This is a preprint of a paper that appeared in IFAC-PapersOnLine, Vol. 50, No. 1, 2017

Decentralized Data-Driven Control of Cooperating Sensor-Carrying UAVs in a Multi-Objective Monitoring Scenario ⋆

Juliane Euler * **Oskar von Stryk** *

* *Technische Universität Darmstadt, Germany (e-mail: {euler,stryk}@sim.tu-darmstadt.de)*

Abstract: For estimating atmospheric dispersion of harmful material, the use of multiple sensor-equipped UAVs for information gathering offers great flexibility, but requires an efficient adaptive sampling strategy that exploits multi-vehicle cooperation. For this purpose, a novel decentralized data-driven online control scheme for cooperating vehicles in multi-objective monitoring scenarios is presented in this paper. In the considered use case, multiple UAVs are to adaptively gather measurements for estimating the parameters of an atmospheric dispersion model. At the same time, they are required to cooperatively patrol predefined checkpoints.
Vehicle-specific optimal waypoints for each UAV are determined by sequential optimum design. Following these waypoints leads to a maximized information gain of the acquired measurements, such that the parameter estimate is iteratively improved. On the other hand, checkpoint allocation as well as trajectory planning is provided by a decentralized model-predictive controller based on a discrete-time mixed-integer linear problem formulation.
By permanent interaction of parameter estimation, waypoint calculation, and cooperative control, a fully optimization-based, yet efficient and adaptive feedback control approach is obtained. Simulations successfully demonstrate its effectiveness.

Keywords: Multi-vehicle systems, Model predictive control of hybrid systems, Networks of sensors and actuators, Decentralized and distributed control, Dynamic Data-Driven Application Systems

1. INTRODUCTION

The use of mobile robotic systems for environmental monitoring and disaster response offers obvious benefits as they can gather important information even if the setting is not easily accessible by humans or even hostile. Mobile sensors, in particular, offer higher local precision, flexibility, and efficiency than stationary sensor networks. However, phenomena like the propagation of airborne contaminants represent large-scale dynamic spatio-temporal processes that can hardly be captured by a single mobile sensor platform. Therefore, cooperative monitoring by a group of sensor-equipped UAVs is investigated in this paper.

A twofold monitoring task is considered that combines adaptive sensing and cooperative patrol. The first objective is the identification of parameters of an atmospheric dispersion model based on measurement data gathered by the sensors. Outcome and efficiency of the measurement process can be maximized using an adaptive path planning approach that accounts for already available information as well as the UAVs' individual physical motion capabilities. This can be achieved by not only using new measurement data to update the parameter estimate, but by feeding the estimate back to the UAVs' motion controller in a way that maximizes the information gain of future measurements. This closed loop between application and measurements is characteristic for so called *Dynamic Data Driven Application Systems* (DDDAS) (Darema, 2004).

Model-based DDDAS in the context of atmospheric dispersion monitoring can be implemented in various ways: Peng and Mohseni (2014) utilize virtual attractor particles to guide multiple UAVs collecting data for a pollutant puff simulation. A loss function combining misclassification and mutual information is minimized by Šmídl and Hofman (2013) to choose the best option from a finite set of possible actions for each UAV. Similarly, Choi and How (2011) employ ensemble forecasting to determine grid-based measurement sequences for multiple sensor platforms, such that the mutual information is maximized. These approaches, however, do not explicitly incorporate the vehicles' motion dynamics.
One possibility to do so is to couple estimation and vehicle control in one complex optimal control problem (OCP) that minimizes an optimality criterion subject to ordinary differential equations representing the vehicle dynamics. In this context, Uciński (2005) provides a comprehensive theoretical background on optimum design for identifying distributed parameter systems. However, online adaptability to the gathered information is not possible in the feedforward OCP approach, the solution of which quickly becomes computationally intensive. In order to overcome these drawbacks, the OCP can repeatedly be solved in a receding horizon fashion, as proposed by Tricaud (2010). This presents an interlaced scheme of trajectory design for

⋆ This research has been supported in parts within the Hessian LOEWE initiative by the NICER project.

a finite time horizon, measurement collection, and data assimilation to update a parameter estimate. Haugen and Imsland (2015) present a similar approach employing an uncertainty measure instead of optimum design. While the feedback loop can be closed that way, both approaches depend on a central instance able to solve the OCP. As communication between a central unit and the vehicles is likely to be limited, decentralized solutions are preferable for real-world applications and offer the additional benefit of being scalable to larger vehicle teams.

None of the existing data-driven control schemes deal with additional mission objectives like cooperative patrol of predefined checkpoints as considered in this paper. A team of UAVs is to perform this typical inherently cooperative task while adapting their motion such that the measurements collected by their onboard sensors are most valuable for the dispersion process estimation. This is achieved by a fully decentralized approach based on a combination of optimum design of sensing trajectories and model-predictive cooperative control based on Mixed Integer Linear Programs (MILP). In contrast to the aforementioned solutions, both aspects are treated as separate problem formulations linked together in a sequential procedure.

Splitting up the OCP and using a MILP-based model-predictive controller has similarly been proposed by Euler et al. (2015) for state estimation of a dispersion process model governed by a partial differential equation. This paper modifies and extends the ideas of Euler et al. (2015) by employing a new optimization-based vehicle-specific sensing strategy and by adding a second mission objective. The full decentralization of the proposed solution makes it flexible and scalable with respect to the type and number of vehicles, and further increases its computational efficiency enabling online application on board each UAV.

2. PROBLEM STATEMENT AND OVERVIEW

Starting point of the considered application scenario is an instantaneous gas release, the evolution of which is to be predicted based on a Gaussian puff model of the dispersion process. This requires an accurate estimate of the model's parameters, which is obtained and repeatedly improved based on measurement data collected by a homogeneous team of n_V sensor-equipped quadrotor UAVs deployed in a working area $G \in \mathbb{R}^3$.

Quadrotor UAVs are modeled as point mass with motion dynamics described by the first order ODE

$$\dot{\mathbf{x}}_{qr} = \begin{pmatrix} v_x \\ v_y \\ v_z \\ u_x \\ u_y \\ u_z \end{pmatrix} = \mathbf{f}_{qr}(\mathbf{x}_{qr}, \mathbf{u}_{qr}) \ , \tag{1}$$

where the state vector $\mathbf{x}_{qr} = (x, y, z, v_x, v_y, v_z)^T$ contains the x/y/z positions and velocities, and the vector of control inputs $\mathbf{u}_{qr} = (u_x, u_y, u_z)^T)$ comprises its x/y/z accelerations. Euler discretization of (1) yields

$$\mathbf{x}_{qr}^{k+1} = \mathbf{x}_{qr}^{k} + \Delta t \cdot \mathbf{f}_{qr}(\mathbf{x}_{qr}^{k}, \mathbf{u}_{qr}^{k}) \ , \tag{2}$$

where Δt is the step size and the superscript k relates to the time step $t_k = k \cdot \Delta t$. The onboard sensor is assumed to recurringly measure every Δt_m seconds.

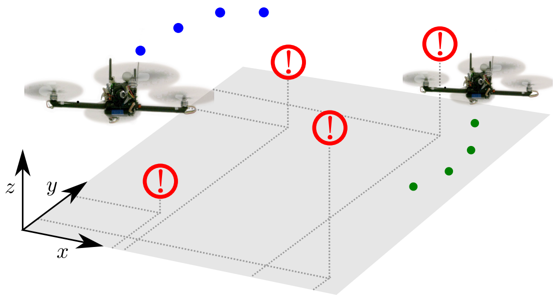

Fig. 1. Illustration of the considered problem. Sensor-equipped quadrotor UAVs recurringly take measurements while following individual, adaptively optimized waypoints. At the same time, points of special interest are to be visited regularly.

The UAVs' one task is to cooperatively patrol n_C a priori defined checkpoints $\in G$ representing locations of special interest, like train stations or hospitals, where accurate local concentration data are required. Checkpoints can be visited in any order at any time. However, in order to increase the incentive for frequent visits, each checkpoint is labeled with a penalty value p_c that increases with each time step t_k according to

$$p_c^{k+1} = p_c^k + \Delta p \ , \quad c = 1, 2, \ldots, n_C \tag{3}$$

while the checkpoint remains unvisited. It is reset to zero when a local measurement has been taken. As soon as the gas concentration at a checkpoint falls below a threshold $c_{\max}$, it is marked "done" and removed from the task.

In order to make the UAVs' flight paths most valuable with respect to the improvement of the parameter estimate, their other task is to follow sequences of vehicle-specific waypoints obtained from a sequential optimum design approach (Section 3). Figure 1 illustrates the scenario combining waypoint following and checkpoint patrol.

A model-predictive controller (MPC) for cooperative vehicle teams provides the control inputs to guide the UAVs along optimized trajectories trading-off the two mission aspects (Section 4). Motion control, waypoint calculation, and checkpoint penalty management is performed by each team member individually. The UAVs are assumed to exchange their current position, their individual measurement data, the checkpoint penalties as well as their current waypoint as long as they are within communication range d_{comm} to each other.

A schematic view of the interaction of parameter estimation, waypoint calculation, and cooperative control in the proposed adaptive sensing loop is shown in Fig. 2.

3. WAYPOINTS FOR OPTIMAL SENSOR MOTION

This section gives a brief overview of the dispersion model, the parameter estimation problem, and the waypoint calculation method as employed in the proposed monitoring solution. A more thorough description, analysis, and discussion of these components is provided in a separate article by the authors (Euler and von Stryk, 2017).

3.1 Gaussian puff model

Gas emissions in the atmosphere can be described by the advection-diffusion equation

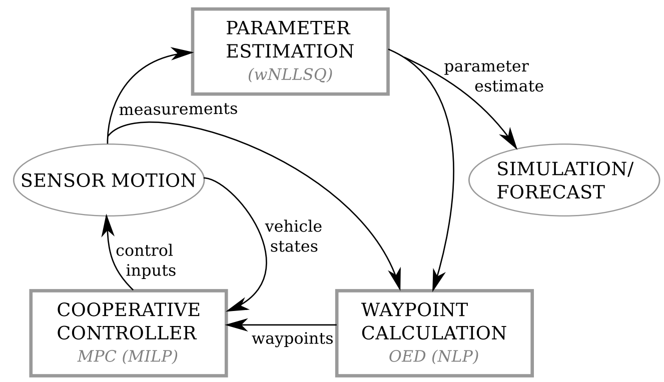

Fig. 2. The proposed dynamic data-driven sensing loop.

$$\frac{\partial C}{\partial t} = -\nabla \cdot \mathbf{q} \ , \tag{4}$$

where $C(\mathbf{x}, t)$ is the concentration in [kg/m^3] at location $\mathbf{x} \in \mathbb{R}^3$ at time t and $\mathbf{q}$ is the mass flux in [kg/m^{2}s]. Under simplifying assumptions detailed by Kathirgamanathan et al. (2002), by Laplace transform one can obtain

$$C(\boldsymbol{\theta}, x, y, z, t) = \frac{Q}{8\pi^{\frac{3}{2}}(K_x K_y K_z)^{\frac{1}{2}} \Delta t^{\frac{3}{2}}} \cdot \\ \cdot e^{-\frac{(\Delta x - u\Delta t)^2}{4K_x \Delta t} - \frac{\Delta y^2}{4K_y \Delta t}} \cdot \left(e^{-\frac{(z-z_0)^2}{4K_z \Delta t}} + e^{-\frac{(z+z_0)^2}{4K_z \Delta t}} \right) \ , \tag{5}$$

as the so called *Gaussian puff solution* of (4). Here, Q [kg] is the total mass of the instantaneous release, (x_0, y_0, z_0) is the location of the source in [m], t_0 [s] the time of the release, $\Delta x = x - x_0$, $\Delta y = y - y_0$, $\Delta t = t - t_0$, $\mathbf{v}_w = (u, 0, 0)$ is the x/y/z wind velocity in [m/s], and $K_x(\mathbf{x}), K_y(\mathbf{x}), K_z(\mathbf{x})$ [m^2/s] are turbulent eddy diffusivities. Further assuming that $K_x = K_y$, and that K_z can be derived from the atmospherical conditions, the vector of unknown parameters for the Gaussian puff solution (5) is $\boldsymbol{\theta} = (Q, K_x, x_0, y_0, z_0, t_0)^T$.

3.2 Parameter Estimation

Let $(\mathbf{p}_i, v_i)$, $i = 1, 2, \ldots, m$, be m data points with measurements v_i defined as

$$v_i = C_i(\boldsymbol{\theta}_{\text{true}}) + \epsilon_i, \qquad i = 1, 2, \ldots, m \ , \tag{6}$$

where $C_i(\boldsymbol{\theta}_{\text{true}}) = C(\boldsymbol{\theta}_{\text{true}}, \mathbf{p}_i)$ as defined by (5) and $\boldsymbol{\theta}_{\text{true}}$ is the true parameter vector to be estimated. $\mathbf{p}_i = (x_i, y_i, z_i, t_i)$ contains the coordinates of the ith measurement in space and time. The measurement error ϵ_i is assumed to be white Gaussian noise with mean $\mathrm{E}(\epsilon_i) = 0$ and variance $\mathrm{var}(\epsilon_i) = \sigma_i^2$. Following the problem definition of Christopoulos and Roumeliotis (2005), the standard deviation σ_i of the error is $\sigma_i = \frac{v_i}{\alpha}$ with a constant signal to noise ratio (SNR) $\alpha^2 = 1000$. It is further assumed that measurement errors are uncorrelated, i.e. $\mathrm{E}(\epsilon_i \epsilon_j) = \mathrm{cov}(\epsilon_i, \epsilon_j) = 0$ for $i \neq j$.

The following *weighted nonlinear least squares* (wNLLSQ) problem is solved to estimate the parameter vector $\boldsymbol{\theta}$:

$$\min_{\boldsymbol{\theta}} \frac{1}{2} \left\| \mathbf{W} \cdot (\mathbf{v} - \mathbf{C}(\boldsymbol{\theta})) \right\|_2^2 \ . \tag{7}$$

$\mathbf{v} = (v_1, v_2, \ldots, v_m)^T$ comprises the concentration measurements at $\mathbf{p}_1, \ldots, \mathbf{p}_m$ and $\mathbf{C}(\boldsymbol{\theta}) = (C_1(\boldsymbol{\theta}), \ldots, C_m(\boldsymbol{\theta}))^T$ combines the corresponding model values. Since the standard deviations of the measurement errors are not identical, each residual $v_i - C_i(\boldsymbol{\theta})$ is weighted by the reciprocal of σ_i. Hence, $\mathbf{W} = \mathrm{diag}(\sigma_1^{-1}, \sigma_2^{-1}, \ldots, \sigma_m^{-1}) \in \mathbb{R}^{m \times m}$.

3.3 Sequential Optimum Design

Let $\bar{\boldsymbol{\theta}}$ be the result of the parameter estimation problem (7). Since the Gaussian puff model (5) is nonlinear in the parameters $\boldsymbol{\theta}$, an optimum design will depend on these parameters and, thus, on the quality of the estimate $\bar{\boldsymbol{\theta}}$. In order to compensate negative effects of bad estimates $\bar{\boldsymbol{\theta}}$ on the design, a sequential design approach (Atkinson et al. (2007)) is employed in this paper.

In a first step, a linearization of (5) by Taylor series expansion about $\bar{\boldsymbol{\theta}}$ combining all measurement data yields

$$\mathbf{C}(\boldsymbol{\theta}) = \mathbf{C}(\bar{\boldsymbol{\theta}}) + \mathbf{J}_C \cdot (\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}) \tag{8}$$

with the Jacobian of $\mathbf{C}(\bar{\boldsymbol{\theta}})$ denoted as

$$\mathbf{J}_C = (\nabla C_1(\bar{\boldsymbol{\theta}}), \nabla C_2(\bar{\boldsymbol{\theta}}), \ldots, \nabla C_m(\bar{\boldsymbol{\theta}}))^T \ .$$

From that the *information matrix* $\mathbf{M} = \mathbf{J}_C^T \mathbf{W}^2 \mathbf{J}_C$ is obtained with $\mathbf{W}$ as in (7). While $\mathbf{M}$ represents information gathered in the past, the question is how the existing measurement sequence can be extended in order to best improve the parameter estimate $\bar{\boldsymbol{\theta}}$. In other words, n_w optimized new measurement locations $\tilde{\mathbf{p}}_{v,1}, \ldots, \tilde{\mathbf{p}}_{v,n_w}$ as waypoints for a UAV v are required in order to obtain an *optimum design* for the linearized model. For this purpose, $\mathbf{M}$ is extended such that

$$\mathbf{M}_{\text{ext}} = \mathbf{M} + \nabla C(\bar{\boldsymbol{\theta}}, \tilde{\mathbf{p}}_{v,1}) \nabla C(\bar{\boldsymbol{\theta}}, \tilde{\mathbf{p}}_{v,1})^T + \\ \cdots + \nabla C(\bar{\boldsymbol{\theta}}, \tilde{\mathbf{p}}_{v,n_w}) \nabla C(\bar{\boldsymbol{\theta}}, \tilde{\mathbf{p}}_{v,n_w})^T \ . \tag{9}$$

That way, $\mathbf{M}_{\text{ext}}$ becomes dependent on the waypoints to be determined. Based on $\mathbf{M}_{\text{ext}}$, a *D-optimality criterion* is minimized subject to the UAV's motion dynamics model:

$$\min_{\tilde{\mathbf{p}}_{v,1}, \tilde{\mathbf{p}}_{v,2}, \cdots \tilde{\mathbf{p}}_{v,n_w}} \quad -\log\det(\mathbf{M}_{\text{ext}} + \mu \mathbf{I}) \tag{10a}$$

$$\text{s.t.} \quad \mathbf{x}_v^{k+1} = \mathbf{x}_v^k + \Delta t \cdot \mathbf{f}_{qr}(\mathbf{x}_v^k, \mathbf{u}_v^k) \tag{10b}$$

$$\tilde{\mathbf{p}}_{v,k+1} = \mathbf{H}\mathbf{x}_v^{k+1} \tag{10c}$$

$$\mathbf{x}_{\min} \leq \mathbf{x}_v^k \leq \mathbf{x}_{\max} \tag{10d}$$

$$\mathbf{u}_{\min} \leq \mathbf{u}_v^k \leq \mathbf{u}_{\max} \ , \tag{10e}$$

where $k = 0, 1, \ldots, n_w - 1$. The term $\mu\mathbf{I}$ with small $\mu > 0$ is used for regularization. Constraint (10c) is not implemented, but stated here to illustrate the extraction of the waypoint coordinates from the overall vehicle state vector. The waypoints can be tailored to the current or a predicted vehicle state by selecting the initial value $\mathbf{x}_v^0$ accordingly. This allows collateral precalculation of new waypoints while the UAV is still collecting data along previously determined waypoints.

Whenever additional measurement data is available, it is incorporated to obtain an updated parameter estimate $\bar{\boldsymbol{\theta}}$ from (7). Based on the new $\bar{\boldsymbol{\theta}}$, the linearization (8) can be recalculated and the sequential design restarts.

Overcoming the dependence of designs for nonlinear models on their parameters by linearizing about a point prior $\bar{\boldsymbol{\theta}}$ and using the Fisher information metric as proposed here is a standard approach in optimum design, which is attractive in its simplicity and computational efficiency. It provides locally optimal designs that maximize the expected information about the parameters. However, prior information about $\boldsymbol{\theta}$ cannot be explicitly incorporated as it would be the case in a more complex Bayesian optimum design formulation. Bayesian designs account for prior distributions of $\boldsymbol{\theta}$, are less prone to clusterization of design points, and aim at minimizing the expected variance of the parameter estimate (cf. e.g. Atkinson et al. (2007)).

4. DECENTRALIZED COOPERATIVE CONTROL

The versatile cooperative model-predictive control (MPC) approach employed in this paper was first presented by Kuhn et al. (2011) and is here adapted to the considered monitoring scenario involving sensor UAVs, waypoints, and checkpoints. Core of the approach is a Mixed Integer Linear Program (MILP) of the general form

$$\min_{U_N} \quad \sum_{k=0}^{N} \mathbf{d}_x^T \mathbf{x}^k + \sum_{k=0}^{N-1} \mathbf{d}_u^T \mathbf{u}^k \tag{11a}$$

$$\text{s.t.} \quad \mathbf{x}^{k+1} = \mathbf{A}\mathbf{x}^k + \mathbf{B}\mathbf{u}^k + \mathbf{b} \tag{11b}$$

$$\mathbf{C}\mathbf{x}^k + \mathbf{D}\mathbf{u}^k \leq \mathbf{c} \tag{11c}$$

$$\mathbf{x}_{\min} \leq \mathbf{x}^k \leq \mathbf{x}_{\max} \tag{11d}$$

$$\mathbf{u}_{\min} \leq \mathbf{u}^k \leq \mathbf{u}_{\max} \ , \tag{11e}$$

where $\mathbf{x} = (\mathbf{x}_c, \mathbf{x}_b)^T$, $\mathbf{x}_c \in \mathbb{R}^{n_c}$, $\mathbf{x}_b \in \{0,1\}^{n_b}$, is the system state vector containing continuous as well as binary variables and $\mathbf{u} \in \mathbb{R}^{n_u}$ is the vector of vehicle control inputs. By choosing the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and $\mathbf{D}$ and the vectors $\mathbf{b}, \mathbf{c}, \mathbf{d}_x$, and $\mathbf{d}_u$ accordingly, a mathematical representation of the cooperative mission is obtained. It combines continuous motion dynamics and discrete decision logic and is used to predict how the system will evolve in future time steps $t+k\cdot\Delta t$, $k = 1, 2, \ldots, N$. Based on the MILP, the model-predictive controller optimizes a sequence of future control inputs $U_N := \{\mathbf{u}^k\}_{k=0}^{N-1}$ for the UAVs such that the performance criterion (11a) is minimized. The first element of the sequence is applied at time t and the procedure repeats with the new system state at time $t+\Delta t$. By updating the control inputs in this manner, an optimization-based closed-loop cooperative control law is obtained.

4.1 Model Constraints

For the considered monitoring task, the MILP constraints (11b)–(11c) are composed of the following elements:

Motion Dynamics Let $\tilde{n}_V \leq n_V$ be the number of UAVs considered in the model. For each quadrotor $v = 1, 2, \ldots, \tilde{n}_V$, the discretized motion dynamics equations (2) are included as problem constraints.

Distances to Waypoints The model contains a linear approximation of the Euclidean distance $d_{vw_v}^k \approx \sqrt{(x_v^k - x_{w_v})^2 + (y_v^k - y_{w_v})^2 + (z_v^k - z_{w_v})^2}$ between a vehicle v and its current waypoint w_v. Minimizing $\sum_{k=0}^{N} d_{vw_v}^k$ in the objective function, leads the UAV to the desired location.

Distances to Checkpoints As for the waypoints, linearized distances d_{vc}^k between a UAV v and a checkpoint c are modeled for each possible (v, c) combination.

Cooperation Logic For each pair (v, c), a binary variable $b_{vc}^k \in \{0, 1\}$ indicates whether or not UAV v measures at checkpoint c at time step k:

$$b_{vc}^k = 1 \quad \Leftrightarrow \quad d_{vc}^k \leq d_{\text{check}} \ , \tag{12}$$

where d_{check} determines how close a UAV has to be in order to complete the checkpoint.

Since it is not important which UAV visits checkpoint c, but that a measurement is eventually taken there by any of the sensors, another binary variable $s_c^k \in \{0,1\}$ is introduced and represents the status of c at time step k:

$$s_c^k = 0 \quad \Leftrightarrow \quad \sum_{v=1}^{\tilde{n}_V} b_{vc}^k \geq 1 \ . \tag{13}$$

Cooperation among the UAVs is realized by minimizing the penalties of those checkpoints currently not visited by any team member. This selection is modeled using the binary variables s_c^k and a set of auxiliary variables $h_c^k \in \mathbb{R}$ that either equal the penalties p_c^k or equal zero, depending on the status of checkpoint c:

$$h_c^k = s_c^k \cdot p_c^k \ . \tag{14}$$

Penalty increase over time is modeled as stated in Sec. 2:

$$p_c^{k+1} = h_c^k + s_c^k \cdot \Delta p \ , \tag{15}$$

where p_c^0 equals the current checkpoint penalty at the time of the controller call.

To enable the formulation of another incentive for checkpoint visits, a variable representing the maximum of all penalties over the prediction horizon N is introduced:

$$p_m \geq p_c^k \ . \tag{16}$$

The constraints (12), (13), (14) are transformed into mixed-integer linear constraints using the modeling rules introduced by Bemporad and Morari (1999). For more details, also on the linearization of distances and a possible model extension for collision avoidance, the reader is referred to previous work by the authors (e.g. Kuhn et al. (2011); Ritter et al. (2014)).

4.2 Optimization Criteria

The performance criterion (11a) has to represent the trade-off between the two conflicting objectives for each UAV:

(1) *Follow the waypoints* for most informative sensing.
(2) *Visit the checkpoints* to provide accurate local concentration data. Select checkpoints according to their priority represented as penalties.

The first objective requires no explicit cooperation as every vehicle follows its individual waypoints which can easily be represented as minimization of

$$J_1 = \sum_{v=1}^{\tilde{n}_V} \sum_{k=0}^{N} d_{vw_v}^k \ . \tag{17}$$

Modeling the second objective, however, has to ensure that no two UAVs visit the same checkpoint and has to account for the checkpoints' different priorities. This is achieved by a combination of three criteria:

$$J_2 = \sum_{v=1}^{\tilde{n}_V} \sum_{c=1}^{n_c} \sum_{k=0}^{N} \frac{1}{n_c} \cdot d_{vc}^k \tag{18a}$$

$$J_3 = \sum_{c=1}^{n_c} \sum_{k=0}^{N} \frac{1}{n_c \cdot p_{\max}^0} \cdot h_c^k \tag{18b}$$

$$J_4 = \frac{N}{p_{\max}^0} \cdot p_m \ , \tag{18c}$$

where $p_{\max}^0 = \max_c p_c^0$. J_2 can be seen as an auxiliary objective preventing the UAVs to drift away from the checkpoints. The factor $\frac{1}{n_c}$ is used to normalize the sum of distances in relation to J_1. J_3 represents the sum of penalties of all unvisited checkpoints over the time horizon N, which is to be minimized. In addition, the maximum penalty p_m denoted as J_4 is to be kept at a minimum. The factors $\frac{1}{n_c \cdot p_{\max}^0}$ in (18b) and $\frac{N}{p_{\max}^0}$ in (18c) normalize J_3 in relation to J_4. Both objectives intend to distribute the UAVs among the checkpoints and reduce the time between consecutive visits.

The four objectives are weighted and combined in the overall performance criterion

$$\min k_1 \cdot J_1 + k_2 \cdot J_2 + k_3 \cdot J_3 + k_4 \cdot J_4 \tag{19}$$

with $k_i \in \mathbb{R}^+$, which is optimized subject to the constraints listed in Sec. 4.1. This weighted sum formulation is used in order to avoid infeasibility of the MILP that is likely to occur for short prediction horizons N if the checkpoint patrol requirements were treated as hard constraints instead.

4.3 Decentralized Data-Driven Control

From a single vehicle's point of view, only an excerpt of the overall problem is relevant for planning its action for the next time steps. This subproblem involves the UAV itself and teammates within its communication range d_{comm}. Therein, the UAVs share their current state $\mathbf{x}_{qr}$ as well as their individual current waypoint. Checkpoint locations, their initial penalty values, and Δp, however, are assumed to be globally known from the mission start. The UAVs individually increment the checkpoints' penalties, but communicate changes in their completion status.

Based on the shared information, for each team member, a model (11) of its local subsystem is set up at runtime at every cycle of its individual MPC loop with control rate Δt. Optimal behavior in terms of (19) for all UAVs in the subsystem is computed, from which only the control input for the currently considered quadrotor is actually applied. As long as the UAVs' communication ranges overlap, this scheme ensures that no two vehicles head for the same checkpoint. Flying along the optimized trajectory, a UAV's onboard sensor takes measurements at rate Δt_m. Every new measurement triggers a recalculation of the UAV's individual parameter estimate $\bar{\boldsymbol{\theta}}$ also including the measurement data of all team mates within communication range. Also the waypoint calculation is performed by each UAV individually. It is frequently updated based on a predicted vehicle state, the UAV's current parameter estimate and all currently known measurement data. The waypoints are, one by one, used by the cooperative controller to provide appropriate motion trajectories, which closes the data-driven sensing loop as depicted in Fig. 2.

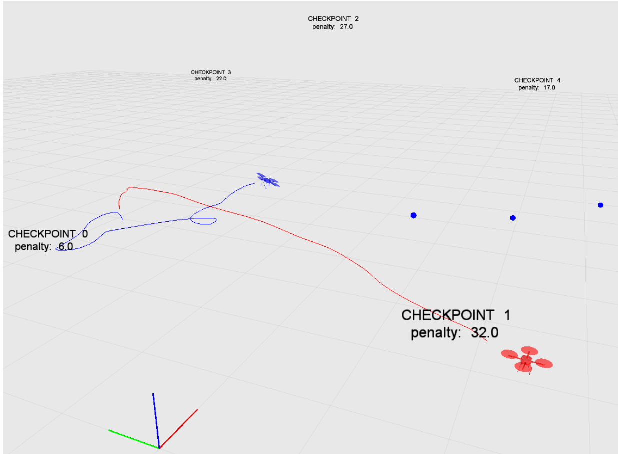

Fig. 3. Snapshot of a simulation involving 2 quadrotors and 5 checkpoints visualized in RVIZ. The blue spheres represent the waypoints calculated for the blue quadrotor. The blue and red lines indicate the quadrotors' motion up to the time of the snapshot.

5. EVALUATION

5.1 Implementation

The simulation framework used to evaluate the proposed solution to the multi-objective monitoring problem is implemented employing the Robot Operating System (ROS [1]) and Gazebo [2] as multi-robot simulator. Each UAV is represented by a collection of several ROS nodes, each responsible for a certain functionality, communicating their inputs and outputs via ROS topics or ROS services. A comprehensive description is omitted here, but a few noteworthy details are given below.

Core of each UAV's motion control is the node implementing the MILP-based MPC introduced in Sec. 4. During each control cycle, the MILP representing the vehicle's individual subproblem is reassembled based on the current state of its local environment. It is solved over a prediction horizon of $N = 10$ using Gurobi (Gurobi Opt. Inc., 2016). In order to obtain an efficient online control strategy, a maximum number of vehicles $\tilde{n}_{V\max} = 3$ in a model may not be exceeded. If there is information on more vehicles available than the defined maximum, only those closest to the controlled UAV are included in the subsystem model. For the MPC, the quadrotor motion is modeled as simple double integrator dynamics (cf. (1)). However, the obtained control inputs are applied to the comprehensive and realistic Gazebo model hector_quadrotor_gazebo (Meyer et al., 2012) simulating the motion of the real quadrotor UAVs. The simulation is visualized using the 3D visualization tool RVIZ, see Fig. 3 for a snapshot.

The node responsible for a UAV's waypoint management initiates a new waypoint calculation whenever the vehicle is heading for the last of the 4 waypoints determined in the previous optimization. In order to keep up a fluent vehicle motion, based on the current sequence of control inputs provided by the MPC, a prediction of the vehicle's state 3 time steps ahead is made and used as initial value for the

[1] `http://www.ros.org`
[2] `http://www.gazebosim.org/`

Table 1. Simulation parameters

parameter	value	unit
domain size	$F = 500$	[m]
initial UAV positions	$\mathbf{x}_0 = (20, 20, 10)$ $\mathbf{x}_1 = (25, 25, 10)$	[m]
maximum velocity	$v_{\max} = 8$	[m/s]
maximum acceleration	$u_{\max} = 3$	[m/s^2]
communication range	$d_{\mathrm{comm}} = 1000$	[m]
sensing rate	$\Delta t_m = 2$	[s]
checkpoint locations and initial penalties	(10, 20, 10, 20) (10, -20, 10, 15) (60, 10, 30, 10) (70, 40, 20, 5) (80, -20, 20, 0)	$[\mathrm{m}, \mathrm{m}, \mathrm{m}, -]$
penalty increase	$\Delta p = 1$	
measurement range	$d_{\mathrm{check}} = 5$	[m]
objective weights	$k_1 = 2, k_2 = 1$ $k_3 = k_4 = 2000$	
true puff parameters	$\boldsymbol{\theta}_{\mathrm{true}} = \begin{pmatrix} Q \\ K_x \\ x_0 \\ y_0 \\ z_0 \\ t_0 \end{pmatrix} = \begin{pmatrix} 1000 \\ 12 \\ 2 \\ 5 \\ 0 \\ -100 \end{pmatrix}$	[kg] [m^2/s] [m] [m] [m] [s]
initial estimate	$\bar{\boldsymbol{\theta}}_0 = \begin{pmatrix} Q \\ K_x \\ x_0 \\ y_0 \\ z_0 \\ t_0 \end{pmatrix} = \begin{pmatrix} 700 \\ 20 \\ 40 \\ 25 \\ 10^{-12} \\ -70 \end{pmatrix}$	[kg] [m^2/s] [m] [m] [m] [s]
diffusivity parameter	$K_z = 0.2113$	$[m^2/s]$
wind speed	$u = 0.5$	[m/s]

new waypoint optimization. SNOPT 7.5 (Gill et al., 2015) is employed to solve problem (10).

Information exchange among the UAVs is routed through a node modeling their communication based on their distances to each other. Communication loss or delays could also be modeled, but is left for future research.

5.2 Results

Simulation runs over a time period of 150 sec were conducted with the framework described in Sec. 5.1 for three different scenario versions: waypoints only, checkpoints only, and the full multi-objective scenario involving both waypoints and checkpoints. For each setup, 2 sensor-equipped UAVs moving within a $[-F, F] \times [-F, F] \times [0, F/2] \subset \mathbb{R}^3$ domain were considered. It is assumed that the simulation starts at time $t = 0$ while the puff release happened at some point in the past. Table 1 summarizes all relevant simulation parameters.

The entire simulation involving two quadrotors and the corresponding recurring solver calls for waypoint generation and vehicle control, respectively, was run on one single machine (Intel®Core™i7-4790K CPU @ 4.00GHz, 16 GB RAM) with real-time factor 1.

By updating the control inputs every Δt seconds, the approximation error between the quadrotor model used in the MILP formulation and the more complex Gazebo model is compensated, permitting the assumption that real quadrotor UAVs could also be controlled in that way. Moreover, it has to be considered that the UAVs do not take measurements at the exact waypoint locations. The MPC guides them such that distances to waypoints are minimized along with other possibly contrasting objectives. Hence, the measurements taken every Δt_m seconds can only be assumed close to the optimized locations.

After each measuremet, a UAV requests its team mates' measurement data and based on the collective information updates the parameter estimate. For each scenario, the RMSE of the parameter estimate over time was evaluated in 10 simulation runs. The results are shown in Fig. 4. As expected, the best error reduction is obtained in the waypoints-only scenario, where optimal parameter estimation is the only objective and the UAVs can concentrate on following their waypoints. In the checkpoints-only scenario, the UAVs focus on cooperative patrol, no waypoints are calculated. Still, measurements are taken every Δt_m seconds and the estimate is updated accordingly. That way, the RMSE is not reduced as much as in the waypoints-only case but still fairly well, possibly due to the checkpoints being conveniently located in the puff's dispersion domain. In the multi-objective scenario, the waypoints deviate the UAVs from heading straight for a checkpoint. Their trajectories are drawn to regions, where more valuable measurements can be gathered. This trade-off is reflected by the RMSE lying between the two single-objective scenarios at the end of the simulated time period. It can be expected, that this effect would be more obvious, if the checkpoints were located further away from the puff's center.

The conflict between the mission objectives is also reflected by the mean and maximum penalty values of the checkpoints depicted in Fig. 5. As can be seen, the maximum penalty over time is significantly lower in the checkpoints-only scenario since the vehicles patrol the checkpoints more quickly.

In Fig. 6, the convergence of the single parameters is plotted over the number of estimate updates. It shows that all of the parameters are estimated comparably well. However, some of the curves are monotonically decreasing, while others start with a large peak before converging to the true parameter value. This peak appears accordingly in Fig. 4 and is due to the low number and low quality of the first measurements taken near the unoptimized initial UAV positions.

6. CONCLUSIONS

An efficient decentralized data-driven control approach for cooperating sensor vehicles was presented in this paper. It combines a new waypoint generation approach for maximizing the informativeness of measurements individually tailored to each vehicle with a versatile decentralized cooperative controller capable of dealing with two conflicting mission objectives in a dynamically changing environment. Simulation results prove the applicability, effectiveness, and efficiency of the proposed approach.

Future research directions include the extension to larger, possibly heterogeneous, vehicle teams as well as the investigation of effects of communication loss and/or delays among the team members.

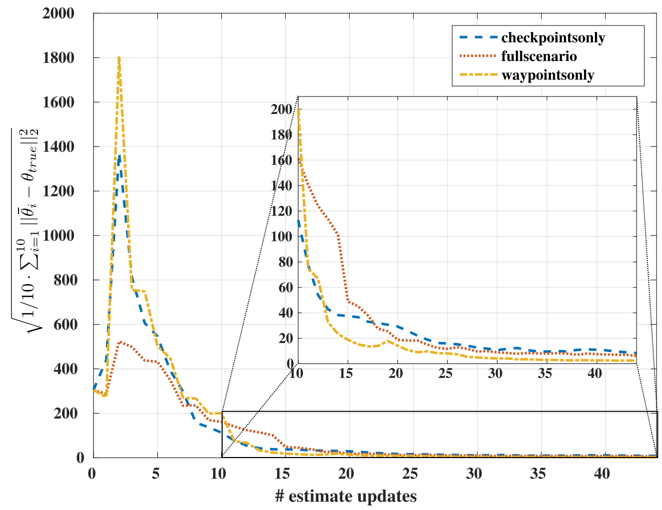

Fig. 4. RMSE in 10 simulation runs for three scenarios: checkpoints only (blue −−), waypoints only (yellow −·), and both combined (red ···).

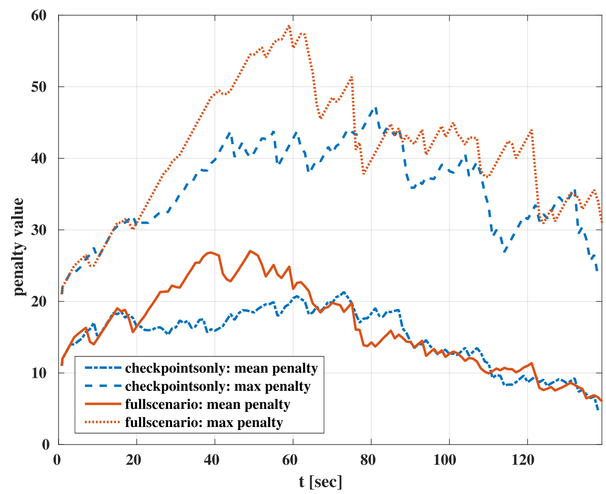

Fig. 5. Mean of the maximum and mean checkpoint penalties in 10 simulation runs for the checkpoints-only and the full scenario.

REFERENCES

Atkinson, A., Donev, A.N., and Tobias, R.D. (2007). *Optimum Experimental Designs, With SAS*. OUP Oxford.

Bemporad, A. and Morari, M. (1999). Control of systems integrating logic, dynamics, and constraints. *Automatica*, 35, 407–427.

Choi, H.L. and How, J.P. (2011). Coordinated Targeting of Mobile Sensor Networks for Ensemble Forecast Improvement. *IEEE Sensors Journal*, 11(3), 621–633.

Christopoulos, V.N. and Roumeliotis, S. (2005). Multi Robot Trajectory Generation for Single Source Explosion Parameter Estimation. In *IEEE ICRA*, 2803–2809.

Darema, F. (2004). DDDAS: A new paradigm for application simulations and measurements. In *ICCS*, volume 3038 of *LNCS*, 662–669. Springer.

Euler, J., Ritter, T., Ulbrich, S., and von Stryk, O. (2015). Centralized Ensemble-Based Trajectory Planning of Cooperating Sensors for Estimating Atmospheric Dispersion Processes. In S. Ravela and A. Sandu (eds.), *DyDESS*, volume 8964 of *LNCS*. Springer, 1 edition.

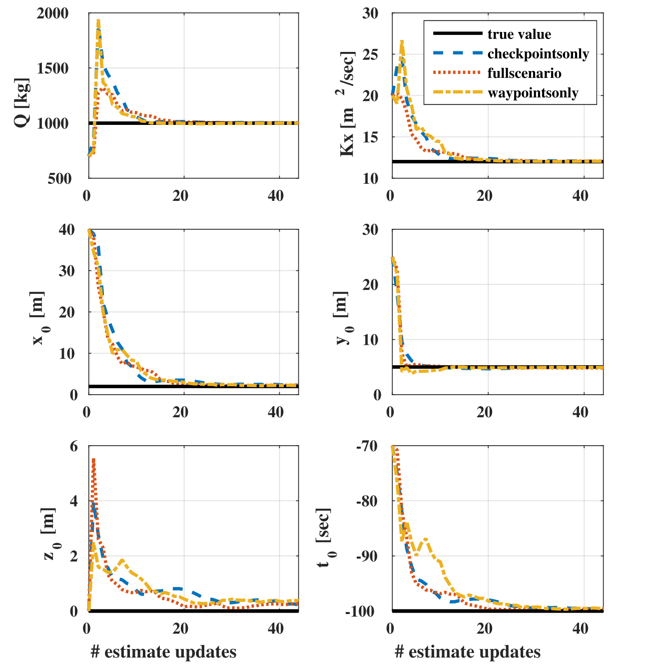

Fig. 6. Mean convergence of the single parameters in 10 simulation runs for the three scenario versions.

Euler, J. and von Stryk, O. (2017). Optimized Vehicle-Specific Trajectories for Cooperative Process Estimation by Sensor-Equipped UAVs. In *IEEE ICRA (accepted)*.

Gill, P.E., Wong, E., Murray, W., and Saunders, M.A. (2015). User's Guide for SNOPT Version 7.5. Technical report, Ctr. Comp. Math., UC San Diego.

Gurobi Opt. Inc. (2016). Gurobi Optimizer Ref. Manual.

Haugen, J. and Imsland, L. (2015). Monitoring an Advection-Diffusion Process Using Aerial Mobile Sensors. *Unmanned Systems*, 3(3), 221–238.

Kathirgamanathan, P., Mckibbin, R., and McLachlan, R. (2002). Source Term Estimation of Pollution from an Instantaneous Point Source. *MODSIM*, 6, 59–67.

Kuhn, J., Reinl, C., and von Stryk, O. (2011). Predictive Control for Multi-Robot Observation of Multiple Moving Targets Based on Discrete-Continuous Linear Models. In *IFAC WC*.

Meyer, J., Sendobry, A., Kohlbrecher, S., Klingauf, U., and von Stryk, O. (2012). Comprehensive Simulation of Quadrotor UAVs Using ROS and Gazebo. In I. Noda, N. Ando, D. Brugali, and J.J. Kuffner (eds.), *IEEE SIMPAR*, volume 7628 of *LNCS*, 400–411. Springer.

Peng, L. and Mohseni, K. (2014). Sensor driven feedback for puff estimation using unmanned aerial vehicles. In *ICUAS*, 562–569.

Ritter, T., Euler, J., Ulbrich, S., and von Stryk, O. (2014). Adaptive Observation Strategy for Dispersion Process Estimation Using Cooperating Mobile Sensors. In *IFAC WC*.

Šmídl, V. and Hofman, R. (2013). Tracking of atmospheric release of pollution using unmanned aerial vehicles. *Atmospheric Environment*, 67, 425–436.

Tricaud, C. (2010). *Optimal Sensing and Actuation Policies for Networked Mobile Agents in a Class of Cyber-Physical Systems*. PhD Thesis.

Uciński, D. (2005). *Optimal Measurement Methods for Distributed Parameter System Identification*. CRC Press.