Sparse Point Source Estimation in Sensor Networks with Gaussian Kernels

Henning Paul
Department of Communications Engineering
University of Bremen
28359 Bremen, Germany
Email: paul@ant.uni-bremen.de

Reiner Jedermann
Institute for Microsensors, -actors and -systems (IMSAS)
University of Bremen
28359 Bremen, Germany
Email: rjedermann@imsas.uni-bremen.de

***Abstract*—In this paper we present a technique for the estimation of point sources in a diffusive environment that can be applied to wireless sensor networks. It is based on methods from the field of sparse recovery, using Gaussian kernels as basis functions. After presenting the underlying physical process, a linear system model is developed out of it. For the estimation of its unknown parameters, solution approaches are presented. The approach is verified by means of numerical simulations. Furthermore, we indicate ways to perform a distributed estimation of the parameters within the network.**

I. Introduction

Sensor networks play an important role in environmental monitoring, e.g., of water quality in water bodies, air pollution [1], or earthquake detection [2]; but also in monitoring of, e.g., chilled food transports [3]. These scenarios have in common that the source of the quantity to be measured, i.e., a pollutant, seismic event, or a heat source, can be assumed point-shaped. Additionally, the propagation of the quantity under measurement follows differential equations, which often represent a diffusion process. A spatially distributed sensor network is able to perform a scalar measurement of the quantity of interest at different positions. Now, the task of this sensor network is to make an estimation of the location of the sources and/or obtain an estimate of the spatial distribution of the quantity of interest, i.e., of its field. In contrast to previous works using Compressed Sensing (CS) techniques for the estimation of diffusion fields, e.g., [4], we will present a flexible framework able to use different basis functions and demonstrate its effectiveness using radial basis functions (RBFs), which are also known as Gaussian kernels.

Based on this framework, we will present in this paper a practical approach to perform an estimation for both location and field, based on the measurements of a sensor network. We will also address the task of distributed estimation within the network of sensors (In-Network Processing).

The remainder of the paper is structured as follows: In Section II, we will present the underlying physical process and derive a system model out of it. In the following, the estimation problem and its solution approach is presented in Section III. Subsequently, results for a centralized estimation in a data fusion center are presented, while in the following section, hints on the distributed estimation are given. Section VI concludes the paper with an outlook of possible future work.

II. System Model

In this paper, we will investigate the problem of estimating a physical quantity in a diffusive process. Therefore, we will first introduce the underlying physics in order to derive a linear system model, which is done in the following.

In general, the solution of any differential equation can be expressed using its specific so-called Green's function, which can be interpreted as the spatio-temporal impulse response of the system. For an n-dimensional diffusion process without boundaries, it reads [5]:

$$G(\mathbf{x},t) = (4\pi D t)^{-\frac{n}{2}} \mathrm{e}^{-\frac{||\mathbf{x}||^2}{4Dt}}, \tag{1}$$

with diffusion constant D, Cartesian coordinate $\mathbf{x} = (x_1, \ldots, x_n)$ and time $t > 0$. If the M sources within the diffusive process are assumed to have a point nature, they can be modeled as spatial delta functions with a temporal envelope $a_m(t)$:

$$s_m(\mathbf{x},t) = a_m(t) \cdot \delta_0(\mathbf{x} - \mathbf{x}_m), \quad m = 1 \ldots M. \tag{2}$$

The entire field $f(\mathbf{x},t)$ can then be obtained by superimposing the spatio-temporal convolution of each of the sources with the Green's function:

$$f(\mathbf{x},t) = \sum_{m=1}^{M} s_m(\mathbf{x},t) \overset{\mathbf{x},t}{*} G(\mathbf{x},t). \tag{3}$$

Taking into account the sifting property of the delta function, only the temporal convolution remains:

$$f(\mathbf{x},t) = \sum_{m=1}^{M} a_m(t) \overset{t}{*} G(\mathbf{x} - \mathbf{x}_m, t). \tag{4}$$

If furthermore the sources are assumed to be activated at $t = 0$ and to maintain the same amplitude a_m afterwards, the convolution integral can be expressed as

$$f(\mathbf{x},t) = \sum_{m=1}^{M} a_m \int_0^t G(\mathbf{x} - \mathbf{x}_m, \tau) d\tau. \tag{5}$$

In this paper, we will consider the case of a 2-dimensional diffusion process, so the particular Green's function in this scenario reads

$$G(\mathbf{x},t) = \frac{1}{4\pi D t} \mathrm{e}^{-\frac{||\mathbf{x}||^2}{4Dt}}. \tag{6}$$

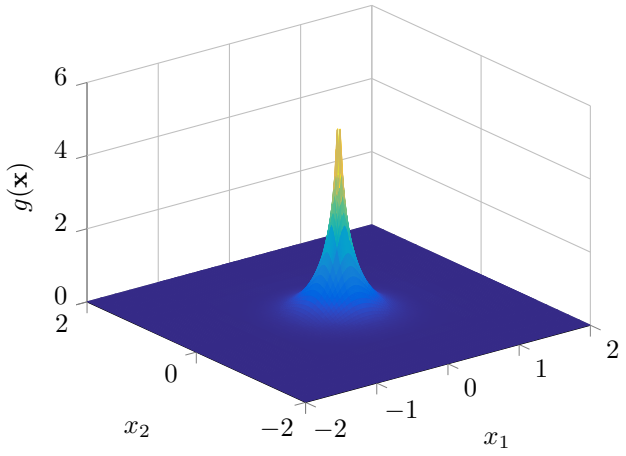

Fig. 1. Shape of the exponential integral function for exemplary values of $t = 1$ and $D = 0.1$

Evaluating the temporal integral in (5), the result

$$\int_0^t G(\mathbf{x}, \tau) d\tau = \frac{1}{4\pi D} \mathrm{E}_1 \left(\frac{||\mathbf{x}||^2}{4Dt} \right), \tag{7}$$

is obtained, where $\mathrm{E}_1(\cdot)$ represents the exponential integral function [6]. It is depicted in Fig. 1 for exemplary values of $t = 1$ and $D = 0.1$. This figure illustrates the characteristic shape of a diffusion process caused by a point source.

Restricting to a single time instant $t = T$ for the time being, the kernel

$$g_T(\mathbf{x}) = \frac{1}{4\pi D} \mathrm{E}_1 \left(\frac{||\mathbf{x}||^2}{4DT} \right), \tag{8}$$

is obtained as basis function for the following considerations.

The true field $f(\mathbf{x})$ can now be approximated by the superposition of $\tilde{M}$ sources at positions $\tilde{\mathbf{x}}_m$:

$$\tilde{f}(\mathbf{x}) = \sum_{m=1}^{\tilde{M}} a_m g_T(\mathbf{x} - \tilde{\mathbf{x}}_m). \tag{9}$$

Please note that the value $\tilde{M}$ for the approximation might differ from the true number M of sources. Using (9), every measurement y_j of the sensor $j = 1, \ldots, J$ located at position $\mathbf{x}_j^{\mathrm{s}}$ can now be described by

$$y_j = \tilde{f}(\mathbf{x}_j^{\mathrm{s}}) + n_j = \sum_{m=1}^{\tilde{M}} a_m g_T(\mathbf{x}_j^{\mathrm{s}} - \tilde{\mathbf{x}}_m) + n_j, \tag{10}$$

where n_j accounts for any measurement error.

Now, the philosophy of [7] is generalized to a two-dimensional grid, i.e., the possible source locations are quantized to $N \cdot L$ possible coordinates and follow the form $\hat{\mathbf{x}}_m = (n_m \Delta_1, \ell_m \Delta_2) + \hat{\mathbf{x}}_0$, with $n_m = 0, \ldots, N-1$, $\ell_m = 0, \ldots, L-1$ and $\hat{\mathbf{x}}_0$ an arbitrary offset. This offset, as well as N, L, Δ_1 and Δ_2, need to be chosen properly and according to the scenario.

Every sensor measurement y_j is now interpreted as superposition of the effect of $N \cdot L$ hypothetical, grid-aligned sources[1] with amplitudes $\hat{a}_m$, $m = 1, \ldots, NL$:

$$y_j = \sum_{m=1}^{N \cdot L} \hat{a}_m g_T(\mathbf{x}_j^{\mathrm{s}} - \hat{\mathbf{x}}_m) + n_j. \tag{11}$$

Combining these equations for all sensors $j = 1, \ldots, J$, the equation system

$$\mathbf{y} = \mathbf{\Phi} \hat{\mathbf{a}} + \mathbf{n} \tag{12}$$

is obtained, with

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_J \end{bmatrix}, \quad \hat{\mathbf{a}} = \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \vdots \\ \hat{a}_{N \cdot L} \end{bmatrix}, \tag{13}$$

noise vector $\mathbf{n}$ and measurement matrix

$$\mathbf{\Phi} = \begin{bmatrix} g_T(\mathbf{x}_1^{\mathrm{s}} - \mathbf{x}_1) & g_T(\mathbf{x}_1^{\mathrm{s}} - \mathbf{x}_2) & \ldots & g_T(\mathbf{x}_1^{\mathrm{s}} - \mathbf{x}_{N \cdot L}) \\ g_T(\mathbf{x}_2^{\mathrm{s}} - \mathbf{x}_1) & g_T(\mathbf{x}_2^{\mathrm{s}} - \mathbf{x}_2) & \ldots & g_T(\mathbf{x}_2^{\mathrm{s}} - \mathbf{x}_{N \cdot L}) \\ \vdots & \vdots & \ddots & \vdots \\ g_T(\mathbf{x}_J^{\mathrm{s}} - \mathbf{x}_1) & g_T(\mathbf{x}_J^{\mathrm{s}} - \mathbf{x}_2) & \ldots & g_T(\mathbf{x}_J^{\mathrm{s}} - \mathbf{x}_{N \cdot L}) \end{bmatrix}. \tag{14}$$

III. Solution of the Equation System

(12) can be solved for $\hat{\mathbf{a}}$, e.g., by a Least Squares (LS) approach. If the measurements y_j possess a common, but unknown offset a_{offset} (which is, e.g., the case for a heat source activated in a diffusive medium of certain temperature), this can be accounted for by augmenting the vector $\hat{\mathbf{a}}$ by this offset coefficient and extending the matrix $\mathbf{\Phi}$ by an all-ones column:

$$\underline{\hat{\mathbf{a}}} = \begin{bmatrix} \hat{\mathbf{a}} \\ \hat{a}_{\mathrm{offset}} \end{bmatrix}, \quad \underline{\mathbf{\Phi}} = \begin{bmatrix} \mathbf{\Phi} & \mathbf{1} \end{bmatrix}, \tag{15}$$

resulting in the modified system equation

$$\mathbf{y} = \underline{\mathbf{\Phi}} \underline{\hat{\mathbf{a}}} + \mathbf{n}. \tag{16}$$

An LS solution ensures minimum error between measured values and reconstructed field at sensor location, but does not consider the field distribution between the sensors. Since the underlying physical process is assumed to only exhibit few sources, i.e., $\tilde{M} \ll N \cdot L$, a sparse estimate for $\hat{\mathbf{a}}$ (or $\underline{\hat{\mathbf{a}}}$, respectively) is expected to show a better performance in the sense of reproducing the original field. If the estimation of $\hat{\mathbf{a}}$ is performed in a sparse fashion, i.e., the sparsity of the estimate is enforced within the estimation criterion, a better recontruction of the field can be achieved. One possible algorithm to perform such an estimation is the Orthogonal Matching Pursuit (OMP) [8] algorithm. This algorithm starts by picking the column from $\mathbf{\Phi}$ with maximum correlation with $\mathbf{y}$, subtracts its contribution to $\mathbf{y}$ from it, picks the next

[1] $N \cdot L$ is significantly larger than M and $\tilde{M}$, the consequences of this fact will be addressed in the next section.

column with maximum correlation to the updated $\mathbf{y}$ and so on, until the prescribed number of $\tilde{M}$ columns have been selected. The corresponding elements of $\hat{\mathbf{a}}$ are obtained using a least squares criterion on the selected columns only. This algorithm is also frequently used in the context of CS, but was originally developed for sparse recovery. Therefore it serves the purpose very well in this application, as will be shown in the results section.

A. Extension into temporal direction

The extension of the approach into temporal direction is straightforward: Instead of restricting to a single time instant $t = T$ in (8), we define kernels for every time instant of interest. If, e.g., the field is sampled at time instants kT with $k = 1 \dots K$, the kernels

$$g_{kT}(\mathbf{x}) = \frac{1}{4\pi D} \mathrm{E}_1 \left(\frac{||\mathbf{x}||^2}{4DkT} \right), \quad k = 1 \dots K \tag{17}$$

are defined, which are then used to construct the multi-time matrix $\mathbf{\Phi}_{\mathrm{MT}}$ shown in (18). Likewise, the measurements at all K time instants are collected in the vector

$$\mathbf{y}_{\mathrm{MT}} = \begin{bmatrix} y_1(T) \\ y_2(T) \\ \vdots \\ y_J(T) \\ y_1(2T) \\ y_2(2T) \\ \vdots \\ y_J(2T) \\ y_1(3T) \\ \vdots \\ y_{J-1}(KT) \\ y_J(KT) \end{bmatrix}. \tag{19}$$

However, in this paper, we will restrict to single-time measurements.

B. Non-point-shaped sources

In reality, sources of, e.g., heat or substances are not exactly point-shaped, but possess a spatial extent. Therefore, the aforementioned exponential integral function (8) might not be a good choice for practical modeling. However, the basis function can be modified to better accomodate for this fact. If $g_T(\mathbf{x})$ as defined in (8) is integrated over a disc shape, another base function is obtained. This function cannot be calculated in closed form and is depicted in Fig. 2 for exemplary values of $D = 0.1$ and a disc radius of 0.1.

It can be seen that this function strongly resembles a Gaussian function, for which one example is depicted in Fig. 3. The similarity is corroborated by Fig. 4 that compares a crosssection of both functions along one axis. If the bandwidth of the Gaussian kernel is adapted (in this case to half the value as in the underlying Green's function), a very good match can be achieved. Therefore, we will in the following use Gaussian functions as basis functions, i.e.,

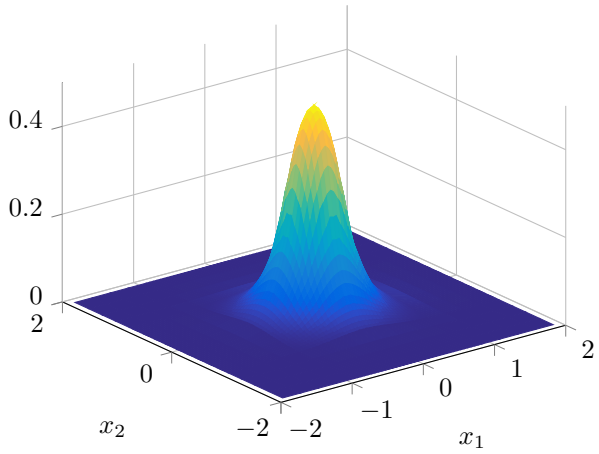

Fig. 2. Shape of the exponential integral function integrated over a disc with radius 0.1 for exemplary values of $t = 1$ and $D = 0.1$

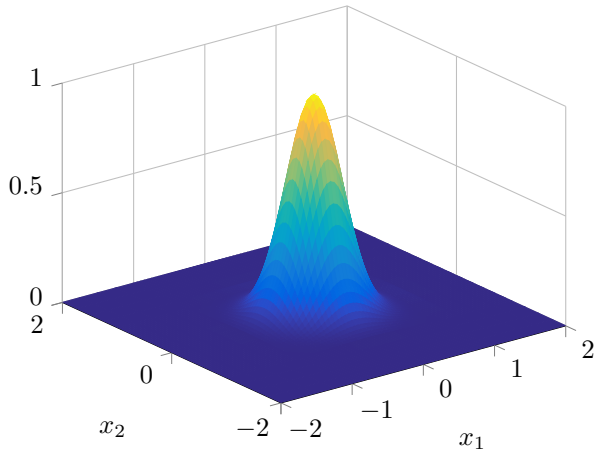

Fig. 3. Shape of the Gaussian function for exemplary values of $t = 1$ and $D = 0.05$

$$g_{kT}(\mathbf{x}) \approx \frac{1}{4\pi D} \mathrm{e}^{-\frac{||\mathbf{x}||^2}{4DkT}}, \quad k = 1..K. \tag{20}$$

Please note that this kind of modeling is also very common in the field of kernel adaptive filtering, e.g., [9]. Also, the Kriging method [3], [10] uses a modified Gaussian kernel to approximate the so-called variogram, i.e., the effect of sources on the measurements.

IV. Performance evaluation

For assessment of the estimators' performance, figures of merit need to be defined. Here, we will use the Reconstruction MSE, defined as mean square error between true field and reconstruction averaged over the whole considered area. This measure is not to mistaken with the Training MSE, which is defined as mean square error between sensor measurement and reconstructed field at sensor location.

$$\mathbf{\Phi}_{\mathrm{MT}} = \begin{bmatrix} g_T(\mathbf{x}_1^{\mathrm{s}} - \mathbf{x}_1) & g_T(\mathbf{x}_1^{\mathrm{s}} - \mathbf{x}_2) & \dots & g_T(\mathbf{x}_1^{\mathrm{s}} - \mathbf{x}_{N\cdot L}) \\ g_T(\mathbf{x}_2^{\mathrm{s}} - \mathbf{x}_1) & g_T(\mathbf{x}_2^{\mathrm{s}} - \mathbf{x}_2) & \dots & g_T(\mathbf{x}_2^{\mathrm{s}} - \mathbf{x}_{N\cdot L}) \\ \vdots & \vdots & \vdots & \vdots \\ g_T(\mathbf{x}_J^{\mathrm{s}} - \mathbf{x}_1) & g_T(\mathbf{x}_J^{\mathrm{s}} - \mathbf{x}_2) & \dots & g_T(\mathbf{x}_J^{\mathrm{s}} - \mathbf{x}_{N\cdot L}) \\ g_{2T}(\mathbf{x}_1^{\mathrm{s}} - \mathbf{x}_1) & g_{2T}(\mathbf{x}_1^{\mathrm{s}} - \mathbf{x}_2) & \dots & g_{2T}(\mathbf{x}_1^{\mathrm{s}} - \mathbf{x}_{N\cdot L}) \\ g_{2T}(\mathbf{x}_2^{\mathrm{s}} - \mathbf{x}_1) & g_{2T}(\mathbf{x}_2^{\mathrm{s}} - \mathbf{x}_2) & \dots & g_{2T}(\mathbf{x}_2^{\mathrm{s}} - \mathbf{x}_{N\cdot L}) \\ \vdots & \vdots & \vdots & \vdots \\ g_{2T}(\mathbf{x}_J^{\mathrm{s}} - \mathbf{x}_1) & g_{2T}(\mathbf{x}_J^{\mathrm{s}} - \mathbf{x}_2) & \dots & g_{2T}(\mathbf{x}_J^{\mathrm{s}} - \mathbf{x}_{N\cdot L}) \\ g_{3T}(\mathbf{x}_1^{\mathrm{s}} - \mathbf{x}_1) & g_{3T}(\mathbf{x}_1^{\mathrm{s}} - \mathbf{x}_2) & \dots & g_{3T}(\mathbf{x}_1^{\mathrm{s}} - \mathbf{x}_{N\cdot L}) \\ \vdots & \vdots & \vdots & \vdots \\ g_{KT}(\mathbf{x}_{J-1}^{\mathrm{s}} - \mathbf{x}_1) & g_{KT}(\mathbf{x}_{J-1}^{\mathrm{s}} - \mathbf{x}_2) & \dots & g_{KT}(\mathbf{x}_{J-1}^{\mathrm{s}} - \mathbf{x}_{N\cdot L}) \\ g_{KT}(\mathbf{x}_J^{\mathrm{s}} - \mathbf{x}_1) & g_{KT}(\mathbf{x}_J^{\mathrm{s}} - \mathbf{x}_2) & \dots & g_{KT}(\mathbf{x}_J^{\mathrm{s}} - \mathbf{x}_{N\cdot L}) \end{bmatrix} \tag{18}$$

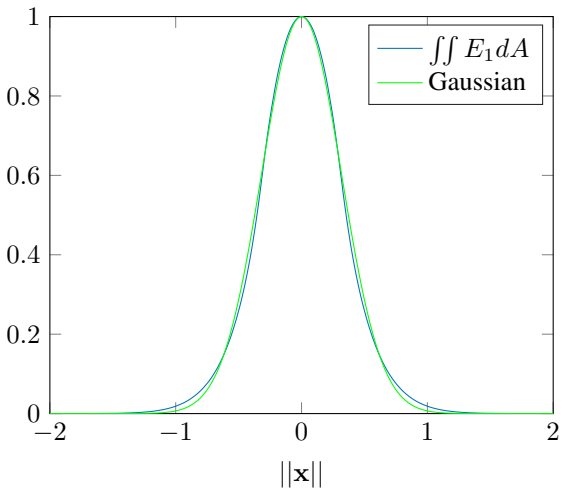

Fig. 4. Comparison of integrated exponential integral function and Gaussian function

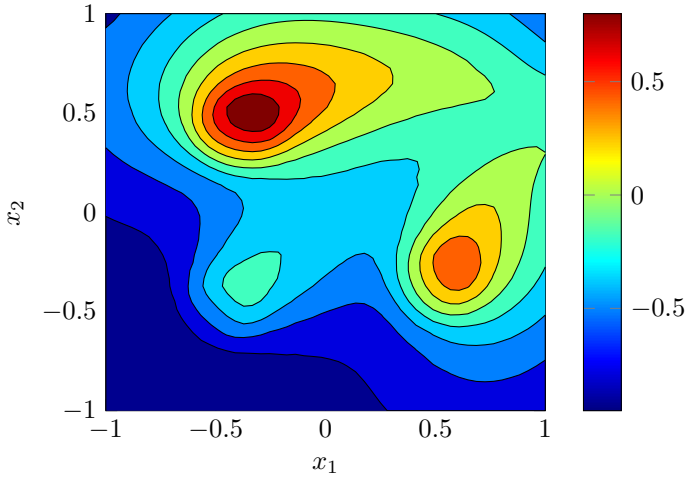

Fig. 5. Diffusion field to be estimated

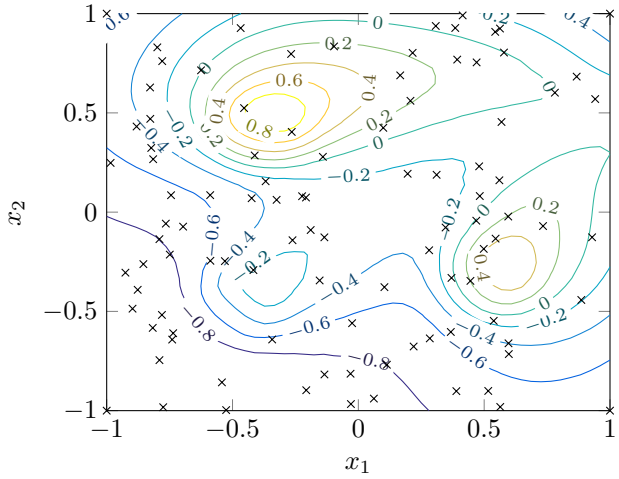

Fig. 6. Diffusion field to be estimated, with sensor locations

Fig. 5 shows an exemplary field with 3 disk shaped sources and a combined diffusion and advection process that has been generated using the COMSOL CFD software [11]. We decided to use this field instead of a pure diffusion problem with point-shaped sources, since it better matches typical real-world scenarios. This field is sampled by a fixed number of randomly placed sensors.

In order to illustrate the difference between Least Squares and OMP estimation, this field is now sampled by $J = 100$ sensors, as it is denoted in Fig. 6 by crosses. Each sensor measurement is superimposed with Gaussian distributed, uncorrelated noise of variance σ_n^2.

The depicted area of size 2×2 is now overlayed with a grid of $N \times L = 100 \times 100$ possible source locations, for which the corresponding amplitudes are subsequently estimated. Note that the equation system (12) (or (16), respectively) is heavily underdetermined, with only $J = 100$ equations for $NL = 10,000$ unknowns.

The LS estimate $\underline{\tilde{\mathbf{a}}}$ of (16) thus is not unique, one possible solution resulting in the minimum squared ℓ_2-norm of the

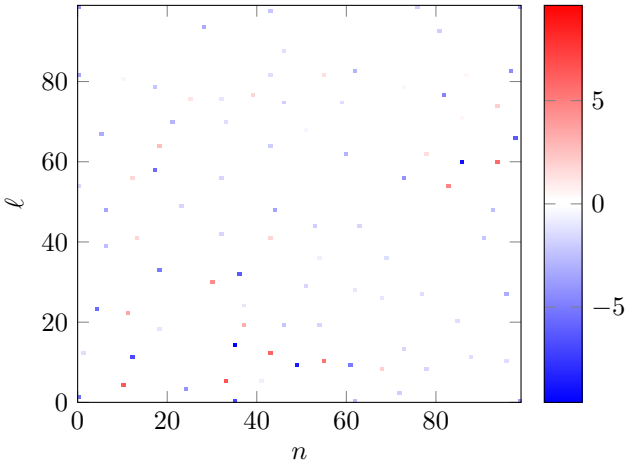

Fig. 7. Estimated coefficients $\tilde{\mathbf{a}}$ for the LS criterion

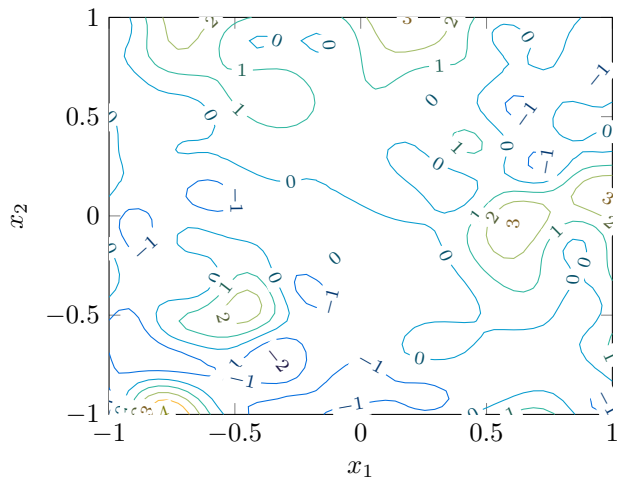

Fig. 8. Estimated field using the LS criterion

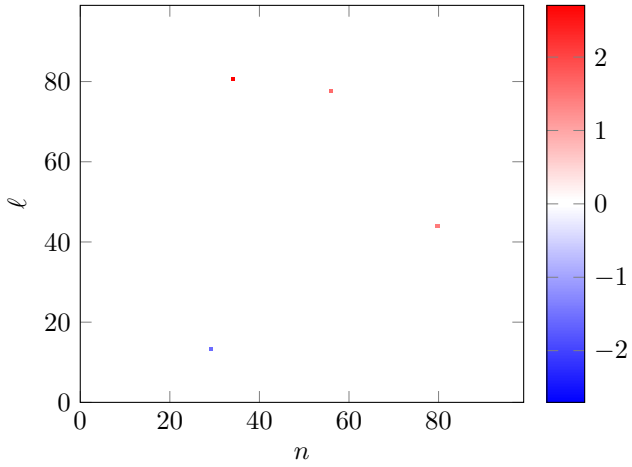

Fig. 9. Estimated coefficients $\tilde{\mathbf{a}}$ using the OMP algorithm for a sparsity of 4

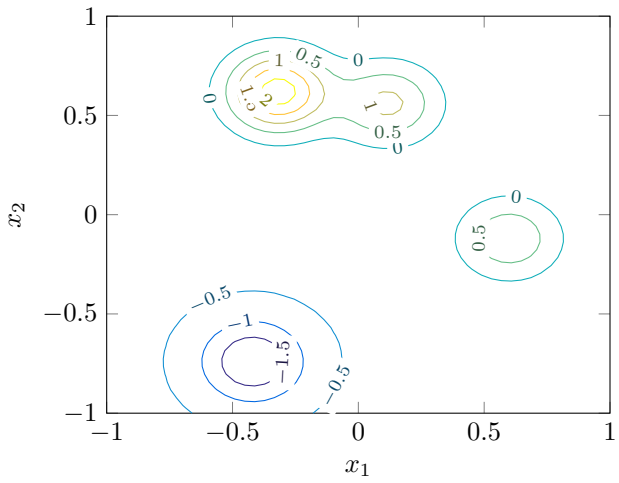

Fig. 10. Estimated field using the OMP algorithm for a sparsity of 4

estimate is given by

$$\tilde{\mathbf{a}} = \underline{\mathbf{\Phi}}^{+}\mathbf{y}, \tag{21}$$

with $(\cdot)^+$ denoting the Moore-Penrose pseudo inverse.

For Gaussian basis functions, $T = 1$, $D = 0.005$ and $\sigma_n^2 = 0.01$, the estimated coefficients $\tilde{\mathbf{a}}$ as shown in Fig. 7 are obtained. The coefficient vector has been reshaped into the 100×100 grid of possible source locations in order to indicate where the effective sources have been placed by the estimation algorithm. Please note that only $J = 100$ coefficients out of $10,000$ are nonzero due to the application of the pseudo inverse. $J - 1 = 99$ of these are depicted in the figure, while the J^{th} nonzero element is $\tilde{a}_{\text{offset}}$, which is also estimated as nonzero, but not shown explicitly.

The reconstructed field according to (9) is shown in Fig. 8. It obviously shows no resemblance to the original field, which also can be quantified by means of the Reconstruction MSE between original field and estimate, which evaluates to 0.7944 in this case.

A. *Sparse estimate*

The sparse estimate obtained by the OMP [8] algorithm for Gaussian basis functions and an exemplary prescribed sparsity of $\tilde{M} = 4$ is depicted in Fig. 9. The resulting recontructed field is shown in Fig. 10. This field follows the shape of the original field in a better way, which also is reflected in a much smaller MSE of 0.1179.

However, the estimation accuracy depends on the choice of D on the one hand and, at least for a sparse estimator, on the prescribed sparsity $\tilde{M}$. Therefore, we will show results of a manual optimization of these parameters for the scenario shown above. While the diffusion field itself was kept constant, sensor locations and individual noise realizations have been generated randomly for every run of the simulation. The presented MSEs are averaged over 100 runs each. Fig. 11 shows the resulting Reconstruction MSE for an alternating optimization of $\tilde{M}$ and D. The minimal value of 0.0319 is achieved at $\tilde{M} = 33$ and $D = 0.005$, the corresponding reconstructed field is shown in Fig. 12.

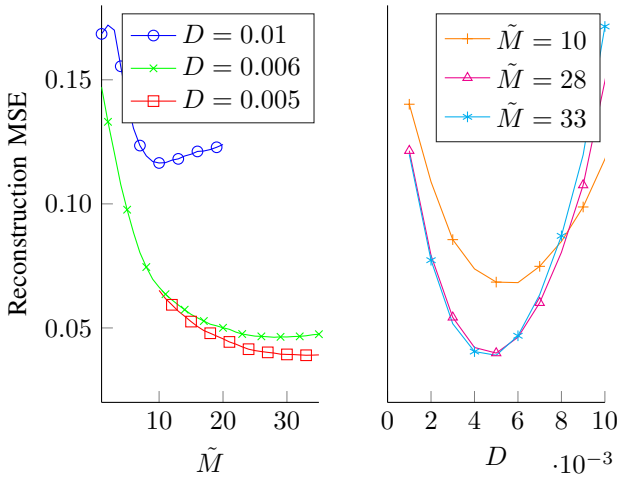

Fig. 11. Reconstruction MSE over $\tilde{M}$ and D using the OMP algorithm

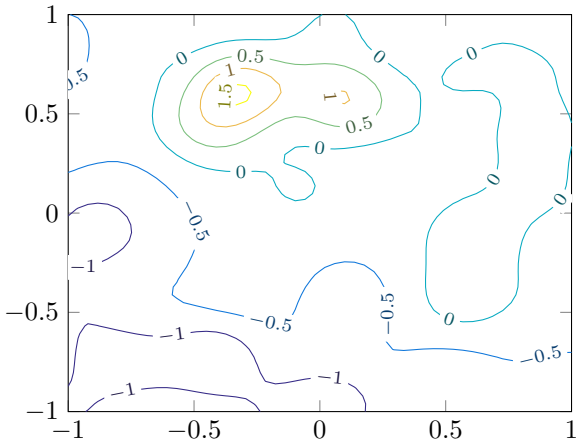

Fig. 12. Estimated field using the OMP algorithm for the optimal parameters of $\tilde{M} = 33$ and $D = 0.005$

V. Alternative Sparse Reconstruction Algorithms

The OMP algorithm requires the explicit specification of the sparsity, i.e., $\tilde{M}$, which needs to be found in a seperate step. Another commonly used method for sparse reconstruction is the Lasso algorithm that stems from the field of statistical regression [12]. Instead of a fixed sparsity, this algorithm uses a tuning parameter λ that implicitly controls the algorithm's "aggressiveness" to find a sparse solution. Exemplary, Fig. 13 shows the resulting Reconstruction MSE and the average value of $\tilde{M}$ for $D = 0.005$ and $0.01 \leq \lambda \leq 0.3$. It can be seen that the average $\tilde{M}$ corresponding to the minimum MSE is approx. 21, which differs from the value of 33 obtained above, and consequently, the achieved MSE of 0.1925 therefore is considerably larger than the MSE obtained above. This is due to the fact that the Lasso algorithm does not consider the Reconstruction MSE, but only the Training MSE, which do not directly relate to each other.

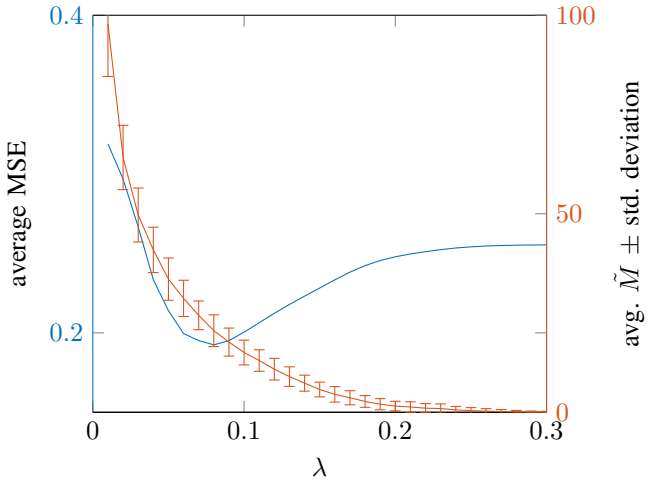

Fig. 13. Reconstruction MSE and mean $\tilde{M}$ over λ for $D = 0.005$ using the Lasso algorithm

A. Outlook on distributed reconstruction

However, for the Lasso, distributed, iterative implementations exist, e.g., the D-Lasso and its variants presented in [13]. This algorithm is able to perform sparse linear regression following the Lasso principle in a distributed fashion within a network. However, this algorithm proved to be very sensitive w.r.t. choice of its parameters, in particular step size and λ [14], so that the performance of the "centralized" Lasso could not be achieved for the scenario investigated here without putting a very high amount of effort into parameterization. In literature, there exist further algorithms for distributed sparse regression that, similar to the Lasso, also feature thresholding functions to promote sparsity of the solution, e.g., [15] and [16]. Unfortunately, these algorithms also share the same sensitivity to parameterization, whose in-depth analysis is beyond the scope of this paper.

VI. Conclusion and Future Work

In this paper, we presented a framework for the estimation of sources in a diffusive environment using Gaussian basis functions and sparse estimation techniques. The use of Gaussian basis functions was motivated through the physical process and verified through numerical simulations.

In order to perform a truly decentralized estimation of the sources, a distributed sparse estimation algorithm needs to be applied. Its design and in particular its parameterization are the subject of future investigations.

Acknowledgment

The work leading to this publication was funded by the German Research Foundation (DFG) under grants Pa2507/1 and Je722/1.

References

[1] J. Ranieri, I. Dokmanic, A. Chebira, and M. Vetterli, "Sampling and reconstruction of time-varying atmospheric emissions," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2012, pp. 3673–3676.

[2] R. Tan, G. Xing, J. Chen, W.-Z. Song, and R. Huang, "Quality-driven volcanic earthquake detection using wireless sensor networks," in *IEEE 31st Real-Time Systems Symposium (RTSS)*, Nov 2010, pp. 271–280.
[3] R. Jedermann and W. Lang, "The minimum number of sensors - interpolation of spatial temperature profiles in chilled transports," in *Wireless Sensor Networks*, ser. Lecture Notes in Computer Science, U. Roedig and C. J. Sreenan, Eds. Springer Berlin Heidelberg, 2009, vol. 5432, pp. 232–246.
[4] M. Rostami, N.-M. Cheung, and T. Quek, "Compressed sensing of diffusion fields under heat equation constraint," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013, pp. 4271–4274.
[5] J. Ranieri and M. Vetterli, "Sampling and reconstructing diffusion fields in presence of aliasing," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013, pp. 5474–5478.
[6] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions*. New York, USA: Dover Publications, Inc., 1968, ch. 5, p. 297.
[7] J. Ranieri, A. Chebira, Y. Lu, and M. Vetterli, "Sampling and reconstructing diffusion fields with localized sources," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011, pp. 4016–4019.
[8] Y. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, Nov 1993, pp. 40–44 vol.1.
[9] S. v. Vaerenbergh, "Kernel methods for nonlinear identification, equalization and separation of signals," Ph.D. dissertation, Universidad de Cantabria.
[10] J. Chilès and P. Delfiner, *Geostatistics - modeling spatial uncertainty*. John Wiley & Sons, New York, 1999.
[11] R. Jedermann and H. Paul, "Accuracy of spatial field reconstruction in wireless sensor networks - what can 100 sensors tell us?" *submitted to Hindawi International Journal of Distributed Sensor Networks*.
[12] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.
[13] G. Mateos, J. Bazerque, and G. Giannakis, "Distributed sparse linear regression," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5262–5276, 2010.
[14] M. Strauch, "Distributed Lasso algorithm for underdetermined linear systems and sparse signals," Studienarbeit, Universität Bremen, 2015.
[15] C. Ravazzi, S. Fosson, and E. Magli, "Distributed soft thresholding for sparse signal recovery," in *2013 IEEE Global Communications Conference (GLOBECOM)*, Dec 2013, pp. 3429–3434.
[16] S. Patterson, Y. Eldar, and I. Keidar, "Distributed compressed sensing for static and time-varying networks," *IEEE Transactions on Signal Processing*, vol. 62, no. 19, pp. 4931–4946, 2014.