range of converters as most converters have quadratic power losses, e.g., the internal combustion engine or the electric machine, which will be shown in Section IV.

\section{Convex Relaxation and Dual Decomposition}

The optimization problem (1) can be a large-scale problem (when \( K \) and \( M \) are large), which is not convex due to the quadratic equality constraint (1b). We propose in this paper to solve (1) by relaxing (1b) and decomposing it into several smaller problems by applying a dual decomposition. We will show below that this allows solving (1) efficiently without sacrificing optimality of the solution.

The problem (1) is not convex due to the quadratic equality constraint (1b). By relaxing this constraint to an inequality constraint, i.e.,

\[
y_{m,k} + \frac{1}{2} q_{m,k} u_{m,k}^2 + f_{m,k} u_{m,k} + e_{m,k} \leq 0, \tag{1b'}
\]

for \( m \in \mathcal{M}, k \in \mathcal{K} \), the relaxed optimization problem

\[
\begin{aligned}
\min_{\{u_{m,k}, y_{m,k}\}} & \quad \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{K}} c_{m} u_{m,k} - d_{m} y_{m,k}, \\
\text{s.t.} & \quad (1b'), (1c) - (1f)
\end{aligned}
\]

is convex. Still, optimization problem (2) cannot be separated due to the complicating constraint (1f). Therefore, we decompose the problem via dual decomposition by introducing the following so-called partial Lagrangian

\[
L(\{u_{m,k}, y_{m,k}, \mu_k\}) = \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{K}} c_{m} u_{m,k} - d_{m} y_{m,k} + \mu_k^T (\Gamma_{m,u} u_{m,k} + \Theta_{m,y} y_{m,k} + \frac{1}{M} v_k), \tag{3}
\]

where \( \mu_k \in \mathbb{R}^N \) is a Lagrange multiplier. Indeed, the partial Lagrangian is obtained by adding the complicating constraints (the constraints that act on more than one subsystem) to the objective function (1a). The partial Lagrange dual function of optimization problem (2) is now given by

\[
g(\{\mu_k\}) = \min_{\{u_{m,k}, y_{m,k}\}} L(\{u_{m,k}, y_{m,k}, \mu_k\}) \tag{4a}
\]

s.t. \( (1b'), (1c) - (1e) \), under the assumption that the infimum is attained. Because it holds that

\[
g_m(\{\mu_k\}) = \min_{\{u_{m,k}, y_{m,k}\}} \sum_{k \in \mathcal{K}} (c_{m} + \Gamma_{m,u}^T \mu_k) u_{m,k} - (d_{m} - \Theta_{m,y}^T \mu_k) y_{m,k} \tag{4b}
\]

s.t. \( (1b'), (1c) - (1e) \), each of the Lagrange dual functions \( g_m(\{\mu_k\}) \) is related to only one of the subsystems, meaning that they can be solved independently. The dual problem of (2) is given by

\[
\max_{\{\mu_k\}} g(\{\mu_k\}) = d^*_CR. \tag{5}
\]

where \( d^*_CR \) is defined as the dual optimal value of (2). The dual problem (5) gives a lower bound on the primal optimal value \( p^*_CR \) of problem (2), which is a relaxation of the primal optimal solution of (1), i.e.,

\[
d^*_CR \leq p^*_CR \leq p^*. \tag{6}
\]

The dual optimal value \( d^*_CR \) will be equal to the primal optimal value \( p^*_CR \), i.e., \( p^*_CR = d^*_CR \), if the constraints satisfy Slater’s constraint qualifications [38]. The conditions under which the convex relaxation is lossless, i.e., \( p^*_CR = p^* \), will be provided in the theorem below. Before presenting this theorem, we will present two lemmas.

\textbf{Lemma 1.} Assume that (1d) has finite bounds and that there exists a feasible point \( \{u_{m,k}, y_{m,k}\} \) for all \( m \in \mathcal{M} \) and for all \( k \in \mathcal{K} \) for (2) with strict inequalities (1b’), (1d) and (1e). Then, there exists a solution \( \{u_{m,k}, y_{m,k}\} \) to (2) and \( \{\mu_k^*\} \), such that \( d^*_CR = p^*_CR > -\infty \).

\textbf{Proof.} The existence of a strictly feasible point \( \{u_{m,k}, y_{m,k}\} \) for all \( m \in \mathcal{M} \) and for all \( k \in \mathcal{K} \) for (2) (i.e., with strict inequalities (1b’), (1d) and (1e)) implies satisfaction of Slater’s constraint qualification (see e.g., [38]). Hence, strong duality holds, i.e., \( d^*_CR = p^*_CR \), provided that the minimum is attained.
It remains to be shown that the minimum in the convex, yet not strictly convex, optimization problem (2) is attained, i.e., \( p_{CR}^* > -\infty \). Because we assume that (1d) has finite bounds, we have that \( y_{m,k} \leq \bar{y}_{m,k} \) with

\[
\bar{y}_{m,k} = \max_{u_{m,k} \in [\underline{y}_{m,k}, \overline{y}_{m,k}]} \left\{ -\frac{1}{2} q_{m,k} u_{m,k}^2 - f_m u_{m,k} - e_{m,k} \right\},
\]

due to (1b'). This maximisation is well defined, because \( q_{m,k} > 0 \). As a result, the objective function of (2) has a finite lower bound, because every term in (2) has a finite lower bound, i.e.,

\[
c_m u_{m,k} - d_m y_{m,k} \geq e_m \overline{u}_{m,k} - d_m \bar{y}_{m,k} > -\infty
\]

because \( e_m, d_m \in \mathbb{R}_+ \). Therefore, if (2) is feasible, \( p_{CR}^* \) is finite. This completes the proof. \( \Box \)

**Lemma 2.** Assume (4b) is feasible and (1d) has finite bounds. If \( \mu_k \) satisfies \( d_m - \Theta_{m,k}^T \mu_k \geq 0 \) for all \( m \in \mathcal{K} \), then the minimum in (4b) is attained, i.e., \( y_m(\{\mu_k\}) > -\infty \). Furthermore, for \( \mu_k \) satisfying \( d_m - \Theta_{m,k}^T \mu_k \geq 0 \), the optimal solution to (4b) satisfies (1b') with equality, i.e., it satisfies (1b), while for \( d_m - \Theta_{m,k}^T \mu_k = 0 \), an optimal solution for (4b) exists that satisfies (1b') with equality.

**Proof.** Because (1d) has finite bounds, we have that the first term in the objective function (4b) is always bounded. Now note that \( y_{m,k} \) only appears in the second term of the objective function and in (1b'). This means that if \( \mu_k \) satisfies \( d_m - \Theta_{m,k}^T \mu_k < 0 \) for some \( k \in \mathcal{K} \), minimising over \( y_{m,k} \) leads to \( y_{m,k} \rightarrow -\infty \) for these \( k \in \mathcal{K} \), which causes the objective function to be unbounded, while (1b') remains satisfied. This proves why \( d_m - \Theta_{m,k}^T \mu_k \geq 0 \) for all \( k \in \mathcal{K} \) is needed.

We will now show that the optimal solution can satisfy (1b') with equality, if \( d_m - \Theta_{m,k}^T \mu_k \geq 0 \) for all \( k \in \mathcal{K} \). To do so, we augment the objective function (4b) with (1b') to obtain the partial Lagrange dual function of problem (4b), which is given by

\[
L(\{u_{m,k}\}, \{\mu_k\}) = \sum_{k \in \mathcal{K}} \left( c_m + \Gamma_{m,k} \mu_k \right) u_{m,k} - \left( d_m - \Theta_{m,k}^T \mu_k \right) y_{m,k} + \nu_{m,k}(y_{m,k} + \frac{1}{2} q_{m,k} u_{m,k}^2 + f_m u_{m,k} + e_{m,k}),
\]

where \( \nu_{m,k} \geq 0 \) is the (scalar) Lagrange multiplier associated with the (scalar-valued) quadratic inequality constraint (1b').

The stationarity conditions with respect to \( y_{m,k}(\{\mu_k\}) \) (one of the necessary conditions for optimality, see, e.g., [38]) is given by

\[
-(d_m - \Theta_{m,k}^T \mu_k) + \nu_{m,k} = 0.
\]

Now, for time instants \( k \in \mathcal{K} \) satisfying \( d_m - \Theta_{m,k}^T \mu_k > 0 \), it holds that \( \nu_{m,k} > 0 \). The positivity of \( \nu_{m,k} \) ensures that inequality (1b') is satisfied as an equality by complementarity slackness [38]. For time instants \( k \in \mathcal{K} \) for which \( d_m - \Theta_{m,k}^T \mu_k = 0 \), we should observe that \( y_{m,k} \) disappears from the objective function and any \( y_{m,k} \) satisfying (1b') leads to the same optimal value, meaning that (1b') with equality is an optimal solution. This completes the proof. \( \Box \)

Lemma 1 provides a condition for which strong duality holds. These conditions (i.e., Slater’s constraint qualification, see, e.g., [38]) are relatively mild and can be satisfied for the numerical example given in Section IV. Lemma 2 states that the minimisation of the dual function, if the minimum is attained, leads to solutions that satisfy (1b), instead of (1b'). These lemmas will be used to prove the main result that we will now present.

**Theorem 3.** Assume that (1d) has finite bounds and that there exists a feasible point \( \{u_{m,k}, y_{m,k}\} \) for all \( m \in \mathcal{M} \) and for all \( k \in \mathcal{K} \) for (2) with strict inequalities (1b'), (1d) and (1e). If \( d_m - \Theta_{m,k}^T \mu_k > 0 \) for all \( m \in \mathcal{M} \) and all \( k \in \mathcal{K} \), then \( \{u_{m,k}, y_{m,k}\} \) is also the solution to the nonconvex problem (1).

**Proof.** The proof follows from Lemma 1 and Lemma 2. If the solution to (5) satisfies \( d_m - \Theta_{m,k}^T \mu_k > 0 \) for all \( m \in \mathcal{M} \) and all \( k \in \mathcal{K} \), the solution to (2) is also a solution to (1). \( \Box \)

The results of this section justify using the dual decomposition to solve the nonconvex optimization problem (1). We will propose to solve the dual problem (5) using a subgradient method as will be shown below. In this approach, we will assume that (1b) holds in every step of the subgradient method. This will lead to an efficient solution.

**C. Maximizing the Lagrange Dual Function**

Maximizing the Lagrange dual function (4) over \( \mu_k \) can be done with a ‘steepest ascent’ method, i.e.,

\[
\mu_k^{+1} = \mu_k^* + \alpha_k^* \left( \sum_{m \in \mathcal{M}} \Gamma_m u_{m,k}^* + \Theta_m y_{m,k}^* + \frac{1}{m} v_k \right),
\]

for all \( k \in \mathcal{K} \) where \( \alpha_k^* \) is a suitably chosen matrix and \( s \in \mathbb{N} \) is the iteration counter. In [31], a diagonal matrix with sufficiently small positive constant step sizes on its diagonal was chosen, such that the Lagrange dual problem will always converge but convergence tended to be slow. A better convergence rate is achieved with a Newton scheme (see e.g., [39]). We will derive this scheme intuitively from a primal feasibility perspective. The idea is to update the dual variables \( \mu_k \) such that for the next iteration primal feasibility for the complicating constraints holds, i.e.,

\[
\sum_{m \in \mathcal{M}} \Gamma_m u_{m,k}^{s+1} + \Theta_m y_{m,k}^{s+1} + \frac{1}{m} v_k = 0.
\]

Furthermore, we can approximate the value for \( u_{m,k}^{s+1} \) and \( y_{m,k}^{s+1} \) linearly by

\[
u_{m,k}^{s+1} \approx u_{m,k}^s + \left( \frac{\partial u_{m,k}^s}{\partial \mu_k} \right)^T (\mu_k^{s+1} - \mu_k^s),
\]

\[
u_{m,k}^{s+1} \approx y_{m,k}^s + \left( \frac{\partial y_{m,k}^s}{\partial \mu_k} \right)^T (\mu_k^{s+1} - \mu_k^s),
\]

where \( \frac{\partial u_{m,k}^s}{\partial \mu_k} \) is a vector with the approximations of the first-order derivatives of \( u_{m,k}(\mu_k) \) with respect to the dual variables \( \mu_k \) at iteration \( s \). Similarly, \( \frac{\partial y_{m,k}^s}{\partial \mu_k} \) is a vector with the approximations of the first-order derivatives of \( y_{m,k}(\mu_k) \) with respect to the dual variables \( \mu_k \) at iteration \( s \). Because of Lemma 2, we can use the fact that (1b') is (can be) satisfied with equality, i.e., (1b) holds in every iteration, so that the
derivatives $y_{m,k}(\mu_k)$ depend on $u^{s_1}_{m,k}(\mu_k)$. By substituting (13) into (12) and solving for $\mu_{k+1}$ we obtain (11) with

$$
\alpha_k = \left( \sum_{m \in M} -\Gamma_m \left( \frac{\partial u^{s_1}_{m,k}}{\partial \mu_k} \right)^T \right)^{-1} \left( \sum_{m \in M} -\Theta_m \left( \frac{\partial y_{m,k}}{\partial \mu_k} \right)^T \right),
$$

which can be obtained by calculating the vector with derivatives for each subsystem in a distributed fashion. Note that this particular choice for $\alpha_k$ might cause the conditions of Lemma 2 to be violated. Furthermore, calculating the derivatives with respect to $\mu_k$ in (13) can be hard, might not even exist due to the presence of constraints or might cause the inverse in (14) to not exist (as the term in the right-hand side might not be full rank). Instead, the derivatives can be approximated by neglecting the state constraints (1e). This might cause that (11) does not converge. In this case, sufficiently small constant step sizes can be chosen as was done in [31] or regularization can be applied, as was proposed in [40] to handle lack of strict convexity. Convergence speed might be significantly slower in this case, as will be shown in the simulation study in Section V. Finally, the dual decomposition algorithm consists of iteratively solving (4) to obtain $\{u^{s_1}_{m,k}, y^{s_1}_{m,k}\}$ and updating the Lagrange multipliers by solving (11) to obtain $\{\mu_{k+1}\}$. In the section below, we will provide methods to efficiently solve the dual functions (4b).

III. EVALUATING THE DUAL FUNCTIONS

Each of the Lagrange dual functions (4b) related to one of the subsystems can be solved separately and can be written as a linearly constrained quadratic program (LCQP) by substituting (1b) into (4b). This is justified by Lemma 2 and gives

$$
g_m(\{\mu_k\}) = \min_{\{u_{m,k}\}} \sum_{k \in K} \mathcal{H}_{m,k} u_{m,k}^2 + \mathcal{F}_{m,k} u_{m,k} + \mathcal{E}_{m,k}
$$

s.t. (1c) - (1e)

with

$$
\mathcal{H}_{m,k} = (d_m - \Theta^T_m \mu_k) q_{m,k},
\mathcal{F}_{m,k} = c_m + \Gamma^T_m \mu_k + (d_m - \Theta^T_m \mu_k) f_{m,k},
\mathcal{E}_{m,k} = (d_m - \Theta^T_m \mu_k) c_{m,k}.
$$

Note that for strict convexity of (15), it is required that $d_m - \Theta^T_m \mu_k > 0$ for all $k \in K$, $m \in M$, which is a slightly more restrictive condition than the condition in Lemma 2 for the partial Lagrange dual function to attain its minimum.

As a result of this substitution, the dual decomposition allows solving the quadratically constrained quadratic program by solving multiple LCQPs iteratively. However, solving a LCQP for a large horizon length $K$ is still very inefficient. Therefore, we introduce two solution methods to solve the optimization problem (15), related to each of the subsystems, efficiently. Both methods use the principle of splitting the horizon $K$ into multiple smaller intervals, where each interval is defined as $K_\ell = \{K_{\ell-1}, \ldots, K_{\ell-1}\}$ with $0 = K_0 < K_1 < \ldots < K_L = K$ and where $\ell \in \mathcal{L} = \{1, \ldots, L\}$ with $\mathcal{L}$ the number of intervals. To decompose the constraints in the optimization problem (15) into smaller optimization problems, we recall that a solution to (1c) satisfies

$$
x_{m,k+1} = A^{k+1}_{m,K_{\ell-1}} \hat{x}_{m,K_{\ell-1}} + \sum_{i \in \{K_{\ell-1}, \ldots, k\}} A_{m,i} (B_{m,u} u_{m,i} + B_{m,w} w_{m,i}),
$$

for all $k \in K_\ell$, where the local initial condition $\hat{x}_{m,K_{\ell-1}}$ at each interval $\ell$ is equal to the final condition at interval $\ell - 1$, i.e.,

$$
x_{m,K_{\ell-1}} = \hat{x}_{m,K_{\ell-1}},
$$

Note that the problem (18) is only coupled in $\ell \in \mathcal{L}$ and $\hat{x}_{m,0} = x_{m,0}$ and $\hat{x}_{m,K} = x_{m,K}$ which follow from the initial and final condition of the full horizon. Using these constraints, we can write the dual function (15) as

$$
g_m(\{\mu_k\}) = \min_{\{u_{m,k}, \hat{x}_{m,K_{\ell-1}}\}} \sum_{\ell \in \mathcal{L}} \sum_{k \in K_{\ell}} \frac{1}{2} \mathcal{H}_{m,k} u_{m,k}^2 + \mathcal{F}_{m,k} u_{m,k} + \mathcal{E}_{m,k}
$$

s.t. (1d), (1e), (17).

Note that the problem (18) is only coupled in $\ell \in \mathcal{L}$ by (17b).

We will introduce two solution methods that can be used to select the intervals $K_\ell$ and the initial state at each interval $\hat{x}_{m,K_{\ell-1}}$. In the first solution method, based on Alternating Direction Method of Multipliers (ADMM), the horizon is split a priori in a fixed number of intervals. For each interval, an optimization problem is solved that takes the initial state at each interval $\ell$ as decision variable. ADMM is a suitable method as the resulting decomposed problem is not strictly convex, yet still convex, as we will show below. In the second solution method, based on the Lagrangian Method, the horizon is split iteratively and the initial state is fixed on the lower or upper state constraint depending on the solution of the state-unconstrained optimization problem. The Lagrangian Method is only applicable to systems with scalar states, while the ADMM method is applicable to systems with multiple states.

A. Horizon Splitting with ADMM

For this method, we define a priori the sets $\bar{K}_\ell = \{K_{\ell-1}, \ldots, K_{\ell-1}\}$, $\ell \in \mathcal{L} = \{1, \ldots, L\}$. This method is similar to the method proposed in [41] where intervals that contain only one time instant, i.e., $\bar{K}_\ell = \{\ell - 1\}$ are used for solving the problem over a short horizon. Contrary to [41], we use intervals containing multiple time instants, thereby making it more applicable for solving the problem over a long horizon as will be demonstrated with the numerical example in Section V. The objective function in (18) is separable in variables related to each interval but is not strictly convex due to the minimization over the local initial state $\hat{x}_{m,K_{\ell-1}}$, which is an essential assumption for the dual decomposition approach taken in the previous section. Lagrangian methods as used in Section II.B, however, assume convexity of the objective function rather than strict convexity. Instead, the partial augmented Lagrangian for problem (18) can be defined
as
\[
\hat{L}(\{u_{m,k}, \bar{x}_m, k_{\ell - 1}, \nu_{m,\ell}\}) = \sum_{\ell \in L} \sum_{k \in K_{\ell}} \frac{1}{2} H_{m,k} u_{m,k}^2 + F_{m,k} u_{m,k} + \nu_{m,\ell}^T (x_{m,k_{\ell - 1}} - \bar{x}_m) + \frac{1}{2} (x_{m,k_{\ell - 1}} - x_m)^T R (x_{m,k_{\ell - 1}} - x_m),
\]
where \( R > 0 \) is a diagonal matrix with penalty parameters on its diagonal. In this expression, we temporarily omit the constraints that are acting only within one interval, i.e., (1d), (1e) and (17a) and will reintroduce them later in the decomposed problem. The partial augmented Lagrange dual function is given by
\[
\hat{g}_m(\{v, \nu\}) = \min_{\{u_{m,k}, \bar{x}_m, k_{\ell - 1}\}} \hat{L}_m(\{u_{m,k}, \bar{x}_m, k_{\ell - 1}, \nu_{m,\ell}\})
\]
s.t. (1d), (1e), (17a),
which can be written as
\[
\sum_{\ell \in L} \hat{g}_m(\nu_{m,\ell - 1}, \nu_{m,\ell}) = \min_{\{u_{m,k}, \bar{x}_m, k_{\ell - 1}\}} \sum_{\ell \in L} \frac{1}{2} H_{m,k} u_{m,k}^2 + F_{m,k} u_{m,k} + \nu_{m,\ell}^T (x_{m,k_{\ell - 1}} - \bar{x}_m) + \frac{1}{2} \bar{x}_m^T R \bar{x}_m - \hat{G} \hat{x}_m, \quad (20a)
\]
with
\[
\hat{g}_m(\nu_{m,\ell - 1}, \nu_{m,\ell}) = \min_{\{u_{m,k}, \bar{x}_m, k_{\ell - 1}\}} \sum_{\ell \in L} \frac{1}{2} H_{m,k} u_{m,k}^2 + F_{m,k} u_{m,k} + \nu_{m,\ell}^T (x_{m,k_{\ell - 1}} - \bar{x}_m) + \frac{1}{2} \bar{x}_m^T R \bar{x}_m - \hat{G} \hat{x}_m, \quad (20b)
\]
s.t. (1d), (1e), (17a),
in which
\[
\hat{G}_{m,\ell} = \nu_{m,\ell - 1}^T - \nu_{m,\ell}^T R - \nu_{m,\ell}^T A_{m,k_{\ell - 1}}, \quad (21a)
\]
\[
\hat{F}_{m,\ell} = F_{m,k} - \nu_{m,\ell}^T A_{m,k_{\ell - 1}} B_{m,u}, \quad (21b)
\]
\[
\hat{E}_{m,\ell} = E_{m,k} - \nu_{m,\ell}^T A_{m,k_{\ell - 1}} - B_{m,u} u_{m,k}, \quad (21c)
\]
for \( \ell \in L \), with \( \nu_{m,L} = 0 \). Expressions (21) are obtained by substituting (17a) for \( k = K_{\ell} - 1 \) into (19), only for the linear part of the equation. This gives a more desirable expression for (21a), i.e., the Hessian matrix is diagonal. As a result, (21a) is not separable due to the term \( x_{m,k_{\ell - 1}} \) in (21a). By minimizing (20) sequentially from interval \( \ell = 1 \) to interval \( \ell = L \), as part of the ADMM algorithm (see, e.g., [38]), the optimization problem (20), can still be solved efficiently.

To maximize the partial augmented Lagrange dual function, we use a ‘steepest ascent’ method, i.e.,
\[
\nu_{m,\ell}^{t+1} = \nu_{m,\ell}^t - R (\bar{x}_m - x_{m,k_{\ell - 1}}) \quad (22)
\]
with \( t \in \mathbb{N} \) the iteration number and for some given initial condition \( \nu_{m,\ell}^0 \) for \( \ell \in L \). Finally, the ADMM algorithm consists of iteratively solving (20) without (17a) to obtain \( \nu_{m,\ell}^t \) for \( \ell \in L \), \( k \in K \) and solving (17a) for \( k = K_{\ell} - 1 \) to obtain \( x_{m,k_{\ell}} \) for \( \ell \in L \), \( k \in K \), followed by an update of the Lagrange multipliers through (22) to obtain \( \nu_{m,\ell}^{t+1} \) for \( \ell \in L \).

### B. Horizon Splitting with the Lagrangian Method

Fixing the interval \( K_{\ell} \) a priori and use ADMM to solve (18) results in a general solution method to solve the LCQP (15). We will also develop an iterative procedure that involves splitting the intervals based on solving (18) for the particular case where \( x_{m,k} \in \mathbb{R} \), i.e., the energy storage device in subsystem \( m \) is a scalar-state system. In Section IV, it will be shown that many components in the CVEM problem can be represented by a scalar-state system and in Section V we will show that a tailored solution method for these components is more favorable with respect to computation time, which also emphasizes the advantage of using the dual decomposition approach to CVEM where each of the dual functions can be solved with the most suitable solution method.

For this method, we initially take only one interval, i.e. the full horizon, so that \( K_{\ell} = \{ K_{\ell - 1}, \ldots, K_{\ell - 1}\} = \{0, \ldots, K - 1\} \) and \( \ell \in L = \{1\} \) and solve (18) without considering the state constraints (1e). The main reason for this is that the problem without (1e) is much easier to solve. Depending on the solution of a state-unconstrained optimization, extra intervals will be added as will be shown later in this section.

First, we define the state unconstrained problem for subsystem \( m \), which is given by (18) without (1e). The Lagrangian of this problem is given by
\[
\hat{L}(\{u_{m,k}, \lambda_{m,k}, \nu_{m,k}\}) = \sum_{\ell \in L} \sum_{k \in K_{\ell}} \frac{1}{2} H_{m,k} u_{m,k}^2 + F_{m,k} u_{m,k} + \nu_{m,\ell}^T (u_{m,k} - \bar{u}_m) + \lambda_{\ell} (A_{m,K_{\ell - 1} - 1} - \bar{x}_m, \nu_{m,k}, B_{m,u} u_{m,k}, \lambda_{m,k} - B_{m,u} u_{m,k}), \quad (23)
\]
with \( \lambda_{\ell} \in \mathbb{R} \), the Lagrange multiplier associated with the constraint (17), \( \nu_{m,k} \in \mathbb{R} \) and \( \lambda_{m,k} \in \mathbb{R} \), the Lagrange multipliers associated with the upper and lower input constraints (1d), respectively. The Karush-Kuhn-Tucker conditions [38] for minimizing the Lagrangian in (23) are given by the first-order necessary optimality condition
\[
\frac{\partial L(\{u_{m,k}, \lambda_{m,k}, \nu_{m,k}\})}{\partial u_{m,k}} = H_{m,k} + F_{m,k} + \nu_{m,k} - \lambda_{m,k} A_{m,K_{\ell - 1} - 1} B_{m,u} = 0, \quad (24)
\]
for all \( k \in K_{\ell}, \ell \in L \), feasibility of the constraint (17) and the complementary slackness conditions for the inequality constraints
\[
\nu_{m,k} (u_{m,k} - \bar{u}_m,0,0) = 0, \quad \lambda_{m,k} (u_{m,k} - \bar{u}_m,0,0) = 0, \quad (25)
\]
for all \( k \in K_{\ell}, \ell \in L \) with \( \nu_{m,k} \geq 0 \) and \( \lambda_{m,k} \geq 0 \). Finding a solution for (24) and (25) simultaneously is difficult and often the solution is found with a shooting method and a bisection algorithm over \( \lambda_{m,k} \), leading to the optimal solution
\[
u_{m,k} = -H_{m,k}^{-1} F_{m,k} + \lambda_{m,k} A_{m,K_{\ell - 1} - 1} B_{m,u} + \nu_{m,k} - \lambda_{m,k} A_{m,K_{\ell - 1} - 1} B_{m,u} + \nu_{m,k} - \lambda_{m,k} A_{m,K_{\ell - 1} - 1} B_{m,u} + \nu_{m,k} - \lambda_{m,k} A_{m,K_{\ell - 1} - 1} B_{m,u} \quad (26)
\]
for all \( k \in K_{\ell}, \ell \in L \) for a given \( \lambda_{m,k} \), \( \nu_{m,k} \) and \( \lambda_{m,k} \). Instead, we propose a procedure that aims, for each interval \( \ell \in L \), at solving
\[
\lambda_{m,k}^{s+1} = (1 - \gamma) \lambda_{m,k}^{s} + \gamma H_{m,k}^{-1} (A_{m,K_{\ell - 1} - 1} \bar{x}_m, K_{\ell - 1} - \bar{x}_m, K_{\ell} + \sum_{k \in K_{\ell}} A_{m,K_{\ell - 1} - 1} (B_{m,u} u_{m,k} - B_{m,u} u_{m,k} - H_{m,k}^{-1} F_{m,k} + \nu_{m,k} - \lambda_{m,k} A_{m,K_{\ell - 1} - 1} B_{m,u})), \quad (27a)
\]
with relaxation parameter $\gamma \in (0, 1]$ and with

$$
\dot{H} = \sum_{i \in \mathcal{K}_\ell} A_{m,i}^{K_i-1-i} B_{m,u} H_{m,i}^{K_i-1-i} B_{m,u},
$$

and

$$
\nu_{k+1}^i = \max \left\{ 0, -H_{m,k} \nu_{m,k} - F_{m,k} - \lambda_{k+1}^i \sum_{i \in \mathcal{K}_\ell} A_{m,i}^{K_i-1-i} B_{m,u}, \right\},
$$

$$
\nu_{k+1} = \max \left\{ 0, H_{m,k} \nu_{m,k} + F_{m,k} + \lambda_{k+1}^i \sum_{i \in \mathcal{K}_\ell} A_{m,i}^{K_i-1-i} B_{m,u}, \right\},
$$

with $t \in \mathbb{N}$ the iteration index and for $\lambda^0 = 0$ and $\nu_0^i = \nu_0^m = 0$ for all $k \in \mathcal{K}$, until $(17)$ is satisfied within some desired tolerance. The expressions in (27) are obtained by substituting (26) into (17) and (1d). If $H_{m,k}$ is strictly positive and if there exists an optimal solution $u^*_m$ for which (1d) and (17) are satisfied, then the solution of (27) will converge to the solution of (18) without (1e) for a well chosen relaxation parameter $\gamma \in (0, 1]$.

For solving the state-constrained optimization problem (15), we use an idea proposed in [16]. In [16] it is proven that when the relation between $\lambda_{K_i}$ and the final state $x_{m,K_i}$ is monotonic (note that it is linear in this paper), the time instant at which the state constrained is violated most $K_i$ is a contact point of the state-constrained solution, i.e., $x_{m,K_i} = \overline{x}_{m,k}$ or $x_{m,K_i} = \underline{x}_{m,k}$. This allows us to add $K_i$ to the set of splitting instances $\{ K_i \}_{i \in \mathcal{L}}$ and we fix $x_{m,k_i}$ at either $\overline{x}_{m,k_i}$ or $\underline{x}_{m,k_i}$ depending on whether the upper or lower bound was violated most and solve smaller optimal control problems subject to initial and terminal constraints. Note that due to the particular projection on the state constraints, this method is only suitable for scalar-state systems, i.e., $x_{m,k} \in \mathbb{R}$.

As a result of splitting the time horizon $\mathcal{K}$ into segments using the method outlined before, we might have a large number of ‘contact points’, which makes the method inefficient. We can prove that the optimal state trajectory equals either $x_{m,k} = \overline{x}_{m,k}$ or $x_{m,k} = \underline{x}_{m,k}$ for all $k \in \mathcal{K}_\ell$ under certain conditions, i.e., it is saturated on the upper or lower bound. This might occur when either $x_{m,K_{\ell-1}} = \overline{x}_{m,K_{\ell-1}}$ and $x_{m,K_\ell} = \overline{x}_{m,K_\ell}$ or $x_{m,K_{\ell-1}} = \underline{x}_{m,K_{\ell-1}}$ and $x_{m,K_\ell} = \underline{x}_{m,K_\ell}$. The following theorem provides conditions, that can be evaluated a priori, for which the optimal state trajectory is saturated on the upper or lower bound.

**Theorem 4.** The optimal state trajectory $x_{m,k}$ satisfies $x_{m,k} = \overline{x}_{m,k}$ for all $k \in \mathcal{K}_\ell$ and the corresponding control input satisfies

$$
\nu_{m,k}^* = \frac{1}{1 - A_m} \left( (1 - A_m) \overline{x}_{m,k} + B_{m,u} w_{m,k} \right)
$$

if $x_{m,K_{\ell-1}} = \overline{x}_{m,K_{\ell-1}}$ and $x_{m,K_\ell} = \overline{x}_{m,K_\ell}$, and if either one of the following conditions holds for all $k \in \mathcal{K}_\ell$:

- $A_m > 0$, $B_{m,u} < 0$, $H_{m,k+1} \overline{x}_{m,k+1} - H_{m,k} \overline{x}_{m,k} < 0$, $F_{m,k+1} - F_{m,k} < 0$
- $A_m > 0$, $B_{m,u} > 0$, $H_{m,k+1} \overline{x}_{m,k+1} - H_{m,k} \overline{x}_{m,k} > 0$, $F_{m,k+1} - F_{m,k} > 0$.

Similarly, the optimal state trajectory $x_{m,k}$ satisfies $x_{m,k} = \underline{x}_{m,k}$ for all $k \in \mathcal{K}_\ell$ and the corresponding control input satisfies

$$
\nu_{m,k}^* = \frac{1}{1 - A_m} \left( (1 - A_m) \underline{x}_{m,k} + B_{m,u} w_{m,k} \right)
$$

if $x_{m,K_{\ell-1}} = \underline{x}_{m,K_{\ell-1}}$ and $x_{m,K_\ell} = \underline{x}_{m,K_\ell}$, and if either one of the following conditions holds for all $k \in \mathcal{K}_\ell$:

- $A_m > 0$, $B_{m,u} < 0$, $H_{m,k+1} \underline{x}_{m,k+1} - H_{m,k} \underline{x}_{m,k} > 0$, $F_{m,k+1} - F_{m,k} > 0$
- $A_m > 0$, $B_{m,u} > 0$, $H_{m,k+1} \underline{x}_{m,k+1} - H_{m,k} \underline{x}_{m,k} < 0$, $F_{m,k+1} - F_{m,k} < 0$.

**Proof.** We only prove the case that $x_{m,k} = \overline{x}_{m,k}$ for all $k \in \mathcal{K}_\ell$ and the proof for $x_{m,k} = \underline{x}_{m,k}$ follows mutatis mutandis. For $w_{m,k}$ to be a feasible solution it needs to satisfy (1d) by definition, such that the Lagrangian of (18) on interval $\ell$ is given by

$$
\hat{L}(\{ u_{m,k}^*, \lambda_{\ell}, \nu_{m,k}^* \}) = \sum_{k \in \mathcal{K}_\ell} \frac{1}{2} H_{m,k} (u_{m,k})^2 + F_{m,k} u_{m,k} + \nu_{k}(x_{m,k} - x_{m,k}) + \lambda_{\ell} (A_{m,K_{\ell-1}} x_{m,K_{\ell-1}} - \overline{x}_{m,K_\ell} + \sum_{i \in \mathcal{K}_\ell} A_{m,i}^{K_i-1-i} (B_{m,w} u_{m,i} + B_{m,u} u_{m,i}) )
$$

(29)

where $x_{m,k}$ is given by (17a) for all $k \in \mathcal{K}_\ell$. In (29), $\nu_{k} \in \mathbb{R}$ is the Lagrange multiplier associated with the lower state constraint (1e). We will show that the lower state is active for all $k \in \mathcal{K}_\ell$ which means that the upper state is inactive and can be left out of the Lagrangian. The Karush-Kuhn-Tucker conditions [38] for minimizing the Lagrangian in (29) are given by the first-order necessary optimality condition

$$
\frac{\partial L}{\partial u_{m,k}} = H_{m,k} u_{m,k}^* + F_{m,k}
$$

$$
+ \lambda_{\ell} \sum_{i \in \mathcal{K}_\ell} A_{m,i}^{K_i-1-i} B_{m,u} - \nu_{k} \sum_{i = K_{\ell-1}}^{k} A_{m,i}^{K_i-1-i} B_{m,u} = 0,
$$

(30)

for all $k \in \mathcal{K}_\ell$, feasibility of the constraint (17) and the complementary slackness conditions for the inequality constraint

$$
\nu_{k}(x_{m,k} - x_{m,k}) = 0,
$$

(31)

for all $k \in \mathcal{K}_\ell$, for a given $\nu_{k} \geq 0$. As $\lambda_{\ell} \sum_{i \in \mathcal{K}_\ell} A_{m,i}^{K_i-1-i} B_{m,u}$ is a constant in (30), we can derive the following relation between the Lagrange multipliers $\nu_{k}$ at time instant $k+1$ and $k$:

$$
\nu_{k+1} = \left( \sum_{i = K_{\ell-1}}^{k+1} A_{m,i}^{K_i-1-i} B_{m,u} \right)^{-1} \nu_{k} \sum_{i = K_{\ell-1}}^{k+1} A_{m,i}^{K_i-1-i} B_{m,u}
$$

$$
+ H_{m,k+1} u_{m,k+1}^* - H_{m,k} u_{m,k}^* + F_{m,k+1} - F_{m,k},
$$

(32)

for $k \in \mathcal{K}_\ell$ and where $\nu_{K_{\ell-1}} > 0$ if $\overline{x}_{m,K_{\ell-1}} = \overline{x}_{m,K_{\ell-1}}$, such that if $A_m > 0$, $B_{m,u} < 0$, $H_{m,k+1} \overline{x}_{m,k+1} - H_{m,k} \overline{x}_{m,k} < 0$ and $F_{m,k+1} - F_{m,k} < 0$ or if $A_m > 0$, $B_{m,u} > 0$, $H_{m,k+1} \overline{x}_{m,k+1} - H_{m,k} \overline{x}_{m,k} > 0$ and $F_{m,k+1} - F_{m,k} > 0$ then there exist a $\nu_{k+1} > 0$ for all $k \in \mathcal{K}_\ell$, such that the first-order optimality conditions and the complementary slackness conditions are satisfied and $u_{m,k}^*$ is optimal for all $k \in \mathcal{K}_\ell$. This completes the proof for $x_{m,k} = \overline{x}_{m,k}$. $\square$
This theorem provides a priori verifiable conditions when the optimal state trajectory is saturated at the lower bound or upper bound for all $k \in \{K_t - 1, \ldots, K_t - 1\}$, respectively. The three preceding results, i.e., i) the solution of the optimal control problem without considering the state constraints, ii) the iterative method for splitting the control problem into smaller ones to incorporate state constraints and iii) conditions for which the optimal solution satisfies $x_{m,k} = \bar{x}_{m,k}$ or $x_{m,k} = \bar{x}_{m,k}$ for all $k \in \mathcal{K}$ allow us to propose the following algorithm for solving (18).

**Algorithm 1.** Take $\mathcal{K}_t = \{0, \ldots, K - 1\}$, $\mathcal{L} = \{1\}$ and let $\tilde{x}_m,0$ and $\tilde{x}_m,\mathcal{K}_t$ be given.

- For each interval $\ell \in \mathcal{L}$, check if the conditions of Theorem 4 are satisfied.
  - If the conditions of Theorem 4 are satisfied, the optimal solution satisfies (28a) or (28b).
  - If the conditions of Theorem 4 are not satisfied, compute the input constrained solution using (27). Then verify
    \begin{align}
    \tilde{\omega}_t &= \max_{k \in \mathcal{K}_t}\{\omega_{m,k} - \omega_{m,k}\}, \\
    \tilde{\omega}_t &= \max_{k \in \mathcal{K}_t}\{\omega_{m,k} - \omega_{m,k}\}.
    \end{align}

* If $\tilde{\omega}_t > 0$ and $\omega_{m,k} > \tilde{\omega}_t$, the lower state constraint is violated more than the upper state constraint and
  \begin{align}
  \tilde{K}_t &= \arg\max_{k \in \mathcal{K}_t}\{\omega_{m,k} - \omega_{m,k}\},
  \end{align}
  is added to the set of contact points $\{\mathcal{K}_t\} \in \mathcal{L}$ and re-ordered, i.e., $0 = K_0 \leq \ldots \leq K_{t-1} \leq \tilde{K}_t \leq K_t \leq K_L$ to define new subsets $\mathcal{K}_t = \{K_{t-1}, \ldots, K_t - 1\}$ and $\tilde{x}_m,\tilde{K}_t = \bar{x}_{m,\mathcal{K}_t}$.

* If $\tilde{\omega}_t > 0$ and $\omega_{m,k} > \tilde{\omega}_t$, the upper state constraint is violated more than the lower state constraint and
  \begin{align}
  \tilde{K}_t &= \arg\max_{k \in \mathcal{K}_t}\{\omega_{m,k} - \omega_{m,k}\},
  \end{align}

is added to the set of contact points $\{\mathcal{K}_t\} \in \mathcal{L}$ and re-ordered, i.e., $0 = K_0 \leq \ldots \leq K_{t-1} \leq \tilde{K}_t \leq K_t \leq K_L$ to define new subsets $\mathcal{K}_t = \{K_{t-1}, \ldots, K_t - 1\}$ and $\tilde{x}_m,\tilde{K}_t = \bar{x}_{m,\mathcal{K}_t}$.

- Repeat until $\max\{\omega_{m,k} - \omega_{m,k}, \omega_{m,k} - \omega_{m,k}\} \leq 0$ for all $k \in \mathcal{K}$.

Similarly as the dual decomposition allows the large-scale optimal control problem to be solved by solving smaller optimal control problems on subsystem level, Algorithm 1 and the ADMM algorithm allow the optimal control problem over a large horizon to be solved through multiple optimal control problems over a smaller horizon. Note that, to ensure convergence of the solution to the dual problem (5), the solution to each dual function obtained with Algorithm 1 or ADMM (22) needs to be converged before proceeding with the maximization in (11). Still, by combining these solution methods, scalability is significantly improved, which will be demonstrated on the complete vehicle energy management problem that we will introduce in the next section.

**IV. APPLICATION TO A VEHICLE POWER NET**

The distributed optimization approach presented in the previous sections will be used to find the optimal solution for energy management of a vehicle with an internal combustion engine (ICE), an electric machine (EM), a high-voltage battery (HVB), a refrigerated semi-trailer (RST), an air supply system (AS), an alternator (ALT), a DCDC converter, a low-voltage battery (LVB) and a climate control system (CCS). The topology is shown in Fig. 2. This topology has three exogenous load signals, i.e., $v_k = [v_{1,k}\ v_{2,k}\ v_{3,k}]^T \in \mathbb{R}^3$, which are the power required to drive a certain drive cycle, the power required for uncontrolled high-voltage auxiliaries and the power required for uncontrolled low-voltage auxiliaries, respectively. These three signals are assumed to be known for every time instant $k \in \mathcal{K}$. Furthermore, we assume that the gearshift strategy is fixed such that the rotational velocity of the drive line $\omega_k$ is known for every time instant $k \in \mathcal{K}$. We also assume that the power losses in the gearbox are negligible, i.e., $y_{gb,k} = u_{gb,k}$, such that nodes connecting the subsystems are given by
\begin{align}
&v_{1,k} - y_{br,k} - y_{ice,k} + u_{em,k} + u_{alt,k} - y_{ccs,k} = 0, \\
&v_{2,k} - y_{em,k} - y_{hvb},k - y_{rst,k} - y_{as,k} + u_{dc,k} = 0, \\
&v_{3,k} - y_{alt,k} - y_{hvb},k - y_{dc,k} = 0,
\end{align}
which we can write in the form of (1f) with
\begin{align}
&\Gamma_{ice} = [0 \ 0 \ 0]^T, \quad \Theta_{ice} = [-1 \ 0 \ 0]^T, \\
&\Gamma_{em} = [1 \ 0 \ 0]^T, \quad \Theta_{em} = [0 \ -1 \ 0]^T, \\
&\Gamma_{alt} = [1 \ 0 \ 0]^T, \quad \Theta_{alt} = [0 \ 0 \ -1]^T, \\
&\Gamma_{dc} = [0 \ 1 \ 0]^T, \quad \Theta_{dc} = [0 \ 0 \ -1]^T, \\
&\Gamma_{br} = [0 \ 0 \ 0]^T, \quad \Theta_{br} = [0 \ -1 \ 0]^T, \\
&\Gamma_{hvb} = [0 \ 0 \ 0]^T, \quad \Theta_{hvb} = [0 \ 0 \ -1]^T, \\
&\Gamma_{rv} = [0 \ 0 \ 0]^T, \quad \Theta_{rv} = [0 \ -1 \ 0]^T, \quad \Theta_{rv} = [0 \ 0 \ -1]^T, \\
&\Gamma_{ccs} = [0 \ 0 \ 0]^T, \quad \Theta_{ccs} = [-1 \ 0 \ 0]^T.
\end{align}

Note that we have replaced the set $\mathcal{M} = \{1, \ldots, M\}$ in the general optimization problem (1) by the set $\mathcal{M} = \{\text{ice, em, hvb, rst, as, ccs, dc, hvb, alt, br}\}$, to better indicate the physical origin of the power flows.

**A. Objective Function**

The objective in energy management is to minimize the fuel consumption, which is equivalent to minimizing the equivalent...
fuel energy, i.e.,
\[
\min_{\{u_{\text{ice},k}\}} \sum_{k \in K} \tau u_{\text{ice},k}
\]
(36)
where the equivalent fuel power is given by \( u_{\text{ice},k} = H_0 \dot{m}_{\text{f},k} \) with \( H_0 \) the lower heating value of the fuel and \( \dot{m}_{\text{f},k} \) the fuel consumption rate at time instant \( k \). This objective function is only defined in variables related to the internal combustion engine. However, we can obtain an objective function that is defined in variables related to every subsystem in the network, i.e., in the form of (1a), which is equivalent to (36) for a specific \( c_m \) and \( d_m \). In particular, this holds if we choose \( c_m \) and \( d_m \) for every subsystem such that \( \sum_{k \in K} c_m u_{m,k} - d_m y_{m,k} \) represents the energy losses in the subsystem. The energy losses in each converter are given by the energy flowing into the converter minus the energy flowing out of the converter where the energy flowing into the converter is indicated with the input and output power, i.e.,
\[
c_m = \tau, \quad d_m = \tau,
\]
(37a)
for \( m \in \{\text{ice, em, alt, dc}, \text{hvb}, \text{lvb}\} \). For the refrigerated semi-trailer, the air supply system, the climate control system and the mechanical brakes, all energy flowing into the subsystem is eventually lost and therefore the energy losses are given for
\[
c_m = 0, \quad d_m = \tau,
\]
(37b)
for \( m \in \{\text{rst, as, ccs, br}\} \). By substituting (37) into (1a), we obtain
\[
\min_{\{u_{m,k}, y_{m,k}\}} \sum_{m \in M} \sum_{k \in K} c_m u_{m,k} - d_m y_{m,k} = \min_{\{u_{m,k}, y_{m,k}\}} \tau (u_{\text{ice},k} - y_{\text{ice},k} + u_{\text{em},k} + u_{\text{alt},k} - y_{\text{ccs},k} - y_{\text{br},k} - y_{\text{em},k} - y_{\text{hvb},k} - y_{\text{r},k} - y_{\text{as},k} + u_{\text{dc},k} - y_{\text{alt},k} - y_{\text{hvb},k} - y_{\text{ice},k} + u_{\text{hvb},k} + u_{\text{vb},k}) \).
\]
(38a)
By using the power balance constraints (34), we can reduce this equation to
\[
\min_{\{u_{\text{ice},k}, u_{\text{hvb},k}, u_{\text{vb},k}\}} \sum_{k \in K} \tau u_{\text{ice},k} - \tau (v_{1,k} + v_{2,k} + v_{3,k}) + \tau (u_{\text{hvb},k} + u_{\text{vb},k}).
\]
(38b)
Moreover, as will be explained in more detail in the next subsection, the high-voltage battery and low-voltage battery satisfy integrator dynamics so that we can write \( \sum_{k \in K} \tau u_{m,k} = x_{m,0} - x_{m,k} \) for \( m \in \{\text{hvb}, \text{lvb}\} \) and further reduce (38b) to
\[
\min_{\{u_{\text{ice},k}\}} \sum_{k \in K} \tau u_{\text{ice},k} - \tau (v_{1,k} + v_{2,k} + v_{3,k}) + x_{\text{hvb},0} - x_{\text{hvb},K} + x_{\text{vb},0} - x_{\text{vb},K}.
\]
(38c)
As the load signals are known for all \( k \in K \) and the initial and final states \( x_{m,0} \) and \( x_{m,K} \) for \( m \in \{\text{hvb}, \text{lvb}\} \) are fixed, the optimal value for \( u_{\text{ice},k} \) for (38c) is equivalent to the optimal value \( u_{\text{ice},k} \) for (36) for all \( k \in K \). Hence, the optimal control problem (1) with objective function (1a) and \( c_m \) and \( d_m \) given by (37) provides the optimal solution for which the fuel consumption is minimized over all \( k \in K \) as is indicated in (36).

B. Subsystem Modeling

The models for the internal combustion engine, the electric machine, the alternator, the DCDC converter, the mechanical brakes the high-voltage battery, the low-voltage battery, the refrigerated semi-trailer, the air supply system and the climate control system will be presented below.

1) Internal combustion engine, electric machine and alternator: The input-output behavior of the internal combustion engine, the electric machine and alternator can be described by the quadratic function (1b) for \( m \in \{\text{ice, em, alt}\} \) where the efficiency coefficients depend on speed, i.e.,
\[
q_{m,k} = q_m(\omega_k), \quad f_{m,k} = f_m(\omega_k), \quad e_{m,k} = e_m(\omega_k),
\]
(39)
for \( m \in \{\text{ice, em, alt}\} \) where \( q_m(\omega_k) \), \( f_m(\omega_k) \) and \( e_m(\omega_k) \) are functions parameterizing the efficiency coefficients as function of drive line speed \( \omega_k \). The input power of the converters are bounded by (1d) for \( m \in \{\text{ice, em, alt}\} \) where the upper and lower bound depend on speed, i.e.,
\[
u_{m,k} - y_{m,k} = \tau m = \tau m(\omega_k).
\]
(40)
for \( m \in \{\text{ice, em, alt}\} \) where \( \tau m(\omega_k) \) and \( \tau \) are functions parameterizing the lower and upper bound as function of drive line speed \( \omega_k \). The input-output behavior is shown in Fig. 3 where the power losses, i.e., \( u_{m,k} - y_{m,k} \) are given for the internal combustion engine, electric machine and alternator for two different speeds. In this figure, we show measured efficiencies of typical components in a truck and the accuracy of the quadratic approximations. We normalized the power losses to avoid sharing confidential information. Still, the figure shows that the quadratic behavior, the lower and upper bound strongly depend on the driveline speed \( \omega_k \). Furthermore, the quadratic assumption on the behavior of converters holds well as the models are close to the measurement data. To be precise, the average root mean square error over all drive line speeds is 2.32 kW, 0.18 kW and 0.05 kW, and the relative error is less than 1.4%, 0.6% and 0.3% for the internal combustion engine, the electric machine and the alternator, respectively. The internal combustion engine, the electric machine and the alternator are subsystems without a (constrained) energy storage. Therefore, constraints on the system dynamics (1c) and state constraints (1e) do not have to be taken into account in these subsystems.

2) DCDC converter and mechanical brakes: The input-output behavior of the DCDC converter and mechanical brakes can be described by the quadratic function (1b) for \( m \in \{\text{dc, br}\} \) with efficiency coefficients \( q_{m,k} \in \mathbb{R}_+ \), \( f_{m,k} \in \mathbb{R} \) and \( e_{m,k} \in \mathbb{R} \) for \( m \in \{\text{dc, br}\} \) which do not depend on speed. The input power is bounded by (1d) for \( m \in \{\text{dc, br}\} \). The DCDC converter and the mechanical brakes are also subsystems without a (constrained) energy storage. Therefore, constraints on the system dynamics (1c) and state constraints (1e) do not have to be taken into account in these subsystems.
3) High- and low-voltage battery: The models of the high- and low-voltage battery are derived from a battery equivalent circuit model, i.e., an open circuit voltage $U_{m,oc}$ for $m \in \{\text{lvb}, \text{lvb}\}$ in series with a resistance $R_{m,i}$ for $m \in \{\text{lvb}, \text{lvb}\}$ (see, e.g., [2]), which lead to an input-output behavior of the converter that can be described by the quadratic function (1b) for $m \in \{\text{lvb}, \text{lvb}\}$ with $q_{m,k} = \frac{R_{m,i}}{U_{m,oc}}$, $f_{m,k} = -1$ and $e_{m,k} = 0$ for $m \in \{\text{lvb}, \text{lvb}\}$. The input power of the high-voltage and low-voltage battery are bounded by (1d) for $m \in \{\text{lvb}, \text{lvb}\}$. The dynamics are given by (1c) for $m \in \{\text{lvb}, \text{lvb}\}$ with $A_m = 1$, $B_{m,w} = 0$ and $B_{m,u} = -\tau$ with $\tau$ being the sample time. Here, the state $x_{m,k}$ represents the energy in the battery at time instant $k$.

4) Refrigerated semi-trailer: The input-output behavior of the converter in the refrigerated semi-trailer can be described by the quadratic function (1b) for $m \in \{ \text{rst} \}$ and the input power of the converter is bounded by (1d) for $m \in \{ \text{rst} \}$. The dynamics of the refrigerated semi-trailer are assumed to satisfy a thermal energy balance (see, e.g., [29]) given by

$$C_{\text{rst}} \frac{d}{dt} T_{\text{rst}} = (u_{\text{rst}} - h(T_{\text{rst}} - \alpha T_{\text{amb}})),$$

where $C_{\text{rst}}$ is the heat capacity of the refrigerated semi-trailer and its contents, $u_{\text{rst}}$ is the thermal power where negative values indicate cooling, $h$ is the heat transfer coefficient between the refrigerated semi-trailer and the environment, $\alpha$ is an insulation coefficient, $T_{\text{amb}}$ is the temperature inside the refrigerated semi-trailer and $T_{\text{amb}}$ is the ambient temperature (which is assumed to be constant). Similar to the battery, we can represent the refrigerated semi-trailer model in terms of stored energy by defining the thermal energy relative to the ambient temperature $x_{\text{rst}} = C_{\text{rst}}(\alpha T_{\text{amb}} - T_{\text{rst}})$. By doing so and by making a forward Euler approximation of (41), the dynamics can be represented by (1c) for $m \in \{ \text{rst} \}$ with

$$A_{\text{rst}} = 1 - \frac{h}{\gamma C_{\text{rst}}}, B_{\text{rst},w} = 0 \text{ and } B_{\text{rst},u} = -\tau \text{ for all } k \in K.$$
with $A_{ccs} = \begin{bmatrix} 1 - \frac{\tau h_0}{C_e} & -\frac{\tau h_0}{C_e} \\ \frac{\tau h_0}{C_e} & 1 - \frac{\tau (h_1 + h_0)}{C_e} \end{bmatrix}$, $B_{ccs,u} = \begin{bmatrix} 0 \\ -\tau \end{bmatrix}$, and $w_{rst,k}$ is the latent heat $Q_l$ at time instant $k$.

The CVEM problem for a vehicle with an internal combustion engine, an electric machine, a high-voltage battery, a refrigerated semi-trailer, an air supply system, an alternator, a DCDC converter, a low-voltage battery and a climate control system is now fully described by the optimal control problem defined in Section II. In particular, the topology of the vehicle is described through (1f) for a given $\Gamma_m$ and $\Theta_m$, the objective function, i.e., minimizing fuel consumption, is described by (1a) by choosing $c_m$ and $d_m$ appropriately and the behavior of each subsystem is fully described by (1b) - (1e) by choosing the efficiency coefficients $g_m,k$, $f_m,k$ and $e_m,k$, the state-space matrices $A_m$, $B_{m,u}$, $B_{m,w}$ describing the dynamics of the subsystems, the upper and lower bound on the inputs $u_{m,k}$ and $\bar{u}_{m,k}$, respectively, and the upper and lower bounds on the states $\bar{x}_{m,k}$ and $\underline{x}_{m,k}$, respectively. This allows the CVEM problem to be solved with the solution methods proposed in Section II and Section III as will be demonstrated in the next section.

V. SIMULATION RESULTS

In this section, we will demonstrate the distributed optimization approach to complete vehicle energy management (CVEM) by using a simulation study. First, we will give the exogenous signals that we used for the simulation study followed by the results that will be discussed in three subsections. In the first subsection, we will analyze the computational performance and comparing it with the state-of-the-art solver CPLEX [37]. In the second subsection, we will discuss the optimal power flows and state trajectories. In the third subsection, the fuel consumption reduction for CVEM will be discussed and in the last subsection, the implications for real-time implementation will be given.

A. Exogenous Inputs

The driving cycle is commonly described by a velocity profile over time. If the gear shift strategy is assumed to be known, the velocity profile can be converted to a power required at the wheels and the engine speed. This set of data is derived for a PAN European driving cycle and shown in Fig. 4. It can be seen that the brake power can reach -1000 kW. However, only a small part of the total braking power can be recovered by the subsystems. Therefore, the power request used as load signal $v_{1,k}$ is limited to the maximum braking power that can be recovered with all subsystems combined. Furthermore, the uncontrolled high-voltage auxiliaries are assumed to be absent such that $v_{2,k} = 0$ kW for all $k \in K$ and the power required from uncontrolled low-voltage auxiliaries is assumed to be constant, i.e., $v_{3,k} = 1.5$ kW for all $k \in K$. A quasi-static approach is generally sufficient for energy management (see, e.g., [3]) and the sample time is chosen according to the sample time of the drive cycle under consideration, which is $\tau = 1$ second, which is smaller than the time constants of the dynamics in the subsystems.

B. Computational Performance

The local optimization problem related to each component defined in (15) can be solved via different solution methods introduced in Section III. In particular, the performance of the ADMM solution method depends on the penalty parameter $R$ in (22) and the interval length $K_{\ell} - K_{\ell-1}$ for $\ell \in L$. For simplicity we assume that the interval length is equal for all intervals $\ell \in L$. The maximum, average and minimum time to compute the solution of the local optimization problem (15) are given in Table I for the high-voltage battery (HVB), the refrigerated semi-trailer (RST), the auxiliary supply system (AS) and the climate control system (CCS) for different values of $K_{\ell} - K_{\ell-1}$ and two horizon lengths $K$. The value of the penalty parameter $R$ in (22) is manually tuned for each $K_{\ell} - K_{\ell-1}$ but is kept constant for different lengths $K$ of the drive cycle.

The computation time required to solve the optimization problem strongly depends on the amount of iterations $t$, as in (22), required for the ADMM method to converge, which depends on the initial guess of the dual variables. We use the dual variables from the previous iteration in the dual decomposition, see (11), as an initial guess and therefore the amount of ADMM iterations reduces as the dual decomposition converges. As a consequence, the maximum time is significantly larger than the minimum time. The main conclusion drawn from this table is that the optimal $K_{\ell} - K_{\ell-1}$ differs per component and it should be chosen neither too small nor too large. Moreover, it does not seem to depend on $K$. For example, the high-voltage battery has the highest performance for $K_{\ell} - K_{\ell-1} = 200$ which is the optimal trade-off between the size and number of QPs.

With the results of Table I, we choose $K_{\ell} - K_{\ell-1} = 200$ for the high-voltage battery, $K_{\ell} - K_{\ell-1} = 50$ for the refrigerated semi-trailer, $K_{\ell} - K_{\ell-1} = 200$ for the air supply system and $K_{\ell} - K_{\ell-1} = 50$ for the climate control system. These results are compared with other solution methods in Table II. This table shows the average computation time to solve the local optimization problem (15) over all iterations $s$ in the dual decomposition. Here, $QP$ indicates the computation time for the dual decomposition, see (11), as an initial guess and therefore the amount of ADMM iterations reduces as the dual decomposition converges. As a consequence, the maximum time is significantly larger than the minimum time. The main conclusion drawn from this table is that the optimal $K_{\ell} - K_{\ell-1}$ differs per component and it should be chosen neither too small nor too large. Moreover, it does not seem to depend on $K$. For example, the high-voltage battery has the highest performance for $K_{\ell} - K_{\ell-1} = 200$ which is the optimal trade-off between the size and number of QPs.

With the results of Table I, we choose $K_{\ell} - K_{\ell-1} = 200$ for the high-voltage battery, $K_{\ell} - K_{\ell-1} = 50$ for the refrigerated semi-trailer, $K_{\ell} - K_{\ell-1} = 200$ for the air supply system and $K_{\ell} - K_{\ell-1} = 50$ for the climate control system. These results are compared with other solution methods in Table II. This table shows the average computation time to solve the local optimization problem (15) over all iterations $s$ in the dual decomposition. Here, $QP$ indicates the computation time for the dual decomposition, see (11), as an initial guess and therefore the amount of ADMM iterations reduces as the dual decomposition converges. As a consequence, the maximum time is significantly larger than the minimum time. The main conclusion drawn from this table is that the optimal $K_{\ell} - K_{\ell-1}$ differs per component and it should be chosen neither too small nor too large. Moreover, it does not seem to depend on $K$. For example, the high-voltage battery has the highest performance for $K_{\ell} - K_{\ell-1} = 200$ which is the optimal trade-off between the size and number of QPs.
TABLE I: Computation time in seconds for ADMM method with different interval lengths.

<table>
<thead>
<tr>
<th>Subsystem</th>
<th>( K = 1000 )</th>
<th>( K = 5000 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>min</td>
<td>av</td>
</tr>
<tr>
<td>HVB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0.09</td>
<td>1.40</td>
</tr>
<tr>
<td>50</td>
<td>0.08</td>
<td>0.69</td>
</tr>
<tr>
<td>100</td>
<td>0.14</td>
<td>0.69</td>
</tr>
<tr>
<td>200</td>
<td>0.16</td>
<td>0.41</td>
</tr>
<tr>
<td>500</td>
<td>0.71</td>
<td>1.73</td>
</tr>
<tr>
<td>RST</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0.09</td>
<td>1.16</td>
</tr>
<tr>
<td>50</td>
<td>0.09</td>
<td>0.72</td>
</tr>
<tr>
<td>100</td>
<td>0.17</td>
<td>0.71</td>
</tr>
<tr>
<td>AS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0.09</td>
<td>1.04</td>
</tr>
<tr>
<td>50</td>
<td>0.09</td>
<td>1.07</td>
</tr>
<tr>
<td>100</td>
<td>0.15</td>
<td>0.6</td>
</tr>
<tr>
<td>200</td>
<td>0.18</td>
<td>0.44</td>
</tr>
<tr>
<td>500</td>
<td>0.99</td>
<td>1.78</td>
</tr>
<tr>
<td>CCS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0.1</td>
<td>0.32</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>0.28</td>
</tr>
<tr>
<td>100</td>
<td>0.12</td>
<td>0.33</td>
</tr>
<tr>
<td>200</td>
<td>0.14</td>
<td>0.39</td>
</tr>
</tbody>
</table>

The LM method (with relaxation parameter \( \gamma = 1 \)) reduces the computation time even further and depends on the number of intervals \( L \) required for the LM method. For \( K = 3000 \), the amount of intervals are 2, 6 and 53 for the high-voltage battery, the air supply system and the refrigerated semi-trailer, respectively. Note that this method cannot be used for the climate control system as this method is only suitable for scalar-state systems, i.e., \( x_{m,k} \in \mathbb{R} \). Due to the absence of state constraints, the optimization problem for the internal combustion engine, electric machine, alternator, DCDC converter and mechanical brakes can be solved explicitly and are not shown in the table. Since the dynamics of the low-voltage battery are similar to the dynamics of the high-voltage battery, we conjecture that LM is also best for the low-voltage battery.

To assess the computational performance of solving the energy management problem for different vehicle configurations, we define six case studies with increasing complexity. These case studies are introduced in Table III. To demonstrate that the conditions in Theorem 1 hold, we show in Table IV the minimum value of the dual variables \( \min_{k \in K} \min_x \{ \mu_{i,k}^* \} \), where \( \mu_{i,k}^* \) denotes the \( i \)-th element of the vector \( \mu_k^* \) at time \( k \in K \) and iteration \( s \), i.e., \( \mu_k^* = [\mu_{1,k}^*, \mu_{2,k}^*, \ldots, \mu_{n,k}^*]^T \in \mathbb{R}^n \). Note that with \( \Theta_m = -1 \) for all \( m \in M \) and \( d_m = \tau \) for all \( m \in M \), the condition in Theorem 1 is satisfied if and only if \( \mu_{i,k}^* > -\tau \) for all \( k \in K \), \( i \in \{1, 2, 3\} \) and \( s \), which is always satisfied as shown in Table IV for \( \tau = 1 \). Theorem 1 is satisfied for all simulations and is not further demonstrated in this section. Moreover, this table shows the reduced iterations of the dual Newton update strategy compared with an update strategy with fixed step sizes, i.e., with \( \alpha_k^s \) being a constant. The Newton strategy always converged which implies that the derivatives in (14) are sufficiently well approximated.

The computation times are given in Table V for each configuration. For these simulations, the optimal control problems related to each subsystem are solved in series which is more straightforward to implement and moreover, this results already in sufficient computational benefits for offline energy management. The computation time of the Distributed Optimization (DO) method introduced in this paper are compared with the computation time of the QCQP solver CPLEX [37]. The CPLEX solver cannot handle quadratic constraints written in vector format and every quadratic constraint needs to be programmed separately. This requires a large amount of assembly time which is not used for solving the actual optimization problem. Therefore, the computation times of CPLEX with and without assembling the optimization problem are given. If we compare only the time required to solve the optimization problem, DO is still 1825 times faster for Case 1 with \( K = 5000 \) and 64 times faster for Case 6 with \( K = 3000 \). Scalability of DO in the horizon length \( K \) is superior compared with CPLEX, even for short horizons, see the results for \( K = 1000 \). Scalability in the number of components is not always better with DO but only in the rare case with small \( K \) and many more components, CPLEX could be better than DO. The flexibility of adding and removing components with CPLEX remains poor though.

### C. Optimal Input and State Trajectories

The optimal power flows as function of time are shown in Fig. 5 for the complete vehicle and with a drive cycle length of only \( K = 3000 \) for clarity. Both, the results from DO, as well as the results from solving the optimization problem with CPLEX are shown. This figure demonstrates that both methods converge to the same solution (within a desired tolerance). Moreover, the fuel consumption of DO is 0.019 % smaller compared with CPLEX, which is negligible. Two important observations can be made from this figure,

TABLE III: Case studies with problem size defined in number of inputs, states and quadratic constraints.

<table>
<thead>
<tr>
<th>Case</th>
<th>Inputs</th>
<th>States</th>
<th>Quadratic constr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>Truck with ICE, EM and a HVB</td>
<td>4K</td>
<td>K</td>
</tr>
<tr>
<td>Case 2</td>
<td>Case 1 with a RST</td>
<td>5K</td>
<td>2K</td>
</tr>
<tr>
<td>Case 3</td>
<td>Case 2 with an AS</td>
<td>6K</td>
<td>3K</td>
</tr>
<tr>
<td>Case 4</td>
<td>Case 3 with a CCS</td>
<td>7K</td>
<td>5K</td>
</tr>
<tr>
<td>Case 5</td>
<td>Case 4 with a ALT and a LVB</td>
<td>9K</td>
<td>6K</td>
</tr>
<tr>
<td>Case 6</td>
<td>Case 5 with a DCDC converter</td>
<td>10K</td>
<td>6K</td>
</tr>
</tbody>
</table>
TABLE IV: Number of iterations and minimum of the dual variables over all iterations for $K = 5000$

<table>
<thead>
<tr>
<th>iter. ($\alpha_i^k$ constant)</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
<th>Case 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>iter. (Newton)</td>
<td>105</td>
<td>122</td>
<td>119</td>
<td>228</td>
<td>221</td>
<td>384</td>
</tr>
<tr>
<td>$\min_{k \in K} \min_i {\mu_i^k}$</td>
<td>0.416</td>
<td>0.430</td>
<td>0.431</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.03</td>
</tr>
<tr>
<td>$\min_{k \in K} \min_i {\mu_i^k}$</td>
<td>0.621</td>
<td>0.622</td>
<td>0.626</td>
<td>0.616</td>
<td>0.623</td>
<td>0.612</td>
</tr>
<tr>
<td>$\min_{k \in K} \min_i {\mu_i^k}$</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>N.A.</td>
<td>0.522</td>
<td>0.628</td>
</tr>
</tbody>
</table>

TABLE V: Computation times of DO and CPLEX in seconds.

<table>
<thead>
<tr>
<th>Case</th>
<th>Solution method</th>
<th>1000</th>
<th>2000</th>
<th>3000</th>
<th>4000</th>
<th>5000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>CPLEX$^1$</td>
<td>8.1</td>
<td>80</td>
<td>300</td>
<td>727</td>
<td>1259</td>
</tr>
<tr>
<td></td>
<td>CPLEX$^2$</td>
<td>46</td>
<td>389</td>
<td>1405</td>
<td>3403</td>
<td>6509</td>
</tr>
<tr>
<td></td>
<td>DO</td>
<td>0.035</td>
<td>0.24</td>
<td>0.21</td>
<td>0.74</td>
<td>0.69</td>
</tr>
<tr>
<td>Case 2</td>
<td>CPLEX$^1$</td>
<td>20.3</td>
<td>369</td>
<td>1186</td>
<td>2355</td>
<td>6464</td>
</tr>
<tr>
<td></td>
<td>CPLEX$^2$</td>
<td>90</td>
<td>924</td>
<td>3085</td>
<td>6946</td>
<td>15422</td>
</tr>
<tr>
<td></td>
<td>DO</td>
<td>0.14</td>
<td>0.99</td>
<td>1.19</td>
<td>1.96</td>
<td>3.25</td>
</tr>
<tr>
<td>Case 3</td>
<td>CPLEX$^1$</td>
<td>45</td>
<td>334</td>
<td>2014</td>
<td>6650</td>
<td>&gt; $10^4$</td>
</tr>
<tr>
<td></td>
<td>CPLEX$^2$</td>
<td>146</td>
<td>1177</td>
<td>4863</td>
<td>13528</td>
<td>&gt; $10^4$</td>
</tr>
<tr>
<td></td>
<td>DO</td>
<td>0.34</td>
<td>1.72</td>
<td>1.69</td>
<td>2.74</td>
<td>7.0</td>
</tr>
<tr>
<td>Case 4</td>
<td>CPLEX$^1$</td>
<td>71</td>
<td>454</td>
<td>3112</td>
<td>&gt; $10^4$</td>
<td>&gt; $10^4$</td>
</tr>
<tr>
<td></td>
<td>CPLEX$^2$</td>
<td>357</td>
<td>2721</td>
<td>10770</td>
<td>&gt; $10^4$</td>
<td>&gt; $10^4$</td>
</tr>
<tr>
<td></td>
<td>DO</td>
<td>4.35</td>
<td>10.5</td>
<td>18.6</td>
<td>24.2</td>
<td>40.1</td>
</tr>
<tr>
<td>Case 5</td>
<td>CPLEX$^1$</td>
<td>76</td>
<td>958</td>
<td>2834</td>
<td>&gt; $10^3$</td>
<td>&gt; $10^3$</td>
</tr>
<tr>
<td></td>
<td>CPLEX$^2$</td>
<td>791</td>
<td>6835</td>
<td>22213</td>
<td>&gt; $10^4$</td>
<td>&gt; $10^4$</td>
</tr>
<tr>
<td></td>
<td>DO</td>
<td>9.6</td>
<td>23.1</td>
<td>43.8</td>
<td>60.9</td>
<td>120.3</td>
</tr>
<tr>
<td>Case 6</td>
<td>CPLEX$^1$</td>
<td>79</td>
<td>965</td>
<td>3285</td>
<td>&gt; $10^4$</td>
<td>&gt; $10^4$</td>
</tr>
<tr>
<td></td>
<td>CPLEX$^2$</td>
<td>1003</td>
<td>8304</td>
<td>28202</td>
<td>&gt; $10^4$</td>
<td>&gt; $10^4$</td>
</tr>
<tr>
<td></td>
<td>DO</td>
<td>11.3</td>
<td>31.5</td>
<td>51.4</td>
<td>73.2</td>
<td>150.7</td>
</tr>
</tbody>
</table>

1) all auxiliaries are used to store (brake) energy and 2) the DCDC converter is generally used to supply the low-voltage auxiliaries, except when free brake energy is available, then the alternator supplies the low-voltage auxiliaries and charges the low-voltage battery. Observation 1 can also be seen from the state trajectories given in Fig. 6 where $\tilde{x}_{hv0}=\tilde{x}_{hv0}$ is the high-voltage battery energy normalized with respect to the maximum battery capacity $E_{hv0}$. $\tilde{x}_{lv0}=\tilde{x}_{lv0}$ is the low-voltage battery energy normalized with respect to the maximum battery capacity $E_{lv0}$. $T_{edc}$ is the air temperature in the refrigerated trailer, $P_{as}$ is the air pressure in the air supply system and $T_{ccs} = [T_w T_{edc}]^T$ is the wall and refrigerant temperature in the climate control system. This figure shows that all state constraints are met, where for the climate control system, only the constraint on the wall temperature is shown.

D. Fuel Consumption Reduction

To analyze the fuel consumption for different parts of the complete drive cycle, the drive cycle is split into three parts. The first part is given by $k \in \{0, \ldots, 19999\}$, the second part by $k \in \{20000, \ldots, 39999\}$ and the third part by $k \in \{40000, \ldots, 55579\}$. The fuel consumption reduction for each of the cases and for each of these drive cycles are given in Table VI. For the first case, the baseline is a non-hybrid truck with the air temperature in the refrigerated semi-trailer kept at its upper bound, the air pressure in the air supply system kept at its lower bound and the temperature in the climate control system kept at its upper bound. For the next cases, the baseline is the previous case to emphasize the potential of each auxiliary. The DCDC converter (Case 6) is the most potential auxiliary for reducing fuel with 0.34%.

Although, the current trend in automotive applications is to electrify the auxiliaries in the vehicle so to have a continuous input set (see, e.g., [44]), still the refrigerated semi-trailer, the air supply system and the climate control system are often attached to the engine via a clutch and, as such, the input set is not continuous. These auxiliaries are switched between an on and off state. Therefore, comparing with a baseline controller

TABLE VI: Fuel consumption reduction results.

<table>
<thead>
<tr>
<th>Case</th>
<th>Part 1</th>
<th>Part 2</th>
<th>Part 3</th>
<th>Complete</th>
<th>Part 1 (switched)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00%</td>
<td>10.93%</td>
<td>2.33%</td>
<td>0.64%</td>
<td>0.00%</td>
</tr>
<tr>
<td>2</td>
<td>0.08%</td>
<td>0.12%</td>
<td>0.03%</td>
<td>0.07%</td>
<td>0.42%</td>
</tr>
<tr>
<td>3</td>
<td>0.02%</td>
<td>0.03%</td>
<td>0.01%</td>
<td>0.02%</td>
<td>0.04%</td>
</tr>
<tr>
<td>4</td>
<td>0.01%</td>
<td>0.04%</td>
<td>0.02%</td>
<td>0.03%</td>
<td>0.21%</td>
</tr>
<tr>
<td>5</td>
<td>0.07%</td>
<td>0.11%</td>
<td>0.02%</td>
<td>0.06%</td>
<td>0.07%</td>
</tr>
<tr>
<td>6</td>
<td>0.34%</td>
<td>0.32%</td>
<td>0.25%</td>
<td>0.34%</td>
<td>0.34%</td>
</tr>
</tbody>
</table>
with continuous input set gives not the full potential of CVEM. The last column in Table VI is therefore added which uses a baseline controller, where the auxiliary is turned on when the state hits the lower bound and turned off when the state hits the upper bound. The fuel reduction for the refrigerated semi-trailer and climate control system with the switched baseline is much higher, i.e., 0.42% and 0.21%, respectively. However, this requires a switched controller with many more switches compared with the baseline. More switches will reduce the life time of the auxiliaries and an optimal trade-off must be found between the number of switches and the fuel reduction.

The fuel reduction for some auxiliaries, e.g., the air supply system and the climate control system, is very low and integration of those auxiliaries into the energy management strategy might not outweigh the additional cost and complexity. However, the main result of the distributed optimization approach presented in this paper is that the energy management problem is decomposed into smaller energy management problems related to each subsystem. Each of the energy management problems on subsystem level is much easier to solve and can be solved with different algorithms, e.g., an ADMM method or a Lagrangian Method. Extensions to more sophisticated models, e.g., by including battery aging or battery thermal dynamics, will be easier and part of future work. Moreover, developing optimal control algorithms for different subsystems can be done in parallel, e.g., thermal management of the internal combustion engine can be included in the internal combustion engine optimization problem, while at the same time thermal management of the high-voltage battery can be included in the high-voltage battery optimization problem.

### E. Implications for Real-time Implementation

The offline solution strategy proposed in this paper provides the optimal solution to the convex approximation of the CVEM problem. The optimal solution, however, is only useful if the convex approximation in Section IV is sufficiently accurate. To validate this convex approximation, we will compare the convex vehicle model with a high-fidelity model of the vehicle developed by the Institute für Kraftfahrzeugen Aachen [45] that is used in [11] to validate an online solution strategy for CVEM. This comparison is illustrated in Fig. 7. The convex vehicle model is validated in an open loop by receiving setpoints that are used in the high-fidelity vehicle model, i.e., the torque request from the gearbox $\tau_{gb}$, the electric machine $\tau_{em}$ and the alternator $\tau_{alt}$ and the engine speed $\omega_{icv}$. The high-fidelity model only allows for on/off control of the refrigerated semi-trailer, the air supply system and the climate control system. These on/off signals, i.e., $\tilde{u}_{rst}$, $\tilde{u}_{ccs}$ are also received by the convex vehicle model. The outputs, i.e., the fuel rate $\dot{m}_f$, the temperature of the air inside the refrigerated semi-trailer $T_{rst}$, the wall temperature of the climate control system $T_{ccs}$, the air pressure $p_{as}$ and the energy stored in the low-voltage battery and low-voltage battery are then compared with the outputs of the high-fidelity simulation model.

The comparison between the state trajectories from the high-fidelity vehicle model and the convex vehicle model is shown in Fig. 8. The state trajectories will diverge because the convex vehicle model is evaluated in an open loop and due to modeling errors and disturbances. Still, the state trajectories show similar behavior. The temperature of the climate control systems seems shows a larger difference as it is more sensitive to disturbances. Still, the fuel consumption is well estimated by the convex vehicle model and is only 2.5% lower than the fuel consumption in the high-fidelity vehicle model after 5000 seconds. This demonstrates that the proposed offline solution strategy provides a valuable benchmark that can be used to validate the performance for online solution strategies. Indeed, the proposed approach is used as a benchmark in [11], where the presented approach is modified to be used for real-time control, i.e., the optimal control problem is defined over a shorter horizon. After dual decomposition, this results in...
Fig. 8: Comparison between the high-fidelity vehicle model and the convex vehicle model.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions that helped us improving the quality of the paper.

REFERENCES


[32] ———, “Complete vehicle energy management with large horizon optimization,” in Proc. 54rd Annu. Conf. on Decision and Control, 2015.


