Extended Rauch-Tung-Striebel Controller

Miroslav Zima, Leopoldo Armesto, Vicent Girbés, Antonio Sala and Václav Šmídl

Abstract— This paper presents a novel controller for nonlinear unconstrained systems, coined as Extended Rauch-Tung-Striebel (ERTS) controller. The controller is derived from a general framework based on the duality between optimal control and estimation established by Todorov. The proposed controller uses Rauch-Tung-Striebel smoother that predicts (filters) future states by linearizing the nonlinear system around predicted states and then applies a backward smoothing. The new controller is applied to solve path following problems of non-holonomic vehicles and compared with the standard LQR controller linearizing the model around the desired trajectory and the iterative LQR (iLQR) controller. The main advantages of ERTS controller with respect to the alternative techniques are good control performance and computational efficiency.

I. INTRODUCTION

This article deals with the duality between optimal control and estimation and particularly with the derivation of a new controller based on the duality for nonlinear systems. The fact, that the covariance matrix of the optimal estimator of a linear system with Gaussian noises and Hessian of the optimal cost-to-go of a linear control problem with a quadratic loss evolves in time under similar Riccati-like equations, is known more than fifty years, [5]. Due to this, both solutions (Kalman Filter, KF, and Linear Quadratic Controller, LQR) have the same form and as a consequence an algorithm computing KF can be used as LQR algorithm. This is known as the (Kalman's) duality between optimal control and estimation for linear-Gaussian systems. This interesting property motivated efforts for development of possible extensions to nonlinear systems, nonetheless, straightforward generalization on nonlinear cases is not known. Theoretical work introducing satisfactory extension was done in [13], where the general (Todorov's) duality between optimal control and estimation is obtained for slightly reformulated optimal control problem based on Kullback-Leibler divergence.

The new general duality applied on LQ problem does not give the same algortihm as the Kalman's approach. It is because the estimation problem is different in both approaches – prediction in Kalman's case and smoothing in Todorov's.

This paper proposes a new controller, coined as Extended Rauch-Tung-Striebel (ERTS) controller, derived from the duality between optimal control and estimation. The new proposed controller is based on the solution of the dual estimation problem given by Rauch-Tung-Striebel (RTS) forward-backward smoother. The computed estimate of next state is then used for the computation of the optimal control. This results in an efficient controller with complexity $\mathcal{O}(N^2)$ in state dimensions. The controller is optimal for LQ systems and the extension to non-linear settings is done by linearizion along predicted trajectory.

The performance of ERTS controller is then illustrated on the (unconstrained) path following problem. The goal of path following problem is to track a robot along a desired path by a control law. Due to wide range of straightforward applications, e.g. motion planning [3] [1], parking [4], overtaking and lane changing [10], and vision-based line following [9], this problem has been studied intensively during last years. However, proposed controllers are commonly strictly specialized on particular tasks or contain "artificial" design parameters (e.g. look-ahead distance) which have to be tuned. The new method is compared with standard LQR approaches and the iterative LQR method (iLQR), [14], obtaining significant improvement in accuracy and time efficiency.

The paper is organized as follows: Section II introduces the duality between estimation and control. Section III particularizes the ideas from previous section and presents ERTS controller. The controller is applied in Section IV on the path following problem and its performance is compared with linearized LQR and iLQR controllers. Conclusions are drawn in Section V.

II. PRELIMINARIES AND PROBLEM STATEMENT

Consider an stochastic nonlinear dynamic system modeled as Markov process with known transition probability depending on the actual state $\mathbf{x}_t$ and the control action $\mathbf{u}_t$

$$\mathbf{x}_{t+1} \sim p(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{u}_t). \tag{1}$$

For an arbitrary stochastic control given by distribution $\pi_t(\mathbf{u}_t)$, the resulting distribution of $\mathbf{x}_{t+1}$ is

$$\mathbf{x}_{t+1} \sim p_\pi(\mathbf{x}_{t+1}|\mathbf{x}_t) = \int_{\mathbb{R}^{n_u}} p(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{u}_t)\pi_t(\mathbf{u}_t)\mathrm{d}\mathbf{u}_t. \tag{2}$$

Let us consider obtaining a *stochastic* controller which optimizes the following expected[1] loss

$$J(\mathbf{x}_0, \bar{\mathbf{s}}_{0:N}, \pi_{0:N-1}) = \mathrm{E}\left\{q_N(\mathbf{x}_N, \bar{\mathbf{s}}_N) + \sum_{t=0}^{N-1} l_t(\mathbf{x}_t, \bar{\mathbf{s}}_t, \pi_t)\right\} \tag{3}$$

---

[1] the expectation is taken over realizations of the random variables $\mathbf{x}_{1:N}$.

where sequence $\bar{\mathbf{s}}_{0:N} \triangleq \bar{\mathbf{s}}_0, \ldots, \bar{\mathbf{s}}_N$ stands for the desired quantities related to states (actually, reference trajectories for some outputs), $q_N(\mathbf{x}_N, \mathbf{s}_N)$ is an arbitrary function for the final cost and the intermediate loss is $l_t(\mathbf{x}_t, \mathbf{s}_t, \pi_t)$.

A. KL-Optimal Control Problem

Similarly as in [13], consider a cost function of the form

$$l_t(\mathbf{x}_t, \mathbf{s}_t, \pi_t) = q_t(\mathbf{x}_t, \mathbf{s}_t) + \text{KL}\big(p_\pi(\mathbf{x}_{t+1}|\mathbf{x}_t)||\bar{p}(\mathbf{x}_{t+1}|\mathbf{x}_t)\big)(4)$$

where $\bar{p}(\mathbf{x}_{t+1}|\mathbf{x}_t)$ is a user-defined *reference dynamics* and $q_t(\mathbf{x}_t, \mathbf{s}_t)$ is an arbitrary function over the state satisfying

$$\int_{\mathbb{R}^{n_s}} e^{-q_t(\mathbf{x}_t, \mathbf{s}_t)} \mathrm{d}\mathbf{s}_t = 1. \tag{5}$$

The KL stands for Kullback-Leibler divergence [7]

$$\text{KL}(p_\pi(\mathbf{x}_{t+1}|\mathbf{x}_t)||\bar{p}(\mathbf{x}_{t+1}|\mathbf{x}_t)) =$$
$$= \int_{\mathbb{R}^{n_x}} \log\left[\frac{p_\pi(\mathbf{x}_{t+1}|\mathbf{x}_t)}{\bar{p}(\mathbf{x}_{t+1}|\mathbf{x}_t)}\right] p_\pi(\mathbf{x}_{t+1}|\mathbf{x}_t) \, \mathrm{d}\mathbf{x}_{t+1} \tag{6}$$

between the system dynamics (1) and the user-defined reference dynamics $\bar{p}$. The optimal stochastic control minimizing (3) is computed as a function of the actual state as $\pi^*(\mathbf{u}_t|\mathbf{x}_t)$.

Note that only the first term in (4) depends on the desired states and KL divergence can be interpreted as "weight" of a control action, [6].

The minimal loss fulfills the well-known Bellman equation

$$J_t = \min_{\pi_t}\left\{l_t(\mathbf{x}_t, \pi_t, \bar{\mathbf{s}}_t) + \int_{\mathbb{R}^{n_x}} J_{t+1} \cdot p_\pi \, \mathrm{d}\mathbf{x}_{t+1}\right\}$$

with $J_N = q_N(\mathbf{x}_N, \bar{\mathbf{s}}_N)$. Here, we used simplified notation $p_\pi(\mathbf{x}_{t+1}|\mathbf{x}_t) \triangleq p_\pi$, $\bar{p}(\mathbf{x}_{t+1}|\mathbf{x}_t) \triangleq \bar{p}$ and $J_t(\mathbf{x}_t, \bar{\mathbf{s}}_{t:N}) \triangleq J_t$.

For the loss function of the form (4), the following holds

$$J_t = q_t(\mathbf{x}_t, \bar{\mathbf{s}}_t) + \min_{\pi_t(\mathbf{u}_t)} \int_{\mathbb{R}^{n_x}}\left[J_{t+1} + \log\frac{p_\pi}{\bar{p}}\right] p_\pi \mathrm{d}\mathbf{x}_{t+1}$$
$$= q_t - \log c + \min_{\pi_t(\mathbf{u}_t)} \text{KL}\left(p_\pi||\frac{1}{c}e^{-J_{t+1}}\bar{p}\right) \tag{7}$$

where the normalizer $c = c(\mathbf{x}_t, \bar{\mathbf{s}}_{t+1:N})$ is equal to

$$c(\mathbf{x}_t, \bar{\mathbf{s}}_{t+1:N}) = \int_{\mathbb{R}^{n_x}} e^{-J_{t+1}(\mathbf{x}_{t+1}, \bar{\mathbf{s}}_{t+1:N})} \bar{p}(\mathbf{x}_{t+1}|\mathbf{x}_t) \mathrm{d}\mathbf{x}_{t+1}.$$

As KL divergence is always non negative and is zero if and only if the distributions are equal almost everywhere [7], this leads to the optimum system behavior with distribution

$$p_\pi(\mathbf{x}_{t+1}|\mathbf{x}_t) \propto e^{-J_{t+1}(\mathbf{x}_{t+1}, \bar{\mathbf{s}}_{t+1:N})} \bar{p}(\mathbf{x}_{t+1}|\mathbf{x}_t). \tag{8}$$

The resulting distribution is a rescaled reference dynamics with a scaling factor favoring states with lower cost-to-go.

If the optimal policy $\pi_t(\mathbf{u}_t)$ exists so that (8) is satisfied, the right-most term in (7) vanishes and, after exponentiating both sides, the optimal cost-to-go satisfies:

$$e^{-J_t(\mathbf{x}_t, \bar{\mathbf{s}}_{t:N})} = e^{-q_t(\mathbf{x}_t, \bar{\mathbf{s}}_t)} c(\mathbf{x}_t) =$$
$$= e^{-q_t(\mathbf{x}_t, \bar{\mathbf{s}}_t)} \int_{\mathbb{R}^{n_x}} e^{-J_{t+1}(\mathbf{x}_{t+1}, \bar{\mathbf{s}}_{t+1:N})} \bar{p}(\mathbf{x}_{t+1}|\mathbf{x}_t) \mathrm{d}\mathbf{x}_{t+1}. \tag{9}$$

The existence of the optimal policy $\pi_t(\mathbf{u}_t)$ satisfying the optimal behavior according to (8) implies that the KL divergence can be minimized to zero. However, this is not guaranteed in general and it has to be studied apart.

B. Dual estimation problem

Assume now a state-space model in the form

$$\mathbf{x}_{t+1} \sim \bar{p}(\mathbf{x}_{t+1}|\mathbf{x}_t) \tag{10}$$
$$\mathbf{y}_t \sim p(\mathbf{y}_t|\mathbf{x}_t) \tag{11}$$

where (10) stands for the reference dynamics from the original control problem and (11) is an observation model.

In [13], the dual estimation problem is defined as backward estimation problem of $\mathbf{x}_{t+1}$ knowing $\mathbf{x}_t$ and the whole observation sequence $\mathbf{y}_{t+1:N}$. The estimate satisfies

$$p(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{y}_{t+1:N}) \propto p(\mathbf{y}_{t+1:N}|\mathbf{x}_{t+1})\bar{p}(\mathbf{x}_{t+1}|\mathbf{x}_t) \tag{12}$$

due to the chain rule and Markov property. Moreover, it holds

$$p(\mathbf{y}_{t+1:N}|\mathbf{x}_{t+1}) = p(\mathbf{y}_{t+1}|\mathbf{x}_{t+1})p(\mathbf{y}_{t+2:N}|\mathbf{x}_{t+1})$$
$$= p(\mathbf{y}_{t+1}|\mathbf{x}_{t+1})\int_{\mathbf{x}_{t+2}} p(\mathbf{y}_{t+2:N}|\mathbf{x}_{t+2})\bar{p}(\mathbf{x}_{t+2}|\mathbf{x}_{t+1})\mathrm{d}\mathbf{x}_{t+2} \tag{13}$$

In order to state the duality between the original control problem and the considered estimation problem, we restrict the observation model (11) to the form

$$p(\mathbf{y}_t|\mathbf{x}_t) = e^{-q_t(\mathbf{x}_t, \mathbf{y}_t)}. \tag{14}$$

where q_t stands for the state penalization from the original control problem. Note that due to the restriction (5), the probability (14) is well defined.

Then, if we write (13) for a particular observation sequence $\mathbf{y}_{t+1:N} \triangleq \bar{\mathbf{s}}_{t+1:N}$, i.e. if we force the observations in the dual estimation problem to the desired states from the original control problem, we obtain

$$p(\bar{\mathbf{s}}_{t+1:N}|\mathbf{x}_{t+1}) =$$
$$= e^{-q_t(\mathbf{x}_{t+1}, \bar{\mathbf{s}}_{t+1})} \int_{\mathbf{x}_{t+2}} p(\bar{\mathbf{s}}_{t+2:N}|\mathbf{x}_{t+1})\bar{p}(\mathbf{x}_{t+2}|\mathbf{x}_{t+1})\mathrm{d}\mathbf{x}_{t+2}.$$

The previous equation has the same form as the equation (9) and due to that duality holds

$$p(\bar{\mathbf{s}}_{t+1:N}|\mathbf{x}_{t+1}) = e^{-J_{t+1}(\mathbf{x}_{t+1}, \bar{\mathbf{s}}_{t+1:N})}. \tag{15}$$

The equality (15) is the mathematical formulation of the general duality between optimal control and estimation [13]. As consequence, the optimal behavior (8) is equal to the smoothed estimate (12)

$$p_\pi(\mathbf{x}_{t+1}|\mathbf{x}_t) = p(\mathbf{x}_{t+1|N}). \tag{16}$$

where we used notation $p(\mathbf{x}_{t+1|N}) \triangleq p(\mathbf{x}_{t+1}|\mathbf{x}_t, \bar{\mathbf{s}}_{t+1:N})$.

C. Deterministic Optimal control

It is well known that the solution to (3) for a *linear* model with Gaussian additive noise and quadratic index can be obtained by solving the *deterministic* LQR problem assuming zero future noise.

Solving (3) for nonlinear plants is not an easy problem, even for deterministic (1). Linearization-based controllers for nonlinear systems are widely known: the model is linearised along a set of points and disturbances are assumed Gaussian additive. Then, the finite-horizon optimal LQR control law

for the resulting linear time-varying plant is proposed to approximate the nonlinear optimal controller.

The basic problem of such approaches is choosing the points in which to linearise the system: the most straightforward approach would suggest linearising around the reference trajectory, on the hope that deviations would be small, and computing the LQR control action. Actually, it is clear that the optimal linearization points would be those of the optimal trajectory; however, as they depend on the to-be-computed control and the control depends on such points, iterative iLQR algorithms are needed (linearise around first trajectory estimate, compute control, compute new trajectory, repeat), see [14]. iLQR improves actual performance with respect to 1-pass reference-trajectory linearization at a much higher computational cost.

D. Problem statement

The result (16) brings a theoretical background for a new controller design based on the transformation of a control problem to a related dual estimation problem and its solution.

The objective of next section will be particularizing the above stochastic control result to Gaussian noises and, using point-wise linearizations, apply them to nonlinear control problems. Such nonlinear control problems will, hence, be solved via an "extended" estimator. A detailed algorithm (to be named ERTS) will be presented and particular cases discussed.

Actually, the most practically relevant particular case will be the deterministic quadratic cost index one, presenting an alternative approach to the proposals discussed in Section II-C. This is made possible by a suitable choice of "artificial" auxiliary probability distributions for both (1) and the probability $\bar{p}$; then, the proposed methodology will make it dual to a stochastic estimation problem which will be solved in just one forward-backward step. In some cases, this will allow approximately solving the deterministic optimal control problem with an accuracy almost that of iLQR without iterations, as a particular path-following kinematic control example will show. In such example, accuracy and computational cost of the different LQR, iLQR, ERTS approaches will be compared.

III. NON-LINEAR QUADRATIC CONTROL VIA ESTIMATION

As discussed above, this section specialises the original KL-optimal control to affine-in-control nonlinear deterministic plants with quadratic cost, proposing a particular algorithm based in the Rauch-Tung-Striebel smoother [11]. We show that there is a related *stochastic* KL optimal control problem for an arbitrary *deterministic* quadratic-cost problem, [13], which, as a consequence, can be solved optimally via the duality.

Consider a nonlinear model $\mathbf{x}_{t+1} = f_t(\mathbf{x}_t) + \mathbf{B}_t\mathbf{u}_t$ for known functions f_t, matrix $\mathbf{B}_t$, and and a quadratic index $J = \frac{1}{2}\mathbf{e}_N^T\mathbf{Q}_N\mathbf{e}_N + \frac{1}{2}\sum_{t=0}^{N-1}(\mathbf{e}_t^T\mathbf{Q}_t\mathbf{e}_t + \mathbf{u}_t^T\mathbf{R}_t\mathbf{u}_t)$, for $\mathbf{e}_t = \mathbf{s}_t - \mathbf{H}_t\mathbf{x}_t$ and known $\mathbf{Q}_t$, $\mathbf{R}_t$, $\mathbf{H}_t$.

In order to apply the KL control, a *fictitious* Gaussian noise will be added, with variance $\mathbf{V}_t$ and, also, a fictitious target

stochastic dynamics will be proposed. Hence, the original dynamics (1) will be restricted to

$$p(\mathbf{x}_{t+1}|\mathbf{x}_t,\mathbf{u}_t) \triangleq \mathcal{N}(f_t(\mathbf{x}_t) + \mathbf{B}_t\mathbf{u}_t, \mathbf{V}_t) \qquad (17)$$

and the reference dynamics in (4) set to

$$\bar{p}(\mathbf{x}_{t+1}|\mathbf{x}_t) \triangleq \mathcal{N}(\bar{f}_t(\mathbf{x}_t), \bar{\mathbf{V}}_t) \qquad (18)$$

Function $\bar{f}_t$, and matrices $\mathbf{V}_t, \bar{\mathbf{V}}_t$ will be actually derived from the cost function, in particular $\bar{\mathbf{V}}_t = \mathbf{B}_t\mathbf{R}_t^{-1}\mathbf{B}_t^T$, and $\mathbf{V}_t$ will be analogous to the Riccati equation solution [13].

The cost function (4) will be stated as the sum of quadratic state-dependent terms

$$q_t(\mathbf{x}_t, \mathbf{s}_t) \triangleq \frac{1}{2}(\mathbf{s}_t - h_t(\mathbf{x}_t))^T\mathbf{W}_t^{-1}(\mathbf{s}_t - h_t(\mathbf{x}_t)) + c_t \quad (19)$$

for $h_t(\mathbf{x}_t) = \mathbf{H}_t\mathbf{x}_t$, $\mathbf{W}_t = \mathbf{H}_t\mathbf{Q}_t^{-1}\mathbf{H}_t^T$ and constant c_t chosen in order to fulfill the restriction (5), plus the KL divergence term between (17) and (18). From well-known KL formulae for Gaussian distributions, this choice of matrices makes (4) identically equal to the above-defined quadratic index. Hence, duality can be used to solve the control problem and, in the nonlinear case, linearization-based smoothers are proposed next.

A. Solution of the dual problem

The estimation problem dual to the previous control problem requires computing the smoothed probability $p(\mathbf{x}_{t+1|N})$, where the desired states $\bar{\mathbf{s}}_{t+1:N}$ from the original control problem are now considered as the observations. The model in the dual problem is defined by (18) and (19) as

$$\mathbf{x}_{t+1} \sim \mathcal{N}(\bar{f}_t(\mathbf{x}_t), \bar{\mathbf{V}}_t) \qquad (20)$$
$$\mathbf{s}_t \sim \mathcal{N}(h_t(\mathbf{x}_t), \mathbf{W}_t). \qquad (21)$$

The control algorithm proposed in this paper uses an approximation of $p(\mathbf{x}_{t+1|N})$ computed by Rauch-Tung-Striebel (RTS) smoother, [11]. The RTS smoother is optimal for linear models with Gaussian noise and it computes the smoothed distribution of $\mathbf{x}_{t+1}$ in two steps:

- forward pass realized by Kalman filter computing $p(\mathbf{x}_\tau|\mathbf{x}_t, \mathbf{y}_{t:\tau})$ for $\tau = t + 1, \ldots, N$
- backward pass computing the $p(\mathbf{x}_\tau|\mathbf{x}_t, \mathbf{y}_{t:N})$ from $p(\mathbf{x}_\tau|\mathbf{x}_t, \mathbf{y}_{t:\tau})$ and $p(\mathbf{x}_{\tau+1}|\mathbf{x}_t, \mathbf{y}_{t:\tau})$ for $\tau = N, \ldots, t+1$

The resulted smoothed distribution $p(\mathbf{x}_{t+1|N})$ of $\mathbf{x}_{t+1}$ is a Gaussian distribution with mean value $\hat{\mathbf{x}}_{t+1|N}$ and covariance matrix $\mathbf{P}_{t+1|N}$, see [11].

If functions $\bar{f}_t, h_t$ are nonlinear, RTS smoother can be used for the linearized model at each trajectory point; however, optimality of the proposed estimate is no longer guaranteed. This is analog to the Extended Kalman filter, and this is why the nonlinear version of the RTS smoother is denoted as ERTS [8].

B. Extended Rauch-Tung-Striebel Non-linear Controller

Once the estimation problem has been solved, duality indicates that the optimal control action should fulfill (16), taking the result of the ERTS smoother, which provides the approximation $p(\mathbf{x}_{t+1|N}) \approx \mathcal{N}(\hat{\mathbf{x}}_{t+1|N}, \mathbf{P}_{t+1|N})$ to the right-hand side term of (16).

In the additive Gaussian case in consideration, both mean and variance should be matched, computing $\mathbf{u}_t$ in (17) to match the mean of the ERTS estimate, and setting $\mathbf{V}_t$ equal to the estimate variance. Thus, as above mentioned, $\mathbf{V}_t$ is related to the standard Ricatti equation solution trough the ERTS algorithm. Hence, the solution is a deterministic control $\hat{\mathbf{u}}_t$ which has to satisfy

$$f_t(\mathbf{x}_t) + \mathbf{B}_t\hat{\mathbf{u}}_t \approx \hat{\mathbf{x}}_{t+1|N} \tag{22}$$

and a fictitious noise variance to

$$\mathbf{V}_t = \mathbf{P}_{t+1|N}. \tag{23}$$

The approximation sign in (22) is necessary because on the right-hand side of (16) we have only an approximation proposed by ERTS smoother[2]. Hence, using the left pseudo-inverse of $\mathbf{B}_t$, denoted as $\boldsymbol{\beta}_t = (\mathbf{B}_t^T\mathbf{B}_t)^{-1}\mathbf{B}_t^T$, we have a control action given by:

$$\hat{\mathbf{u}}_t = \boldsymbol{\beta}_t(\hat{\mathbf{x}}_{t+1|N} - f_t(\mathbf{x}_t)) \tag{24}$$

as the proposal for the nonlinear control law[3].

C. The algorithm of the ERTS controller

The resulting control algorithm, denoted as Extended Rauch-Tung-Striebel (ERTS) controller, is composed from two parts: 1) computing $\hat{\mathbf{x}}_{t+1|N}$ via ERTS smoother, and 2) obtaining the approximation of the optimal control (24). The algorithm of the ERTS controller is summarized in Algorithm 1, using process noise $\bar{\mathbf{V}}_t = \mathbf{B}_t\mathbf{R}_t^{-1}\mathbf{B}_t^T$ and measurement noise $\mathbf{W}_t = \mathbf{H}_t\mathbf{Q}_t^{-1}\mathbf{H}_t^T$. The state $\mathbf{x}_t$ is assumed to be known, so the proposed controller is an state feedback one.

Matrices $\mathbf{A}_{\tau-1}, \mathbf{P}_{\tau|\tau-1}, \mathbf{P}_{\tau|\tau}$ and vectors $\hat{\mathbf{x}}_{\tau+1|\tau}$ computed during the prediction step are stored in the memory and used during the backward smoothing step. It should be noted that our backward smoothing step does *not* compute matrices $\mathbf{P}_{\tau|T}$: expression (23) is only needed in the formal problem solution, but the value of $\mathbf{V}_t$ is not needed to solve (22).

D. Unscented Rauch-Tung-Striebel Controller

Similarly, we coin the controller based on the solution of the dual problem proposed by RTS smoother with Uscented Kalman filter (UKF) and smoother (UKS) [12] as Unscented Rauch-Tung-Striebel (URTS) controller. The URTS controller's structure is the identical to ERTS controller's structure. During the prediction step with UKF, a set of σ-points are computed based on augmented state using current estimate and zero-mean noises. In the smoothing step, a new

---

Algorithm 1 ERTS Controller

1: Initialization
2: $\hat{\mathbf{x}}_{t|t} = \mathbf{x}_t$, $\mathbf{P}_{t|t} = \mathbf{0}$
3: Prediction
4: **for** $\tau = t+1, \dots, N$ **do**
5: $\quad \hat{\mathbf{x}}_{\tau|\tau-1} = \bar{f}_{\tau-1}(\hat{\mathbf{x}}_{\tau-1|\tau-1})$
6: $\quad$ Linearization
7: $\quad\quad \mathbf{A}_{\tau-1} = \frac{\partial \bar{f}_{\tau-1}}{\partial \mathbf{x}_{\tau-1}}|\mathbf{x}_{\tau-1} = \hat{\mathbf{x}}_{\tau-1|\tau-1}$
8: $\quad\quad \mathbf{H}_{\tau} = \frac{\partial h_{\tau}}{\partial \mathbf{x}_{\tau}}|\mathbf{x}_{\tau} = \hat{\mathbf{x}}_{\tau|\tau-1}$
9: $\quad \mathbf{P}_{\tau|\tau-1} = \mathbf{A}_{\tau-1}\mathbf{P}_{\tau-1|\tau-1}\mathbf{A}_{\tau-1}^T + \mathbf{B}_t\mathbf{R}_t^{-1}\mathbf{B}_t^T$
10: $\quad \mathbf{K}_{\tau} = \mathbf{P}_{\tau|\tau-1}\mathbf{H}_{\tau}^T(\mathbf{H}_{\tau}\mathbf{P}_{\tau|\tau-1}\mathbf{H}_{\tau}^T + \mathbf{H}_t\mathbf{Q}_t^{-1}\mathbf{H}_t^T)^{-1}$
11: $\quad \mathbf{P}_{\tau|\tau} = (\mathbf{I} - \mathbf{K}_{\tau}\mathbf{H}_{\tau})\mathbf{P}_{\tau|\tau-1}$
12: $\quad \hat{\mathbf{x}}_{\tau|\tau} = \hat{\mathbf{x}}_{\tau|\tau-1} + \mathbf{K}_{\tau}(\bar{\mathbf{s}}_{\tau} - h_{\tau}(\hat{\mathbf{x}}_{\tau|\tau-1}))$
13: **end for**
14: Smoothing
15: **for** $\tau = N-1, N-2, \dots, t+1$ **do**
16: $\quad \mathbf{L}_{\tau} = \mathbf{P}_{\tau|\tau}\mathbf{A}_{\tau}^T\mathbf{P}_{\tau+1|\tau}^{-1}$
17: $\quad \hat{\mathbf{x}}_{\tau|N} = \hat{\mathbf{x}}_{\tau|\tau} + \mathbf{L}_{\tau}(\hat{\mathbf{x}}_{\tau+1|N} - \hat{\mathbf{x}}_{\tau+1|\tau})$
18: **end for**
19: $\mathbf{u}_t^* = (\mathbf{B}_t^T\mathbf{B}_t)^{-1}\mathbf{B}_t^T(\hat{\mathbf{x}}_{t+1|N} - f_t(\mathbf{x}_t))$

---

set of σ-points must be computed as well as weights for mean $W_{m,i}$ and covariance $W_{m,i}$ following the ideas of [12]. The remainder of steps are identical to the ERTS controller.

IV. APPLICATION TO KINEMATIC CONTROL OF NON-HOLONOMIC WHEELED ROBOTS

In this section, a path following problem for non-holonomic wheeled robots is stated and solved by the previously presented ERTS controller. Some other standard control techniques from literature are tested for comparison.

A. Robot model and cost index specification

A vehicle state $\mathbf{x}_t = (x_t, y_t, \phi_t, v_t, \kappa_t)^T$ in time instant t is characterized by coordinates (x_t, y_t), angle (ϕ_t), velocity (v_t) and curvature (κ_t) and it evolves through input $\mathbf{u}_t = (a_t, \epsilon_t)^T$ given by acceleration (a_t) and curvature derivative (ϵ_t) as

$$\mathbf{x}_{t+1} = \mathbf{A}(\mathbf{x}_t)\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \mathbf{v}_t \tag{25}$$

with noise $\mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, 10^{-6}\mathbf{I})$ and state-dependent transition matrix $\mathbf{A}(\mathbf{x}_t)$ and control-to-state matrix $\mathbf{B}$

$$\mathbf{A}(\mathbf{x}_t) = \begin{pmatrix} 1 & 0 & 0 & \cos\phi_t\Delta t & 0 \\ 0 & 1 & 0 & \sin\phi_t\Delta t & 0 \\ 0 & 0 & 1 & \kappa_t\Delta t & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \Delta t & 0 \\ 0 & \Delta t \end{pmatrix},$$

Δt is the simulation step (Euler integration).

The observation model is simply stated as,

$$\mathbf{y}_t = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \mathbf{x}_t + \boldsymbol{\psi}_t \tag{26}$$

with noise $\boldsymbol{\psi}_t \sim \mathcal{N}(\mathbf{0}, 10^{-6}\mathbf{I})$. The interpretation is that only the position and orientation of the vehicle are measured, other states have to be estimated via an observer.

The aim is to drive a vehicle around a desired path, with a given reference speed of $\bar{v}_t$. The reference positions, $\bar{x}_t, \bar{y}_t$, orientation $\bar{\phi}_t$ and curvature $\bar{\kappa}_t$ are taken from the path itself, based on the reference speed and the simulation time for a

---

[2]Indeed, the relation becomes equality for linear $\bar{f}_t, h_t$.

[3]If the KL control problem were set with user-defined $\mathbf{V}_t$ instead of (23), a variant of the approach resulting in a probability distribution over control actions would ensue, see [16].

given simulation step. In order to evaluate the performance of a control algorithm, we introduce a quadratic loss

$$(\mathbf{x}_N - \bar{\mathbf{x}}_N)^T \mathbf{Q}_t (\mathbf{x}_N - \bar{\mathbf{x}}_N) + \sum_{t=0}^{N-1} \left((\mathbf{x}_t - \bar{\mathbf{x}}_t)^T \mathbf{Q}_t (\mathbf{x}_t - \bar{\mathbf{x}}_t) + \mathbf{u}_t^T \mathbf{R}_t \mathbf{u}_t \right)$$

(27)

with penalization matrices $\mathbf{Q}_t = \text{diag}(100, 100, 1, 1, 1)$, $\mathbf{R}_t = \text{diag}(1, 1)$. The values of the matrices has been taken with respect to some preliminary simulations and the well known interpretation of the penalization which states that the true values are kept on the desired ones with precision given by $\mathbf{Q}_t^{-1}$ using zero control with precision $\mathbf{R}_t^{-1}$, [6].

B. Estimation

Extended Kalman Filter (EKF) is used for the estimation of the state. The estimate is Gaussian with mean $\hat{\mathbf{x}}_t = (\hat{x}_t, \hat{y}_t, \hat{\phi}_t, \hat{v}_t, \hat{\omega}_t)^T$ and covariance matrix $\mathbf{P}_t$. The linearized transition matrix is

$$\mathbf{A}_t = \frac{\partial \mathbf{A}}{\partial \mathbf{x}_t} \Big|_{\hat{\mathbf{x}}_t} = \begin{pmatrix} 1 & 0 & -\hat{v}_t \sin \hat{\phi} \Delta t & \cos \hat{\phi}_t \Delta t & 0 \\ 0 & 1 & \hat{v}_t \cos \hat{\phi} \Delta t & \sin \hat{\phi}_t \Delta t & 0 \\ 0 & 0 & 1 & \hat{\kappa}_t \Delta t & \hat{v}_t \Delta t \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

(28)

In the linear case, this would amount to consider separation principle: synthesizing the feedback control law with the max-likelihood point estimate from an observer. In the nonlinear case, however, separation principle is only an approximation, as well known.

Note importantly, that the EKF estimator has no conceptual relation to the ERTS controller discussed in this work: the estimated variance from the (causal) EKF filter is *not* used in the (non-causal) ERTS smoother because the optimal control algorithm assumes initial zero variance.

C. Comparison of control strategies

In order to assess the advantages of the proposed duality-based ERTS control, comparison of three strategies will be made out in later simulations, as follows.

1) ERTS controller: The presented algorithm is used for KL optimal control problem given by (17)-(19) with $h_t = \mathbf{I}$, $\bar{\mathbf{s}} = \bar{\mathbf{x}}_t$, $\mathbf{W}_t = \mathbf{Q}_t^{-1}$ and the reference dynamics set to the passive original dynamics. The evaluation of the KL-divergence (6) results to the quadratic (27). The initialization of the dual estimation is $\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_t$, $\mathbf{P}_{t|t} = \mathbf{0}$. Note that ERTS controller also uses the linearized transition matrix (28).

2) LQR controller linearized along the desired path: A common extension of the standard LQR controller is done by linearization of the system along a desired path. This method was used for example in [15] where the method is combined with Rapidly-exploring Random Trees (RRT).

3) Iterative LQR method: An iterative extension of the classical LQR algorithm was presented in [14]. The iterative Linear-Quadratic (iLQR) method constructs an affine feedback control law, obtained by minimizing of a quadratic approximation to the optimal cost-to-go function. Single iteration of the algorithm consists of:

- application of the control computed in the previous step,
- linearization of the system and cost-to-go function around the computed trajectory,
- calculation of the optimal perturbation of the control for the linearized system via LQR-like recursion.

The algorithm ends when the relative improvement of the simulated cost-to-go is bellow fixed threshold $\eta = 0.001$ or maximum number of iterations $N_{\max} = 100$ is reached.

The method is able to solve nonlinear problems with control constraints and non-convex cost. Under certain assumptions, convergence to a local minimum is guaranteed, see [14]. The implementation mentioned there is used.

D. Experiment parameters

The target path is defined as the connection line of the following sequence of (x, y) points: $\{(0, 0); (2, 0); (2, 6); (6, -4); (-4, 10); (10, 10); (10, 4); (-3, 4); (-3, 0)\}$. The reference speed is $\bar{v}_t = 5\text{m.s}^{-1}$ and the discretization step $\Delta t = 0.05s$. No useful information about the initial position and orientation is available (the initial state is initialized randomly) and the initial velocity $v_0 = 0\text{m.s}^{-1}$.

The tracking problem will be set up with a receding horizon policy with horizon $N_0 = 20$ samples. That assumes a particular segment of the path known in advance; this is a reasonable assumption in many optimal path-following scenarios in practice. If an unexpected event or a supervision level suddenly changes the path (moving obstacles, task changes...), recalculation is needed and hence, a handful of past control actions would have been actually non-optimal for the new path, whatever the choice of control algorithm.

E. Results and discussion

Figure 1 shows a result of a single simulation. It can be seen that each of the controllers is able to reach and to keep the vehicle on the desired path. The realized loss and the time spent during each iteration of algorithms are listed in Table I. The loss is practically the same for iLQR and ERTS, however, the time spent is more than 7x higher for iLQR which makes on-line usage impractical except for very slow sampling rates. The fastest algorithm is linearized LQR, however ERTS is only 2x slower and provides a control with significantly better path following dynamics. Particularly, when the state is far away from the reference path, the differences between the LQR (naively linearized at the reference) and the iLQR and ERTS become larger.

The algorithm of ERTS is approximately 2x slower than linearized LQR because Algorithm 1 consist of two loops, one from t to N and other from N to t, meanwhile the linearized LQR iterates only from N to t. The operations are comparable in both cases. On the other hand, iLQR iterates until convergence using LQR-like recursions. In simulation, iLQR converged after ~ 13 iterations, which is in accordance with the time spent listed in Table I.

As a conclusion, ERTS controller is able to successfully operate outside the desired path, being several times faster and approximately equally accurate in comparison to iLQR.
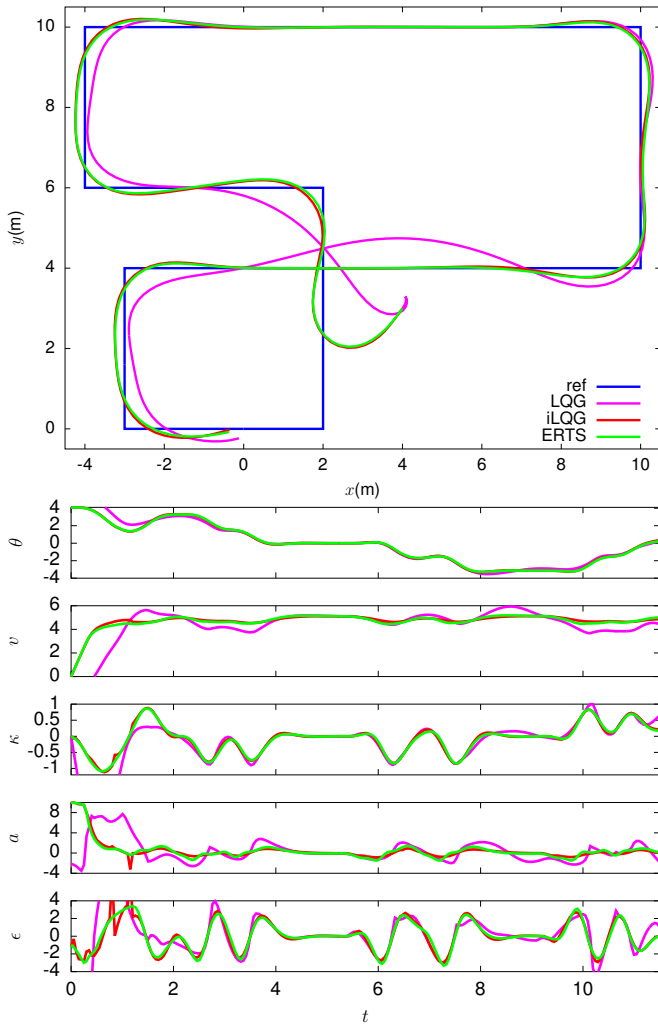
Fig. 1. Qualitative results of tracking along the desired path with the initial position outside the desired path. The displayed quantities are: coordinates (x_t, y_t), angle ϕ_t[rad], velocity (v_t)[m.s^{-1}], curvature κ_t[rad.m^{-1}], acceleration a_t[rad.m^{-2}], and derivative of curvature ϵ_t[rad.m^{-1}s^{-1}].

algorithm	realized cost	time per iter.	time per sim.
LQR	74947	0.01	2.38
iLQR	16396	0.14	31.59
ERTS	16284	0.02	4.46

TABLE I

THE TOTAL COST-TO-GO AND THE COMPUTATIONAL TIME[S].

V. CONCLUSIONS AND FUTURE WORKS

A new controller, coined as Extended Rauch-Tung-Striebel (ERTS) controller, was presented on the bases of the duality between control and estimation. The ERTS controller solves the control task via the transformation of the original problem to the dual estimation problem; the dual problem uses future reference states as observations and noise variances are obtained from the primal control cost index definition. The dual problem is solved by computationally efficient Rauch-Tung-Striebel smoother for linearized system.

In the linear case, the algorithm produces an exact equivalent to the well-known LQR control. This motivates the extended linearization-based approach suggested here.

The performance of the controller was studied on the path-following problem of a 5th-order vehicle and compared with linearized LQR and iterative LQR controller. ERTS proposes nearly the same control as iLQR controller after convergence, but in a fraction of the time. Hence, the combined accuracy and reduced computational cost makes ERTS an interesting solution for on-line control *with receding horizon policies.*

Future work will consider other nonlinear estimation paradigms such as unscented Uscented Kalman filter and particle-filter approaches for the dual estimation problem. We expect that these approaches would propose better performance than ERTS controller when more severe nonlinearities are present. Moreover, as the assumption on Gaussian noises in the dual estimation problem implies that both state and control action are unconstrained, the ERTS will fail on problems with hard constraints. On the other hand, a particle-filter based controller does not need such assumption. An implementation of a particle controller and case studies on problems with constraints are left for future work.

REFERENCES

[1] L. Armesto, V. Girbés, M. Vincze, S. Olufs, and P. Muñoz-Benavent. Mobile robot obstacle avoidance based on quasi-holonomic smooth paths. *Advances in Autonomous Robotics*, pages 244–255, 2012.

[2] A. Doucet, S. Godsill, and C. Andrieu. On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and computing*, 10(3):197–208, 2000.

[3] V. Girbés, L. Armesto, J. Tornero, and J.E. Solanes. Continuous-curvature kinematic control for path following problems. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pages 4335–4340. IEEE, 2011.

[4] K. Jiang, DZ Zhang, and LD Seneviratne. A parallel parking system for a car-like robot with sensor guidance. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 213(6):591–600, 1999.

[5] R.E. Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960.

[6] M. Kárnỳ. Towards fully probabilistic control design. *Automatica*, 32(12):1719–1722, 1996.

[7] S. Kullback and R.A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

[8] CT Leondes, JB Peller, and EB Stear. Nonlinear smoothing theory. *Systems Science and Cybernetics, IEEE Transactions on*, 6(1):63–71, 1970.

[9] M. Manz, F. von Hundelshausen, and H.J. Wuensche. A hybrid estimation approach for autonomous dirt road following using multiple clothoid segments. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, pages 2410–2415. IEEE, 2010.

[10] I. Papadimitriou and M. Tomizuka. Fast lane changing computations using polynomials. In *American Control Conference, 2003. Proceedings of the 2003*, volume 1, pages 48–53. IEEE, 2003.

[11] H.E. Rauch, F. Tung, and CT Striebel. Maximum likelihood estimates of linear dynamic systems. *AIAA journal*, 3(8):1445–1450, 1965.

[12] S. Sarkka. Unscented rauch–tung–striebel smoother. *Automatic Control, IEEE Transactions on*, 53(3):845–849, 2008.

[13] E. Todorov. General duality between optimal control and estimation. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 4286–4292. IEEE, 2008.

[14] E. Todorov and W. Li. A generalized iterative lqg method for locally-optimal feedback control of constrained nonlinear stochastic systems. In *American Control Conference, 2005. Proceedings of the 2005*, pages 300–306. IEEE, 2005.

[15] J. Van Den Berg, P. Abbeel, and K. Goldberg. Lqg-mp: Optimized path planning for robots with motion uncertainty and imperfect state information. *The International Journal of Robotics Research*, 30(7):895–913, 2011.

[16] M. Zima. Duality between estimation and optimal control. Master's thesis, Czech Technical University, 2013.