On feasibility, stability and performance in distributed model predictive control

Pontus Giselsson and **Anders Rantzer**
Department of Automatic Control LTH
Lund University
Box 118, SE-221 00 Lund, Sweden
{pontusg,rantzer}@control.lth.se

Abstract

We present a stopping condition to the duality based distributed optimization algorithm presented in [1] when used in a distributed model predictive control (DMPC) context. To enable distributed implementation, the optimization problem has neither terminal constraints nor terminal cost that has become standard in model predictive control (MPC). The developed stopping condition guarantees a prespecified performance, stability, and feasibility with finite number of algorithm iterations. Feasibility is guaranteed using a novel adaptive constraint tightening approach that gives the same feasible set as when no constraint tightening is used. Stability and performance of the proposed DMPC controller without terminal cost or terminal constraints is shown based on a controllability parameter for the stage costs. To enable quantification of the control horizon necessary to ensure stability and the prespecified performance, we show how the controllability parameter can be computed by solving a mixed integer linear program (MILP).

Index Terms

Distributed model predictive control, performance guarantee, stability, feasibility

I. Introduction

Model predictive control (MPC) is an optimization based control technology for input and state constrained systems. The idea behind MPC is to, in every time step, minimize some cost function based on predictions of future states while respecting state and control constraints. A control trajectory is obtained from the optimization and the first control action from this trajectory is applied to the plant. In the following samples the procedure is repeated with the latest state measurement as initial condition to the state predictions. For a thorough description of MPC, see [2], [3]. There exist a variety of methods to prove stability for system controlled by MPC, see [4] for a survey of such methods and [3] for further material. As pointed out in [4], common 'ingredients' in these stability proofs are the use of a terminal cost and/or a terminal constraint set in the optimization problem together with a terminal controller that controls the system to the origin once the terminal constraint set is reached. These 'ingredients' are then used to, in various ways, prove that the optimal value function to the optimization problem is a Lyapunov function for the system.

The methods to prove stability in standard MPC [4] are not directly applicable in DMPC formulations where a centralized optimization problem is solved in distributed fashion. Such distributed optimization algorithms often require the cost function to be separable and the constraints to be sparse. This is not the case for the terminal cost or the terminal constraints in standard MPC [4]. Further, the terminal controller that is commonly used to show stability in standard MPC [4] needs to be decentralized or distributed in the context of DMPC. Such stabilizing controllers do not exist for all constrained linear systems [5]. One approach to overcome the aforementioned problems to prove stability in DMPC is to solve local optimization problems sequentially that take neighboring interaction and solutions into account. This is done in [6] for linear systems and in [7] for nonlinear systems. In [8] a DMPC scheme is presented in which stability is proven by adding a constraint to the optimization problem that requires a reduction of an explicit control Lyapunov function. In [9], [10] stability is guaranteed for systems satisfying a certain matching condition and if the coupling interaction is small enough. None of the above methods solves a centralized MPC problem and worse global performance is expected than if using an appropriate centralized MPC controller.

To achieve the same performance in DMPC as in centralized MPC, a centralized problem formulation needs to be considered and solved in distributed fashion. In [11] a centralized MPC problem is solved in distributed fashion and stability is guaranteed in every algorithm iteration. A drawback to this method is that full model knowledge is assumed in each node. Some methods in the DMPC literature rely on duality theory to solve a centralized MPC problem in distributed fashion. In [12], [13], [14], [15] a (sub)gradient algorithm is used to solve the dual problem while the algorithm in [16], [17] is based on the smoothing technique presented in [18]. The only stability proof is given in [14], [15] where the terminal constraint is set to the origin which is very restrictive and requires long control horizons. Other distributed MPC formulations have been presented in [19] where the DMPC controller is based on a cooperative game and [20], [21] for dynamically decoupled systems. See also [22] for a recent survey of distributed and hierarchical MPC methods.

In this paper, a centralized optimization problem is solved in distributed fashion using the distributed accelerated gradient method presented in [1]. We present a stopping condition for this optimization algorithm that guarantees feasibility, stability, and prespecified performance of the closed loop system. However, the stopping conditions are not restricted to the optimization algorithm in [1] but any (distributed) optimization algorithm that produce dual feasible points can be used. The stated optimization problem has neither terminal cost nor terminal constraint set. Stability for MPC without terminal constraints and terminal cost has previously been treated in, e.g., [24]. Further results were reported in [25] where it was shown how to compute the minimal control horizon necessary to achieve stability and a prespecified performance. The results in [25] rely on relaxed dynamic programming that was originally presented in [26] and extended to MPC in [27], and on a controllability assumption on the stage costs. The parameters in the controllability assumption in [25] may be very difficult to compute for a given system. In this paper we take a similar approach as in [25], but we specify a different controllability parameter than in [25]. We show, through an explicit expression, how this parameter relates to the performance of the closed loop system. We also show how, for systems with linear dynamics and linear constraints, the controllability parameter can be computed by solving a mixed integer linear program (MILP). This makes the stabilizing control horizon practically computable. We will see that one benefit of not using terminal constraints is that the region of attraction can be increased significantly compared to standard MPC.

Previous work on MPC where a suboptimal solution to the optimization problem is enough to prove stability has been reported in [28], [29], [30]. These rely on that the terminal constraint set can be reached also for suboptimal solutions, which can be used to show closed loop stability. For MPC without a terminal constraint set, stability was shown in [31] for incomplete optimization. The optimization algorithm is terminated early when a certain decrease in the cost has been obtained. However, they do not provide any guarantees that this decrease is achievable in each step. In this paper we use a different decrease condition than in [31] which enables a priori guarantees that the condition will hold with finite number of algorithm iterations in every time step.

An issue associated with duality-based optimization is that primal feasibility cannot be guaranteed before convergence of the optimization algorithm. Such feasibility problems have previously been addressed in [32] using a constraint tightening approach. Constraint tightening can be used to generate feasible solutions but complicates stability analysis. The reason is that the optimal value function without constraint tightening is used to show stability, while the actual optimization is performed with constraint tightening. This problem is overcome in [32] by assuming that the difference between the optimal value functions with and without constraint tightening is bounded by a constant. However, to actually compute such a constant may be difficult. In this paper we instead use a novel adaptive constraint tightening approach that ensures feasibility w.r.t. the original constraint set with a finite number of algorithm iterations. We introduce a condition for the adaptation that bounds the difference between the optimal value functions with and without constraint tightening. This makes it possible to prove stability without stating additional assumptions.

The paper is organized as follows. In Section II we introduce the problem and present the distributed optimization algorithm in [1]. In Section III the stopping condition is presented and feasibility, stability, and performance is analyzed. Section IV is devoted to computation of the controllability parameter. A numerical example that shows the efficiency of the proposed stopping condition is presented in Section V. Finally, in Section VI we conclude the paper.

II. Problem setup and preliminaries

We consider linear dynamical systems of the form

$$x_{t+1} = Ax_t + Bu_t, \qquad\qquad x_0 = \bar{x} \tag{1}$$

where $x_t \in \mathbb{R}^n$ and $u_t \in \mathbb{R}^m$ denote the state and control vectors at time t and the pair (A, B) is assumed controllable. We introduce the following state and control variable partitions

$$x_t = [(x_t^1)^T, (x_t^2)^T, \ldots, (x_t^M)^T]^T, \tag{2}$$

$$u_t = [(u_t^1)^T, (u_t^2)^T, \ldots, (u_t^M)^T]^T \tag{3}$$

where the local variables $x_t^i \in \mathbb{R}^{n_i}$ and $u_t^i \in \mathbb{R}^{m_i}$. The A and B matrices are partitioned accordingly

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1M} \\ \vdots & \ddots & \vdots \\ A_{M1} & \cdots & A_{MM} \end{pmatrix}, \qquad B = \begin{pmatrix} B_{11} & \cdots & B_{1M} \\ \vdots & \ddots & \vdots \\ B_{M1} & \cdots & B_{MM} \end{pmatrix}.$$

These matrices are assumed to have a sparse structure, i.e., some $A_{ij} = 0$ and $B_{ij} = 0$ and the neighboring interaction is defined by the following sets

$$\mathcal{N}_i = \{j \in \{1, \ldots, M\} \mid \text{if } A_{ij} \neq 0 \text{ or } B_{ij} \neq 0\}$$

for $i = 1, \ldots, M$. This gives the following local dynamics

$$x_{t+1}^i = \sum_{j \in \mathcal{N}_i} \left(A_{ij} x_t^j + B_{ij} u_t^j \right), \qquad x_0^i = \bar{x}_i$$

for $i = 1, \ldots, M$. The local control and state variables are constrained, i.e., $u^i \in \mathcal{U}_i$ and $x^i \in \mathcal{X}_i$. The constraint sets, $\mathcal{X}_i, \mathcal{U}_i$ are assumed to be bounded polytopes containing zero in their respective interiors and can hence be represented as

$$\mathcal{X}_i = \{x^i \in \mathbb{R}^{n_i} \mid C_x^i x^i \leq d_x^i\}, \qquad \mathcal{U}_i = \{u^i \in \mathbb{R}^{m_i} \mid C_u^i u^i \leq d_u^i\} \tag{4}$$

where $C_x^i \in \mathbb{R}^{n_{c_{x^i}} \times n_i}$, $C_u^i \in \mathbb{R}^{n_{c_{u^i}} \times m_i}$, $d_x^i \in \mathbb{R}_{>0}^{n_{c_{x^i}}}$ and $d_u^i \in \mathbb{R}_{>0}^{n_{c_{u^i}}}$. We also denote the total number of linear inequalities describing all constraint sets by $n_c := \sum_{i=1}^M \left(n_{c_{x^i}} + n_{c_{u^i}} \right)$. The global constraint sets are defined from the local ones through

$$\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_M, \qquad \mathcal{U} = \mathcal{U}_1 \times \ldots \times \mathcal{U}_M.$$

We use a separable quadratic stage cost

$$\ell(x, u) = \sum_{i=1}^M \ell_i(x^i, u^i) = \frac{1}{2} \left(\sum_{i=1}^M (x^i)^T Q_i x^i + (u^i)^T R_i u^i \right) \tag{5}$$

where $Q_i \in \mathbb{S}_{++}^{n_i}$ and $R_i \in \mathbb{S}_{++}^{m_i}$ for $i = 1, \ldots, M$ and $\mathbb{S}_{++}^n$ denotes the set of symmetric positive definite matrices in $\mathbb{R}^{n \times n}$. The optimal infinite horizon cost from initial state $\bar{x} \in \mathcal{X}$ is defined by

$$\begin{aligned} V_\infty(\bar{x}) := \quad \min_{x,u} \quad & \sum_{t=0}^{\infty} \ell(x_t, u_t) \\ \text{s.t.} \quad & x_t \in \mathcal{X} \quad , \quad u_t \in \mathcal{U} \\ & x_{t+1} = A x_t + B u_t \\ & x_0 = \bar{x}. \end{aligned} \tag{6}$$

Such infinite horizon optimization problems are in general intractable to solve exactly. A common approach is to solve the problem approximately in receding horizon fashion. To this end we introduce the predicted state and control sequences $\{z_\tau\}_{\tau=0}^{N-1}$ and $\{v_\tau\}_{\tau=0}^{N-1}$ and the corresponding stacked vectors

$$\mathbf{z} = [z_0^T, \ldots, z_{N-1}^T]^T, \qquad \mathbf{v} = [v_0^T, \ldots, v_{N-1}^T]^T \tag{7}$$

where z_τ and v_τ are predicted states and controls τ time steps ahead. The predicted state and control variables z_τ, v_τ are partitioned into local variables as in (2) and (3) respectively. We also introduce the following stacked local vectors

$$\mathbf{z}_i = [(z_0^i)^T, \ldots, (z_{N-1}^i)^T]^T, \qquad \mathbf{v}_i = [(v_0^i)^T, \ldots, (v_{N-1}^i)^T]^T. \tag{8}$$

Further, we introduce the tightened state and control constraint sets

$$(1-\delta)\mathcal{X}_i = \{x^i \in \mathbb{R}^{n_i} \mid C_x^i x^i \leq (1-\delta)d_x^i\}, \tag{9}$$

$$(1-\delta)\mathcal{U}_i = \{u^i \in \mathbb{R}^{m_i} \mid C_u^i u^i \leq (1-\delta)d_u^i\} \tag{10}$$

where $\delta \in (0,1)$ decides the amount of relative constraint tightening. The following optimization problem is solved in the DMPC controller for the current state $\bar{x} \in \mathbb{R}^n$

$$\begin{aligned} V_N^\delta(\bar{x}) := \min_{\mathbf{z}_t, \mathbf{v}_t} \quad & \sum_{\tau=0}^{N-1} \ell(z_\tau, v_\tau) \\ \text{s.t.} \quad & z_\tau \in (1-\delta)\mathcal{X},\ \tau = 0, \ldots, N-1 \\ & v_\tau \in (1-\delta)\mathcal{U},\ \tau = 0, \ldots, N-1 \\ & z_{\tau+1} = Az_\tau + Bv_\tau,\ \tau = 0, \ldots, N-2 \\ & z_0 = \bar{x}. \end{aligned} \tag{11}$$

Such optimization problems can be solved in distributed fashion using, i.e., the alternating direction of multipliers method [23] or dual ascent [23]. In this work we have chosen to use the recently developed distributed method presented in [1] which is an accelerated dual ascent method which has superior convergence properties $O(1/k^2)$ compared to the classical dual ascent method which achieves $O(1/k)$. For distribution purposes, we have neither a terminal cost nor a terminal constraint set in the optimization problem (11).

Next, we present the distributed optimization algorithm in [1]. We stack all decision variables into one vector

$$\mathbf{y} = [z_0^T, \ldots, z_{N-1}^T, v_0^T, \ldots, v_{N-1}^T]^T \in \mathbb{R}^{(n+m)N}. \tag{12}$$

The optimization problem (11) can more compactly be written as

$$\begin{aligned} V_N^\delta(\bar{x}) := \min_{\mathbf{y}} \quad & \tfrac{1}{2}\mathbf{y}^T\mathbf{H}\mathbf{y} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{y} = \mathbf{b}\bar{x} \\ & \mathbf{C}\mathbf{y} \leq (1-\delta)\mathbf{d} \end{aligned} \tag{13}$$

where $\mathbf{H} \in \mathbb{S}_{++}^{(n+m)N}, \mathbf{A} \in \mathbb{R}^{n(N-1)\times(n+m)N}, \mathbf{b} \in \mathbb{R}^{n(N-1)\times n}, \mathbf{C} \in \mathbb{R}^{n_c N\times(n+m)N}$ and $\mathbf{d} \in \mathbb{R}_{>0}^{Nn_c}$ are built accordingly. The separable structure of the cost function (5) and constraint sets (4) gives block diagonal $\mathbf{H}$ and $\mathbf{C}$-matrices. Further, the matrix $\mathbf{A}$ is sparse since it is composed of sparse matrices A, B and I that define the linear dynamic constraints (1). The dual problem to (13) is created by introducing dual variables $\boldsymbol{\lambda} \in \mathbb{R}^{n(N-1)}$ for the equality constraints and dual variables $\boldsymbol{\mu} \in \mathbb{R}_{\geq 0}^{Nn_c}$ for the inequality constraints. The dual problems becomes

$$\max_{\boldsymbol{\lambda},\boldsymbol{\mu}\geq 0} \min_{\mathbf{y}} \frac{1}{2}\mathbf{y}^T\mathbf{H}\mathbf{y} + \boldsymbol{\lambda}^T(\mathbf{A}\mathbf{y} - \mathbf{b}\bar{x}) + \boldsymbol{\mu}^T(\mathbf{C}\mathbf{y} - (1-\delta)\mathbf{d}) \tag{14}$$

which, as shown in [1], can explicitly be written as

$$\max_{\boldsymbol{\lambda},\boldsymbol{\mu}\geq 0} -\frac{1}{2}(\mathbf{A}^T\boldsymbol{\lambda} + \mathbf{C}^T\boldsymbol{\mu})^T\mathbf{H}^{-1}(\mathbf{A}^T\boldsymbol{\lambda} + \mathbf{C}^T\boldsymbol{\mu}) - \boldsymbol{\lambda}^T\mathbf{b}\bar{x} - \boldsymbol{\mu}^T\mathbf{d}(1-\delta). \tag{15}$$

We define the dual function for initial condition $\bar{x} \in \mathbb{R}^n$ as

$$D_N^\delta(\bar{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) := -\frac{1}{2}(\mathbf{A}^T\boldsymbol{\lambda} + \mathbf{C}^T\boldsymbol{\mu})^T\mathbf{H}^{-1}(\mathbf{A}^T\boldsymbol{\lambda} + \mathbf{C}^T\boldsymbol{\mu}) - \boldsymbol{\lambda}^T\mathbf{b}\bar{x} - \boldsymbol{\mu}^T\mathbf{d}(1-\delta). \tag{16}$$

It was in [1] shown that the smallest Lipschitz constant to ∇D_N^δ is $L = \|[\mathbf{A}^T, \mathbf{C}^T]^T\mathbf{H}^{-1}[\mathbf{A}^T, \mathbf{C}^T]\|$ and that (13) can be solved by the following accelerated dual gradient method

$$\mathbf{y}^k = -\mathbf{H}^{-1}(\mathbf{A}^T\boldsymbol{\lambda}^k + \mathbf{C}^T\boldsymbol{\mu}^k) \tag{17}$$

$$\bar{\mathbf{y}}^k = \mathbf{y}^k + \frac{k-1}{k+2}(\mathbf{y}^k - \mathbf{y}^{k-1}) \tag{18}$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \frac{k-1}{k+2}(\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k-1}) + \frac{1}{L}(\mathbf{A}\bar{\mathbf{y}}^k - \mathbf{b}\bar{x}) \tag{19}$$

$$\boldsymbol{\mu}^{k+1} = \max\left(0, \boldsymbol{\mu}^k + \frac{k-1}{k+2}(\boldsymbol{\mu}^k - \boldsymbol{\mu}^{k-1}) + \frac{1}{L}(\mathbf{C}\bar{\mathbf{y}}^k - \mathbf{d}\,(1-\delta))\right). \tag{20}$$

Due to the structure of the matrices $\mathbf{H}, \mathbf{C}, \mathbf{A}$ this algorithm can be implemented in distributed fashion where communication between subsystems i and j takes place if $j \in \mathcal{N}_i$ or $i \in \mathcal{N}_j$, see [1] for details.

In the following section we present a stopping condition to algorithm (17)-(20) when solving (11) that guarantees feasibility, stability, and a prespecified performance of the DMPC scheme.

However, the stopping condition is not developed exclusively for the presented algorithm. It is directly applicable to any (distributed) optimization algorithm that produces dual feasible iterations that converge to the optimal dual variables.

A. Notation

We denote by $\mathbb{R}$ the set of real numbers, $\mathbb{R}_{\geq c}$ the set of real numbers $d \geq c$ and $\mathbb{R}_{>c}$ the set of real numbers $d > c$. We denote by $\mathbb{S}^n_{++} \subset \mathbb{R}^{n\times n}$ the set of real symmetric positive definite matrices. Further $\mathbb{N}_{\geq T}$ is the set of natural numbers $t \geq T$. The norm $\|\cdot\|$ refers to the Euclidean norm or the induced Euclidean norm unless otherwise is specified and $\langle\cdot,\cdot\rangle$ refers to the inner product in Euclidean space. The norm $\|x\|_T = \sqrt{x^T T x}$. The interior of a set $\mathcal{X}$ is denoted $\mathrm{int}(\mathcal{X})$. The optimal value function with original constraint set, i.e. $V_N^0(\bar{x})$, is denoted $V_N(\bar{x})$. The optimal state and control sequences to (11) for initial value x and constraint tightening δ are denoted $\{z_\tau^*(x,\delta)\}_{\tau=0}^{N-1}$ and $\{v_\tau^*(x,\delta)\}_{\tau=0}^{N-1}$ respectively and the optimal solution to the equivalent problem (13) by $\mathbf{y}^*(x,\delta)$. The state and control sequences for iteration k in (17)-(20) are denoted $\{z_\tau^k(x,\delta)\}_{\tau=0}^{N-1}$ and $\{v_\tau^k(x,\delta)\}_{\tau=0}^{N-1}$ respectively. We drop the initial state and constraint tightening arguments (x,δ) when no ambiguities can arise.

B. Definitions and assumptions

We adopt the convention that $V_N^\delta(\bar{x}) = \infty$ for states $\bar{x} \in \mathbb{R}^n$ that result in (13) being infeasible. We define by $\mathbb{X}_\infty$ the set for which (6) is feasible. We also define the minimum of the stage-cost ℓ for fixed x

$$\ell^*(x) := \min_{u\in\mathcal{U}} \ell(x,u) = \frac{1}{2}x^T Q x.$$

Further, κ is the smallest scalar such that $\kappa Q - A^T Q A \succeq 0$. The state sequence resulting from applying $\{v_\tau\}_{\tau=0}^{N-1}$ to (1) is denoted by $\{\xi_\tau\}_{\tau=0}^{N-1}$, i.e.,

$$\xi_{\tau+1} = A\xi_\tau + Bv_\tau, \qquad \xi_0 = \bar{x}. \tag{21}$$

We introduce $\boldsymbol{\xi} = [(\xi_0)^T, \ldots, (\xi_{N-1})^T]^T$ and define the primal cost

$$P_N(\bar{x},\mathbf{v}) := \begin{cases} \sum_{\tau=0}^{N-1} \ell(\xi_\tau, v_\tau) & \text{if } \boldsymbol{\xi} \in \mathcal{X}^N \text{ and } \mathbf{v} \in \mathcal{U}^N, \text{ and (21) holds} \\ \infty & \text{else} \end{cases} \tag{22}$$

where $\mathcal{X}^N$ and $\mathcal{U}^N$ are the state and control constraints for the full horizon. We also introduce the shifted control sequence $\mathbf{v}_s = [(v_1)^T, \ldots, (v_{N-1})^T, 0^T]^T$. We have $P_N(\bar{x}, \mathbf{v}^k) \geq V_N(\bar{x})$ and $P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k) \geq V_N(A\bar{x} + Bv_0^k)$ for every algorithm iteration k. We denote by $\{\xi_\tau^k\}_{\tau=0}^{N-1}$ the state sequence that satisfies (21) using controls $\{v_\tau^k\}_{\tau=0}^{N-1}$. The definition of the cost (22) implies

$$P_N(\bar{x}, \mathbf{v}^k) = P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k) + \ell(\bar{x}, v_0^k) - \ell^*(A\xi_{N-1}^k) \tag{23}$$

if $v_0^k \in \mathcal{U}$, $\bar{x} \in \mathcal{X}$ and $A\xi_{N-1}^k \in \mathcal{X}$.

III. Stopping condition

Rather than finding the optimal solution in each time step in the MPC controller, the most important task is to find a control action that gives desirable closed loop properties such as stability, feasibility, and a desired performance. Such properties can sometimes be ensured well before convergence to the optimal solution. To benefit from this observation, a stopping condition is developed that allows the iterations to stop when the desired performance, stability, and feasibility can be guaranteed. Before the stopping condition is introduced, we briefly go through the main ideas below.

A. Main ideas

The distributed nature of the optimization algorithm makes it unsuitable for centralized terminal costs and terminal constraints. Thus, stability and performance need to be ensured without these constructions. We define the following infinite horizon performance for feedback control law ν

$$V_\infty^\nu(\bar{x}) = \sum_{t=0}^{\infty} \ell(x_t, \nu(x_t)) \tag{24}$$

where $x_{t+1} = Ax_t + B\nu(x_t)$ and $x_0 = \bar{x}$. For a given performance parameter $\alpha \in (0, 1]$ and control law ν it is known (c.f. [26], [27], [25]) that the following decrease in the optimal value function

$$V_N^0(x_t) \geq V_N^0(Ax_t + B\nu(x_t)) + \alpha\ell(x, \nu(x_t)) \tag{25}$$

for every $t \in \mathbb{N}_{\geq 0}$ gives stability and closed loop performance according to

$$\alpha V_\infty^\nu(\bar{x}) \leq V_\infty(\bar{x}). \tag{26}$$

Analysis of the control horizon N needed for an MPC control law without terminal cost and terminal constraints such that (25) holds, is performed in [27], [25] and also in this paper. Once a control horizon N is known such that (25) is guaranteed, the performance result (26) relies on computation of the optimal solution to the MPC optimization problem in every time step. An exact optimal solution cannot be computed and the idea behind this paper is to develop stopping conditions that enable early termination of the optimization algorithm with maintained feasibility, stability, and performance guarantees. The idea behind our stopping condition is to compute a lower bound to $V_N^0(x)$ through the dual function $D_N^0(x, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k)$ and an upper bound to the next step value function $V_N^0(Ax + Bv_0^k)$ through a feasible solution $P_N(Ax + Bv_0^k, \mathbf{v}_s^k)$. If at iteration k the following test is satisfied

$$D_N^0(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) \geq P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k) + \alpha\ell(\bar{x}, v_0^k) \tag{27}$$

the performance condition (25) holds since

$$\begin{aligned} V_N^0(\bar{x}) \geq D_N^0(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) &\geq P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k) + \alpha\ell(\bar{x}, v_0^k) \\ &\geq V_N^0(A\bar{x} + Bv_0^k) + \alpha\ell(\bar{x}, v_0^k). \end{aligned}$$

This implies that stability and the performance result (26) can be guaranteed with finite algorithm iterations k by using control action v_0^k.

The test (27) includes computation of $P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k)$ which is a feasible solution to the optimization problem in the following step. A feasible solution cannot be expected with finite number of iterations k for duality-based methods since primal feasibility is only guaranteed in the limit of iterations. Therefore we introduce tightened state and control constraint sets $(1-\delta)\mathcal{X}$, $(1-\delta)\mathcal{U}$ with $\delta \in (0,1)$ and use these in the optimization problem. By generating a state trajectory $\{\xi_\tau^k\}_{\tau=0}^{N-1}$ from the control trajectory $\{v_\tau^k\}_{\tau=0}^{N-1}$ that satisfies the equality constraints (21), we will see that $\{\xi_\tau^k\}_{\tau=0}^{N-1}$ satisfies the original inequality constraints with finite number of iterations. Thus, a primal feasible solution $P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k)$ can be generated after a finite number of algorithm iterations k. However, since the optimization now is performed over a tightened constraint set, the dual function value $D_N^\delta(\bar{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ is not a lower bound to $V_N^0(\bar{x})$ and cannot be used directly in the test (27) to ensure stability and the performance specified by (26). In the following lemma we show a relation between the dual function value when using the tightened constraint sets and the optimal value function when using the original constraint sets.

Lemma 1: For every $\bar{x} \in \mathbb{R}^n$, $\boldsymbol{\lambda} \in \mathbb{R}^{n(N-1)}$ and $\boldsymbol{\mu} \in \mathbb{R}^{Nn_c}_{\geq 0}$ we have that

$$V_N^0(\bar{x}) \geq D_N^\delta(\bar{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) - \delta \boldsymbol{\mu}^T \mathbf{d}. \tag{28}$$

Proof. From the definition of the dual function (16) we get that

$$D_N^\delta(\bar{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = D_N^0(\bar{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) + \delta \mathbf{d}^T \boldsymbol{\mu}.$$

By weak duality we get

$$V_N^0(\bar{x}) \geq D_N^0(\bar{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = D_N^\delta(\bar{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) - \delta \mathbf{d}^T \boldsymbol{\mu}. \tag{29}$$

This completes the proof. □

The presented lemma enables computation of a lower bound to $V_N^0(\bar{x})$ at algorithm iteration k that depends on $\delta \boldsymbol{\mu}^T \mathbf{d}$. By adapting the amount of constraint tightening δ to satisfy

$$\delta (\boldsymbol{\mu}^k)^T \mathbf{d} \leq \epsilon \ell^*(\bar{x}) \tag{30}$$

for some $\epsilon > 0$ and use this together with the following test

$$D_N^\delta(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) \geq P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k) + \alpha \ell(\bar{x}, v_0^k) \tag{31}$$

we get from Lemma 1 and if (30) and (31) holds that

$$\begin{aligned} V_N^0(\bar{x}) &\geq D_N^\delta(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) - \delta(\boldsymbol{\mu}^k)^T \mathbf{d} \geq P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k) + \alpha \ell(\bar{x}, v_0^k) - \epsilon \ell^*(\bar{x}) \\ &\geq V_N^0(A\bar{x} + Bv_0^k) + (\alpha - \epsilon)\ell(\bar{x}, v_0^k). \end{aligned}$$

This is condition (25), which guarantees stability and performance specified by (26) if $\alpha > \epsilon$.

B. Stopping conditions

From the discussion in the previous section we conclude that two parameters need to be specified in the stopping condition. The first is the performance parameter $\alpha \in (0, 1]$ which guarantees closed loop performance as specified by (26). The larger α, the better performance is guaranteed but a longer control horizon N will be needed to guarantee the specified performance. The second is an initial constraint tightening parameter, which we denote by $\delta_{\text{init}} \in (0, 1]$, from which the constraint tightening parameter δ will be adapted (reduced), to satisfy (30). A generic value of δ_{init} is $\delta_{\text{init}} = 0.2$, i.e., 20% initial constraint tightening. Also a third parameter needs to be set. It is the relative optimality tolerance $\epsilon > 0$ where $\epsilon < \alpha$. The effect of this parameter on the algorithm is smaller than the effect of the other parameters and it is generically chosen to satisfy $\epsilon \in [0.01, 0.001]$.

Algorithm 1: **Stopping condition**

Input: $\bar{x}$

Set: $k = 0,\ l = 0,\ \delta = \delta_{\text{init}}$

Initialize algorithm (17)-(20) with:

$\boldsymbol{\lambda}^0 = \boldsymbol{\lambda}^{-1} = 0, \boldsymbol{\mu}^0 = \boldsymbol{\mu}^{-1} = 0$ and $\mathbf{y}^0 = \mathbf{y}^{-1} = 0$.

Do

  If $D_N^\delta(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) \geq P_N(\bar{x}, \mathbf{v}^k) - \frac{\epsilon}{l+1}\ell^*(\bar{x})$

   or $\delta \mathbf{d}^T \boldsymbol{\mu}^k > \epsilon \ell^*(\bar{x})$

    Set $\delta \leftarrow \delta/2$ // reduce constraint tightening

    Set $l \leftarrow l + 1$

    Set $k = 0$ // reset step size and iteration counter

  End

  Run Δk iterations of (17)-(20)

  Set $k \leftarrow k + \Delta k$

Until $D_N^\delta(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) \geq P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k) + \alpha \ell(\bar{x}, v_0^k)$ **and**

   $\delta \mathbf{d}^T \boldsymbol{\mu}^k \leq \epsilon \ell^*(\bar{x})$

Output: v_0^k

Except for the initial condition $\bar{x}$, Algorithm 1 is always identically initialized and follows a deterministic scheme. Thus, for fixed initial condition the same control action is always computed. This implies that Algorithm 1 defines a static feedback control law, which we denote by ν_N. We get the following closed loop dynamics

$$x_{t+1} = Ax_t + B\nu_N(x_t), \qquad x_0 = \bar{x}. \tag{32}$$

The objective of this section is to present a theorem stating that the feedback control law ν_N is well defined on $\text{int}(\mathbb{X}_N^0)$ where

$$\mathbb{X}_N^\delta := \{\bar{x} \in \mathbb{R}^n \mid V_N^\delta(\bar{x}) < \infty \text{ and } Az_{N-1}^*(\bar{x}, 0) \in \text{int}(\mathcal{X})\} \tag{33}$$

which satisfies $\mathbb{X}_N^{\delta_1} \subseteq \mathbb{X}_N^{\delta_2}$ for $\delta_1 > \delta_2$. First, however we state the following definition.

Definition 1: The constant Φ_N is the smallest constant such that the optimal solution $\{z_\tau^*(\bar{x},0)\}_{\tau=0}^{N-1}$, $\{v_\tau^*(\bar{x},0)\}_{\tau=0}^{N-1}$ to (11) for every $\bar{x} \in \mathbb{X}_N^0$ satisfies

$$\ell^*(z_{N-1}^*(\bar{x},0)) \leq \Phi_N \ell(\bar{x}, v_0^*(\bar{x},0)) \tag{34}$$

for the chosen control horizon N.

In Section IV a method to compute Φ_N is presented.

Remark 1: In [24], [25] an exponential controllability on the stage costs is assumed, i.e., that for $C \geq 1$ and $\sigma \in (0,1)$ the following holds for $\tau = 0, \ldots, N-1$

$$\ell^*(z_\tau^*(\bar{x},0), v_\tau^*(\bar{x},0)) \leq C\sigma^\tau \ell(\bar{x}, v_0^*(\bar{x},0)). \tag{35}$$

This implies $\Phi_N \leq C\sigma^{N-1}$.

We also need the following lemmas that are proven in Appendix-A, Appendix-B and Appendix-C respectively to prove the upcoming theorem.

Lemma 2: Suppose that $\epsilon > 0$ and $\delta \in (0,1]$. For every $\bar{x} \in \mathbb{X}_N^\delta$ we have for some finite k that

$$D_N^\delta(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) \geq P_N(\bar{x}, \mathbf{v}^k) - \epsilon \ell^*(\bar{x}). \tag{36}$$

Lemma 3: Suppose that $\epsilon > 0$ and $\delta \in (0,1]$. For every $\bar{x} \in \mathbb{X}_N^\delta$ and algorithm iteration k such that (36) holds we have for $\tau = 0, \ldots, N-1$ that

$$\frac{1}{2}\left\| \begin{bmatrix} \xi_\tau^k(\bar{x},\delta) \\ v_\tau^k(\bar{x},\delta) \end{bmatrix} - \begin{bmatrix} z_\tau^*(\bar{x},0) \\ v_\tau^*(\bar{x},0) \end{bmatrix} \right\|_H^2 \leq \epsilon \ell^*(\bar{x}) + \delta(\boldsymbol{\mu}^k)^T \mathbf{d}$$

where $H = \text{blkdiag}(Q, R)$.

Lemma 4: Suppose that $\epsilon > 0$ and $\delta \in (0,1]$. For $\bar{x} \in \mathbb{X}_N^0$ but $\bar{x} \notin \mathbb{X}_N^\delta$ we have that $\delta(\boldsymbol{\mu}^k)^T \mathbf{d} > \epsilon \ell^*(\bar{x})$ with finite k.

We are now ready to state the following theorem, which is proven in Appendix-D.

Theorem 1: Assume that $\epsilon > 0$, $\delta_{\text{init}} \in (0,1]$ and

$$\alpha \leq 1 - \epsilon - \kappa(\sqrt{2\epsilon} + \sqrt{\Phi_N})^2(\sqrt{2\epsilon}+1)^2. \tag{37}$$

Then the feedback control law ν_N, defined by Algorithm 1, is well defined for every $\bar{x} \in \text{int}(\mathbb{X}_\text{N}^0)$. Further

$$V_N^0(\bar{x}) \geq V_N^0(A\bar{x} + B\nu_N(\bar{x})) + (\alpha - \epsilon)\ell(\bar{x}, \nu_N(\bar{x})). \tag{38}$$

holds for every $\bar{x} \in \text{int}(\mathbb{X}_{\text{N}}^0)$.

Corollary 1: Suppose that $\alpha \leq 1 - \kappa\Phi_N$ and that $\nu_N^*(\bar{x}) = v_0^*(\bar{x}, 0)$. Then

$$V_N^0(\bar{x}) \geq V_N^0(A\bar{x} + B\nu_N^*(\bar{x})) + \alpha\ell(\bar{x}, \nu_N^*(\bar{x})).$$

holds for every $\bar{x} \in \mathbb{X}_N^0$.

Proof. For every $\bar{x} \in \mathbb{X}_N^0$ we have

$$\begin{aligned} V_N^0(\bar{x}) &= \sum_{\tau=0}^{N-1} \ell(z_\tau^*, u_\tau^*) + \ell(Az_{N-1}^*, 0) - \ell(Az_{N-1}^*, 0) \\ &\geq V_N^0(A\bar{x} + B\nu_N^*(\bar{x})) + \ell(\bar{x}, v_0^*) - \ell(Az_{N-1}^*, 0) \\ &\geq V_N^0(A\bar{x} + B\nu_N^*(\bar{x})) + \ell(\bar{x}, v_0^*) - \kappa\ell(z_{N-1}^*, 0) \\ &\geq V_N^0(A\bar{x} + B\nu_N^*(\bar{x})) + (1 - \kappa\Phi_N)\ell(\bar{x}, v_0^*) \end{aligned}$$

where the first inequality holds since $Az_{N-1}^* \in \mathcal{X}$ by construction of $\mathbb{X}_N^0$, the second due to the definition of κ and the third due to the definition of Φ_N. □

Remark 2: By setting $\epsilon = 0$ in Theorem 1 we get $\alpha \leq 1 - \kappa\Phi_N$ as in Corollary 1.

We have proven that the feedback control law is well defined on $\text{int}(\mathbb{X}_N^0)$. The topic of the following section is to analyze feasibility, stability, and performance of the proposed feedback controller.

C. Feasibility, stability and performance

The following proposition shows one-step feasibility when using the feedback control law ν_N.

Proposition 1: Suppose that α satisfies (37). For every $x_t \in \text{int}(\mathbb{X}_N^0)$ we have that $x_{t+1} = Ax_t + B\nu_N(x_t) \in \mathcal{X}$.

Proof. From Theorem 1 we have that $\nu_N(x_t)$ is well defined and from Algorithm 1 we have that $P_N(x_{t+1}, \mathbf{v}_s^k) < \infty$ which, by definition, implies that $x_{t+1} \in \mathcal{X}$. □

The proposition shows that x_{t+1} is feasible if the control law $\nu_N(x_t)$ is well defined. We define the recursively feasible set as the maximal set such that

$$\mathbb{X}_{\text{rf}} = \{x \in \mathcal{X} \mid Ax + B\nu_N(x) \in \mathbb{X}_{\text{rf}}\} \tag{39}$$

In the following theorem we show that $\mathbb{X}_{\text{rf}}$ is the region of attraction and that the control law ν_N achieves a prespecified performance as specified by (24).

Theorem 2: Suppose that $\alpha > \epsilon$ satisfies (37). Then for every initial condition $\bar{x} \in \mathbb{X}_{\text{rf}}$ we have that $\|x_t\| \to 0$ as $t \to \infty$ and that the closed loop performance satisfies

$$(\alpha - \epsilon) V_\infty^{\nu_N}(\bar{x}) \leq V_\infty(\bar{x}). \tag{40}$$

Further, $\mathbb{X}_{\text{rf}}$ is the region of attraction.

Proof. From the definition of $\mathbb{X}_{\text{rf}}$ we know that $\bar{x} = x_0 \in \mathbb{X}_{\text{rf}}$ implies $x_t \in \mathbb{X}_{\text{rf}}$ for all $t \in \mathbb{N}_{\geq 0}$. This implies that $\nu_N(x_t)$ is well defined and that (38) holds for all x_t, $t \in \mathbb{N}_{\geq 0}$. In [27, Proposition 2.2] it was shown using telescope summation that (38) implies (40). Further, since the stage cost ℓ satisfies [25, Assumption 5.1] we get from [25, Theorem 5.2] that $\|x_t\| \to 0$ as $t \to \infty$.

What is left to show is that $\mathbb{X}_{\text{rf}}$ is the region of attraction. Denote by $\mathbb{X}_{\text{roa}}$ the region of attraction using ν_N. We have above shown that $\mathbb{X}_{\text{rf}} \subseteq \mathbb{X}_{\text{roa}}$. We next show that $\mathbb{X}_{\text{roa}} \subseteq \mathbb{X}_{\text{rf}}$ by a contradiction argument to conclude that $\mathbb{X}_{\text{rf}} = \mathbb{X}_{\text{roa}}$. Assume that there exist $\bar{x} \in \mathbb{X}_{\text{roa}}$ such that $\bar{x} \notin \mathbb{X}_{\text{rf}}$. If $\bar{x} \in \mathbb{X}_{\text{roa}}$ the closed loop state sequence $\{x_t\}_{t=0}^\infty$ is feasible in every step (and converges to the origin) and consequently $\{Ax_t + B\nu_N(x_t)\}_{t=0}^\infty$ is feasible in every step. This is exactly the requirement to have $\bar{x} \in \mathbb{X}_{\text{rf}}$, which is a contradiction. Thus $\mathbb{X}_{\text{rf}} \subseteq \mathbb{X}_{\text{roa}} \subseteq \mathbb{X}_{\text{rf}}$ which implies that $\mathbb{X}_{\text{rf}} = \mathbb{X}_{\text{roa}}$.

This completes the proof. □

Remark 3: The lack of terminal constraint sets implies that recursive feasibility cannot be guaranteed. However, to actually guarantee recursive feasibility in presence of disturbances, robust MPC formulations need to be considered. These can be fairly restrictive and have a rather small region of attraction. In the examples we will see that the region of attraction can be significantly enlarged by not using terminal constraints.

To guarantee a priori that the control law ν_N achieves the performance (40) specified by α, we need to find a control horizon N such that the corresponding controllability parameter Φ_N satisfies (37). This requires the computation of controllability parameter Φ_N which is the topic of the next section.

IV. Offline controllability verification

The stability and performance results in Theorem 2 rely on Definition 1. For the results to be practically meaningful it must be possible to compute Φ_N in Definition 1. In this section we will show that this can be done by solving a mixed integer linear program (MILP). For desired performance specified by α, we get a requirement on the controllability parameter through (37) for Theorem 1 and Theorem 2 to hold. We denote by Φ_α the largest controllability parameter such that Theorem 1 and Theorem 2 holds for the specified α. This parameter is the one that gives equality in (37), i.e., satisfies

$$\alpha = 1 - \epsilon - \kappa(\sqrt{2\epsilon} + \sqrt{\Phi_\alpha})^2(\sqrt{2\epsilon} + 1)^2 \tag{41}$$

for the desired performance α and optimality tolerance ϵ. The parameters α and ϵ must be chosen such that $\Phi_\alpha > 0$. The objective is to find a control horizon N such that the corresponding controllability parameter Φ_N satisfies $\Phi_N \leq \Phi_\alpha$. First we show that for long enough control horizon N there exist a $\Phi_N \leq \Phi_\alpha$.

Lemma 5: Assume that α and ϵ are chosen such that $\Phi_\alpha > 0$ where Φ_α is implicitly defined in (41). Then there exists control horizon N and corresponding controllability parameter $\Phi_N \leq \Phi_\alpha$.

Proof. Since $\mathbb{X}_{\text{rf}}$ is the region of attraction we have $\mathbb{X}_{\text{rf}} \subseteq \mathbb{X}_\infty$. This in turn implies that (13) is feasible for every control horizon $N \in \mathbb{N}_{\geq 1}$ due to the absence of terminal constraints. We have

$$V_N(\bar{x}) = \sum_{\tau=0}^{N-2} \ell(z_\tau^*, v_\tau^*) + \ell(z_{N-1}^*, v_{N-1}^*) \geq V_{N-1}(\bar{x}) + \ell(z_{N-1}^*, v_{N-1}^*).$$

Since the pair (A, B) is assumed controllable and since (13) has neither terminal constraints nor terminal cost we have for some finite M that $M \geq V_\infty(\bar{x}) \geq V_N(\bar{x}) \geq V_{N-1}(\bar{x})$. Thus the sequence $\{V_N(\bar{x})\}_{N=0}^\infty$ is a bounded monotonic increasing sequence which is well known to be convergent. Thus, for $N \geq \bar{N}$ where $\bar{N}$ is large enough the difference $V_N(\bar{x}) - V_{N-1}(\bar{x})$ is arbitrarily small. Especially $\ell(z_{N-1}^*, v_{N-1}^*) = \ell^*(z_{N-1}^*) \leq V_N(\bar{x}) - V_{N-1}(\bar{x}) \leq \Phi_\alpha \ell(\bar{x}, v_0^*)$ since $\Phi_\alpha > 0$. That is, for long enough control horizon $N \geq \bar{N}$, $\Phi_N \leq \Phi_\alpha$. This completes the proof. □

The preceding Lemma shows that there exists a control horizon N such that $\Phi_N \leq \Phi_\alpha$ if $\Phi_\alpha > 0$ for the chosen performance α and tolerance ϵ. The choice of performance parameter α gives requirements on how ϵ can be chosen to give $\Phi_\alpha > 0$. Larger ϵ requires smaller Φ_α to satisfy (41) which in turn requires longer control horizons N since Φ_N must satisfy $\Phi_N \leq \Phi_\alpha$. In the following section we address the problem of how to compute the control horizon N and corresponding Φ_N such that the desired performance specified by α can be guaranteed.

A. *Exact verification of controllability parameter*

In the following proposition we introduce an optimization problem that tests if the controllability parameter Φ_N corresponding to control horizon N satisfies $\Phi_N \leq \Phi_\alpha$ for the desired performance specified by α. Before we state the proposition, the following matrices are introduced

$$T = \text{blkdiag}(0, \ldots, 0, -Q, \Phi_\alpha R, 0, \ldots, 0, -R)$$
$$S = \text{blkdiag}(0, \ldots, 0, I, 0, \ldots, 0)$$

where Q and R are the cost matrices for states and inputs and Φ_α is the required controllability parameter for the chosen α. Recalling the partitioning (12) of $\mathbf{y}$ implies that

$$\mathbf{y}^T T \mathbf{y} = v_0^T R v_0 - z_{N-1}^T Q z_{N-1} - v_{N-1}^T R v_{N-1}$$
$$S\mathbf{y} = z_{N-1}$$

Proposition 2: Assume that $\Phi_\alpha > 0$ satisfies (41) for the chosen performance parameter α and optimality tolerance ϵ. Further assume that the control horizon N is such that

$$\begin{aligned} 0 = \min_{\bar{x}} \;& \frac{1}{2}\left(\Phi_\alpha \bar{x}^T Q \bar{x} + \mathbf{y}^T T \mathbf{y}\right) \\ \text{s.t. } & \bar{x} \in \mathbb{X}_N^0 \\ & \mathbf{y} = \arg\min V_N^0(\bar{x}) \end{aligned} \tag{42}$$

then $\Phi_N \leq \Phi_\alpha$.

Proof. First we note that $\bar{x} = 0$ gives $\mathbf{y} = 0$ and $\Phi_\alpha \bar{x}^T Q \bar{x} + \mathbf{y}^T T \mathbf{y} = 0$, i.e., we have that 0 is always a feasible solution. Further, (42) implies for every $\bar{x} \in \mathbb{X}_N^0$ that

$$0 \leq \Phi_\alpha \bar{x}^T Q \bar{x} + \mathbf{y}^T T \mathbf{y} = \Phi_\alpha \ell(\bar{x}, v_0^*) - \ell(z_{N-1}^*, v_{N-1}^*) = \Phi_\alpha \ell(\bar{x}, v_0^*) - \ell^*(z_{N-1}^*)$$

since $v_{N-1}^* = 0$. This is exactly the condition in Definition 1. Since Φ_N is the smallest such constant, we have $\Phi_N \leq \Phi_\alpha$ for the chosen control horizon N and desired performance α and optimality tolerance ϵ. □

The optimization problem (42) is a bilevel optimization problem with indefinite quadratic cost (see [33] for a survey on bilevel optimization). Such problems are in general NP-hard to solve. The problem can, however, be rewritten as an equivalent MILP as shown in the following proposition which is a straightforward application of [34, Theorem 2].

Proposition 3: Assume that Φ_α satisfies (41) for the chosen performance parameter α and optimality tolerance ϵ. If the control horizon N is such that the following holds

$$0 = \min \; -\frac{1}{2}\left(d_x^T \mu^{U1} + d_x^T \mu^{U2} + \mathbf{d}^T \mu^{UL1}\right) \tag{43}$$

$$\text{s.t. } \beta_i^L \in \{0,1\} \;,\; \beta_i^{U1} \in \{0,1\} \;,\; \beta_i^{U2} \in \{0,1\}$$

Upper level

Primal and dual feasibility

$$\begin{aligned} & C_x \bar{x} - d_x - s^x = 0 \\ & s^x \leq 0 \;,\; \mu^{U1} \geq 0 \\ & C_x A S \mathbf{y} - d_x - s^z = 0 \\ & s^z \leq 0 \;,\; \mu^{U2} \geq 0 \end{aligned}$$

Stationarity

$$\begin{aligned} & \Phi_\alpha Q \bar{x} + (C_x)^T \mu^{U1} - \mathbf{b}^T \lambda^{UL2} = 0 \\ & T\mathbf{y} + H^T \lambda^{UL1} + \mathbf{A}^T \lambda^{UL2} + \mathbf{C}^T \mu^{UL1} + (C_x A S)^T \mu^{U2} = 0 \\ & \mathbf{A} \lambda^{UL1} = 0 \\ & \mathbf{C} \lambda^{UL1} - \mu^{UL2} = 0 \end{aligned}$$

Complementarity

$$\begin{aligned} & \beta_i^L = 1 \Rightarrow \mu_i^{UL2} = 0 \;,\; \beta_i^L = 0 \Rightarrow \mu_i^{UL1} = 0 \\ & \beta_i^{U1} = 1 \Rightarrow s_i^x = 0 \;,\; \beta_i^{U1} = 0 \Rightarrow \mu_i^{U1} = 0 \\ & \beta_i^{U2} = 1 \Rightarrow s_i^z = 0 \;,\; \beta_i^{U2} = 0 \Rightarrow \mu_i^{U2} = 0 \end{aligned}$$

Lower level

Primal and dual feasibility

$$\begin{aligned} & \mathbf{A}\mathbf{y} - \mathbf{b}\bar{x} = 0 \\ & \mathbf{C}\mathbf{y} - \mathbf{d} - s = 0 \\ & s \leq 0 \;,\; \mu^L \geq 0 \end{aligned}$$

Stationarity

$$H\mathbf{y} + \mathbf{A}^T \lambda^L + \mathbf{C}^T \mu^L = 0$$

Complementarity

$$\beta_i^L = 1 \Rightarrow s_i = 0 \;,\; \beta_i^L = 0 \Rightarrow \mu_i^L = 0$$

then $\Phi_\alpha \geq \Phi_N$.

Proof. The set $\mathbb{X}_N^0$ can equivalently be written as

$$\mathbb{X}_N^0 = \{x \in \mathbb{R}^n \mid \mathbf{A}\mathbf{y}^*(x,0) = \mathbf{b}x, \mathbf{C}\mathbf{y}^*(x,0) \leq \mathbf{d}, C_x AS\mathbf{y}^*(x,0) \leq d_x, C_x x \leq d_x\}. \tag{44}$$

We express the set $\mathbb{X}_N^0$ in (42) using (44). The equivalence between the optimization problems (43) and (42) is established in [34, Theorem 2]. The remaining parts of the proposition follow by applying Proposition 2. □

The transformation from (42) to (43) is done by expressing the lower level optimization problem in (42) by its sufficient and necessary KKT conditions to get a single level indefinite quadratic program with complementarity constraints. The resulting indefinite quadratic program with complementarity constraints can in turn be cast as a MILP to get (43).

Remark 4: Although MILP problems are NP-hard, there are efficient solvers available such as CPLEX and GUROBI. There are also solvers available for solving the bilevel optimization problem (42) directly, e.g., the function *solvebilevel* in YALMIP, [35].

If the chosen control horizon N is not long enough for $\Phi_N \leq \Phi_\alpha$, different heuristics can be used to choose a new longer horizon to be verified. One heuristic is to assume exponential controllability as in Remark 1, i.e., that there exist constants $C \geq 1$ and $\sigma \in (0,1)$ such that

$$C\sigma^\tau \ell(\bar{x}, v_0^k) \geq \ell(z_\tau^k, v_\tau^k) \tag{45}$$

for all $\tau = 0, \ldots, N-1$. The C and σ-parameters should be determined using the optimal solution $\mathbf{y}$ to (13) for the x that minimized (43) in the previous test. Under the assumption that (45) holds as N increases, a new guess on the control horizon N can be computed by finding the smallest N such that $C\sigma^{N-1} \leq \Phi_\alpha$.

B. *Controllability parameter estimation*

The test in Proposition 3 verifies if the control horizon N is long enough for the controllability assumption to hold for the required controllability parameter Φ_α. Thus, an initial guess on the control horizon is needed. A guaranteed lower bound can easily be computed by solving (13) for a variety of initial conditions $\bar{x}$ and compute the worst controllability parameter, denoted by $\hat{\Phi}_N$, for these sample points. If the estimated controllability parameter $\hat{\Phi}_N \geq \Phi_\alpha$, we know that the control horizon need to be increased for (43) to hold. If instead $\hat{\Phi}_N \leq \Phi_\alpha$ the control horizon N might serve as a good initial guess to be verified by (43).

Remark 5: For large systems (43) may be too complex to verify the desired performance. In such cases the heuristic method mentioned above can be used in conjunction with an adaptive horizon scheme. The adaptive scheme keeps the horizon fixed for all time-steps until the controllability assumption does not hold. Then, the control horizon is increased to satisfy the assumption and kept at the new level until the controllability assumption does not hold again. Eventually the control horizon will be large enough for $\Phi_N \leq \Phi_\alpha$ and the horizon need not be increased again.

V. Numerical example

We evaluate the efficiency of the proposed distributed feedback control law ν_N by applying it to a randomly generated dynamical system with sparsity structure. The random dynamics matrix is scaled such that the magnitude of the largest eigenvalue is 1.1, i.e., the system is unstable. The system has 3 subsystems with 5 states and 1 input each, i.e., 15 states and 3 inputs in all. All state and input variables are upper and lower bounded by random numbers in the intervals $[0.5, 1.5]$ and $[-0.15 - 0.05]$ respectively. The stage cost is chosen to be

$$\ell_i(x_i, u_i) = x_i^T x_i + u_i^T u_i$$

for $i = 1, 2, 3$. We have chosen two different suboptimality parameters $\alpha_1 = 0.01$ and $\alpha_2 = 0.5$. We need to find control horizon $N(\alpha_i)$ such that the controllability parameter $\Phi_{N(\alpha_i)} \leq \Phi_{\alpha_i}$ for $i = 1, 2$. To compute Φ_{α_i} the optimality tolerance ϵ need to be chosen and κ need to be computed where κ is the smallest constant such that $\kappa Q \succeq A^T Q A$. We have chosen $\epsilon = 0.005$ and we have found that $\kappa = 1.22$. Using (41) we get $\Phi_{\alpha_1} = 0.51$ and $\Phi_{\alpha_2} = 0.22$. This implies that we need to find a control horizon $N(0.01)$ such that $\Phi_{N(0.01)} \leq 0.51$ and $N(0.5)$ such that $\Phi_{N(0.5)} \leq 0.22$. Verification by solving the MILP in (43) gives that $N(0.01) = 6$ and $N(0.5) = 9$.

The efficiency of the optimization algorithm (17)-(20) is investigated in [1]. The focus of this section is to evaluate the efficiency of the proposed adaptive constraint tightening approach in Algorithm 1. Further, we analyze the region of attraction for Algorithm 1, which is based on an optimization problem without terminal constraint. We compare the region of attraction to the region of attraction in standard MPC where a terminal constraint set is used. The terminal constraint set is computed as the maximal positive invariant set (see [36]) which in our example is a polytope defined by 288 linear inequality constraints.

TABLE I

EXPERIMENTAL RESULTS FOR DIFFERENT PERFORMANCE REQUIREMENTS α AND DIFFERENT INITIAL CONSTRAINT TIGHTENINGS $\delta_{\rm init}$ IN THE DMPC-CONTROLLER. ALSO, THE REGION OF ATTRACTION (R.O.A.) FOR THE DMPC-CONTROLLER IS COMPARED TO THE REGION OF ATTRACTION IN CENTRALIZED MPC WITH TERMINAL CONSTRAINT SET.

Algorithm comparison, $\alpha = 0.01$, $N = 6$					
	ϵ	$\delta_{\rm init}$	avg. # iters	avg. δ	R.o.A.
Alg. 1	0.005	0.0001	278.2	0.0001	82.4 %
Alg. 1	0.005	0.001	155.6	0.001	82.4 %
Alg. 1	0.005	0.01	66.6	0.01	82.4 %
Alg. 1	0.005	0.05	36.9	0.047	82.4 %
Alg. 1	0.005	0.1	35.6	0.056	82.4 %
Alg. 1	0.005	0.2	35.3	0.064	82.4 %
Alg. 1	0.005	0.5	35.3	0.080	82.4 %
CMPC	-	-	-	-	0.9 %

Algorithm comparison, $\alpha = 0.5$, $N = 9$					
	ϵ	$\delta_{\rm init}$	avg. # iters	avg. δ	R.o.A.
Alg. 1	0.005	0.0001	403.2	0.0001	92.2 %
Alg. 1	0.005	0.001	199.0	0.001	92.2 %
Alg. 1	0.005	0.01	82.5	0.01	92.2 %
Alg. 1	0.005	0.05	61.3	0.026	92.2 %
Alg. 1	0.005	0.1	60.6	0.030	92.2 %
Alg. 1	0.005	0.2	60.1	0.035	92.2 %
Alg. 1	0.005	0.5	59.8	0.042	92.2 %
CMPC	-	-	-	-	9.7 %

Table I presents the results obtained when the algorithm is running with different suboptimality parameters, $\alpha_1 = 0.01$ and $\alpha_2 = 0.5$. The first column specifies the stopping condition used, Alg. 1 refers to Algorithm 1 and CMPC refers to a centralized MPC-formulation with terminal constraints which is solved by a centralized solver. The second column specifies the tolerance ϵ and the third column specifies the initial constraint tightening $\delta_{\rm init}$.

Columns four, five and six contain the simulation results. The results are obtained by simulating the system with 10000 randomly chosen initial conditions that are drawn from a uniform distribution on $\mathcal{X}$. Column four contains the mean number of iterations needed and column five presents the average constraint tightening δ used at termination of Algorithm 1. The final column shows the fraction (in $\%$) of initial conditions that where steered to the origin using the different methods, i.e., an estimate of the region of attraction.

We see that the adaptive constraint tightening approach gives considerably less iterations for a larger initial tightening. However, for more than 10% initial constraint tightening ($\delta_{\text{init}} = 0.1$), the number of iterations is not significantly affected. It is remarkable to note that 50% initial constraint tightening ($\delta_{\text{init}} = 0.5$) is as efficient as, e.g., 5% ($\delta_{\text{init}} = 0.05$) considering that more reductions in the constraint tightening need to be performed. This indicates early detection of infeasibility. In the final column we have estimated the region of attraction, $\mathbb{X}_{\text{rf}}$. We see that, for the considered example, there is a huge improvement in the region of attraction using our method without terminal constraints compared to classical MPC (CMPC) with terminal constraints.

VI. Conclusions and future work

We have equipped the duality-based distributed optimization algorithm in [1], when used in a DMPC context, with a stopping condition that guarantees a prespecified performance, stability and feasibility. We have used an optimization problem without terminal constraints and have shown how to verify stability and a prespecified performance. Further, we have developed an adaptive constraint tightening approach that enables us to generate a feasible solution w.r.t. the original constraint set with finite number of iterations. The numerical example shows that the region of attraction can be significantly enlarged when no terminal constraint set is used compared to when using (the maximal positive invariant) terminal constraint set as in standard MPC. Further, the numerical example shows that the adaptive constraint tightening approach can significantly reduce the number of iterations needed to guarantee feasibility, stability, and the prespecified performance.

VII. Acknowledgements

The authors were supported by the Swedish Research Council through the Linnaeus center LCCC.

References

[1] P. Giselsson, M. D. Doan, T. Keviczky, B. De Schutter, and A. Rantzer, “Accelerated gradient methods and dual decomposition in distributed model predictive control,” *Automatica*, submitted.

[2] J. M. Maciejowski, *Predictive control with constraints*. Essex, England: Prentice Hall, 2002.

[3] J. Rawlings and D. Mayne, *Model Predictive Control: Theory and Design*. Nob Hill Publishing, 2009.

[4] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789 – 814, Jun. 2000.

[5] N. Sandell, P. Varaiya, M. Athans, and M. Safonov, "Survey of decentralized control methods for large scale systems," *IEEE Transactions on Automatic Control*, vol. 23, no. 2, pp. 108–128, Apr. 1978.

[6] A. Richards and J. How, "Robust distributed model predictive control," *International Journal of Control*, vol. 80, no. 9, pp. 1517–1531, Sep. 2007.

[7] W. Dunbar, "Distributed receding horizon control of dynamically coupled nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 52, no. 7, pp. 1249–1263, Jul. 2007.

[8] A. J. R.M. Hermans, M. Lazar, "Almost decentralized lyapunov-based nonlinear model predictive control," in *Proceedings of the 2010 American Control Conference*, Baltimore, July 2010, pp. 3932 – 3938.

[9] D. Jia and B. Krogh, "Distributed model predictive control," in *Proceedings of the 2001 American Control Conference*, Arlington, VA, Jun. 2001, pp. 2767–2772.

[10] E. Camponogara, D. Jia, B. Krogh, and S. Talukdar, "Distributed model predictive control," *IEEE Control Systems Magazine*, vol. 22, no. 1, pp. 44–52, Feb. 2002.

[11] A. N. Venkat, I. A. Hiskens, J. B. Rawlings, and S. J. Wright, "Distributed MPC strategies with application to power system automatic generation control," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 6, pp. 1192–1206, Nov. 2008.

[12] R. Negenborn, B. De Schutter, and J. Hellendoorn, "Multi-agent model predictive control for transportation networks: Serial versus parallel schemes," *Engineering Applications of Artificial Intelligence*, vol. 21, no. 3, pp. 353–366, Apr. 2008.

[13] Y. Wakasa, M. Arakawa, K. Tanaka, and T. Akashi, "Decentralized model predictive control via dual decomposition," in *Proceedings of the 47th IEEE Conference on Decision and Control*, Cancun, Mexico, Dec. 2008, pp. 381 – 386.

[14] D. Doan, T. Keviczky, I. Necoara, M. Diehl, and B. De Schutter, "A distributed version of Han's method for DMPC using local communications only," *Control Engineering and Applied Informatics*, vol. 11, no. 3, pp. 6–15, 2009.

[15] M. D. Doan, T. Keviczky, and B. De Schutter, "An improved distributed version of Han's method for distributed MPC of canal systems," in *12th symposium on Large Scale Systems: Theory and Applications*, Villeneuve d'Ascq, France, July 2010.

[16] I. Necoara and J. Suykens, "Application of a smoothing technique to decomposition in convex optimization," *IEEE Transactions on Automatic Control*, vol. 53, no. 11, pp. 2674 –2679, Dec. 2008.

[17] I. Necoara, D. Doan, and J. A. K. Suykens, "Application of the proximal center decomposition method to distributed model predictive control." in *Proceedings of the 47th IEEE Conference on Decision and Control*, Mexico, Dec. 2008, pp. 2900–2905.

[18] Y. Nesterov, "Smooth minimization of non-smooth functions," *Math. Program.*, vol. 103, no. 1, pp. 127–152, May 2005.

[19] J. M. Maestre, D. Muoz de la Pea, and E. F. Camacho, "Distributed model predictive control based on a cooperative game," *Optimal Control Applications and Methods*, vol. 32, no. 2, pp. 153–176, 2011.

[20] W. Dunbar and R. Murray, "Distributed receding horizon control for multi-vehicle formation stabilization," *Automatica*, vol. 42, no. 4, pp. 549–558, Apr. 2006.

[21] T. Keviczky, F. Borrelli, and G. Balas, "Decentralized receding horizon control for large scale dynamically decoupled systems," *Automatica*, vol. 42, no. 12, pp. 2105–2115, Dec. 2006.

[22] R. Scattolini, "Architectures for distributed and hierarchical model predictive control - A review," *Journal of Process Control*, vol. 19, no. 5, pp. 723 – 731, May 2009.

[23] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[24] G. Grimm, M. J. Messina, S. E. Tuna, and A. R. Teel, "Model predictive control: for want of a local control Lyapunov function, all is not lost," *Automatic Control, IEEE Transactions on*, vol. 50, no. 5, pp. 546–558, 2005.

[25] L. Grüne, "Analysis and design of unconstrained nonlinear MPC schemes for finite and infinite dimensional systems," *SIAM Journal on Control and Optimization*, vol. 48, no. 8, pp. 4938–4962, 2009.

[26] B. Lincoln and A. Rantzer, "Relaxing dynamic programming," *IEEE Transactions on Automatic Control*, vol. 51, no. 8, pp. 1249–1260, Aug. 2006.

[27] L. Grüne and A. Rantzer, "On the infinite horizon performance of receding horizon controllers," *IEEE Transactions on Automatic Control*, vol. 53, no. 9, pp. 2100–2111, Oct. 2008.

[28] L. Chisci, A. Lombardi, and E. Mosca, "Dual receding horizon control of constrained discrete-time systems," *European Journal of Control*, no. 2, pp. 278–285, 1996.

[29] P. O. M. Scokaert, D. Q. Mayne, and J. B. Rawlings, "Suboptimal model predictive control (feasibility implies stability)," *IEEE Transactions on Automatic Control*, vol. 44, no. 3, pp. 648–654, Mar. 1999.

[30] M. Diehl, R. Findeisen, F. Allgower, H. Bock, and J. Schloder, "Nominal stability of real-time iteration scheme for nonlinear model predictive control," *Control Theory and Applications, IEE Proceedings* -, vol. 152, no. 3, pp. 296 – 308, may 2005.

[31] L. Grüne and J. Pannek, "Analysis of unconstrained NMPC schemes with incomplete optimization," in *Proceedings of the 8th IFAC Symposium on Nonlinear Control Systems – NOLCOS 2010*, 2010, pp. 238–243.

[32] M. Doan, T. Keviczky, and B. De Schutter, "A dual decomposition-based optimization method with guaranteed primal feasibility for hierarchical MPC problems," in *Proceedings of the 18th IFAC World Congress*, Milan, Italy, Aug.–Sep. 2011, pp. 392–397.

[33] B. Colson, P. Marcotte, and G. Savard, "Bilevel programming: A survey," *4OR: A Quarterly Journal of Operations Research*, vol. 3, no. 2, pp. 87–107, 2005.

[34] C. Jones and M. Morari, "Approximate Explicit MPC using Bilevel Optimization," in *European Control Conference*, Budapest, Hungary, Aug. 2009.

[35] J. Löfberg, "Yalmip : A toolbox for modeling and optimization in MATLAB," in *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[36] E. Gilbert and K. Tan, "Linear systems with state and control constraints: the theory and application of maximal output admissible sets," *IEEE Transactions on Automatic Control*, vol. 36, no. 9, pp. 1008 –1020, sep 1991.

[37] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course (Applied Optimization)*, 1st ed. Springer Netherlands, 2003.

[38] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[39] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sciences*, vol. 2, no. 1, pp. 183–202, Oct 2009.

[40] A. Nedić and A. Ozdaglar, "Approximate primal solutions and rate analysis for dual subgradient methods," *SIAM J. on Optimization*, vol. 19, no. 4, pp. 1757–1780, Feb. 2009.

A. *Proof for Lemma 2*

We divide the proof into two parts, the first for $\bar{x} = 0$ and the second for $\bar{x} \neq 0$. For $\bar{x} = 0$ we have at iteration $k = 0$ that $\mathbf{y}^0 = 0$ which is the optimal solution. Hence (36) holds for $k = 0$ since all terms are 0 and $0 = A\xi_{N-1}^0 \in \mathcal{X}$.

Next, we show the result for $\bar{x} \neq 0$. Whenever (13) is feasible we have convergence in primal variables [1, Theorem 1]. This together with the linear relation through which ξ is defined (21) gives $\xi_\tau^k \to z_\tau^*$ for $\tau = 0, \ldots, N-1$ as $k \to \infty$. We have $z_\tau^* \in (1-\delta)\mathcal{X}$ and since $(1-\delta)\mathcal{X} \subset \mathcal{X}$ for every $\delta \in (0, 1]$ this implies that there exists finite k_0^x such that $\xi_\tau^k \in \mathcal{X}$ for all $k \geq k_0^x$. Equivalent convergence reasoning holds for v_τ^k. Together this implies that there exists finite k_0^P such that $P_N(\bar{x}, \mathbf{v}^k) < \infty$ and that $P_N(\bar{x}, \mathbf{v}^k) \to V_N^\delta(\bar{x})$ for all $k \geq k_0^P$. Together with convergence in dual function value [1, Theorem 1] gives that

$$D_N^\delta(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) \geq P_N(\bar{x}, \mathbf{v}^k) - \epsilon \ell^*(\bar{x})$$

holds with finite k since $\ell^*(\bar{x}) > 0$ and $\epsilon > 0$. This concludes the proof. □

B. *Proof for Lemma 3*

We introduce $\mathbf{y}^k = [(\boldsymbol{\xi}^k(\bar{x}, \delta))^T (\mathbf{v}^k(\bar{x}, \delta))^T]^T$, where $\boldsymbol{\xi}^k(\bar{x}, \delta)$ and $\mathbf{v}^k(\bar{x}, \delta)$ satisfies the dynamic equations (21). Whenever (36) holds we have that $\xi_\tau^k(\bar{x}, \delta) \in \mathcal{X}$ and $v_\tau^k(\bar{x}, \delta) \in \mathcal{U}$ for $\tau = 0, \ldots, N-1$. We also introduce $\mathbf{y}^* = [(\boldsymbol{z}^*(\bar{x}, 0))^T (\mathbf{v}^*(\bar{x}, 0))^T]^T$. This implies

$$\begin{aligned}
\frac{1}{2}(\mathbf{y}^k - \mathbf{y}^*)^T \mathbf{H} (\mathbf{y}^k - \mathbf{y}^*) &= \frac{1}{2}(\mathbf{y}^k)^T \mathbf{H} \mathbf{y}^k - \frac{1}{2}(\mathbf{y}^*)^T \mathbf{H} \mathbf{y}^* - \langle H\mathbf{y}^*, \mathbf{y}^k - \mathbf{y}^* \rangle \\
&\leq P_N(\bar{x}, \mathbf{v}^k) - V_N^0(\bar{x}) \leq D_N^\delta(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) + \epsilon \ell^*(\bar{x}) - V_N^0(\bar{x}) \\
&\leq \delta(\boldsymbol{\mu}^k)^T \mathbf{d} + \epsilon \ell^*(\bar{x})
\end{aligned}$$

where the first inequality comes from the first order optimality condition [37, Theorem 2.2.5] and by definition of V_N^0 and P_N. The second inequality is due to (36) and the last inequality follows from Lemma 1. Further, since $\mathbf{H} = \text{blkdiag}(Q, \ldots, Q, R, \ldots, R)$ we have for $\tau = 0, \ldots, N-1$ that

$$\frac{1}{2} \left\| \begin{bmatrix} \xi_\tau^k(\bar{x}, \delta) \\ v_\tau^k(\bar{x}, \delta) \end{bmatrix} - \begin{bmatrix} z_\tau^*(\bar{x}, 0) \\ v_\tau^*(\bar{x}, 0) \end{bmatrix} \right\|_H^2 \leq \frac{1}{2}(\mathbf{y}^k - \mathbf{y}^*)^T \mathbf{H} (\mathbf{y}^k - \mathbf{y}^*) \leq \delta(\boldsymbol{\mu}^k)^T \mathbf{d} + \epsilon \ell^*(\bar{x})$$

where $H = \text{blkdiag}(Q, R)$, whenever (36) holds. This completes the proof. □

C. Proof for Lemma 4

Since $x \in \mathbb{X}_N^0$ but $x \notin \mathbb{X}_N^\delta$ we have that $V_N^0(\bar{x}) < \infty$ and $V_N^\delta(\bar{x}) = \infty$. Further, from the strong theorem of alternatives [38, Section 5.8.2] we know that since $V_N^\delta(\bar{x}) = \infty$ for the current constraint tightening δ the dual problem is unbounded. Hence there exist $\boldsymbol{\lambda}_f$, $\boldsymbol{\mu}_f$ such that

$$\delta \boldsymbol{\mu}_f^T \mathbf{d} \geq D_N^\delta(\bar{x}, \boldsymbol{\lambda}_f, \boldsymbol{\mu}_f) - V_N^0(\bar{x}) \geq 2\epsilon\ell^*(\bar{x}) \tag{46}$$

where Lemma 1 is used in the first inequality. Further, the convergence rate in [39, Theorem 4.4] for algorithm (17)-(20) is

$$D_N^\delta(\bar{x}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) - D_N^\delta(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) \leq \frac{2L}{(k+1)^2} \left\| \begin{bmatrix} \boldsymbol{\lambda}^* \\ \boldsymbol{\mu}^* \end{bmatrix} - \begin{bmatrix} \boldsymbol{\lambda}^0 \\ \boldsymbol{\mu}^0 \end{bmatrix} \right\|^2.$$

By inspecting the proof to [39, Theorem 4.4] (and [39, Lemma 2.3, Lemma 4.1]) it is concluded that the optimal point $\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*$ can be changed to any feasible point $\boldsymbol{\lambda}_f, \boldsymbol{\mu}_f$ and the convergence result still holds, i.e.,

$$D_N^\delta(\bar{x}, \boldsymbol{\lambda}_f, \boldsymbol{\mu}_f) - D_N^\delta(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) \leq \frac{2L}{(k+1)^2} \left\| \begin{bmatrix} \boldsymbol{\lambda}_f \\ \boldsymbol{\mu}_f \end{bmatrix} - \begin{bmatrix} \boldsymbol{\lambda}^0 \\ \boldsymbol{\mu}^0 \end{bmatrix} \right\|^2.$$

That is, there exists a feasible pair $(\boldsymbol{\lambda}_f, \boldsymbol{\mu}_f)$ such that with finite k we have

$$D_N^\delta(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) > D_N^\delta(\bar{x}, \boldsymbol{\lambda}_f, \boldsymbol{\mu}_f) - \epsilon\ell^*(\bar{x}). \tag{47}$$

This implies

$$\delta \mathbf{d}^T \boldsymbol{\mu}^k \geq D_N^\delta(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) - V_N^0(\bar{x}) > D_N^\delta(\bar{x}, \boldsymbol{\lambda}_f, \boldsymbol{\mu}_f) - V_N^0(\bar{x}) - \epsilon\ell^*(\bar{x}) \geq \epsilon\ell^*(\bar{x})$$

where Lemma 1 is used in the first inequality, (47) in the second inequality and (46) in the final inequality. This completes the proof. □

D. Proof for Theorem 1

To prove the assertion we need to show that the do loop will exit for every $\bar{x} \in \text{int}(\mathbb{X}_N^0)$. For every point $\bar{x} \in \text{int}(\mathbb{X}_N^0)$ there exists $\bar{\delta} \in (0,1)$ such that $\frac{\bar{x}}{1-\delta} \in \text{int}(\mathbb{X}_N^0)$. Since $\text{int}(\mathbb{X}_N^0) \subseteq \mathbb{X}_N^0$, we have that $V_N^0(\frac{\bar{x}}{1-\delta}) < \infty$ and the optimal solution $\mathbf{y}(\frac{\bar{x}}{1-\delta}, 0)$ satisfies $\mathbf{A}\mathbf{y}^*(\frac{\bar{x}}{1-\delta}, 0) = \mathbf{b}\frac{\bar{x}}{1-\delta}$ and $\mathbf{C}\mathbf{y}^*(\frac{\bar{x}}{1-\delta}, 0) \leq \mathbf{d}$. We create the following vector

$$\bar{\mathbf{y}}(\bar{x}) := (1-\bar{\delta})\mathbf{y}^*(\frac{\bar{x}}{1-\bar{\delta}}, 0) \tag{48}$$

which satisfies

$$\mathbf{A}\bar{\mathbf{y}}(\bar{x}) = \mathbf{A}\mathbf{y}^*(\frac{\bar{x}}{1-\bar{\delta}}, 0)(1-\bar{\delta}) = \mathbf{b}\bar{x}\frac{1-\bar{\delta}}{1-\bar{\delta}} = \mathbf{b}\bar{x} \tag{49}$$

$$\mathbf{C}\bar{\mathbf{y}}(\bar{x}) = \mathbf{C}\mathbf{y}^*(\frac{\bar{x}}{1-\bar{\delta}}, 0)(1-\bar{\delta}) \leq \mathbf{d}(1-\bar{\delta}). \tag{50}$$

Hence, by definition (33) of $\mathbb{X}_N^{\delta}$ we conclude that for every $\bar{x} \in \text{int}(\mathbb{X}_N^0)$ there exist $\bar{\delta} \in (0,1)$ such that $\bar{x} \in \mathbb{X}_N^{\bar{\delta}}$. This implies that for every $\bar{x} \in \text{int}(\mathbb{X}_N^0)$ we have that either $\bar{x} \in \mathbb{X}_N^{\delta}$ for the current constraint tightening $\delta \in (0,1)$ or $\bar{x} \notin \mathbb{X}_N^{\delta}$ but $\bar{x} \in \mathbb{X}_N^0$. Thus, from Lemma 2 and Lemma 4 we conclude that either the do loop is terminated or δ is reduced and l is increased for every $\bar{x} \in \text{int}(\mathbb{X}_N^0)$ with finite number of algorithm iterations k.

To guarantee that the do loop will terminate for every $\bar{x} \in \text{int}(\mathbb{X}_N^0)$, we need to show that the conditions in the do loop will hold for small enough δ and with finite k. That is, we need to show that the following two conditions will hold.

1) For small enough δ, i.e., large enough l, we have that

$$\delta(\boldsymbol{\mu}^k)^T\mathbf{d} \leq \epsilon\ell^*(\bar{x}) \tag{51}$$

where $\delta = 2^{-l}\delta_{\text{init}}$ holds for every algorithm iteration k.

2) For small enough δ, i.e., large enough l, the condition

$$D_N^{\delta}(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) \geq P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k) + \alpha\ell(\bar{x}, v_0^k) \tag{52}$$

with α satisfying (37) holds with finite k whenever

$$D_N^{\delta}(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) \geq P_N(\bar{x}, \mathbf{v}^k) + \frac{\epsilon}{l+1}\ell(\bar{x}, v_0^k) \tag{53}$$

holds.

We start by showing argument 1. From the convergence rate of the algorithm [1] it follows that there exists $\underline{D} > -\infty$ such that $D_N^{\delta}(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) \geq \underline{D}$ for every algorithm iteration $k \geq 0$. This is used below where we extend the result from [40, Lemma 1] to handle the presence of equality constraints. For algorithm iteration $k \geq 0$, $\bar{x} \in \mathrm{int}(\mathbb{X}_N^0)$ and $\delta \leq \bar{\delta}/2$ we have

$$\begin{aligned}\underline{D} \leq D_N^\delta(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) &= \inf_{\mathbf{y}} \frac{1}{2}\mathbf{y}^T\mathbf{H}\mathbf{y} + (\boldsymbol{\lambda}^k)^T(\mathbf{A}\mathbf{y} - \mathbf{b}\bar{x}) + (\boldsymbol{\mu}^k)^T(\mathbf{C}\mathbf{y} - (1-\delta)\mathbf{d}) \\ &\leq \frac{1}{2}(\bar{\mathbf{y}}(\bar{x}))^T\mathbf{H}\bar{\mathbf{y}}(\bar{x}) + (\boldsymbol{\lambda}^k)^T(\mathbf{A}\bar{\mathbf{y}}(\bar{x}) - \mathbf{b}\bar{x}) + (\boldsymbol{\mu}^k)^T(\mathbf{C}\bar{\mathbf{y}}(\bar{x}) - (1-\delta)\mathbf{d}) \\ &\leq (1-\bar{\delta})^2 V_N^0(\frac{\bar{x}}{1-\bar{\delta}}) + (\boldsymbol{\mu}^k)^T(\mathbf{C}\bar{\mathbf{y}}(\bar{x}) - (1-\bar{\delta})\mathbf{d}) + (\boldsymbol{\mu}^k)^T\mathbf{d}(\delta - \bar{\delta}) \\ &\leq V_N^0(\frac{\bar{x}}{1-\bar{\delta}}) + (\boldsymbol{\mu}^k)^T\mathbf{d}(\delta - \bar{\delta}) \\ &\leq V_N^0(\frac{\bar{x}}{1-\bar{\delta}}) - \frac{1}{2}(\boldsymbol{\mu}^k)^T\mathbf{d}\bar{\delta}\end{aligned}$$

where the equality is by definition, the second inequality holds since any vector $\bar{\mathbf{y}}(\bar{x})$ is gives larger value than the infimum, the third and fourth inequalities are due to (48), (49) and (50) and since $(1-\bar{\delta}) \in (0,1)$ and the final inequality holds since $\delta \leq \bar{\delta}/2$. This implies that

$$(\boldsymbol{\mu}^k)^T\mathbf{d} \leq \frac{2(V_N^0(\frac{\bar{x}}{1-\bar{\delta}}) - \underline{D})}{\bar{\delta}}$$

which is finite. We denote by l_d the smallest l such that $\bar{\delta} \geq 2^{-l_d}\delta_{\text{init}}$. Since $\delta = 2^{-l}\delta_{\text{init}}$ this implies that

$$\delta(\boldsymbol{\mu}^k)^T\mathbf{d} \leq \delta\frac{2(V_N^0(\frac{\bar{x}}{1-\bar{\delta}}) - \underline{D})}{\bar{\delta}} \leq 2^{-l}\delta_{\text{init}}\frac{2(V_N^0(\frac{\bar{x}}{1-\bar{\delta}}) - \underline{D})}{2^{-l_d}\delta_{\text{init}}} \leq 2^{-l+l_d+1}(V_N^0(\frac{\bar{x}}{1-\bar{\delta}}) - \underline{D}) \to 0 \tag{54}$$

as $l \to \infty$. Especially, with finite l we have that (51) holds for every algorithm iteration k. This proves argument 1.

Next we prove argument 2. We start by showing for large enough but finite l that $P^N(A\bar{x} + B\nu_N(\bar{x}), \mathbf{v}_s^k)$ is finite whenever (53) holds. From the definition of P_N and $\mathbf{v}_s^k$ we have that $P^N(A\bar{x} + B\nu_N(\bar{x}), \mathbf{v}_s^k)$ is finite whenever $P_N(\bar{x}, \mathbf{v}_s^k)$ is finite and if $A\xi_{N-1}^k(\bar{x}, \delta) \in \mathcal{X}$. For algorithm iteration k such that (53) holds we have

$$\begin{aligned}\|A(\xi_{N-1}^k(\bar{x},\delta) - z_{N-1}^*(\bar{x},0))\|^2 &\leq \frac{\|A\|^2}{\lambda_{\min}(H)}\|\xi_{N-1}^k(\bar{x},\delta) - z_{N-1}^*(\bar{x},0)\|_H^2 \\ &\leq \frac{2\|A\|^2}{\lambda_{\min}(H)}(\delta(\boldsymbol{\mu}^k)^T\mathbf{d} + \frac{\epsilon}{l+1}\ell^*(\bar{x})) \\ &\leq \frac{2\|A\|^2}{\lambda_{\min}(H)}\left(2^{-l+l_d+1}(V_N^0(\frac{\bar{x}}{1-\bar{\delta}}) - \underline{D}) + \frac{\epsilon}{l+1}\ell^*(\bar{x})\right) \to 0\end{aligned} \tag{55}$$

as $l \to \infty$ where $H = \text{blkdiag}(Q, R)$ and the smallest eigenvalue $\lambda_{\min}(H) > 0$ since H is positive definite. The first inequality follows from Cauchy-Schwarz inequality and Courant-Fischer-Weyl min-max principle, the second inequality comes from Lemma 3 and the third comes from (54). By definition of $\mathbb{X}_N^\delta$ we have $Az_{N-1}^*(\bar{x}, 0) \in \text{int}(\mathcal{X})$ which through (55) implies that $A\xi_{N-1}^k(\bar{x}, \delta) \in \mathcal{X}$ for some large enough by finite l, i.e., small enough δ, and for algorithm iteration k such that (53) holds.

What is left to show is that (52) holds for every $\alpha \leq 1 - 2\epsilon - \kappa(\sqrt{2\epsilon} + \sqrt{\Phi_N})^2(\sqrt{2\epsilon} + 1)^2$ for large enough but finite l whenever (53) holds. From Lemma 3 and (54) we know for large enough l and any algorithm iteration k such that (53) holds that

$$\begin{aligned}\frac{1}{2}\left\| \begin{bmatrix} \xi_\tau^k \\ v_\tau^k \end{bmatrix} - \begin{bmatrix} z_\tau^* \\ v_\tau^* \end{bmatrix} \right\|_H^2 &\leq \delta(\boldsymbol{\mu}^k)^T \mathbf{d} + \frac{\epsilon}{l+1}\ell^*(\bar{x}) \\ &= 2^{-l}\delta_{\text{init}}(\boldsymbol{\mu}^k)^T \mathbf{d} + \frac{\epsilon}{l+1}\ell^*(\bar{x}) \leq 2\epsilon\ell^*(\bar{x})\end{aligned}$$

for any $\tau = 0, \ldots, N-1$, where $H = \text{blkdiag}(Q, R)$. Taking the square-root and applying the reversed triangle inequality gives

$$\left| \left\| \begin{bmatrix} \xi_\tau^k \\ v_\tau^k \end{bmatrix} \right\|_H - \left\| \begin{bmatrix} z_\tau^* \\ v_\tau^* \end{bmatrix} \right\|_H \right| \leq \left\| \begin{bmatrix} \xi_\tau^k \\ v_\tau^k \end{bmatrix} - \begin{bmatrix} z_\tau^* \\ v_\tau^* \end{bmatrix} \right\|_H \leq 2\sqrt{\epsilon\ell^*(\bar{x})}. \tag{56}$$

This implies that

$$\begin{aligned}\left\| \begin{bmatrix} \xi_{N-1}^k \\ v_{N-1}^k \end{bmatrix} \right\|_H &\leq \left\| \begin{bmatrix} z_{N-1}^* \\ v_{N-1}^* \end{bmatrix} \right\|_H + 2\sqrt{\epsilon\ell^*(\bar{x})} = \sqrt{2}\sqrt{\ell(z_{N-1}^*, v_{N-1}^*)} + 2\sqrt{\epsilon\ell^*(\bar{x})} \\ &\leq \sqrt{2\Phi_N}\sqrt{\ell(z_0^*, v_0^*)} + 2\sqrt{\epsilon\ell^*(\bar{x})} \leq (\sqrt{2\Phi_N} + 2\sqrt{\epsilon})\sqrt{\ell(z_0^*, v_0^*)} \\ &= (\sqrt{\Phi_N} + \sqrt{2\epsilon})\left\| \begin{bmatrix} z_0^* \\ v_0^* \end{bmatrix} \right\|_H \leq (\sqrt{\Phi_N} + \sqrt{2\epsilon})\left(\left\| \begin{bmatrix} \xi_0^k \\ v_0^k \end{bmatrix} \right\|_H + 2\sqrt{\epsilon\ell^*(\bar{x})}\right) \\ &\leq (\sqrt{\Phi_N} + \sqrt{2\epsilon})(1 + \sqrt{2\epsilon})\left\| \begin{bmatrix} \xi_0^k \\ v_0^k \end{bmatrix} \right\|_H\end{aligned}$$

where we have used (56), $z_0^* = \xi_0^k = \bar{x}$, $\|[z^T v^T]^T\|_H = \sqrt{z^T Q z + v^T R v} = \sqrt{2\ell(z, v)}$ and Definition 1. Squaring both sides gives through the definition of κ that

$$\frac{1}{\kappa}\ell^*(A\xi_{N-1}^k) \leq \ell^*(\xi_{N-1}^k) = \ell(\xi_{N-1}^k, v_{N-1}^k) \leq (\sqrt{\Phi_N} + \sqrt{2\epsilon})^2(1 + \sqrt{2\epsilon})^2\ell(\xi_0^k, v_0^k). \tag{57}$$

We get for large enough l and for k such that (53) holds that

$$\begin{aligned}
D_N^\delta(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) &\geq P_N(\bar{x}, \mathbf{v}^k) - \frac{\epsilon}{l+1}\ell^*(\bar{x}) \geq P_N(\bar{x}, \mathbf{v}^k) - \epsilon\ell^*(\bar{x}) \\
&= P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k) + (1-\epsilon)\ell(\xi_0^k, v_0^k) - \ell^*(A\xi_{N-1}^k) \\
&\geq P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k) + \left(1 - \epsilon - \kappa(\sqrt{\Phi_N} + \sqrt{2\epsilon})^2(1+\sqrt{2\epsilon})^2\right)\ell(\bar{x}, v_0^k) \\
&\geq P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k) + \alpha\ell(\bar{x}, v_0^k) \qquad (58)
\end{aligned}$$

where the first inequality comes from (53), the second since $l \geq 0$, the equality is due to (23), the third inequality comes from (57), and the final inequality comes from (37). This concludes the proof for argument 2. Thus, the do loop will terminate with finite l and k which implies that $\nu_N(\bar{x})$ is well defined for every $\bar{x} \in \mathrm{int}(\mathbb{X}_N^0)$.

Finally, to show (38) we have that

$$\begin{aligned}
V_N^0(\bar{x}) &\geq D_N^\delta(\bar{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) - \delta\mathbf{d}^T\boldsymbol{\mu}^k \\
&\geq P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k) - \epsilon\ell^*(\bar{x}) + \alpha\ell(\bar{x}, v_0^k) \\
&\geq V_N^0(A\bar{x} + Bv_0^k) + (\alpha - \epsilon)\ell(\bar{x}, v_0^k)
\end{aligned}$$

where the first inequality comes from Lemma 1, the second from (51) and (52), and the third holds since $P_N(A\bar{x} + Bv_0^k, \mathbf{v}_s^k) \geq V_N(A\bar{x} + Bv_0^k)$ and by definition of ℓ^*. This concludes the proof. □