An Efficient Algorithm for Tube-based Robust Nonlinear Optimal Control with Optimal Linear Feedback

Florian Messerer[1] and Moritz Diehl[1,2]

***Abstract*— We propose an algorithm for solving tube-based robust nonlinear optimal control problems based on the approximate propagation of ellipsoidal uncertainty tubes. Crucially, the algorithm does not only optimize the nominal control trajectory, but the decision variables include linear feedback gains for each time step. In consequence, the resulting trajectories do not suffer from the unrealistically large uncertainty sets of open-loop robust trajectories, but are able to approximately capture the feedback behavior implicit to model predictive control. The proposed algorithm iterates by alternatingly performing a Riccati recursion and solving a perturbed nominal optimal control problem. We provide a theoretical analysis of the local convergence behavior and demonstrate its basic applicability on the example problem of controlling a towing kite.**

I. INTRODUCTION

Model predictive control (MPC) is an advanced method for the control of dynamical systems [27]. At each time instance, the control input is obtained by solving an optimal control problem (OCP), in which a model of the system is used to plan a trajectory over a given time horizon. This especially allows for the inclusion of explicit state and control constraints into the plan. Due to model-plant mismatch and disturbances this plan will never be perfect, such that already after one time step the model prediction will be different from the real state of the system. MPC reacts to this deviation by resolving the OCP based on the new information. However, in standard, i.e., nominal, MPC this form of feedback is only implicit: no model of the uncertainty is taken into account. In terms of constraint satisfaction this can be devastating. If the originally planned trajectory was already at the edge of a constraint, a small perturbation will suffice to cause constraint violation. Keeping a heuristically chosen fixed safety distance from every constraint seems an easy solution, but the choice of distance is arbitrary while still not allowing for guarantees of constraint satisfaction.

Robust model predictive control (RMPC) takes into account the uncertainty explicitly [21], [27]. Under the assumption of bounded uncertainty sets, it plans trajectories that will be feasible for every possible disturbance. Two major approaches for this are scenario-tree based RMPC

[1] Department of Microsystems Engineering (IMTEK), University of Freiburg, 79110 Freiburg, Germany {florian.messerer, moritz.diehl}@imtek.uni-freiburg.de

[2] Department of Mathematics, University of Freiburg, 79104 Freiburg, Germany

This research was supported by the German Federal Ministry for Economic Affairs and Energy (BMWi) via DyConPV (0324166B) and by DFG via Research Unit FOR 2401 and project 424107692. The authors would like to thank Sergio Lucia and Marco Molnar for inspiring discussions.

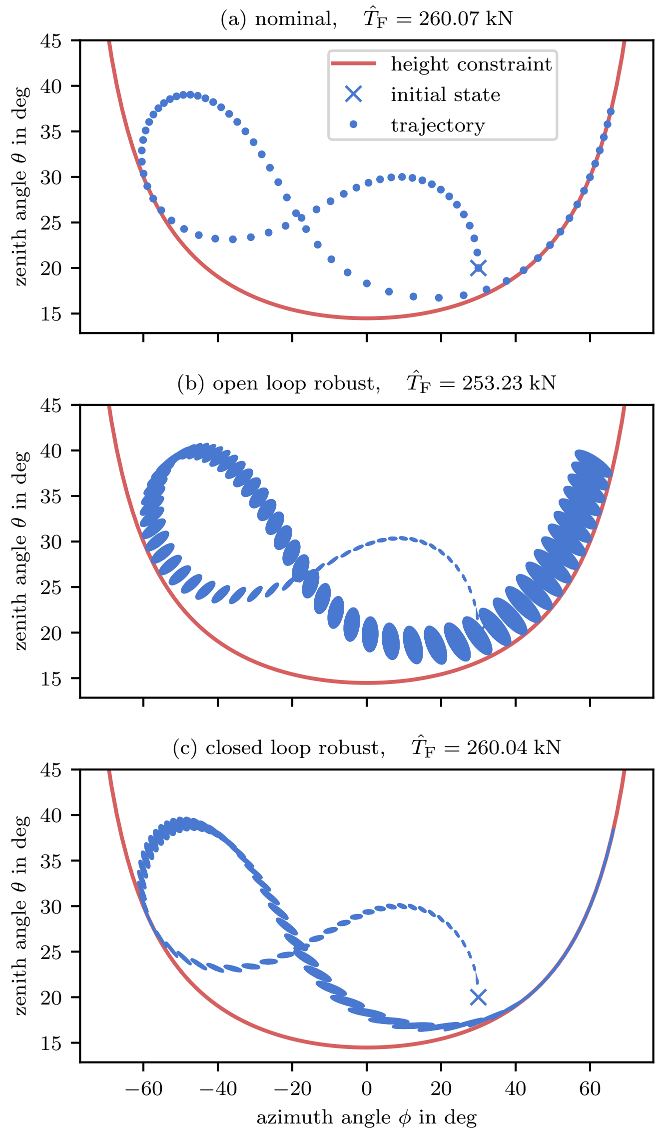

Fig. 1. Trajectories and resulting average thrust $\hat{T}_F$ for the towing kite problem detailed in Sec. V. The trajectories are obtained by solving (a) the nominal OCP, (b) the robustified OCP without including feedback in the prediction ($K = 0$), (c) the robustified OCP with optimal linear feedback at every time point. While the nominal OCP ignores the uncertainty, such that the slightest pertubation could lead to constraint violation, the open loop robust solution suffers from a growing uncertainty tube and needs to keep a large backoff. By taking into account future feedback, the closed loop robust OCP obtains a trajectory similar to the nominal one, with the important difference that a small backoff is kept to ensure constraint satisfaction.

[3], [6], [16] and tube-based RMPC [15], [18], [20], [25], [26]. Based on discrete uncertainty sets, scenario-tree based RMPC spans up a tree of every possible scenario, one for each realization of uncertainty. As the planned control trajectory will be different for every branch of the tree, this approach inherently considers the possibility of future feedback. However, the tree grows exponentially in the horizon length, such that issues of computational tractability arise already for moderate horizon lengths. Tube-based approaches on the other hand usually consider continuous uncertainty sets, and instead of a finite collection of trajectories they plan a tube: for each time point the tube defines a connected set in the state space in which the true state will be contained for every possible disturbance trajectory [4]. If only an open-loop control trajectory is planned, this tube will quickly grow. This can have a detrimental effect on controller performance, and even lead to infeasibility, as the trajectory will keep an unrealistically high distance from constraints. By incorporating the possibility of future feedback in the planned trajectory, this growth of uncertainty can be counteracted. Ideally, this feedback would correspond to the solution of the optimal control for each future time instance. However, this creates an infinite recursion, and is impractical even for a truncated recursion. Instead, it can be approximated by a linear feedback law [19]. One possibility is to use a constant feedback gain, precomputed, e.g., at a target steady state. The resulting optimal control problem is then still equivalent in structure to the robust open-loop problem, but with reduced growth of uncertainty. An alternative is to also optimize this feedback gain, by including an individual feedback gain for every time instant in the decision variables. It is well known that optimization over state feedback policies can significantly reduces the conservatism present in open-loop robust MPC [22], and was shown to even lead to convex optimization problems in the special case of constrained linear systems with additive disturbances [2], [10]. For a discussion of stability in this context we refer to [14]. The difference between solving a nominal OCP, an open loop robust OCP in which no feedback is considered, and a closed loop robust OCP with optimal linear feedback is illustrated in Fig. 1. However, due to the additional degrees of freedom, this problem is in general challenging to solve.

In this paper we propose an algorithm for the solution of robust nonlinear OCP based on ellipsoidal tubes that include linear feedback gains in the decision variables. The new algorithm exploits the specific problem structure to uncouple the nominal trajectory and the uncertainty part. This allows one to alternate between the solution of a perturbed nominal OCP and a feedback problem based on the current nominal trajectory. We show that the feedback problem can be efficiently solved via a Riccati recursion. The combination of these two ideas leads to an efficient algorithm for solving robust OCP of the given structure. Under standard assumptions on regularity, we prove that the proposed algorithm will converge to a solution of the original problem if initialized sufficiently close to the solution and if the uncertainty is not too large. Further, we illustrate the behavior of the algorithm for the example of controlling a towing kite. Similar ideas of uncoupling the uncertainty part have been proposed in [8], [31], but with the important difference that the feedback gains were not decision variables. Open-loop covariance predictions based on linearization along trajectories are also used in [12] to perform a simplified backoff computation, which can be interpreted as one iteration of the more recent zero-order robust MPC method from [31].

A. Notation and preliminaries

For vectors $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ we denote by $(x, y) := [x^\top, y^\top]^\top$ their vertical concatenation. For a matrix $A \in \mathbb{R}^{m\times n}$, the operation $\mathrm{vec}(A)$ means the vertical concatenation of its columns, such that a vector $z \in \mathbb{R}^{mn}$ is obtained. The square root of a vector $\sqrt{x}$ is to be understood elementwise, and $\mathrm{diag}(x)$ returns a diagonal matrix with x as its diagonal. The set of all positive semi-definite matrices of a given size is $\mathbb{S}^n_+$. Each matrix $Q \in \mathbb{S}^n_+$ defines an ellipsoid $\mathcal{E}(Q) := \{Q^{\frac{1}{2}}x \mid x \in \mathbb{R}^n, x^\top x \leq 1\}$, centered at the origin. For a vector valued function $f\colon \mathbb{R}^n \to \mathbb{R}^q$, we denote by $\nabla f(x)$ the gradient, which is defined as the transpose of the Jacobian $\frac{\partial f(x)}{\partial x}$. For a function with multiple arguments, $g\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^q$, we use $\nabla f(x, y)$ to denote the vertically concatenated gradient w.r.t. both arguments, or specify the argument as $\nabla_x f(x, y)$. For concepts from numerical optimization, such as the First Order Necessary Conditions (FONC), Linear Independence Constraint Qualification (LICQ), the Karush-Kuhn-Tucker (KKT) conditions, and Second Order Sufficient Condition (SOSC) we refer to [23]. For a nonlinear program (NLP) with decision variable $x \in \mathbb{R}^n$, we define the active set $\mathcal{A}(x)$ to denote the index set of all inequality constraints active at x, and correspondingly inactive set $\mathcal{I}(x)$ for the inactive ones.

II. APPROXIMATELY ROBUST OPTIMAL CONTROL WITH ELLIPSOIDAL TUBES

In the following, we describe the considered problem set-up in detail and define the robustified optimal control problem we want to solve. We consider uncertain discrete-time nonlinear dynamical systems of the form

$$x_0 = \bar{\bar{x}}_0, \quad x_{k+1} = f_k(x_k, u_k, w_k), \quad k = 0, \dots, N-1, \tag{1}$$

with state $x_k \in \mathbb{R}^{n_x}$, control $u_k \in \mathbb{R}^{n_u}$ and perturbations $w_k \in \mathbb{R}^{n_w}$. The perturbations $w = (w_0, \dots, w_{N-1})$ are drawn collectively from an ellipsoidal set, $w \in \mathcal{E}(\sigma^2 I)$, and scaled by the uncertainty parameter $\sigma \geq 0$. We note that choosing $\sigma^2 I$ as the ellipsoid matrix is more general than it seems at the first glance as transformations of w_k can always be considered to be part of the dynamics function $f_k(x_k, u_k, w_k)$.

The corresponding nominal optimal control problem, ignoring the uncertainty by setting $w_k = 0$ everywhere, takes the form

$$\min_{\bar{x},\bar{u}} \quad \sum_{k=0}^{N-1} l_k(\bar{x}_k,\bar{u}_k) + E(\bar{x}_N) \tag{2a}$$

$$\text{s.t.} \quad \bar{x}_0 = \bar{\bar{x}}_0, \tag{2b}$$

$$\bar{x}_{k+1} = f_k(\bar{x}_k,\bar{u}_k,0), \quad 0 \le k < N, \tag{2c}$$

$$0 \ge h_k(\bar{x}_k,\bar{u}_k), \quad 0 \le k < N, \tag{2d}$$

$$0 \ge h_N(\bar{x}_N), \tag{2e}$$

where $\bar{x} = (\bar{x}_0,\dots,\bar{x}_N)$, $\bar{u} = (\bar{u}_0,\dots,\bar{u}_{N-1})$ denotes the nominal trajectory, and $h_k\colon \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_{h_k}}$, $k = 0,\dots,N-1$, and $h_N\colon \mathbb{R}^{n_x} \to \mathbb{R}^{n_{h_N}}$ are the stage resp. terminal constraints.

To robustify this problem against perturbations w_k, we need to first describe and propagate the uncertainty tube. Following [9] and [13], we approximate it by ellipsoidal sets $\mathcal{E}(P_k)$ around the nominal trajectory. Crucially, instead of planning only open loop trajectories, we assume that at each $k = 1,\dots,N-1$, linear feedback with gain $K_k \in \mathbb{R}^{n_u \times n_x}$ is applied,

$$u_k = \bar{u}_k + K_k(x_k - \bar{x}_k), \tag{3}$$

reacting to deviations from the nominal trajectory. Starting at $P_0 = 0$, the state uncertainty ellipsoids are propagated as

$$\begin{aligned} P_{k+1} &= (A_k + B_k K_k) P_k (A_k + B_k K_k)^\top + \sigma^2 \Gamma_k \Gamma_k^\top \\ &=: \psi_k(\bar{x}_k,\bar{u}_k,P_k,K_k,\sigma), \end{aligned} \tag{4}$$

$k = 0,\dots,N-1$, based on the sensitivities of the dynamics evaluated at the nominal trajectory as given by

$$\begin{aligned} &A_k := \nabla_x f_k(\bar{x}_k,\bar{u}_k,0)^\top, \quad B_k := \nabla_u f_k(\bar{x}_k,\bar{u}_k,0)^\top, \\ &\Gamma_k := \nabla_w f_k(\bar{x}_k,\bar{u}_k,0)^\top, \quad k = 0,\dots,N-1. \end{aligned} \tag{5}$$

We point out that due to $P_0 = 0$ the value of K_0 is irrelevant. However, to avoid ambiguity while keeping notation simple, we set $K_0 = 0$ everywhere and remove it from the decision variables. Finally, approximating robust constraint satisfaction by local linearization of the inequality constraints, cf. [9], [13], we obtain the robustified optimal control problem

$$\min_{\substack{\bar{x},\bar{u},\\K,P,\beta}} \quad \sum_{k=0}^{N-1} l_k(\bar{x}_k,\bar{u}_k) + E(\bar{x}_N) \tag{6a}$$

$$\text{s.t.} \quad \bar{x}_0 = \bar{\bar{x}}_0, \tag{6b}$$

$$\bar{x}_{k+1} = f_k(\bar{x}_k,\bar{u}_k,0), \quad 0 \le k < N, \tag{6c}$$

$$P_0 = 0, \tag{6d}$$

$$P_{k+1} = \psi_k(\bar{x}_k,\bar{u}_k,P_k,K_k,\sigma), \quad 0 \le k < N, \tag{6e}$$

$$0 \ge h_k(\bar{x}_k,\bar{u}_k) + \sqrt{\beta_k + \epsilon}, \quad 0 \le k < N, \tag{6f}$$

$$0 \ge h_N(\bar{x}_N) + \sqrt{\beta_N + \epsilon}, \tag{6g}$$

$$\beta_k = H_k(\bar{x}_k,\bar{u}_k,P_k,K_k), \quad 0 \le k < N, \tag{6h}$$

$$\beta_N = H_N(\bar{x}_N,P_N), \tag{6i}$$

with $P = (P_0,\dots,P_N)$, $K = (K_1,\dots,K_{N-1})$, $K_0 = 0$, $\beta = (\beta_0,\dots,\beta_N)$, $\beta_k \in \mathbb{R}^{n_{h_k}}$, $k = 0,\dots,N$, and where, for $k = 0,\dots,N-1$ and $i = 1,\dots,n_{h_k}$, the functions

$$\begin{aligned} &H_k^i(\bar{x}_k,\bar{u}_k,P_k,K_k) = \\ &\qquad \nabla h_k^i(\bar{x}_k,\bar{u}_k)^\top \begin{bmatrix} I \\ K_k \end{bmatrix} P_k \begin{bmatrix} I \\ K_k \end{bmatrix}^\top \nabla h_k^i(\bar{x}_k,\bar{u}_k), \end{aligned} \tag{7}$$

$$H_N^i(\bar{x}_N,P_N) = \nabla h_N^i(\bar{x}_N)^\top P_N \nabla h_N^i(\bar{x}_N) \tag{8}$$

define the backoff that has to kept from every inequality constraint in dependence of the uncertainty in the relevant direction. Further, to ensure differentiability at all feasibile points, we add some small offset $\epsilon = \mathbf{1}\varepsilon$ under each square root, with $\varepsilon > 0$ and $\mathbf{1}$ a vector of ones, effectively resulting in a minimal backoff $\sqrt{\varepsilon}$ to be kept from each constraint. While it would be straight-forward to eliminate β from (6) it will be relevant for the proposed algorithm to keep the problem in the given form.

Remark 1. As the propagation of the uncertainty tube and the robustification of the inequality constraints are based on linearization, only approximate robustness can be claimed. Without the linearization error, i.e., for affine dynamics and constraints, one would obtain an exact and tight robustification with respect to the given assumptions on the uncertainty.

Remark 2. For clarity of exposition we restrict ourselves to a robust framework. However, the problem formulation (6) can also be given a stochastic interpretation. Here, under the assumption of Gaussian i.i.d. noise, the P_k correspond to covariance matrices with corresponding confidence ellipsoids, and the inequality constraints to (single) chance constraints, cf. [8], [12], again approximate within linearization.

III. SEQUENTIAL INEXACT ROBUST OPTIMIZATION (SIRO)

In this section, we derive the proposed algorithm and provide a theoretical analysis of its local convergence properties. To be able to describe and analyze the algorithm on a higher level of abstraction, we summarize the robustified problem (6) as

$$\min_{y,M,\beta} \quad f(y) \tag{9a}$$

$$\text{s.t.} \quad g(y) = 0, \tag{9b}$$

$$h(y) + \sqrt{\beta + \epsilon} \le 0, \tag{9c}$$

$$\sigma^2 H(y,M) - \beta = 0, \tag{9d}$$

where $y = (\bar{x},\bar{u}) \in \mathbb{R}^{n_y}$ contains the variables associated with the nominal trajectory, $g\colon \mathbb{R}^{n_y} \to \mathbb{R}^{n_g}$ are the nominal dynamics, and $h\colon \mathbb{R}^{n_y} \to \mathbb{R}^{n_h}$ the nominal part of the inequality constraints. The uncertainty part $\sigma^2 H(y,M)$, with $H\colon \mathbb{R}^{n_y} \times \mathbb{R}^{n_M} \to \mathbb{R}^{n_h}$, is obtained by eliminating P from (6d) and (6e) and stacking the right-hand sides of (6h) and (6i), such that $M = (\mathrm{vec}(K_1),\dots,\mathrm{vec}(K_{N-1})) \in \mathbb{R}^{n_M}$ is the remaining variable. We define the Lagrangian of (9) as

$$\begin{aligned} &\mathcal{L}(y,M,\beta,\lambda,\mu,\eta) = f(y) + \lambda^\top g(y) \\ &\quad + \mu^\top (h(y) + \sqrt{\beta + \epsilon}) + \eta^\top (\sigma^2 H(y,M) - \beta), \end{aligned} \tag{10}$$

with λ, μ, η the corresponding multipliers of appropriate dimension. Further, we define the shorthand

$$\hat{\mathcal{L}}(y, \lambda, \mu) = f(y) + \lambda^\top g(y) + \mu^\top h(y). \tag{11}$$

A. Algorithm

Consider the KKT conditions of (9), given by

$$\nabla_y \hat{\mathcal{L}}(y, \lambda, \mu) + \sigma^2 \nabla_y H(y, M)\eta = 0, \tag{12a}$$
$$\sigma^2 \nabla_M H(y, M)\eta = 0, \tag{12b}$$
$$\tfrac{1}{2}\operatorname{diag}(\beta + \epsilon)^{-\frac{1}{2}}\mu - \eta = 0, \tag{12c}$$
$$g(y) = 0, \tag{12d}$$
$$0 \leq \mu \perp h(y) + \sqrt{\beta + \epsilon} \leq 0, \tag{12e}$$
$$\sigma^2 H(y, M) - \beta = 0. \tag{12f}$$

We can interpret (12b) as the FONC of

$$\min_{M} \ \eta^\top H(y, M). \tag{13}$$

Under the assumption that (13) is strictly convex, its solution is the unique root of (12b) in M given y, η, and (12b) implicitly defines the function

$$m(y, \eta) = \arg\min_{M} \ \eta^\top H(y, M), \tag{14}$$

such that

$$\nabla_M H(y, m(y, \eta))\eta = 0. \tag{15}$$

Using this definition to eliminate M from (12), we obtain the reduced system

$$\nabla_y \hat{\mathcal{L}}(y, \lambda, \mu) + \sigma^2 \nabla_y H(y, m(y, \eta))\eta = 0, \tag{16a}$$
$$\tfrac{1}{2}\operatorname{diag}(\beta + \epsilon)^{-\frac{1}{2}}\mu - \eta = 0, \tag{16b}$$
$$g(y) = 0, \tag{16c}$$
$$0 \leq \mu \perp h(y) + \sqrt{\beta + \epsilon} \leq 0, \tag{16d}$$
$$\sigma^2 H(y, m(y, \eta)) - \beta = 0, \tag{16e}$$

which is equivalent to the KKT conditions (12). The core idea of the proposed algorithm is to freeze the uncertainty parts of (16) at some given $\bar{y}$, $\bar{\eta}$ and $\bar{M} = m(\bar{y}, \bar{\eta})$, obtaining

$$\nabla_y \hat{\mathcal{L}}(y, \lambda, \mu) + \sigma^2 \nabla_y H(\bar{y}, \bar{M})\bar{\eta} = 0, \tag{17a}$$
$$\tfrac{1}{2}\operatorname{diag}(\beta + \epsilon)^{-\frac{1}{2}}\mu - \eta = 0, \tag{17b}$$
$$g(y) = 0, \tag{17c}$$
$$0 \leq \mu \perp h(y) + \sqrt{\beta + \epsilon} \leq 0, \tag{17d}$$
$$\sigma^2 H(\bar{y}, \bar{M}) - \beta = 0. \tag{17e}$$

Defining $\bar{c} = \nabla_y H(\bar{y}, \bar{M})\bar{\eta}$ and $\bar{H} = H(\bar{y}, \bar{M})$, we can interpret (17) as the KKT conditions of the perturbed nominal problem

$$\mathbb{P}(\sigma, \bar{z}): \quad \begin{aligned} \min_{y, \beta} \quad & f(y) + \sigma^2 \bar{c}^\top y && \text{(18a)} \\ \text{s.t.} \quad & g(y) = 0, && \text{(18b)} \\ & h(y) + \sqrt{\beta + \epsilon} \leq 0, && \text{(18c)} \\ & \sigma^2 \bar{H} - \beta = 0, && \text{(18d)} \end{aligned}$$

where $\bar{z} = (\bar{y}, \bar{\beta}, \bar{\lambda}, \bar{\mu}, \bar{\eta})$. While it would be straight-forward to eliminate β in (18), we abstain from doing so as this will simplify the notation of the analysis. The algorithm then iterates by sequentially solving this perturbed nominal problem and updating the perturbations at the obtained solution. As the sensitivities of the frozen parts are neglected, the algorithm can be interpreted as an inexact Newton method for the solution of (16).

B. Local Convergence Analysis

Denote by $z = (y, \beta, \lambda, \mu, \eta)$ the primal dual variables of (18). Then, for a given σ and current iterate z_k, the next iterate is obtained as a KKT point of $\mathbb{P}(\sigma, z_k)$. Since (18) is in general not convex, this point is not necessarily unique. To obtain a well-defined iteration operator

$$z_{k+1} = z^{\text{sol}}(z_k), \tag{19}$$

we assume that necessary precautions are taken, e.g., by always picking the solution which has the least distance to z_k in some norm. In a slight abuse of notation we define $m(z) := m(y, \eta)$. We start by establishing a relationship between stationary points of (19) and KKT points of the original robust problem (9), which we want to solve.

Assumption 3. Assume that $H(y, M) \geq 0$ for all $y \in \mathbb{R}^{n_y}$, $M \in \mathbb{R}^{n_M}$, and that (13) is strictly convex for all $\eta \in \mathbb{R}^{n_h}_+$ that correspond to a KKT point of $\mathbb{P}(\sigma, \bar{z})$ for $\sigma > 0$ and arbitrary $\bar{z}$.

Lemma 4. *Let Ass. 3 hold, z_* be a fixed point of* (19) *with associated $M_* = m(z_*)$, and $\sigma \geq 0$. Then (z_*, M_*) is a KKT point of* (9)*. On the other hand, let (z_*, M_*) be a KKT point of* (9)*, and $\sigma > 0$. Then z_* is a stationary point of* (19) *and $M_* = m(z_*)$.*

Proof. For the first statement, it follows from the premise that the KKT conditions (17) of the perturbed nominal problem (18) hold at $z = \bar{z} = z_*$, and (12b) holds by construction of M_*. It follows that the KKT conditions (12) of the original problem (9) hold at $z = z_*$. For the second statement, it follows from the premise that the KKT conditions (12) of (9) hold at z_*. We point out that for $\sigma > 0$, M_* is uniquely defined from (12b) given z_*, due to Ass. 3. As this corresponds to the definition of $m(z)$, we have that $M_* = m(z_*)$. It then follows that the KKT conditions (17) of (18) hold at $z = \bar{z} = z_*$, such that z_* is a stationary point of (19). □

Before continuing the convergence analysis for fixed σ, we want to develop an understanding of regularity of the perturbed nominal problem $\mathbb{P}(\sigma, \bar{z})$ as given in (18). It turns out that this is strongly related to regularity of the nominal problem $\mathbb{P}(0, \bar{z})$, which for $\sigma = 0$ is actually independent of $\bar{z}$. As a side node we point out that due to the squaring $\mathbb{P}(\sigma, \bar{z})$ is symmetric in σ, in the sense that an identical problem is obtained if one uses $-\sigma$ instead of a given σ. Therefore the problem is well-defined also for negative σ, even if they are not meaningful in the context of uncertainty. We can then state the following lemma, which is due to a more general result by Robinson [28, Thrm 2.1].

Lemma 5. *Let $\hat{z}_*$ be a KKT point of $\mathbb{P}(0, \hat{z}_*)$, at which LICQ, strict complementarity and SOSC hold. Then there exist open neighborhoods $\mathcal{N}_z$ and $\mathcal{N}_z'$ of $\hat{z}_*$, $\mathcal{N}_\sigma$ of 0, as well as a continuously defined function $Z\colon \mathcal{N}_\sigma \times \mathcal{N}_z' \to \mathcal{N}_z$ such that $\hat{z}_* = Z(0, \hat{z}_*)$. Further, for $\sigma \in \mathcal{N}_\sigma$ and $\bar{z} \in \mathcal{N}_z'$, the perturbed problem $\mathbb{P}(\sigma, \bar{z})$ has the locally unique solution $Z(\sigma, \bar{z}) \in \mathcal{N}_z$, at which LICQ, strict complementarity and SOSC hold.*

We now consider $\mathbb{P}(\sigma, \bar{z})$ for a fixed $\sigma > 0$. To establish contraction of the iteration map (19), we will need the following assumptions.

Assumption 6. Assume (z_*, M_*) is a local minimizer of the robust problem (9) for a given $\sigma > 0$.

Assumption 7. Assume that LICQ, strict complementarity and SOSC of $\mathbb{P}(\sigma, z_*)$ hold at z_*.

Remark 8. To be able to consider only a fixed $\sigma > 0$ in the statement and proof of the convergence theorem, we require Ass. 7 directly. However, Lemma 5 tells us that Ass. 7 follows from regularity of the nominal problem for any σ that is sufficiently small.

Before stating the main theorem, we need the following lemma, which is a standard property of problems with strict complementarity [28].

Lemma 9. *Let Ass. 6 and 7 hold for z_*. Then, for $\bar{z}$ sufficiently close to z_*, the active set of $\mathbb{P}(\sigma, \bar{z})$ is stable and corresponds to $\mathcal{A}_* = \mathcal{A}(z_*)$, the active set of $P(\sigma, z_*)$. Analogously, the same holds for the inactive set $\mathcal{I}_* = \mathcal{I}(z_*)$.*

Theorem 10. *Let Ass. 3, 6, and 7 hold for z_*, M_*. Then z_* is a fixed point of* (19) *and the iteration sequence $\{z_k\}_{k=0}^\infty$ defined by* (19) *converges q-linearly to z_*, if initialized sufficiently close to z_* and if σ is sufficiently small. Further, the corresponding sequence of $M_k = m(z_k)$ converges r-linearly to M_*.*

Proof. That z_* is fixed point under these assumptions is the content of Lemma 4, which also means that z_* is a solution to $\mathbb{P}(\sigma, z_*)$. A consequence of Lemma 9 is that for local convergence analysis we can disregard the inactive constraints, and regard the active inequality constraints as equality constraints. For the inactive multipliers we know that $\mu_{\mathcal{I}_*} = 0$, which due to (17b) implies $\eta_{\mathcal{I}_*} = 0$, and we collect the remaining variables in $\zeta = (y, \beta, \lambda, \mu_{\mathcal{A}_*}, \eta_{\mathcal{A}_*})$, with $\bar{\zeta}$, ζ_*, ζ_k defined correspondingly. The locally relevant information of the KKT conditions (17) can be summarized in the residual map

$$r(\zeta, \bar{\zeta}) = \begin{bmatrix} \nabla \hat{\mathcal{L}}(y, \lambda, \mu) + \sigma^2 \nabla_y H(\bar{y}, m(\bar{z}))\bar{\eta} \\ [\frac{1}{2}\operatorname{diag}(\beta + \epsilon)^{-\frac{1}{2}}\mu - \eta]_{\mathcal{A}_*} \\ g(y) \\ [\operatorname{diag}(\mu)(h(y) + \sqrt{\beta + \epsilon})]_{\mathcal{A}_*} \\ \sigma^2 H(\bar{y}, m(\bar{z})) - \beta \end{bmatrix}, \tag{20}$$

where the relevant equations are selected line-wise, as denoted by $[\,\cdot\,]_{\mathcal{A}_*}$. The components of μ, $\bar{z}$, $\bar{\eta}$ corresponding to inactive constraints are understood the be given by 0. Locally, the algorithm iterates by finding ζ_{k+1} such that $r(\zeta_{k+1}, \zeta_k) = 0$. We have that $r(\zeta_*, \zeta_*) = 0$ and the residual map is differentiable in both arguments and and we denote the derivatives at z_* by

$$R_* = \left.\frac{\partial r(\zeta, \bar{\zeta})}{\partial \zeta}\right|_{\zeta = \bar{\zeta} = \zeta_*}, \; E_* = \left.\frac{\partial r(\zeta, \bar{\zeta})}{\partial \bar{\zeta}}\right|_{\zeta = \bar{\zeta} = \zeta_*}. \tag{21}$$

R_* in fact corresponds to the KKT matrix of $\mathbb{P}(\sigma, z_*)$ at z_* and due to LICQ and SOSC is invertible [23, Lem. 16.1]. Since every term in $r(\zeta, \bar{\zeta})$ depending on $\bar{\zeta}$ is multiplied by σ^2, there exists E_*', which does not depend on σ, such that $E_* = \sigma^2 E_*'$. Then we know from the implicit function theorem that there exist open neighborhoods $\mathcal{N}_\zeta$ and $\mathcal{N}_\zeta'$ of ζ_* as well as a continously differentiable function $\zeta^{\mathrm{sol}}\colon \mathcal{N}_\zeta' \to \mathcal{N}_\zeta$, such that $\zeta^{\mathrm{sol}}(\zeta_*) = \zeta_*$. For $\bar{\zeta} \in \mathcal{N}_\zeta'$ a locally unique solution of $r(\zeta, \bar{\zeta}) = 0$ is given by $\zeta^{\mathrm{sol}}(\bar{\zeta})$. Further,

$$\left.\frac{\partial \zeta^{\mathrm{sol}}(\bar{\zeta})}{\partial \bar{\zeta}}\right|_{\bar{\zeta} = \zeta_*} = -R_*^{-1} E_* = -\sigma^2 R_*^{-1} E_*' =: Z_*. \tag{22}$$

Locally, the method iterates as $\zeta_{k+1} = \zeta^{\mathrm{sol}}(\zeta_k)$, which corresponds to the ζ-part of $z^{\mathrm{sol}}(z_k)$ as defined in (19). A first order taylor approximation of $\zeta^{\mathrm{sol}}(\zeta_k)$ at ζ_* yields

$$\zeta_{k+1} - \zeta_* = Z_*(\zeta_k - \zeta_*) + \mathcal{O}(\|\zeta_k - \zeta_*\|^2). \tag{23}$$

Convergence of $\{\zeta_k\}_{k=0}^\infty$ to ζ_* for ζ_0 sufficiently close to ζ_* is then determined by the spectral radius $\rho(Z_*)$ [24]. If $\rho(Z_*) = \sigma^2 \rho(R_*^{-1} E_*') < 1$, the iterates will converge q-linearly to ζ_* with asymptotic rate $\rho(Z_*)$. This condition clearly holds for σ sufficiently small. Due to stability of the active set, this result directly transfers to the iterates in z, and the sequence $\{z_k\}_{k=0}^\infty$ as defined by (19) converges q-linearly to z_*, if initialized sufficiently close to z_*.

Now consider the corresponding $M_k = m(z_k)$. Due to strict convexity of (13) we have that $m(z)$ is continuously differentiable for all z. This means for each bounded neighborhood $\mathcal{N}_z''$ of z_*, there exists $\ell \geq 0$ such that $\|m(z) - m(z_*)\| \leq \ell \|z - z_*\|$ for $z \in \mathcal{N}_z''$. Now choose $\mathcal{N}_z''$ such that it includes all z_k, $k \geq 0$, which is always possible under the given assumptions, as we have already established convergence of $\{z_k\}_{k=0}^\infty$. Then it holds that $\|M_k - M_*\| = \|m(z_k) - m(z_*)\| \leq \ell \|z_k - z_*\|$, i.e., $\{\|M_k - M_*\|\}_{k=0}^\infty$ is upper bounded by a q-linearly converging sequence, which corresponds to r-linear convergence. □

IV. A RICCATI RECURSION FOR EFFICIENT SIRO

In the following, we describe the application of the proposed algorithm to the robustified optimal control problem (6). This will result in alternatingly solving a perturbed nominal optimal control problem and a performing Riccati recursion, which allows us to efficiently solve the uncertainty minimization problem (13).

We start by specifying the optimization problem (13) associated with the uncertainty part. For (6), after a few manipulations, the corresponding problem can be stated as

$$\min_{K,P} \quad \sum_{k=0}^{N-1} \operatorname{Tr}\left(C_k \begin{bmatrix} I \\ K_k \end{bmatrix} P_k \begin{bmatrix} I \\ K_k \end{bmatrix}^\top\right) + \operatorname{Tr}(C_N P_N) \tag{24a}$$

$$\text{s.t.} \qquad P_0 = 0, \tag{24b}$$

$$P_{k+1} = \psi_k(\bar{x}_k, \bar{u}_k, P_k, K_k, \sigma), \quad 0 \le k < N, \tag{24c}$$

where $K = (K_1, \ldots, K_{N-1})$, $K_0 = 0$. The cost matrices are given by

$$C_k = \nabla h_k(\bar{x}_k, \bar{u}_k) \operatorname{diag}(\eta_k) \nabla h_k(\bar{x}_k, \bar{u}_k)^\top, \tag{25}$$

$$C_N = \nabla h_N(\bar{x}_N) \operatorname{diag}(\eta_N) \nabla h_N(\bar{x}_N)^\top, \tag{26}$$

where η_k, $k = 0, \ldots, N$, are the Lagrange multipliers associated with (6h) resp. (6i). The exact form of (13) can be obtained by eliminating P in (24). We further split the stage cost matrices into block components,

$$C_k = \begin{bmatrix} C_k^x & C_k^{xu} \\ C_k^{ux} & C_k^u \end{bmatrix}, \tag{27}$$

such that $C_k^x \in \mathbb{S}_+^{n_x}$, $C_k^u \in \mathbb{S}_+^{n_u}$ and $C_k^{ux} = C_k^{xu\top}$ of corresponding dimension, for $k = 0, \ldots, N-1$. It turns out that (24) can be analytically solved via a Riccati recursion.

Lemma 11. *Let $C_N \succeq 0$, $C_k \succeq 0$, $C_k^x \succeq 0$, $C_k^u \succ 0$, for $k = 0, \ldots, N-1$. Further, let $P_k \succ 0$ for all feasible P_k, $k = 1, \ldots, N$. Then the solution to* (24) *is uniquely defined by the Riccati recursion*

$$S_N = C_N, \tag{28a}$$

$$K_k^* = -(C_k^u + B_k^\top S_{k+1} B_k)^{-1} (C_k^{ux} + B_k^\top S_{k+1} A_k), \tag{28b}$$

$$S_k = C_k^x + A_k^\top S_{k+1} A_k + (C_k^{xu} + A_k^\top S_{k+1} B_k) K_k^*, \tag{28c}$$

with $k = N-1, \ldots, 1$, followed by a linear Lyapunov matrix forward simulation

$$P_0^* = 0, \qquad P_{k+1}^* = \psi_k(\bar{x}_k, \bar{u}_k, P_k^*, K_k^*, \sigma), \tag{29}$$

for $k = 0, \ldots, N-1$, where $K_0^* = K_0 = 0$.

Proof. The proof goes by showing structural equivalence to the associated discrete time finite-horizon stochastic linear quadratic regulator (sLQR) problem, for which the solution is known. This switch to a stochastic frame work might be surprising, but we only make an argument from the algebraic structure of the problem, not from its interpretation. The sLQR problem is given by

$$\min_{\pi(\cdot)} \quad \mathop{\mathbb{E}}_{\omega}\Big\{ \sum_{k=0}^{N-1} \begin{bmatrix} \Delta x_k^\pi(\omega) \\ \pi_k(\Delta x_k^\pi(\omega)) \end{bmatrix}^\top C_k \begin{bmatrix} \Delta x_k^\pi(\omega) \\ \pi_k(\Delta x_k^\pi(\omega)) \end{bmatrix} + \Delta x_N^\pi(\omega)^\top C_N \Delta x_N^\pi(\omega) \Big\} \tag{30}$$

where, $\Delta x_0^\pi(\omega) = 0$ and, for $k = 0, \ldots, N-1$,

$$\Delta x_{k+1}^\pi(\omega) = A_k \Delta x_k^\pi(\omega) + B_k \pi(\Delta x_k^\pi(\omega)) + \Gamma_k \omega_k, \tag{31}$$

and $\omega = (\omega_0, \ldots, \omega_{N-1}) \sim \mathcal{N}(0, \sigma^2 I)$ is normally distributed noise. Here, $\pi(\cdot) = (\pi_1(\cdot), \ldots, \pi_{N-1}(\cdot))$ denotes a control policy, which gives the control input as a function of the current state, $\Delta u_k = \pi_k(\Delta x_k)$. As a standard result from linear control theory [29], the unique optimal policy for the given problem is linear feedback, $\pi_k(\Delta x) = K_k^* \Delta x$, with control gains K_k^* defined by the Riccati recursion (28). Incorporating this knowledge, we can rephrase (30) as

$$\min_{K} \quad \mathop{\mathbb{E}}_{\omega}\Big\{ \sum_{k=0}^{N-1} \begin{bmatrix} \Delta x_k^K(\omega) \\ K_k \Delta x_k^K(\omega)) \end{bmatrix}^\top C_k \begin{bmatrix} \Delta x_k^K(\omega) \\ K_k \Delta x_k^K(\omega) \end{bmatrix} + \Delta x_N^K(\omega)^\top C_N \Delta x_N^K(\omega) \Big\} \tag{32}$$

where $\Delta x_0^K(\omega) = 0$ and, for $k = 0, \ldots, N-1$,

$$\Delta x_{k+1}^K(\omega) = (A_k + B_k K_k) \Delta x_k^K(\omega) + \Gamma_k \omega_k, \tag{33}$$

which has the same solution as (30). Noting that

$$\begin{bmatrix} x_k \\ K_k x_k \end{bmatrix}^\top C_k \begin{bmatrix} x_k \\ K_k x_k \end{bmatrix} = \operatorname{Tr}\left(C_k \begin{bmatrix} I \\ K_k \end{bmatrix} x_k x_k^\top \begin{bmatrix} I \\ K_k \end{bmatrix}^\top\right),$$

we can transcribe (32) to

$$\min_{K} \quad \mathop{\mathbb{E}}_{\omega}\Big\{ \sum_{k=0}^{N-1} \operatorname{Tr}\left(C_k \begin{bmatrix} I \\ K_k \end{bmatrix} \Delta x_k^K(\omega) \Delta x_k^K(\omega)^\top \begin{bmatrix} I \\ K_k \end{bmatrix}^\top\right) + \operatorname{Tr}\left(C_N \Delta x_N^K(\omega) \Delta x_N^K(\omega)^\top\right) \Big\}. \tag{34}$$

Introducing $P_0 = 0$ and, for $k = 0, \ldots, N-1$,

$$P_{k+1} = (A_k + B_k K_k) P_k (A_k + B_k K_k)^\top + \sigma^2 \Gamma_k \Gamma_k^\top, \tag{35}$$

it turns out that

$$\mathop{\mathbb{E}}_{\omega}\{\Delta x_k^K(\omega) \Delta x_k^K(\omega)^\top\} = P_k \tag{36}$$

for $k = 0, \ldots, N$, i.e., P_k are the covariance matrices of $\Delta x_k^K(\omega)$, defining ellipsoidal confidence levels. Substituting this in (34) and interpreting the confidence ellipsoids as robust ellipsoids instead – which is possible as we are only making an argument from the algebraic structure – the resulting problem is the same as (24), which concludes the proof. □

After deriving the solution of the uncertainty minimization problem (13) associated with the robustified optimal control problem, the corresponding perturbed nominal problem (18) alongside the perturbations remains to be be specified. Denote by $\bar{K}$, $\bar{P}$ the solution to (24) for a given nominal trajectory $\bar{x}$, $\bar{u}$, $\bar{\eta}$. The backoffs can then be obtained as

$$\bar{b}_k = \sqrt{H_k(\bar{x}_k, \bar{u}_k, \bar{P}_k, \bar{K}_k) + \epsilon}, \; k = 0, \ldots, N-1, \tag{37a}$$

$$\bar{b}_N = \sqrt{H_N(\bar{x}_k, \bar{P}_k) + \epsilon}. \tag{37b}$$

The gradient correction, using the notation of the general formulation and $\bar{M}$ corresponding to $\bar{K}$, is $\bar{c} = \nabla_y H(\bar{y}, \bar{M})\bar{\eta}$, and the $\bar{c}_0, \ldots, \bar{c}_N$ can be obtained by splitting $\bar{c}$ into its corresponding components. Thus, the algorithm (28) to (29) for obtaining $\bar{K}$ and $\bar{P}$ needs to be differentiated with

Algorithm 1 SIRO for tube-based robust optimal control

Input: Initial guess $\bar{x}$, $\bar{u}$, η, λ, μ
(e.g. solution to nominal OCP)
repeat
  $K \leftarrow$ `riccatiRecursion`$(\bar{x}, \bar{u}, \eta)$ // (28)
  $P \leftarrow$ `lyapunovForward`$(\bar{x}, \bar{u}, K)$ // (29)
  if `isStationary`$(\bar{x}, \bar{u}, K, P, \eta, \lambda, \mu)$ // (12a)
    break
  end if
  $\bar{b}, \bar{c} \leftarrow$ `getPerturbation`$(\bar{x}, \bar{u}, \eta, P, K)$ // (37)
  $\bar{x}, \bar{u}, \lambda, \mu \leftarrow$ `solvePerturbedOCP`$(\bar{c}, \bar{b})$ // (38)
  $\eta \leftarrow \frac{1}{2}\,\mathrm{diag}(\bar{b})^{-1}\mu$ // (17b)
return: $\bar{x}, \bar{u}, P, K$

respect to its input arguments $\bar{x}$ and $\bar{u}$, cf. (24) to (26). However, note that to obtain $\bar{c}^\top = \bar{\eta}^\top \nabla_y H(y, \bar{M})^\top$, it is not necessary to compute the full Jacobian $\nabla_y H(y, \bar{M})^\top$. Rather, it is sufficient to use one pass of the backward mode of Algorithmic Differentiation seeded by $\bar{\eta}$ [11]. Due to the optimal control structure, there may be further exploitable structure hidden in the computation of $\bar{c}$, which we do not discuss here. The perturbed nominal problem is then

$$\min_{\bar{x}, \bar{u}} \quad \sum_{k=0}^{N-1} l_k(\bar{x}_k, \bar{u}_k) + \bar{c}_k^\top \begin{bmatrix} \bar{x}_k \\ \bar{u}_k \end{bmatrix} + E(\bar{x}_N) + \bar{c}_N^\top \bar{x}_N \tag{38a}$$

$$\text{s.t.} \qquad \bar{x}_0 = \bar{\bar{x}}_0, \tag{38b}$$

$$\bar{x}_{k+1} = f_k(\bar{x}_k, \bar{u}_k, 0), \qquad 0 \le k < N, \tag{38c}$$

$$0 \ge h_k(\bar{x}_k, \bar{u}_k) + \bar{b}_k, \quad 0 \le k < N, \tag{38d}$$

$$0 \ge h_N(\bar{x}_N) + \bar{b}_N. \tag{38e}$$

Opposed to (9), here the β are eliminated to highlight the optimal control problem structure. In consquence, solutions to (38) do not include the multipliers η of the eliminated constraints needed for updating the perturbations. However, this is not a problem as given a solution to (38), they can be easily obtained via (17b). The resulting procedure is sketched in Algorithm 1.

V. NUMERICAL EXAMPLE

We now demonstrate the algorithm for the control of a towing kite. For the formulation of the control problem we follow [17], with the model taken from [5], [7].

For constant tether length L the position of the kite is defined by angles θ and ϕ, and its orientation by ψ. Subsuming them in state $x = (\theta, \phi, \psi)$, the continous time dynamics are given as

$$\dot{\theta} = \frac{(\bar{v}_0 + w_v)E(u)\cos\theta}{L}\Big(\cos\psi - \frac{\tan\theta}{E(u)}\Big) + w_\theta \tag{39a}$$

$$\dot{\phi} = \frac{-(\bar{v}_0 + w_v)E(u)\cos\theta\sin\psi}{L\sin\theta} + w_\phi \tag{39b}$$

$$\dot{\psi} = \frac{(\bar{v}_0 + w_v)E(u)\cos\theta}{L}u + \dot{\phi}\cos\theta + w_\psi \tag{39c}$$

TABLE I
PARAMETER VALUES FOR THE KITE MODEL .

L (m)	E_0	$\tilde{c}$	$\bar{v}_0$ ($\frac{\mathrm{m}}{s}$)	$h_{\min}$ (m)	$\check{u}$	$\rho \cdot A$ ($\frac{\mathrm{kg}}{\mathrm{m}}$)
400	5	0.028	10	100	10	300

where the glide ratio is given by $E(u) = E_0 - \tilde{c}u^2$, and the control input is the steering deflection u. We assume some perturbations in every state, but the main source of uncertainty is the unknown apparent wind speed v_0 (referenced to the boat), expressed via deviations from an assumed nominal $\bar{v}_0$. The overall perturbation vector is $w = (w_\theta, w_\phi, w_\psi, w_v)$ with $w \in \mathcal{E}(W)$, where $W = \mathrm{diag}(\varepsilon_\theta^2, \varepsilon_\phi^2, \varepsilon_\psi^2, \varepsilon_v^2)$ and $\varepsilon_\theta = \varepsilon_\phi = \varepsilon_\psi = 10^{-4}$ and $\varepsilon_v = 1$ m/s. The discrete time dynamics are obtained by integrating (39) with one step of a Runge-Kutta integrator of fourth-order, while keeping u and w piecewise constant, for a time step of $\Delta t = 0.3$ s. The horizon length is $N = 80$. The control goal is the maximization of the nominal achieved thrust,

$$l_k(x, u) = -\tfrac{1}{2}\rho\bar{v}_0^2 A\cos^2\theta(E(u)+1)\sqrt{E^2(u)+1} \tag{40}$$

with air density ρ and kite area A. The controls are bounded and the kite has to stay above a minimal height $h_{\min}$,

$$-\check{u} \le u_k \le \check{u}, \qquad k = 0, \dots, N-1, \tag{41}$$

$$h_{\min} \le L\sin\theta_k\cos\theta_k, \qquad k = 1, \dots, N, \tag{42}$$

which defines the stage and terminal constraints. The values for all parameters are given in Table I.

The algorithm is implemented via the Python interface of CasADi [1] and the perturbed nominal problems are solved with IPOPT [30]. The initial guess is obtained by solving the nominal problem without perturbation. For comparison we also solve the nominal problem, as well as the open loop robust OCP, in which the possibility of feedback is not included in the plan. The solution can be obtained by a modification of the algorithm: instead of solving the Riccati recursion, the feedback is always set to $K = 0$, which is in principle similar to variants of [8], [31]. The trajectories are visualized in Fig. 1. Note how the closed loop robust solution does not suffer from a growing uncertainty set. Instead it achieves low uncertainty where necessary, i.e., in the vicinity of the inequality constraint.

We now investigate the local convergence behavior. To this end, we solve the robustified OCP for three different levels of uncertainty, $\sigma \in \{0.5, 1, 2\}$, where $\sigma = 1$ corresponds to the perturbation parameters reported earlier in the text. The resulting convergence behavior, again with initialization at the unperturbed nominal problem, is illustrated in Fig. 2. The algorithm is considered converged when the max-norm of the KKT conditions (12) is smaller than 10^{-3}. As predicted in Theorem 10 the convergence rate appears to be linear. Further, larger values of σ slow down the convergence, as would also be expected. We perform no rigorous benchmarking of CPU times, as the focus was on a proof of concept implementation. However, we can state

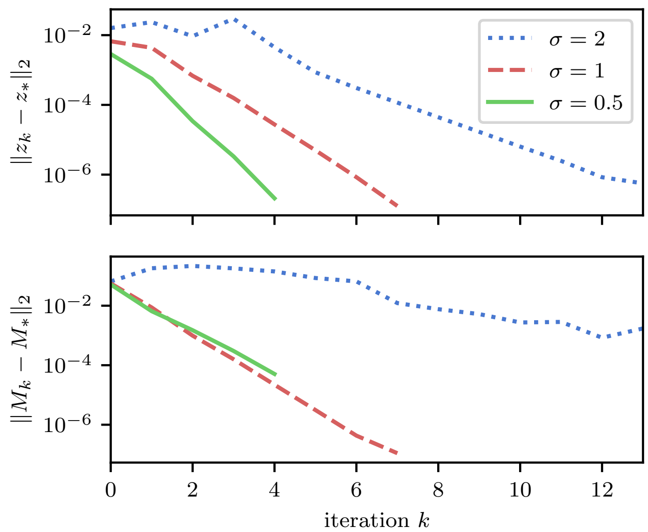

Fig. 2. Convergence behavior of SIRO for different levels of uncertainty σ. Top: primal-dual variables z of the perturbed nominal problem. Bottom: the eliminated variable M, here corresponding to the feedback matrices K. The convergence rate is linear, with slower convergence for higher uncertainty. The final iterate (excluded from plot) is used as proxy for the true solution.

that after initialization to the nominal solution the algorithm took roughly 0.5 s for $\sigma = 0.5$ and 1.5 s for $\sigma = 2$, on a standard personal notebook, while there is still ample room for improvement. The solution of each perturbed nominal OCP consistently needed 14 iterations of IPOPT.

VI. CONCLUSIONS AND FUTURE WORK

This paper presented a novel algorithm for the solution of robustified optimal control problems that involve linear feedback gains as optimization variables. Important next steps are an investigation of its resulting performance when used in an MPC framework as well as an efficient implementation to explore its real-time potential.

REFERENCES

[1] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl. CasADi – a software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*, 11(1):1–36, 2019.

[2] A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004.

[3] D. Bernardini and A. Bemporad. Scenario-based model predictive control of stochastic constrained linear systems. In *Proceedings of the IEEE Conference on Decision and Control (CDC)*, pages 6333–6338, 2009.

[4] D. Bertsekas and I. Rhodes. On the minimax reachability of target sets and target tubes. *Automatica*, 7:233–247, 1971.

[5] S. Costello, G. François, and D. Bonvin. Real-time optimization for kites. *Proceedings of the IFAC World Congress*, 46(12):64–69, 2013.

[6] D. M. de la Pena, A. Bemporad, and T. Alamo. Stochastic programming applied to model predictive control. In *Proceedings of the IEEE Conference on Decision and Control (CDC)*, volume 2, pages 1361–1366, 2005.

[7] M. Erhard and H. Strauch. Control of towing kites for seagoing vessels. *IEEE Transactions on Control Systems Technology*, 21(5):1629–1640, 2012.

[8] X. Feng, S. D. Cairano, and R. Quirynen. Inexact Adjoint-based SQP Algorithm for Real-Time Stochastic nonlinear MPC. In *Proceedings of the IFAC World Congress*, 2020.

[9] J. Gillis and M. Diehl. A positive definiteness preserving discretization method for nonlinear Lyapunov differential equations. In *Proceedings of the IEEE Conference on Decision and Control (CDC)*, 2013.

[10] P. J. Goulart, E. C. Kerrigan, and J. M. Maciejowski. Optimization over state feedback policies for robust control with constraints. *Automatica*, 42:523–533, 2006.

[11] A. Griewank and A. Walther. *Evaluating Derivatives*. SIAM, 2 edition, 2008.

[12] L. Hewing, J. Kabzan, and M. N. Zeilinger. Cautious model predicitive control using gaussian process regression. *IEEE Transaction on Control Systems Technology*, 28(6):2736–2743, 2020.

[13] B. Houska and M. Diehl. Robustness and stability optimization of power generating kite systems in a periodic pumping mode. In *Proceedings of the IEEE Multi-conference on Systems and Control (MSC)*, pages 2172–2177, Yokohama, Japan, 2010.

[14] D. Limon et al. Input-to-state stability: A unifying framework for robust model predictive control. In F. A. L. Magni, D. M. Raimondo, editor, *Nonlinear Model Predictive Control. Lecture Notes in Control and Information Sciences*, volume 384. Springer, Berlin, Heidelberg, 2009.

[15] B. T. Lopez, J. E. Slotine, and J. P. How. Dynamic tube MPC for nonlinear systems. *Proceedings of the American Control Conference (ACC)*, 2019.

[16] S. Lucia, J. A. E. Andersson, H. Brandt, M. Diehl, and S. Engell. Handling uncertainty in economic nonlinear model predictive control: A comparative case study. *Jounal of Process Control*, 24:1247–1259, 2014.

[17] S. Lucia and S. Engell. Control of towing kites under uncertainty using robust economic nonlinear model predictive control. *Proceedings of the European Control Conference (ECC)*, 2014.

[18] D. Mayne, E. Kerrigan, E. J. van Wyk, and P. Falugi. Tube-based robust nonlinear model predictive control. *International Journal of Robust and Nonlinear Control*, 21:1341–1353, 2011.

[19] D. Q. Mayne, E. C. Kerrigan, and P. Falugi. Robust model predictive control: advantages and disadvantages of tube-based methods. *Proceedings of the IFAC World Congress*, 2011.

[20] D. Q. Mayne, M. M. Seron, and S. V. Rakovic. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica*, 41:219–224, 2005.

[21] M. Morari. Robust Process Control. *Chem. Eng. Res. Des.*, 65:462–479, 1987.

[22] Z. Nagy and R. Braatz. Open-loop and closed-loop robust optimal control of batch processes using distributional and worst-case analysis. *Journal of Process Control*, 14:411–422, 2004.

[23] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer, 2 edition, 2006.

[24] A. M. Ostrowski. *Solutions of Equations in Euclidean and Banach Spaces*. Academic Press, New York and London, 1973.

[25] S. Rakovic, B. Kouvaritakis, M. Cannon, C. Panos, and R. Findeisen. Parameterized tube model predictive control. *Automatic Control, IEEE Transactions on*, 57(11):2746–2761, Nov 2012.

[26] S. Raković, W. S. Levine, and B. Açıkmeşe. Elastic tube model predictive control. *Proceedings of the American Control Conference (ACC)*, 2016.

[27] J. B. Rawlings, D. Q. Mayne, and M. M. Diehl. *Model Predictive Control: Theory, Computation, and Design*. Nob Hill, 2nd edition, 2017.

[28] S. Robinson. Perturbed Kuhn-Tucker points and rates of convergence for a class of nonlinear programming algorithms. *Mathematical Programming*, 7:1–16, 1974.

[29] R. F. Stengel. *Optimal Control and Estimation*. Dover, 1986.

[30] A. Wächter and L. T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106(1):25–57, 2006.

[31] A. Zanelli, J. Frey, F. Messerer, and M. Diehl. Zero-order robust nonlinear model predictive control with ellipsoidal uncertainty sets. *Proceedings of the IFAC Conference on Nonlinear Model Predictive Control (NMPC)*, 2021.