Distributed Greedy Sensor Scheduling for Model-based Reconstruction of Space-Time Continuous Physical Phenomena

Marco F. Huber, Achim Kuwertz, Felix Sawo, and Uwe D. Hanebeck

Intelligent Sensor-Actuator-Systems Laboratory (ISAS),
Institute for Anthropomatics,
Universität Karlsruhe (TH), Germany.
{marco.huber, uwe.hanebeck}@ieee.org, {kuwertz, sawo}@ira.uka.de

Abstract – *A novel distributed sensor scheduling method for large-scale sensor networks observing space-time continuous physical phenomena is introduced. In a first step, the model of the distributed phenomenon is spatially and temporally decomposed leading to a linear probabilistic finite-dimensional model. Based on this representation, the information gain of sensor measurements is evaluated by means of the so-called covariance reduction function. For this reward function, it is shown that the performance of the greedy sensor scheduling is at least half that of the optimal scheduling considering long-term effects. This finding is the key for distributed sensor scheduling, where a central processing unit or fusion center is unnecessary, and thus, scaling as well as reliability is ensured. Hence, greedy scheduling in combination with a proposed hierarchical communication scheme requires only local sensor information and communication.*

Keywords: Sensor scheduling, distributed estimation, Kalman filter, submodular functions.

1 Introduction

Recent developments in wireless communication and sensor technologies facilitate the usage of large-scale sensor networks for monitoring space-time continuous physical phenomena. Examples of such phenomena, i.e., physical quantities varying in time and space, include temperature distributions, pollution concentrations, or fluid flows. By densely deploying large numbers of sensor nodes inside a phenomenon, these quantities can be cooperatively acquired. Individual sensors, however, are only able to capture distributed quantities at discrete locations and points of time. For feasibly quantifying the phenomena not only at sensor locations and measurement times, a reconstruction has to be performed. By exploiting background information in form of a physical model, a *model-based reconstruction* [1, 2] approach as stated in Sec. 3 can be employed to completely estimate a phenomenon, even for a limited number of

discrete measurements. By using Bayesian estimation techniques, this stochastic approach inherently considers noise and modeling uncertainties.

With a large number of sensors available, the best phenomenon reconstruction is obtained by employing all sensors at each measurement time. However, acquired measurement values have to be exchanged among the nodes. In wireless networks, this may lead to message collisions and retransmissions, which unnecessarily deplete the limited energy resources of the sensor nodes. A possible reduction of sensor usage improves the overall life-time of a sensor network, while on the other hand, causes estimation accuracy to decrease. Thus, a trade-off between resource utilization and estimation quality has to be found, which is addressed by sensor scheduling. *Sensor scheduling*, also known as sensor management, aims at determining an optimal sensor schedule, i.e., a time sequence of sensor nodes scheduled for future measurements, to maximize the information gain in presence of resource restrictions, and therefore can be seen as some kind of time-multiplexing on the wireless network.

In Sec. 4 of this paper, we consider sensor scheduling as a stochastic control problem, assuming the monitored phenomenon to be given as a linear stochastic system and assuming a covariance-based reward function to be employed for rating sensor schedules. Then, the optimal schedule can be determined a priori and independently of actual measurements, which basically is possible by an exhaustive off-line tree search [3]. However, even for a fixed time horizon N, this off-line search is NP-hard [4] and thus, in general computationally infeasible. To deal with the computational complexity, approximate algorithms have been proposed in literature, including pruning techniques [5] and greedy/myopic algorithms [6, 7, 8], which calculate only a one-step ahead solution.

Wireless networks of self-sufficient sensor nodes benefit from a flexible network structure, allowing for dynamic node relocation and network extensions. To fully exploit this advantage, distributed control and processing schemes

with local operations are favorable over centralized approaches. Thus, this paper proposes a *distributed greedy sensor selection* for model-based phenomenon reconstruction over a finite time horizon. It is shown that greedy sensor scheduling is well suited for distributed implementation, as it requires only a minimum amount of computation and simply relies on local sensor information.

If one sensor per time step is scheduled, a greedy selection furthermore leads to a constant quality approximation. By using the so-called *covariance reduction*, a specific scalar reward function on the covariance of state estimates, greedy scheduling rewards can be proven to be at least half of the optimal solution. This finding, motivated by the quality bounds stated in [9], extends existing theoretical results in many ways: 1.) the bound holds for dynamic systems with multivariate states and measurement values, which is contrary to [9, 7, 10] employing similar variance-based reward functions for scalar static systems. 2.) Further bounds for non-scalar dynamic systems, as the one proposed in [8] for mutual information, are typically computationally infeasible for high-dimensional state/measurement-spaces due to calculating matrix determinants. Instead, the employed covariance reduction reward function utilizes the computationally cheap trace operation. 3.) Instead of a centralized realization, this paper further proposes a distributed greedy sensor scheduling scheme based on *hierarchically structured communication* in order to reduce the scheduling and communication overhead. The effectiveness of the proposed distributed greedy sensor scheduling is demonstrated via simulation in Sec. 5.

2 Problem Formulation

The main goal is to design a distributed scheduling algorithm for the energy-efficient but accurate reconstruction of a space-time continuous physical phenomenon via a sensor network. A large number of such phenomena can be described by means of a set of linear partial differential equations. For simplicity, a one-dimensional diffusion equation and its solution $p(z,t)$ are used within this paper, given by

$$\frac{\partial\, p(z,t)}{\partial\, t} - c\, \frac{\partial^2 p(z,t)}{\partial\, z^2} - s(z,t) = \mathbb{L}(\, p(z,t)\,) = 0\,, \quad (1)$$

describing for example the temperature at a certain location z and time t. The known source term is denoted by $s(z,t)$, the diffusion coefficient by the constant c. As solution domain, the interval $\Omega = \{z\,|\,0 \le z \le L\}$ is used, assuming the Neumann boundary conditions $\frac{\partial\, p(\,z=0,t\,)}{\partial z} = g_N^1$ and $\frac{\partial\, p(\,z=L,t\,)}{\partial z} = g_N^2$. To simplify Bayesian estimation, the partial differential equation is converted into a finite state-space form. Herein, the phenomenon is characterized by a linear stochastic discrete-time system model, which represents arising uncertainties as additive white Gaussian noise. By incorporating these models within a Kalman filter, the Riccati equation is obtained, serving as the basis for sensor scheduling.

Given a set of sensors $\mathcal{S}$ and a fixed estimation time horizon N, a sensor schedule $\mathcal{A}_{0:N-1} = (\mathcal{A}_0, \mathcal{A}_1, \ldots, \mathcal{A}_{N-1})$ represents a time sequence of sensor subsets $\mathcal{A}_k \subseteq \mathcal{S}$ allocated for measurements at time step t_k. By applying a sensor schedule, the initial estimation uncertainty, denoted by the covariance matrix $\mathbf{C}_0^x$, is reduced to a remaining uncertainty $\mathbf{C}_N^x(\mathcal{A}_{0:N-1})$, which reflects the estimation error for this sensor schedule. Thanks to the linear Gaussian system resulting from the conversion, the potential reduction of the uncertainty by a given sensor schedule can be determined a priori and independently of actual measurements by employing a covariance-based reward function V. Sensor scheduling then aims at finding the optimal schedule

$$\mathcal{A}_{0:N-1}^* = \arg\max_{\mathcal{A}_{0:N-1}} V(\,\mathcal{A}_{0:N-1}\,)$$

that maximizes the reduction of the uncertainty. Considering computational and energy constraints on sensor nodes, this paper proposes a greedy/myopic algorithm for determining approximate schedules $\alpha_{0:N-1}$ of single sensor nodes α_k per time step. This provably yields schedules of constant approximation quality compared to the optimal single-sensor schedule $\alpha_{0:N-1}^*$.

3 Model-based Reconstruction

In model-based reconstruction [2], techniques for state estimation are combined with background knowledge about space-time continuous phenomena. Given in form of a physical model, this knowledge facilitates a continuous reconstruction on the basis of discrete space/time measurements. By converting the physical model into a finite-dimensional state-space form, the phenomenon can be characterized by a finite-dimensional stochastic process with linear transition functions (see Sec. 3.1 and 3.2). Reconstruction can then be carried out via the well-known Kalman filter (see Sec. 3.3).

3.1 Conversion of Continuous Phenomena

The conversion of the physical model, given as a set of linear partial differential equations, into a finite state-space form is achieved by following a numeric methodology known as Galerkin formulation [11]. First, the solution domain is spatially decomposed by employing methods like the finite-element-method or the finite-spectral method. Then, the resulting model is temporally discretized. Finally, the space-time continuous physical phenomenon $p(z,t_k)$ is represented as a finite state vector $\underline{x}_k$ at discrete time t_k.

Spatial Decomposition Using the Galerkin formulation, the solution $p(z,t)$ of the partial differential equation (1) is replaced by the finite approximation

$$\tilde{p}(\,z,t\,) = \sum_{j=0}^{N_{dof}-1} x_j(t)\, \Phi_j(z) \approx p(\,z,t\,)\,,$$

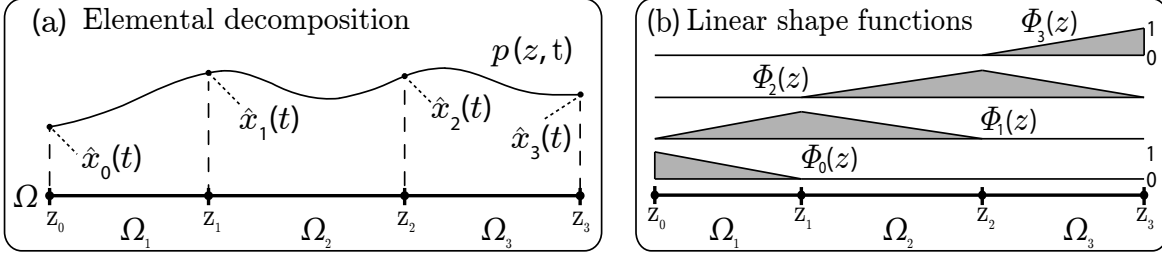where N_{dof} represents the degree of freedom. This approximation separates the space-dependent analytic *shape*

Figure 1: The solution $p(z,t)$ of the partial differential equation is element-wise approximated by the shape functions $\Phi_j(z)$ and scaling coefficient $x_j(t)$ for $N_{dof} = 4$.

functions $\Phi_j(z)$ from time-dependent *scaling factors* $x_j(t)$. Fig. 1 illustrates the approximation for piece-wise linear shape functions. In order to determine the scaling factors $x_j(t)$, a number of conditions has to be set. In the Galerkin formulation, these conditions aim at minimizing the approximation residual $\mathbb{L}(\tilde{p}(z,t))$, and therefore are given by weighted integrals

$$\int_\Omega \Phi_j(z)\, \mathbb{L}(\,\tilde{p}(z,t)\,)\,\mathrm{d}z \;=\; 0\,,\; \forall\, j \in [\,0, 1, \ldots, N_{dof}-1\,]$$

over the solution domain Ω. The choice of the set of shape functions then determines the numeric method employed in the discretization process, decomposing the solution domain Ω into disjoint *elements* Ω_g. Finally, the scaling factors $x_j(t)$ are collected in the *state vector* $\underline{x}(t) := [x_0(t), x_1(t), \ldots, x_{N_{dof}-1}(t)]^T$. As a result of the spatial decomposition, the set of ordinary differential equations

$$\mathbf{M}\,\underline{\dot{x}}(t) \;=\; \mathbf{M}\,\underline{\hat{s}}(t) \;-\; c\,\mathbf{D}\,\underline{x}(t) \;+\; \underline{b}^*(t)$$

evolves, characterized by the mass matrix $\mathbf{M}$ with $M_{ij} := \int_\Omega \Phi_i(z)\Phi_j(z)\mathrm{d}z$ and the diffusion matrix $\mathbf{D}$ with $D_{ij} := \int_\Omega \frac{\partial \Phi_i(z)}{\partial z}\frac{\partial \Phi_j(z)}{\partial z}\mathrm{d}z$. As detailed in [2], the vector $\underline{\hat{s}}(t)$ represents the discrete space source term, and $\underline{b}^*(t)$ summarizes the applied boundary conditions.

Temporal Discretization Following the spatial decomposition, the resulting model has to be discretized in time, leading to the desired finite-dimensional state-space model. Within this paper, implicit discretization methods are preferred, as they offer unconditional numeric stability. Employing a Crank-Nicolson discretization [2] leads to a discrete-time state-space description defined by

$$\mathbf{A}_k \;:=\; (\,\mathbf{M} + 0.5\,\Delta t\,\mathbf{D}\,)^{-1}\,(\,\mathbf{M} - 0.5\,\Delta t\,\mathbf{D}\,)\,,$$

$$\mathbf{B}_k \;:=\; \Delta t\,(\,\mathbf{M} + 0.5\,\Delta t\,\mathbf{D}\,)^{-1}\,,$$

where $\mathbf{A}_k$ represent the system matrix, $\mathbf{B}_k$ denotes the input matrix, and Δt is the length of a time step.

3.2 System and Measurement Equation

In state-space form, the physical phenomena are characterized by two equations. The *system equation* relates phenomenon states at different points of time, thereby describing the physical behavior of the phenomenon according to

$$\underline{x}_{k+1} \;=\; \mathbf{A}_k\,\underline{x}_k \;+\; \mathbf{B}_k(\,\underline{\hat{u}}_k + \underline{w}_k\,)\,,$$

where $\underline{w}_k$ is white zero-mean Gaussian noise with covariance $\mathbf{C}_k^w$ subsuming system uncertainties as well as input noise. The augmented input vector $\underline{\hat{u}}_k$ contains the boundary conditions besides the input values. The *measurement equation* relates a measurement $\hat{y}_k^i$ by sensor $\alpha_k = i$ at time step t_k to the phenomenon state $\underline{x}_k$ as

$$\hat{y}_k^i \;=\; \mathbf{H}_k^i\,\underline{x}_k \;+\; \underline{v}_k^i\,,$$

where $\underline{v}_k^i$ is white zero-mean Gaussian measurement noise with covariance $\mathbf{C}_k^{v,i}$. The measurement matrix $\mathbf{H}_k^i := [\,\Phi_0(z_i),\,\Phi_1(z_i),\,\ldots,\,\Phi_{N_{dof}-1}(z_i)\,]^T$ is determined by evaluating the shape functions at location z_i of sensor i. To recalculate the phenomenon from an estimated state $\underline{x}_k = [\,(\underline{x}_k)_0, (\underline{x}_k)_1, \ldots, (\underline{x}_k)_{N_{dof}-1}\,]^T$, phenomenon values $\tilde{p}(\,z_i, t_k\,)$ are related to the current state via the *output equation*

$$\tilde{p}(\,z_i, t_k\,) \;=\; \sum_{j=0}^{N_{dof}-1} (x_k)_j\,\Phi_j(z_i)\,. \tag{2}$$

3.3 Reconstruction via Kalman Filter

As conversion result, a linear Gaussian system is obtained. For these systems, the Kalman filter ensures optimal results in state estimation. By representing the phenomenon state as Gaussian random vector $\underline{x}_k$, this estimator recursively updates the state mean $\underline{\hat{x}}_k$ and covariance $\mathbf{C}_k^x$ in two steps. In the prediction step or *time update*, the system state is propagated to the next time step t_{k+1} according to

$$\underline{\hat{x}}_{k+1}^p \;=\; \mathbf{A}_k\,\underline{\hat{x}}_k^e \;+\; \mathbf{B}_k\,\underline{\hat{u}}_k\,,$$

$$\mathbf{C}_{k+1}^p \;=\; \mathbf{A}_k\,\mathbf{C}_k^e\,\mathbf{A}_k^T \;+\; \mathbf{B}_k\,\mathbf{C}_k^w\,\mathbf{B}_k^T\,. \tag{3}$$

The filter step or *measurement update* fuses the propagated state estimate with the information resulting from a measurement $\underline{\hat{y}}_k^i$ of sensor $i \in \mathcal{S}$ according to

$$\underline{\hat{x}}_k^e \;=\; \underline{\hat{x}}_k^p \;+\; (\,\underline{\hat{y}}_k^i - \mathbf{H}_k^i\,\underline{\hat{x}}_k^p\,)\,,$$

$$\mathbf{C}_k^e \;=\; \mathbf{C}_k^p \;-\; \mathbf{K}_k^i\,\mathbf{H}_k^i\,\mathbf{C}_k^p \;=:\; f_k^i(\,\mathbf{C}_k^p\,)\,, \tag{4}$$

where $\mathbf{K}_k^i = \mathbf{C}_k^p (\mathbf{H}_k^i)^T (\mathbf{C}_k^{v,i} + \mathbf{H}_k^i \mathbf{C}_k^p (\mathbf{H}_k^i)^T)^{-1}$ denotes the Kalman gain for sensor i. Combining both update steps into a single equation allows for compactly characterizing the evolution of system uncertainty for a sensor $i \in \mathcal{S}$ by the recursive discrete-time *Riccati equation*

$$r_k^i(\mathbf{C}_k^x) := \mathbf{C}_{k+1}^x(i) =$$
$$\mathbf{A}_k \mathbf{C}_k^x \mathbf{A}_k^T + \mathbf{B}_k \mathbf{C}_k^w \mathbf{B}_k^T - \mathbf{A}_k \mathbf{K}_k^i \mathbf{H}_k^i \mathbf{C}_k^x \mathbf{A}_k^T . \quad (5)$$

The evolution of an initial covariance $\mathbf{C}_0^x$ under a sensor sequence $\alpha_{0:N-1}$ can then be stated as $\mathbf{C}_N^x(\alpha_{0:N-1}) := r_{N-1}^{\alpha_{N-1}}(r_{N-2}^{\alpha_{N-2}}(\dots r_0^{\alpha_0}(\mathbf{C}_0^x)\dots))$. Finally, two important properties of the Riccati equation are given, originally stated in [12].

Theorem 1 (Concavity and Monotonicity)
Given positive definite noise matrices $\mathbf{C}_k^{v,i}$, the Riccati equation (5) *is concave and monotone. For two positive semi-definite matrices $\mathbf{C}_k^1$ and $\mathbf{C}_k^2$ and $i \in \mathcal{S}$, monotonicity displays as relation $\mathbf{C}_k^1 \preceq \mathbf{C}_k^2 \Rightarrow r_k^i(\mathbf{C}_{k+1}^1) \preceq r_k^i(\mathbf{C}_{k+1}^2)$, where $\mathbf{C}^1 \preceq \mathbf{C}^2$ abbreviates a positive semi-definite matrix difference $\mathbf{C}^1 - \mathbf{C}^2 \succeq 0$.*

4 Distributed Sensor Scheduling

A distributed approach to scheduling in a sensor network is advantageous over a central control and processing structure for many reasons. It allows to reduce communication and computational overhead, and improves the flexibility, reliability, and scalability of the system. In this section, the performance of a greedy scheduling approximation is examined and an energy-efficient distributed sensor scheduling based on greedy sensor selection is proposed.

4.1 Key Idea

As a major design choice, this paper employs a greedy approach to sensor scheduling, described in Algorithm 1. Though being an approximate method, a constant factor bound can be proven for the greedy approach by choosing a *submodular* reward function. Furthermore, the greedy approach reduces the problem of finding a long-term sensor schedule to an on-line sensor selection per time step. This by purpose fits a distributed implementation by reducing the knowledge mutually required by the sensor nodes. In designing an appropriate algorithm, the degree of distribution determines the information being either acquired by communication or derived from pre-installed knowledge. Here, two extremes can be distinguished. 1.) Each sensor acts independently based on the complete knowledge of the important measurement parameters of all other sensors, i.e., their current positions, measurement matrices $\mathbf{H}$, and noises $\mathbf{C}^v$. 2.) Sensors only have local knowledge of their own important parameters. In the first extreme, sensor scheduling can be performed individually by each sensor, resulting in a high computational load, whereas the second approach requires a lot of communication for negotiation. As a trade-off, a distributed sensor scheduling approach following a hierarchical

---

Algorithm 1, Greedy Sensor Scheduling.

1: **for all** time steps t_k with $k \in [0, 1, \dots, N-1]$ **do**
2:     **for all** sensors $i \in \mathcal{S}$ **do**
3:         Calculate covariance reduction $V_k(i) = \mathrm{trace}(\mathbf{C}_{k+1}^P - r_k^i(\mathbf{C}_k^x))$ for sensor i ;
4:     Determine maximum reduction $V_k^{max} := V_k(i^g) = \max_{i \in \mathcal{S}}(V_k(i))$;
5:     Schedule best sensor $\alpha_k := i^g$ for next measurement: $\alpha_{0:k} := (\alpha_{0:k-1}, \alpha_k)$;

---

selection and communication scheme is proposed, based on the grouping of sensors into disjoint subsets.

4.2 Reward Function

The reward function is used to judge the quality of the current state estimate in terms of stochastic uncertainty. For this purpose, a scalar value is assigned to the current state covariance, representing the possible reduction of the estimation error. This section introduces a theoretically important reward function based on matrix traces.

Covariance Reduction As reward function in greedy sensor scheduling, a scalar measure of the reduction in state covariance by a given measurement is chosen. At any time step t_k, given a prior covariance $\mathbf{C}_k^x$, the covariance reduction by a non-empty sensor subset $\mathcal{A} = \{\alpha_0, \alpha_1, \dots, \alpha_{M-1}\} \subseteq \mathcal{S}$ is given by

$$V_k(\mathcal{A}) := \mathrm{trace}\left(\mathbf{C}_{k+1}^P - (\mathbf{A}_k \mathbf{C}_\mathcal{A} \mathbf{A}_k^T + \mathbf{B}_k \mathbf{C}_k^w \mathbf{B}_k^T)\right).$$
$$(6)$$

The matrix $\mathbf{C}_k^P$ denotes the result of applying k time update steps (3) on the initial covariance $\mathbf{C}_0^x$ and $\mathbf{C}_\mathcal{A} := f_k^{\alpha_{M-1}}(f_k^{\alpha_{M-2}}(\dots f_k^{\alpha_0}(\mathbf{C}_k^x)\dots))$ abbreviates the repeated application of M measurement updates (4) on $\mathbf{C}_k^x$ within one time step, using each sensor in $\mathcal{A}$. At each time step, a greedy algorithm then chooses the sensor subset resulting in the highest covariance reduction. To account for multiple measurements at the same time step, a conditioned covariance reduction is defined. The reduction of covariance through a sensor subset $\mathcal{A}_1 \subseteq \mathcal{S}$ at time step t_k, conditioned on the event that at the same time step subset $\mathcal{A}_0 \subseteq \mathcal{S}$ was already used for measurements, is given by

$$V_k(\mathcal{A}_1 \mid \mathcal{A}_0) := V_k(\mathcal{A}_1 \cup \mathcal{A}_0) - V_k(\mathcal{A}_0) . \quad (7)$$

Submodularity A function $g(\cdot)$ mapping sets to real numbers is submodular if

$$g(\mathcal{C} \cup \mathcal{A}) - g(\mathcal{A}) \geq g(\mathcal{C} \cup \mathcal{B}) - g(\mathcal{B}), \quad \forall \mathcal{B} \supseteq \mathcal{A}$$

holds for arbitrary subsets $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$. By replacing $g(\cdot)$ with a reward function for sensor scheduling, submodularity can be interpreted as the property that by utilizing additional measurements, the information value of so far unused measurements is reduced. This property serves as a motivation for using few sensors per time step. In the following, the

submodularity of the covariance reduction reward function is shown.

Theorem 2 (Submodularity of Covariance Reduction)
The covariance reduction reward function, given by (6)*, is submodular.*

PROOF. For brevity, this proof omits all time indices. If only one sensor $i \in \mathcal{S}$ performs a measurement at an arbitrary time step, the conditioned covariance reduction for an arbitrary sensor subset $\mathcal{A} \subseteq \mathcal{S}$ results in

$$V(\{i\} \mid \mathcal{A}) = V(\mathcal{A} \cup \{i\}) - V(\mathcal{A})$$
$$= \operatorname{trace}(\mathbf{A}\,\mathbf{K}^i_{\mathcal{A}}\,\mathbf{H}^i\,\mathbf{C}_{\mathcal{A}}\,\mathbf{A}^T)\,,$$

where $\mathbf{K}^i_{\mathcal{A}}$ abbreviates the Kalman gain based on the covariance $\mathbf{C}_{\mathcal{A}}$ and sensor i. The product $\mathbf{KHC}$ in general describes the reduction in variance per measurement update. Following from the monotonicity and concavity of the Riccati equation stated in Theorem 1, this reduction is smaller in the sense of a positive semi-definite ordering, the smaller the current covariance $\mathbf{C}$ is. Since applying more measurement updates further reduces the prior covariance, the covariance reduction of a measurement update based on a sensor subset becomes smaller with the growing cardinality of this subset. Hence, for a joint sensor subset $\mathcal{C} \cup \mathcal{A}$ the inequality $\mathbf{K}^i_{\mathcal{A}}\mathbf{H}^i\mathbf{C}_{\mathcal{A}} \succeq \mathbf{K}^i_{(\mathcal{C} \cup \mathcal{A})}\mathbf{H}^i\mathbf{C}_{(\mathcal{C} \cup \mathcal{A})}$ holds, which can be easily proven using the Woodbury identity [13]. As a consequence, the product $\mathbf{A}\,(\mathbf{K}^i_{\mathcal{A}}\,\mathbf{H}^i\,\mathbf{C}_{\mathcal{A}} - \mathbf{K}^i_{(\mathcal{C} \cup \mathcal{A})}\,\mathbf{H}^i\,\mathbf{C}_{(\mathcal{C} \cup \mathcal{A})})\,\mathbf{A}^T$ is positive semi-definite, and thus, the inequality $V(\{i\} \mid \mathcal{A}) - V(\{i\} \mid \mathcal{C} \cup \mathcal{A}) \geq 0$ holds, which can be equivalently recast to

$$V(\mathcal{C} \cup \mathcal{A}) - V(\mathcal{A}) \geq V(\mathcal{C} \cup \mathcal{A} \cup \{i\}) - V(\mathcal{A} \cup \{i\})\,.$$

By continuously iterating the previous equation over all sensors $i \in \mathcal{B} \setminus \mathcal{A}$, the submodularity of the covariance reduction at a single time step can be proven. □

Corollary 1 *The covariance reduction reward function is non-decreasing.*
PROOF. For all sensor subsets $\mathcal{A} \subset \mathcal{C} \subseteq \mathcal{S}$ holds

$$V(\mathcal{C}) - V(\mathcal{A}) = \operatorname{trace}\big(\mathbf{A}\,(\mathbf{C}_{\mathcal{A}} - \mathbf{C}_{\mathcal{C}})\,\mathbf{A}^T\big) \geq 0\,,$$

because of $\mathbf{C}_{\mathcal{A}} \succeq \mathbf{C}_{\mathcal{C}}$. □

Remark 1 *These results also hold when replacing the matrix trace in the reward function by the determinant or the maximum eigenvalue.*

Theorem 2 and Corollary 1 allow proving a quality bound for greedy scheduling in case of static systems or multiple sensors per time step. As [9] states, the reconstruction quality resulting in this setting is at least the $1 - 1/e$-fold of the respective optimal schedule. The next section extends these results to greedy scheduling in dynamic systems, using only one sensor per time step.

4.3 Greedy Sensor Scheduling

In order to rate the performance of a sensor schedule $\mathcal{A}_{0:N-1}$, that is, a time sequence of sensor subsets $\mathcal{A}_k \subseteq \mathcal{S}$, a reward function considering multiple time steps is required. Therefore, the multi-step *covariance reduction*

$$V(\mathcal{A}_{0:N-1}) := \operatorname{trace}\big(\mathbf{C}^P_N - \mathbf{C}^x_N(\mathcal{A}_{0:N-1})\big)$$

describes the overall reduction in state covariance by a given sensor sequence as difference between the expected uncertainty after N prediction steps (3) and the uncertainty achieved through the sensor sequence. By using this reward function, a constant approximation quality can be proven for greedy sensor scheduling, based on the submodularity of the covariance reduction at each time step.

Theorem 3 (Greedy Sensor Scheduling)
In case of temporally independent measurements, the greedy Algorithm 1 leads to a schedule $\alpha^g_{0:N-1}$ of single sensors with a total covariance reduction that is at least half the covariance reduction of the optimal schedule $\alpha^*_{0:N-1}$, i.e.,

$$V(\alpha^*_{0:N-1}) \leq 2 \cdot V(\alpha^g_{0:N-1})\,.$$

PROOF. In analogy to [8], where N was set to two and $V(\cdot)$ was the mutual information objective, it follows

$V(\alpha^*_{0:N-1})$

$\overset{(a)}{\leq} V(\alpha^*_{0:N-1}, \alpha^g_{0:N-1})$

$\overset{(b)}{=} V(\alpha^g_0) + V(\alpha^*_0 \mid \alpha^g_0) +$

$\displaystyle \sum_{k=1}^{N-1}\big(V(\alpha^g_k \mid \alpha^*_{0:k-1}, \alpha^g_{0:k-1}) + V(\alpha^*_k \mid \alpha^*_{0:k-1}, \alpha^g_{0:k})\big)$

$\overset{(c)}{\leq} V(\alpha^g_0) + \underbrace{V(\alpha^*_0)}_{\leq V(\alpha^g_0)} +$

$\displaystyle \sum_{k=1}^{N-1}\big(V(\alpha^g_k \mid \alpha^g_{0:k-1}) + \underbrace{V(\alpha^*_k \mid \alpha^g_{0:k-1})}_{\leq V(\alpha^g_k \mid \alpha^g_{0:k-1})}\big)$

$\overset{(d)}{\leq} 2 \cdot V(\alpha^g_{0:N-1})\,,$

where (a) results from Corollary 1, (b) and (d) follow from (7), and (c) results from the submodularity of $V(\cdot)$. □

4.4 Distributed Scheduling Scheme

In order to cope with the distributed nature of both the phenomenon and the sensor network, implementing the greedy sensor scheduling listed in Algorithm 1 as a distributed algorithm is advantageous. Without a central processing station, the central communication overhead is omitted, and no single point of failure exists. In distributed scheduling, each sensor only knows its own measurement parameters and can calculate its own potential covariance reduction independently from other sensors for each time step.
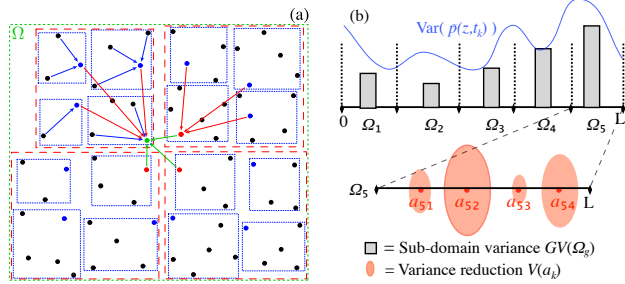
Figure 2: (a) Hierarchical communication scheme for distributed greedy sensor scheduling. Colored boxes and nodes indicate sub-domains and representative nodes, respectively. (b) Hierarchical sensor selection strategy for a one-dimensional phenomenon.

Thus, the computational load is fairly distributed among all sensor nodes. Furthermore, the sensor network can easily be adapted to changes in its topology, since each sensor only needs to update its own position data. A distributed scheduling approach hence allows for a flexible and even mobile network.

To deal with the remaining communication overhead resulting from information exchange during distributed sensor selection, a hierarchical, tree-like *communication scheme* is proposed. This scheme employs directed communication, where all communication is directed towards the center of the solution domain Ω in order to avoid energy-consuming broadcasts. For this purpose, the global solution domain Ω is hierarchically divided in so-called *sub-domains* (see Fig. 2 (a)). Within each sub-domain, one of the contained sensor nodes acts as representative. Communication in and between sub-domains for scheduling purposes is carried out over representative sensor nodes only. It is important to note that the selection of appropriate sub-domains and representative sensor nodes is outside the scope of this paper and thus, devoted to future work. A straightforward way for sub-domain selection would be to exploit the disjoint elements Ω_g used for the conversion in Section 3.1. For representative nodes, selecting sensor nodes that are next to the center of Ω is a straightforward choice.

For determining the next sensor node to perform a measurement, at first the individually calculated covariance reductions of the sensor nodes are compared within the sub-domains on the lowest level (blue sub-domains in Figure 2 (a)). The best of all low-level decisions is then selected on the next higher sub-domain (red sub-domains in Figure 2 (a)). This is repeated until the selection of the best sensor for the global solution domain Ω (green domain in Figure 2 (a)) is performed. Obviously, this procedure guarantees the previously proven bound of greedy sensor scheduling *and* is fully distributed, as the best sensor is scheduled for measurement in a scalable manner involving only local computation and communication.

Completing the distributed approach, the scheduled sensor node performs a central filter step (4) and time update (3). The updated state estimate is then propagated to all sensor nodes by exploiting the proposed communication tree in reversed order. According to [14], this distributed reconstruction followed by a propagation of the updated state estimate is advantageous over a distribution of the measurement value itself by improved robustness and efficiency.

4.5 Possible Extensions

The greedy sensor selection can be approximated by a *hierarchical greedy selection* to further reduce the computational and communication load. This strategy constitutes a compromise between the two extremes stated in Section 4.1. For this purpose, individual sensors are grouped into continuous disjoint subsets Ω_g, which then become subject to selection on behalf of the contained nodes. For choosing between groups, an integral objective function can be employed, approximating the potential covariance reduction of its sensors by an easy-to-compute measure. Here, the current total variance $GV(\Omega_g) = \sum_{z \in \Omega_g} \mathrm{Var}\big(\tilde{p}(z, t_k)\big)$ over the corresponding sub-domain Ω_g is utilized.

The hierarchical sensor selection then comprises two phases, depicted in Figure 2 (b). In a high-level selection, the best sensor group is chosen according to the approximate measure (Ω_5 in Figure 2 (b)). In the second phase, the low-level selection determines the best sensor of the chosen group by the presented greedy algorithm. Hence, individual covariance reductions only have to be calculated for the sensors of the selected group and communicated between them, omitting all other sensors. This reduces the computational load, as well as the main part of communication activities, by the number of defined groups. In both phases, the aforementioned tree-like communication scheme applies. By employing an approximate measure, the quality bound (3) for greedy scheduling is no longer valid for the hierarchical greedy scheduling approach. Finding a similar bound for this approach is devoted to future research.

5 Simulation Results

As proof of concept, this section presents performance results for the proposed greedy and hierarchical greedy sensor selection strategies. Here, the greedy strategies are compared against 1.) an optimal schedule resulting from exhaustive tree-search, 2.) an approximate schedule arising from a model-predictive control approach and 3.) a randomly selected schedule. For simulating a temperature distribution, the one-dimensional diffusion equation (1), constrained by Neumann boundary conditions, is decomposed by the finite-spectral method into five elements with seven polynomials each. After a Crank-Nicolson time discretization, the resulting state-space model characterizes the phenomenon state by $\underline{x}_k \in \mathbb{R}^{31}$.

For quantifying the scheduling performance, the reconstructed phenomenon is discretely represented as random

|	Greedy		Optimal	
Set	Quality in K^2	Time	Quality in K^2	Time
1	3.699e+8	0.06 s	3.298e+8	73 s
2	4.663e+8	0.05 s	3.992e+8	13 s

Table 1: Reconstruction quality and runtime for optimal and greedy sensor scheduling approaches over a time horizon $N = 3$, simulated with two different sensor sets.

vector $\underline{\boldsymbol{p}}_k$ at equidistant sampling points z_i, and the reconstruction quality $g(\underline{\boldsymbol{p}}_k)$ is given by the trace of $\text{Cov}(\underline{\boldsymbol{p}}_k)$. Depending merely on the sampling points, a discrete sampling matrix $\left(\Phi_j(z_i)\right)_{ij} =: \boldsymbol{\Phi}_z$ can be determined a priori, with $\underline{\boldsymbol{p}}_k = \boldsymbol{\Phi}_z \underline{\boldsymbol{x}}_k$ and $\text{Cov}(\underline{\boldsymbol{p}}_k) = \boldsymbol{\Phi}_z \mathbf{C}_k^x \boldsymbol{\Phi}_z^T$, according to (2). Determining the best sensor on the basis of $\underline{\boldsymbol{p}}_k$ instead of $\underline{\boldsymbol{x}}_k$ increases greedy reconstruction quality. By exploiting trace properties, it holds $g(\underline{\boldsymbol{p}}_k) = \text{trace}(\mathbf{C}_k^x \boldsymbol{\Phi}^*)$, where $\boldsymbol{\Phi}^* = \boldsymbol{\Phi}_z^T \boldsymbol{\Phi}_z$ is of the same dimensions as $\mathbf{C}_k^x$ and can be calculated in advance. As can be easily proven, the quality bound for greedy scheduling (3) still holds for this extension.

For the simulation, the initial estimation uncertainty $\mathbf{C}_0^x$ is defined by a diagonal matrix with only certain areas set to a non-zero variance value of 10^{10}. This initial setting requires and rewards adaptation in scheduling. The input noise is given as $\mathbf{C}_k^w = \mathbf{I}$, with a further additive system uncertainty of $\mathbf{I}$. The measurement noise is modeled as $\mathbf{C}_k^{v,i} = \mathbf{I}$. For evaluating the impact of node density and node location, two differing sets of sensors are utilized to simulate a sensor network. Sensor set 1 consists of 41 nodes, placed equidistantly over the given sample domain. Set 2 consists of 23 randomly placed sensor nodes. As simulation software, Matlab_R2008a® is used on an Intel® Core™ 2 Duo with 2.2 GHz. In the following, the given reconstruction quality denotes the average measured quality per phenomenon sampling point, i.e., the average variance over the elements of $\underline{\boldsymbol{p}}_k$, in Kelvin^2.

At first, the performance of greedy scheduling over a fixed time horizon of length $N = 3$ is evaluated. Here, the greedy schedule is compared to an *optimal sensor schedule*, determined by exhaustively searching the tree of all possible sensor schedules $\alpha_{0:N-1}$ for the one leading to a maximum reduction in covariance. Table 1 displays the resulting reconstruction quality and runtime for both sensor sets. The greedy approach thereby performs in close range to the optimal reconstruction quality, by far tighter than the theoretical bound (3). At the same time, the greedy approach is much less time consuming, which also can be seen as a demonstration of energy efficiency of the greedy scheduling approach, if computations for scheduling are considered only.

Another common approach to sensor scheduling is known as *receding horizon control*, or model-predictive control [15]. Within this approximate approach, the estimation time horizon N is reduced to a finite scheduling horizon
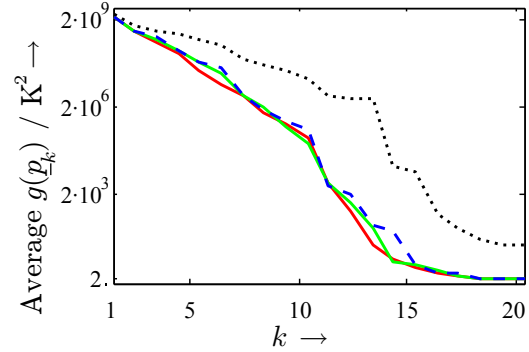
(a) Resulting reconstruction quality over time as average variance of $\underline{\boldsymbol{p}}_k$.
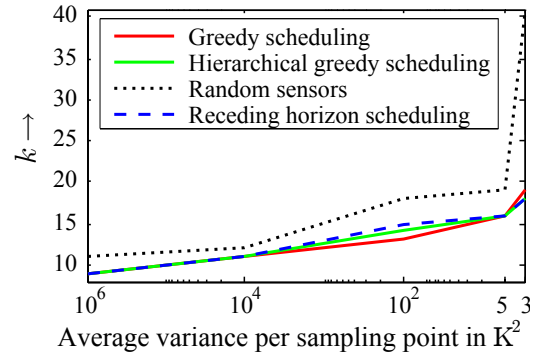
(b) Number of required time steps to reach pre-defined quality bounds.

Figure 3: Performance of the greedy scheduling strategies compared to receding horizon control and an exemplary random sensor selection.

$P \leq N$, and open-loop feedback scheduling is employed for sensor selection. Fig. 3(a) depicts the temporal evolution of reconstruction quality for the proposed greedy scheduling strategies compared to receding horizon control with $P = 3$. As can be seen, both greedy strategies offer a promising performance, exceeding the receding horizon approach in average. In runtime, and consequently energy consumption, the greedy strategies with 0.5 seconds outperform the receding horizon approach with 1342 seconds by four orders of magnitude.

As a final demonstration of greedy scheduling performance, several bounds on absolute reconstruction quality are considered. Fig. 3(b) depicts the number of time steps required by each scheduling approach to reach the given reconstruction quality. It can be seen that the greedy scheduling approaches again perform at least as good as the more complex receding horizon scheduling. Moreover, only the greedy approaches are able to converge below an average variance of $2.5\ K^2$.

6 Conclusions and Future Work

When performing sensor scheduling for model-based reconstruction of space-time continuous physical phenomena,

one has to balance between computation and communication costs. Determining optimal sensor schedules requires global knowledge. Furthermore, the consideration of long-term effects when performing scheduling in a distributed fashion is impractical from a computation and communication perspective. Thanks to the following novelties, merely local computation and communication is required, which in turn results in energy savings.

1. The proposed greedy scheduling algorithm marks a practical way towards distributed scheduling and reconstruction, since merely a fraction of the computations compared to non-myopic algorithms is required.

2. Though being suboptimal, a tight bound on the performance of the proposed greedy scheduling algorithm compared to the optimal solution can be guaranteed for the employed covariance reduction reward function. This extends existing findings to dynamic systems and multivariate states for this specific reward function.

3. With the proposed hierarchical communication scheme, sensor nodes merely require local information for determining the next sensor. This decision is achieved by comparing individually determined covariance reductions and by communicating these scheduling results between neighboring domains.

4. Thanks to the model-based reconstruction approach, physical background knowledge about the considered space-time continuous phenomenon can be systematically exploited.

Future work is devoted to the selection of convenient domains, which are required for the hierarchical communication and the hierarchical greedy selection schemes. Proving performance guarantees for the hierarchical greedy selection is also the subject of future work.

References

[1] A. J. Baker, *Finite Element Computational Fluid Mechanics*. Taylor and Francis, 1983.

[2] F. Sawo, K. Roberts, and U. D. Hanebeck, "Bayesian Estimation of Distributed Phenomena using Discretized Representations of Partial Differential Equations," in *Proceedings of the 3rd International Conference on Informatics in Control, Automation and Robotics (ICINCO 2006)*, Setúbal, Portugal, Aug. 2006, pp. 16–23.

[3] L. Meier, III, J. Peschon, and R. M. Dressler, "Optimal Control of Measurement Subsystems," *IEEE Transactions on Automatic Control*, vol. 12, no. 5, pp. 528–536, Oct. 1967.

[4] M. F. Huber and U. D. Hanebeck, "Priority List Sensor Scheduling using Optimal Pruning," in *Proceedings of the 11th International Conference on Information Fusion (Fusion 2008)*, Cologne, Germany, Jul. 2008.

[5] V. Gupta, T. Chung, B. Hassibi, and R. M. Murray, "Sensor Scheduling Algorithms Requiring Limited Computation," in *Proceedings of IEEE Conference on Acoustics, Speech and Signal Processing*, Montreal, Canada, May 2004, pp. 825–828.

[6] F. Zhao, J. Shin, and J. Reich, "Information-Driven Dynamic Sensor Collaboration," *Signal Processing Magazine, IEEE*, vol. 19, no. 2, pp. 61–72, Mar. 2002.

[7] A. Krause, E. Horvitz, A. Kansal, and F. Zhao, "Toward Community Sensing," *International Conference on Information Processing in Sensor Networks (IPSN 2008)*, pp. 481–492, April 2008.

[8] J. L. Williams, J. W. Fisher, and A. S. Willsky, "Performance Guarantees for Information Theoretic Sensor Resource Management," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 3, pp. III–933–III–936, April 2007.

[9] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An Analysis of Approximations for Maximizing Submodular Set Functions—I," in *Mathematical Programming*, vol. 14, no. 1. Springer Berlin/Heidelberg, Dec. 1978, pp. 265–294.

[10] A. Das and D. Kempe, "Algorithms For Subset Selection in Linear Regression," in *Proceedings of the 40th annual ACM Symposium on Theory of Computing (STOC 2008)*. New York, NY, USA: ACM, 2008, pp. 45–54.

[11] G. Karniadakis and S. Sherwin, *Spectral/hp Element Methods for Computational Fluid Dynamics*. Oxford University Press, 2005.

[12] V. Gupta, T. H. Chung, B. Hassibi, and R. M. Murray, "On a Stochastic Algorithm for Sensor Scheduling," in *Proceedings of the 16th IFAC World Congress*, Prague, Czech Republic, Jul. 2005.

[13] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. The Johns Hopkins University Press, Baltimore, 1996.

[14] L.-L. Ong, T. Bailey, B. Upcroft, and H. Durrant-Whyte, "Decentralised Particle Filtering for Multiple Target Tracking in Wireless Sensor Networks," in *Proceedings of the 11th International Conference on Information Fusion (Fusion 2008)*, Cologne, Germany, Jul. 2008.

[15] M. Morari and J. H. Lee, "Model Predictive Control: Past, Present and Future," 1997. [Online]. Available: citeseer.ist.psu.edu/morari97model.html