Gölz, Martin; Zoubir, Abdelhak M.; Koivunen, Visa

Improving Inference for Spatial Signals by Contextual False Discovery Rates

Published in:
2022 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2022 - Proceedings

DOI:
10.1109/ICASSP43922.2022.9747596

Published: 01/01/2022

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

IMPROVING INFERENCE FOR SPATIAL SIGNALS BY CONTEXTUAL FALSE DISCOVERY RATES

Martin Gölz[†]*, Abdelhak M. Zoubir*[†] *and Visa Koivunen*[*]

[†]Signal Processing Group, Technische Universität Darmstadt, 64283 Darmstadt, Germany
[*]Department of Signal Processing and Acoustics, Aalto University, 02150 Espoo, Finland

ABSTRACT

A spatial signal is monitored by a large-scale sensor network. We propose a novel method to identify areas where the signal behaves interestingly, anomalously, or simply differently from what is expected. The sensors pre-process their measurements locally and transmit a local summary statistic to a fusion center or a cloud. This saves bandwidth and energy. The fusion center or cloud computes a spatially varying empirical Bayes prior on the signal's spatial behavior. The spatial domain is modeled as a fine discrete grid. The contextual local false discovery rate is computed for each grid point. A decision on the local state of the signal is made for each grid point, hence, many decisions are made simultaneously. A multiple hypothesis testing approach with false discovery rate control is used. The proposed procedure estimates the areas of interesting signal behavior with higher precision than existing methods. No tuning parameters have to be defined by the user.

***Index Terms*—** Sensor networks, multiple hypothesis testing, local false discovery rate, spatial inference, information fusion

1. INTRODUCTION

A spatial signal or phenomenon is a physical process that varies as a function of location. Such phenomena are usually observed using multiple sensors in distinct spatial locations. Spatial phenomena are encountered in a variety of application domains such as radio spectrum in radar and wireless, meteorology, environmental monitoring, seismology, and agriculture. Practical examples include the monitoring of air humidity, temperature or precipitation across a certain area, emission, pollen or ozone level maps, as well as seismic activity measured at specific locations to predict earthquakes and spectrum maps that indicate the state of the radio frequency bands in different locations.

In this work, we aim to detect and localize the regions in which the state of the observed spatial phenomenon differs from its nominal state [1]. The differences are associated with spatial regions of interesting, different or anomalous behavior. A deviation from the nominal state could be, for example, a change in occupancy of the radio spectrum, an unusual increase in seismic activity indicating an impending volcanic eruption or an anomalous decrease in air oxygen saturation inside a building caused by a fire. Spatial phenomena typically vary smoothly in space. If the phenomenon deviates from its nominal state at a certain location, similar deviations at nearby locations are likely. The state of the spatial signal is monitored by a large-scale sensor network. The problem is illustrated in Fig. 1.

Heterogeneous large-scale sensor networks are a key enabler in the development of the Internet of Things (IoT) [2]–[5]. Cheap sensors with wireless connectivity are deployed in all kind of devices from vehicles to household appliances. Modern-age sensor networks can easily comprise hundreds or thousands of nodes of heterogeneous sensing capability. Advanced inference and learning methods are needed to extract relevant information from the large volume of observed heterogeneous data. Joint processing of the raw sensor-level data at a fusion center (FC) or in the cloud produces a severe overhead of intra-network communication. The resulting bandwidth requirements lead to a congested spectrum. In addition, the continuous streaming of local information consumes energy and reduces the lifespan of a battery-powered sensor system [6], [7].

We recently [8] proposed to learn the spatial region of interesting anomalous or different behavior by multiple hypothesis testing (MHT). The observation area is discretized using a fine spatial grid. Per grid point, a decision on the local state of the phenomenon is made. The nominal state is defined in terms of the *null hypothesis* H_0. Where H_0 does not hold, the phenomenon is in one of M alternate states. The *alternative* H_1 summarizes all alternate states. Sensors in distinct locations are positioned at a sparse subset of grid points. To make efficient use of the spectrum and energy resources, each sensor performs local computations to condense its observations into a local summary statistic. This is an MHT approach, since a large number of decisions is made simultaneously. The false discovery rate (FDR) [9] is controlled at a desired level to guarantee the accurate localization of interesting spatial regions. In [8], we proposed a novel method to efficiently estimate the local false discovery rate (lfdr) [10]–[13] for each sensor. The lfdr is equivalent to the posterior probability of the null hypothesis H_0. Since spatial signals vary smoothly in space, we proposed to interpolate the lfdr at grid points in between sensors to make decisions at locations where no sensor is located. Traditional FDR controlling methods [9], [14]–[17] conduct each test with data that is independent of, or only partially correlated with the data used for other tests [18].

In this paper, we propose an empirical Bayes method that uses the contextual lfdr (clfdr) with spatially varying prior probabilities for H_0. The developed approach makes it more likely to choose the alternative H_1 in spatial regions where multiple sensors observe similarly interesting or anomalous behavior. This improves the detection power over [8] while false positive control is maintained.

Similar ideas have been pursued in a limited number of works in the open literature. However, [19], [20] suffer from high computational complexity even for a moderate number of grid points and sensors. The methods in [21], [22] are computationally demanding and base upon a data model that can be too rigid to capture the probability model under H_1 [8]. The problem of FDR control for large-scale spatial data is also considered in [23].

Our contributions are as follows. We extend our previous work [8] by introducing an empirical Bayesian method to account for spatially varying prior probabilities for H_0. We then propose two novel methods, sensor lfdr smoothing (SLS) and screened null sensor smoothing (SNS), to learn the spatially varying prior information from the data. By employing appropriately estimated spatially varying priors, the detection power is increased significantly. We additionally propose a generally applicable criterion to automatically select the best suitable spatial prior from a set of candidates. Consequently, the SLS and SNS methods do not require the adjustment of any tuning parameters. This makes both methods easy to use. Our

The work of M. Gölz is supported by the German Research Foundation (DFG) under grant ZO 215/17-2. E-mail: goelz@spg.tu-darmstadt.de.

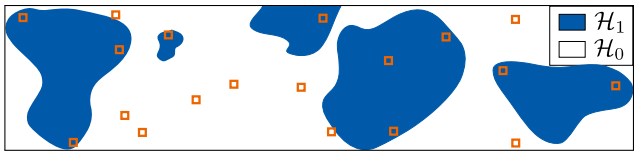


Fig. 1. A spatial phenomenon, squares indicate sensors. The unknown region of interesting/anomalous/different behavior $\mathcal{H}_1$ consists of five spatially continuous subregions (blue).

simulations confirm a significant increase in detection power while maintaining false positive control for the proposed methods.

2. PROBLEM FORMULATION

The observation area is discretized by a fine grid of Q grid points in a 2D-plane. The location of grid point $q \in \{1, \dots, Q\} = [Q]$ is denoted by the coordinate vector $\mathbf{c}_q = [x_q, y_q]^\top$. Sensors are placed at $N \leq Q$ grid points. To keep the notation simple, we use the same ordering in the indices of sensors and grid points, i.e., sensor $n \in [N]$ is located at grid point $q = n$. The state of the phenomenon at location $\mathbf{c}_q, q \in [Q]$ is defined in terms of the unknown true local hypothesis $H_q \in \{\mathrm{H}_0, \mathrm{H}_1\}$. If the observed spatial phenomenon is in its nominal state at $\mathbf{c}_q$, the null hypothesis H_0 holds. If not, the alternative H_1 is correct. We aim to identify the region where the phenomenon behaves as expected, $\mathcal{H}_0 = \{q \in [Q] : H_q = \mathrm{H}_0\}$, and where it does not, $\mathcal{H}_1 = \{q \in [Q] : H_q = \mathrm{H}_1\}$. $\mathcal{H}_0$ and $\mathcal{H}_1$ are mutually exclusive sets of grid points. A decision $\hat{H}_q$ between null and alternative is made for each $q \in [Q]$. If $\hat{H}_q = \mathrm{H}_1$, a *discovery* is made. The estimates $\hat{\mathcal{H}}_1$ and $\hat{\mathcal{H}}_0$ are found by grouping the grid points according to their $\hat{H}_q$, $q \in [Q]$.

Assume that per grid point $q \in [Q]$, the *local null probability* $\Pr(H_q = \mathrm{H}_0)$ is known. Then, estimating $\mathcal{H}_1$ by

$$\hat{\mathcal{H}}_1 = \underset{\mathcal{H} \subseteq [Q]}{\operatorname{argmax}} \left\{ |\mathcal{H}| : \frac{1}{|\mathcal{H}|} \cdot \sum_{q \in \mathcal{H}} \Pr(H_q = \mathrm{H}_0) \leq \alpha_{\mathrm{FDR}} \right\} \quad (1)$$

results in an estimate $\hat{\mathcal{H}}_1$ that contains at most a fraction α_{FDR} of false positives, i.e., grid points where $\hat{H}_q = \mathrm{H}_1 \,|\, H_q = \mathrm{H}_0$. Thus, the estimator in Eq. (1) controls the FDR [9]

$$\mathrm{FDR} = \mathrm{E}\left[\frac{V}{\max(R,1)}\right] = \mathrm{E}\left[\frac{\sum_{q \in \mathcal{H}_0} \mathbb{1}\{\hat{H}_q = \mathrm{H}_1\}}{\max\left(\sum_{q=1}^{Q} \mathbb{1}\{\hat{H}_q = \mathrm{H}_1\}, 1\right)}\right] \quad (2)$$

at level α_{FDR}. V is the number of false discoveries, R the number of all discoveries, $V/\max(R,1)$ the false discovery proportion (FDP) and $\mathbb{1}\{\cdot\}$ the indicator function. The local null probabilities are unknown in practice and need to be estimated from the data.

3. PROPOSED METHOD

We estimate $\hat{\mathcal{H}}_1$ by Eq. (1). We compute the clfdr for each sensor location. The clfdr is found from the local summary statistics, the probability models under the null and the alternative, as well as the spatially varying prior probability of H_0. While the model under H_0 is known, the model under H_1 has to be estimated from the data. The local prior has to be estimated as well. A local prior allows to quantify the local state of the phenomenon more accurately than a spatially constant prior. This increases the detection power. Finally, the sensor clfdr's are interpolated to enable decision making at grid points where no sensors are located. We begin this section by introducing the probability models for the local summary statistics. We proceed with the definition of the clfdr and propose its estimator. Algorithm 1 details the proposed method, referred to as clfdr-sMoM.

We use p-values as local summary statistics. Other options include likelihood ratios, empirical likelihoods or any other sufficient statistic. A discussion on the requirements for suitable local summary statistics is provided in [8]. A p-value quantifies the support for H_0 at the node, small values indicating less support. To compute the p-values, each sensor must know or learn [24], [25] its local data model under H_0. p-values enable the fusion of information from heterogeneous sensors with locally different data models.

The set of p-values $\mathcal{P}^N = \{p_1, \dots, p_N\}$ is available at the FC. The probability models under the null and the alternative are

$$\begin{aligned} H_n = \mathrm{H}_0 : &\quad P_n \sim f_{P|\mathrm{H}_0}(p), \\ H_n = \mathrm{H}_1 : &\quad P_n \sim f_{P_n|\mathrm{H}_1}(p), \end{aligned} \quad (3)$$

with $f_{P|\mathrm{H}_0}(p)$ and $f_{P_n|\mathrm{H}_1}(p)$ the p-value probability density function (PDF) at the location of sensor $n \in [N]$ under H_0 and H_1, respectively. The p-values follow a uniform distribution at all sensors where H_0 is in place, $f_{P|\mathrm{H}_0}(p) = \mathcal{U}[0,1]$. Any kind of interesting, different or anomalous behavior is defined in terms of the alternative H_1. Thus, $f_{P_n|\mathrm{H}_1}(p)$ varies across sensors $n \in [N]$.

The clfdr is the local null probability given the p-value, i.e., $\mathrm{clfdr}_n(p_n) = \Pr(H_n = \mathrm{H}_0 \,|\, p_n)$ and will be used as the local null probabilities of the sensors in Eq. (1). It is defined as [26]

$$\mathrm{clfdr}_n(p) = \frac{\pi_{n,0} f_{P|\mathrm{H}_0}(p)}{\pi_{n,0} f_{P|\mathrm{H}_0}(p) + (1 - \pi_{n,0}) f_{P_n|\mathrm{H}_1}(p)}, \quad (4)$$

with $\pi_{n,0}$ being the prior probability for H_0 at sensor $n \in [N]$.

In practice, the $\pi_{n,0}$ and $f_{P_n|\mathrm{H}_1}(p)$ are unknown $\forall\, n \in [N]$ and need to be estimated from the data. Per sensor $n \in [N]$, a single p-value p_n is available. Thus, estimating the individual $f_{P_n|\mathrm{H}_1}(p)$ is infeasible. Instead, the idea of MHT is to replace the local models $f_{P_n|\mathrm{H}_1}(p)$ with the model for all local summary statistics observed under H_1, $f_{P|\mathrm{H}_1}(p) = N^{-1} \sum_{n=1}^{N} f_{P_n|\mathrm{H}_1}(p)$ [13]. Estimating $f_{P|\mathrm{H}_1}(p)$ is often difficult, but doable. The estimated clfdr is

$$\hat{\mathrm{clfdr}}_n(p) = \frac{\hat{\pi}_{n,0} f_{P|\mathrm{H}_0}(p)}{\hat{\pi}_{n,0} f_{P|\mathrm{H}_0}(p) + (1 - \hat{\pi}_{n,0}) \hat{f}_{P|\mathrm{H}_1}(p)} \quad (5)$$

We propose to determine $\hat{f}_{P|\mathrm{H}_1}(p)$ with the fast and efficient estimator from our previous work [8]. Finding estimates $\hat{\pi}_{n,0}$, $\forall\, n \in [N]$ that accurately capture the spatial variation of the signal is not straightforward. Note, that using a constant $\hat{\pi}_{n,0} = \hat{\pi}_0 \,\forall\, n \in [N]$ in Eq. (5) recovers the lfdr, as used in [8, Algorithm 1].

Deploying clfdr's with a locally varying prior instead of lfdr's with a fix prior offers great potential to increase the detection power significantly [21]–[23], [26], [27]. In Sec. 3.1, we propose two approaches to learn a spatially varying prior from the data. We model $\pi_{n,0} \,\forall\, n \in [N]$ as a spatially smooth function $\pi_{n,0} = g\big(\{\mathbf{c}_m\}_{m=1}^N; \{p_m\}_{m=1, m\neq n}^N\big)$. $g(\cdot)$ depends on the sensor location $\mathbf{c}_n$ and the locations and p-values of all other sensors.

3.1. Spatial prior estimation

Both methods proposed in this section apply kernel smoothing [28] in a two-dimensional space. A kernel is a positive real-valued function $K(d_{q,r}; b)$ of the Euclidean distance $d_{q,r} = ||\mathbf{c}_q - \mathbf{c}_r||$ between two grid point locations and a scaling parameter $b > 0$. The maximum of the kernel is $K(0; b) = 1$. Its value typically decreases with increasing $d_{q,r}$. The rate of decrease depends on the shape of the kernel and the value of b, with smaller b resulting in a faster decrease. We work with a set $\mathcal{K}$ of kernel functions $K(\cdot)$, containing the most widely used kernel functions with bounded and unbounded support, see Table 1. We propose a method to automatically select the kernel and the scaling parameter in Sec. 3.2.

3.1.1. Sensor lfdr smoothing (SLS)

We propose to use the Nadaraya-Watson kernel weighted average [29], [30] of the lfdr's estimated for the other sensors,

$$\hat{\pi}_{n,0}^{\mathrm{SLS}}(K(\cdot); b) = \frac{\sum_{\substack{m=1 \\ m \neq n}}^{N} K(d_{n,m}; b) \cdot \hat{\mathrm{lfdr}}(p_m)}{\sum_{\substack{m=1 \\ m \neq n}}^{N} K(d_{n,m}; b)}. \quad (6)$$

The $\hat{\mathrm{lfdr}}(p_n), n \in [N]$ are found with [8, Algorithm 1].

3.1.2. Screened null sensor smoothing (SNS)

In this section, we propose a two-step prior estimator for inference with clfdrs. This procedure is based on a method used to compute p-value weights in [26].

During the initial screening step, a screened subset $\mathcal{N}_S = \{n \in [N] : p_n \geq \tau\}$ of sensors is selected. The threshold τ is chosen

Algorithm 1 clfdr-sMoM: Estimation of $\hat{\mathcal{H}}_1$ with clfdr's

Input: $\mathcal{P}^N = \{p_1, \dots, p_N\}$, $\mathcal{K}$, $\mathcal{B}$, α_{FDR}
Output: $\hat{\mathcal{H}}_1$

Determine $\hat{f}_{P|\mathrm{H}_1}(p)$ with lfdr-sMoM [8, Alg. 1]
for $K(\cdot) \in \mathcal{K}$ **do**
  for $b \in \mathcal{B}$ **do**
    Compute $\forall\, n \in [N]$
      a: $\hat{\pi}^{\mathrm{SLS}}_{n,0}(K(\cdot); b)$ (clfdr-sMoM-SLS) or
      b: $\hat{\pi}^{\mathrm{SNS}}_{n,0}(K(\cdot); b)$ (clfdr-sMoM-SNS)
Determine $(K^*(\cdot), b^*) = \theta^* = \underset{\theta=(K(\cdot)\in\mathcal{K}, b\in\mathcal{B})}{\operatorname{argmax}} c(\theta)$
Find $\hat{\pi}_{n,0}$ from $\hat{\pi}^{\mathrm{SLS}}_{n,0}(K^*(\cdot); b^*)$ or $\hat{\pi}^{\mathrm{SNS}}_{n,0}(K^*(\cdot); b^*)\, \forall\, n \in [N]$
Compute $\hat{\mathrm{clfdr}}_n = \hat{\mathrm{clfdr}}_n(p_n)\, \forall\, n \in [N]$ from Eq. (5)
For $N < q \leq Q$, find $\hat{\mathrm{clfdr}}_q$ by interpolation [8, Eq. (27)]
Estimate $\hat{\mathcal{H}}_1$ with $\Pr(H_q = \mathrm{H}_0) = \hat{\mathrm{clfdr}}_q\, \forall\, q \in [Q]$ in Eq. (1)

such that $\mathcal{N}_S$ only contains nodes at which H_0 holds with high confidence. A scheme to automatically select τ was proposed in [26]. The screening step is followed by computing

$$\hat{\pi}^{\mathrm{SNS}}_{n,0}(K(\cdot); b) = \frac{\sum_{\substack{m\in\mathcal{N}_S \\ m\neq n}} K(d_{n,m}; b)}{(1-\tau)\sum_{\substack{m=1 \\ m\neq n}}^{N} K(d_{n,m}; b)}. \tag{7}$$

The numerator in Eq. (7) contains the empirical count of the number of sensors where H_0 holds, weighted by their distance to sensor $n \in [N]$. The denominator represents the expected number of sensors at which H_0 is in place, weighted by their distance to sensor $n \in [N]$. For $K(\cdot) = 1$, the spatially constant standard estimate [14] for the proportion of true alternatives $\sum_{n=1}^{N} \mathbb{1}\{p_n \geq \tau\}/N$ is recovered. There is a subtle difference between $\hat{\pi}^{\mathrm{SNS}}_{n,0}(K(\cdot); b)$ and the estimator in [26]: In this work, we estimate the *prior* local null probabilities. Thus, sensor n is excluded from the sums in $\hat{\pi}^{\mathrm{SNS}}_{n,0}(K(\cdot); b)$, $n \in [N]$.

SNS estimates the prior from a screened subset of sensor p-values. This can cause stability issues, if the sensor network is composed of a small number of sensors.

3.2. A criterion for selecting the best prior

SLS and SNS both require the definition of a certain kernel function $K(d; b)$ and a bandwidth parameter b. A generally optimal kernel/scale parameter combination does not exist. The practitioner may not necessarily possess the required a priori knowledge to confidently select a kernel and a scale parameter for a certain application upfront. Hence, we propose a criterion to automate the selection of the best prior from a set of candidates.

Let the z-score at sensor $n \in [N]$ be denoted by $z_n = \Phi^{-1}(p_n)$. $\Phi(p)$ is the standard normal cumulative distribution function (CDF). Due to the one-to-one mapping between p-values and z-scores, the PDF of the z-scores can be expressed in terms of the p-value PDF. Under H_0, the z-scores follow the standard normal PDF, $f_{Z|\mathrm{H}_0}(z) = \phi(z)$. The estimate for the PDF under the alternative is $\hat{f}_{Z|\mathrm{H}_1}(z) = \hat{f}_{P|\mathrm{H}_1}(\Phi(z)) \cdot \phi(z)$. Let $\hat{\pi}_{n,0}(\theta)$ denote the local prior estimate for sensor $n \in [N]$ that depends on a parameter θ. The estimated sensor likelihood function (LF) in terms of the z-score, marginalized over the state of the phenomenon, is then

$$l_n(\theta) = \hat{\pi}_{n,0}(\theta) f_{Z|\mathrm{H}_0}(z_n) + (1 - \hat{\pi}_{n,0}(\theta)) \hat{f}_{Z|\mathrm{H}_1}(z_n). \tag{8}$$

We define the weighted average over marginalized sensor likelihoods

$$c(\theta) = \sum_{n=1}^{N} w_n l_n(\theta), \tag{9}$$

where the w_n's represent a set of pre-defined weights such that $\sum_{n=1}^{N} w_n = 1$. The likelihoods $l_n(\theta)\, \forall\, n \in [N]$ quantify how well a local prior estimate fits to the data. Thus, we propose to maximize $c(\theta)$ to identify the value of θ for which $\hat{\pi}_{q,0}(\theta)$ fits best to the observations. The $w_n, n \in [N]$ can be used to weight the importance of individual sensors, for example, to increase the influence of data from sensors that are known to provide particularly trustworthy local summary statistics.

Table 1. The set $\mathcal{K}$ of kernel functions used in this work.

Name	$K(d; b)$
Gaussian	$(2\pi b^2)^{-1/2} \exp(-d^2/b^2)$
Exponential	$(2b)^{-1} \exp(-d/2b)$
Epanechnikov	$3/(4b) \cdot \left(1 - d^2/b^2\right) \cdot \mathbb{1}\{d \leq b\}$
Linear	$b^{-1}\left(1 - d/b\right) \cdot \mathbb{1}\{d \leq b\}$
Cosine	$\cos\left(\pi d/(2b)\right) \cdot \mathbb{1}\{d \leq b\}$

While we use p-values to compute the clfdr's, z-scores are used to determine the best kernel/scale parameter combination for the prior. The reasoning is as follows. $f_{P|\mathrm{H}_1}(p)$ is typically large for a small p. Thus, balancing $c(\theta)$ for p-values such that it is not dominated by a couple of very small p-values requires careful selection of the w_n. For z-scores and with uniform weights $w_n = N^{-1}\, \forall\, n \in [N]$, all sensors play a similar role in determining θ.

In SLS and SNS, $\theta = (K(\cdot), b)$ is a tuple of the Kernel function and the scale parameter. We observed that it suffices to select a b that lies within a certain interval of suitable bandwidth parameters. We deploy a grid search over $\mathcal{K}$ and a coarse bandwidth grid $\mathcal{B}$ to identify a tuple of suitable prior parameters.

4. SIMULATION RESULTS

In this section, we analyze the performance of the proposed method on simulated radio frequency electromagnetic field data. Interesting behavior refers to spectrum occupancy, i.e., presence of a signal at a sensor. Thus, the region of anomalous behavior is composed of all sensors at which a signal is decided to be present. The signals are simulated with the nonuniform sampling method from [31]. More details on the data generation are given in [8].

The observation area is discretized by a 100×100 grid. We provide results for three scenarios. Differing propagation conditions lead to differently shaped signal regions in urban and suburban environments. In a suburban environment, the dominance of the line-of-sight (LOS) propagation path leads to $\mathcal{H}_1$ with smooth boundaries, whereas the large number of scatterers in an urban environment causes non-LOS (NLOS) propagation and rough region boundaries.

Sc. A: 2 sources in a suburban environment, low transmission power.

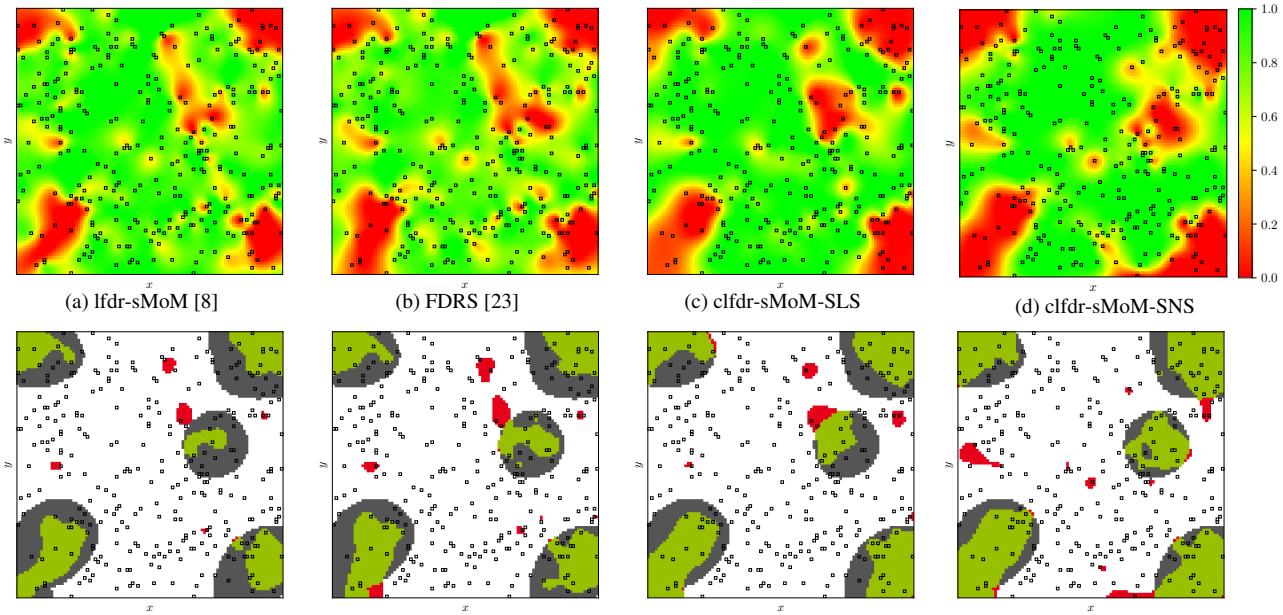
Sc. B: 8 sources in a suburban environment, low transmission power.

Sc. C: 1 source in an urban environment, high transmission power

The regions of anomalous behavior are conceptually very different for these scenarios, which poses different challenges. In Sc. A, the local subregions are small and rare. Thus, the nominal FDR is easily violated if a prior estimator is overly optimistic and associates a too large region with high signal probability. In Sc. B, H_1 covers a large fraction of the observed area, but the region of anomalous behavior is composed of many small locally continuous subregions. This tests the prior's capability to capture a short-term decrease in the local null probability. The phenomenon in Sc. C exhibits a large, spatially continuous region of anomaly, for which a good prior estimate leads to an increased detection power in a large area.

We also vary the size of the sensor network. We compare results for $N \in 300, 1000, 3000$ sensors. $N = 300$ is an extreme example, because p-values are available for only 3% of the grid points. A small number of sensors means that the probability models are estimated from a small number of observations. We use the same coarse bandwidth grid $\mathcal{B} = \{0.1 + j/2\}_{j=1}^{40}$ for clfdr-sMoM-SLS and clfdr-sMoM-SNS in all scenarios. FDR and detection power are determined by averaging over 200 realizations of each experiment. The code to reproduce the result is available on github.

Competitors: Only lfdr-based or clfdr-based MHT methods that scale well with network size should be applied to perform inference with large-scale sensor networks. Among the lfdr-based methods, the method of moment-based approach from [8] (lfdr-sMoM) provides the best performance. Hence, we compare lfdr-sMoM to clfdr-sMoM-SLS and clfdr-sMoM-SNS proposed in this work and the also



(a) lfdr-sMoM [8] (b) FDRS [23] (c) clfdr-sMoM-SLS (d) clfdr-sMoM-SNS

Fig. 2. Sc. B with $N = 300$ sensors. First row: The estimated lfdr's (a) or clfdr's (b) - (d) for an exemplary run. Second row: The detection patterns for $\alpha_{\mathrm{FDR}} = 0.1$. Green, red and gray indicate true, false and missed discoveries. Boxes mark sensor locations. The clfdr's capture the underlying structure better as indicated by a larger number of true positives.

clfdr-based false discovery rate smoothing (FDRS) from [23].

Table 2 provides the empirical FDR levels and the detection power for the typical nominal FDR level $\alpha_{\mathrm{FDR}} = 0.1$, i.e., at most 10% of the grid points in $\hat{\mathcal{H}}_1$ should be false discoveries. The FDR is strictly controlled for all scenarios and network sizes for lfdr-sMoM and the proposed clfdr-sMoM-SLS. With clfdr-sMoM-SNS, the nominal level is violated for Sc. B with $N = 300$ sensors. This is due to the initial screening step: If the number of sensors in the network is already small, further reducing the data used for the estimation of the spatially varying prior by the screening step may yield an unstable estimate. In fact, the observed FDR violation is caused by a few realizations of the experiment during which the FDP is considerably higher than usually. FDRS breaks down completely for Sc. A. While the empirical FDR decreases with increasing N, the nominal level is violated substantially for all considered network sizes. We conclude that FDRS can overestimate the regions of low null probability in scenarios where $\mathcal{H}_1$ is small.

Table 2 shows that the appropriate incorporation of a spatially varying null prior leads to massive gains in detection power. With clfdr-sMoM-SLS, the detection power is about 50% higher than with a spatially constant prior. With clfdr-sMoM-SNS, the gain reaches almost 100%. FDRS provides a lower gain, despite violating the nominal FDR in some cases.

In Fig. 2, we provide the estimated lfdr's and clfdr's to illustrate that the gain in detection power for clfdr-sMoM-SLS and clfdr-sMoM-SLS is caused by an improvement in the estimated spatial signal structure. In Fig. 3, the empirical FDR and the detection power are plotted versus the nominal α_{FDR} for Sc. C. The plots underline that our methods maintain FDR control for a wide range of nominal levels. In addition, we see that a substantial gain in detection power for all α_{FDR} is obtained.

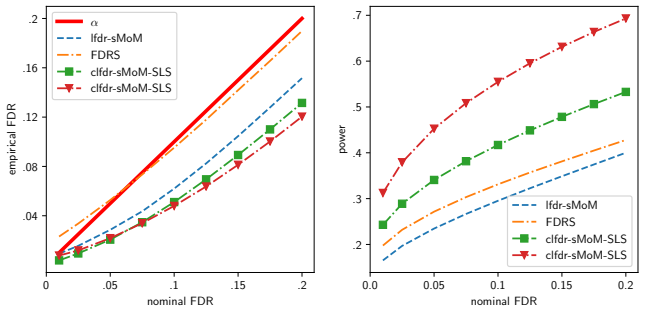
Both proposed methods, clfdr-sMoM-SLS and clfdr-sMoM-SNS result in significantly more accurate estimates $\hat{\mathcal{H}}_1$ than existing methods. For a sensor network with small N, we recommend clfdr-sMoM-SLS to guarantee strict FDR control. If the sensor network comprises more nodes, we recommend clfdr-sMoM-SNS, as it provides the maximal gain in terms of detection power.

5. CONCLUSION

We have proposed a novel method to identify the regions of anomalous, interesting or different behavior in spatial phenomena. The method is efficient in terms of computational complexity, bandwidth

Table 2. FDR and detection power for all considered scenarios and network sizes with nominal FDR level $\alpha_{\mathrm{FDR}} = 0.1$. The FDR is strictly controlled by lfdr-sMoM and the proposed clfdr-sMoM-SLS. clfdr-sMoM-SNS violates α_{FDR} in Sc. C when N is very small, but in general provides the highest power. FDRS breaks down for Sc. A.

		$N = 300$		$N = 1000$		$N = 3000$	
		FDR	Power	FDR	Power	FDR	Power
Sc. A	lfdr-sMoM [8]	.039	.117	.023	.138	.028	.163
	FDRS [23]	**.415**	.239	**.324**	.294	**.264**	.377
	clfdr-sMoM-SLS	.031	.211	.022	.260	.026	.283
	clfdr-sMoM-SNS	.109	.224	.034	.320	.013	.404
Sc. B	lfdr-sMoM [8]	.070	.263	.050	.286	.052	.302
	FDRS [23]	.082	.257	.057	.312	.047	.378
	clfdr-sMoM-SLS	.052	.375	.038	.410	.038	.431
	clfdr-sMoM-SNS	**.156**	.494	.068	.560	.030	.604
Sc. C	lfdr-sMoM [8]	.079	.283	.062	.295	.060	.297
	FDRS [23]	**.149**	.303	.095	.331	.074	.388
	clfdr-sMoM-SLS	.063	.406	.051	.417	.055	.418
	clfdr-sMoM-SNS	.088	.518	.048	.555	.028	.577



(a) Empirical vs. nominal FDR (b) Detection power vs. nominal FDR

Fig. 3. Sc. C with $N = 1000$ sensors. The FDR is controlled at all levels except for FDRS at small α_{FDR}. The proposed clfdr-sMoM-SLS and clfdr-sMoM-SNS provide the largest detection power.

and energy consumption. We incorporated a spatially varying prior into a multiple hypothesis testing method. This boosts the detection power in comparison to existing work. The prior has to be estimated such that the underlying spatial structure of the phenomenon of interest is captured with high fidelity. Otherwise, the constraint on false positives is violated. Our results demonstrate that a high-quality local prior can be learned from the data fully autonomously.

6. REFERENCES

[1] R. Nowak and U. Mitra, "Boundary estimation in sensor networks: Theory and methods," in *Information Processing in Sensor Networks*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 80–95.

[2] E. Arias-de-Reyna, P. Closas, D. Dardari, and P. M. Djuric, "Crowd-based learning of spatial fields for the internet of things: From harvesting of data to inference," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 130–139, Sep. 2018.

[3] F. K. Teklehaymanot, M. Muma, and A. M. Zoubir, "Diffusion-based Bayesian cluster enumeration in distributed sensor networks," in *Proc. 2018 IEEE Statist. Signal Process. Workshop (SSP)*, IEEE, Jun. 2018.

[4] M. R. Leonard, M. Stiefel, M. Fauss, and A. M. Zoubir, "Robust sequential testing of multiple hypotheses in distributed sensor networks," in *Proc. 2018 IEEE Int. Conf. Acoust. Speech Signal Process.*, IEEE, Apr. 2018.

[5] T. Hasija, M. Gölz, M. Muma, P. J. Schreier, and A. M. Zoubir, "Source enumeration and robust voice activity detection in wireless acoustic sensor networks," in *Proc. 53rd Asilomar Conf. Signals Syst. Comput.*, IEEE, Nov. 2019.

[6] X. Wang, G. Li, and P. K. Varshney, "Detection of sparse signals in sensor networks via locally most powerful tests," *IEEE Signal Process. Lett.*, vol. 25, no. 9, pp. 1418–1422, Sep. 2018.

[7] E. Nitzan, T. Halme, and V. Koivunen, "Bayesian methods for multiple change-point detection with reduced communication," *IEEE Trans. Signal Process.*, vol. 68, pp. 4871–4886, 2020.

[8] M. Gölz, A. M. Zoubir, and V. Koivunen, "Multiple hypothesis testing framework for spatial signals," Aug. 27, 2021. arXiv: 2108.12314.

[9] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *J. Roy. Statist. Soc. Ser. B*, vol. 57, no. 1, pp. 289–300, 1995.

[10] B. Efron, R. Tibshirani, J. D. Storey, and V. Tusher, "Empirical Bayes analysis of a microarray experiment," *J. Amer. Statist. Assoc.*, vol. 96, no. 456, pp. 1151–1160, 2001.

[11] B. Efron, "Local false discovery rates," Tech. Rep., 2005.

[12] ——, "Microarrays, empirical Bayes and the two-groups model," *Statist. Sci.*, vol. 23, no. 1, pp. 1–22, Feb. 2008.

[13] ——, *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction*, ser. Institute of Mathematical Statistics Monographs. Cambridge, UK: Cambridge University Press, 2010.

[14] J. D. Storey, "A direct approach to false discovery rates," *J. Roy. Statist. Soc. Ser. B (Statist. Methodology)*, vol. 64, no. 3, pp. 479–498, Aug. 2002.

[15] Y. Benjamini and R. Heller, "False discovery rates for spatial signals," *J. Amer. Statist. Assoc.*, vol. 102, no. 480, pp. 1272–1281, 2007.

[16] C. Zhang, J. Fan, and T. Yu, "Multiple testing via FDRL for large-scale imaging data," *Ann. Stat.*, vol. 39, no. 1, pp. 613–642, Feb. 2011.

[17] A. Chouldechova, "False discovery rate control for spatial data," Ph.D. dissertation, Stanford University, 2014.

[18] Y. Benjamini and D. Yekutieli, "The control of the false discovery rate in multiple testing under dependency," *Ann. Stat.*, vol. 29, no. 4, pp. 1165–1188, 2001.

[19] W. Sun, B. J. Reich, T. T. Cai, M. Guindani, and A. Schwartzman, "False discovery control in large-scale spatial multiple testing," *J. Roy. Statist. Soc. Ser. B*, vol. 77, no. 1, pp. 59–83, Apr. 2014.

[20] H. Shu, B. Nan, and R. Koeppe, "Multiple testing for neuroimaging via hidden markov random field," *Biometrics*, vol. 71, no. 3, pp. 741–750, May 2015.

[21] T. Halme, M. Gölz, and V. Koivunen, "Bayesian multiple hypothesis testing for distributed detection in sensor networks," in *Proc. 2019 IEEE Data Sci. Workshop*, IEEE, Jun. 2019.

[22] M. Gölz, M. Muma, T. Halme, A. M. Zoubir, and V. Koivunen, "Spatial inference in sensor networks using multiple hypothesis testing and Bayesian clustering," in *Proc. 27th Eur. Signal Process. Conf.*, Sep. 2019, pp. 1–5.

[23] W. Tansey, O. Koyejo, R. A. Poldrack, and J. G. Scott, "False discovery rate smoothing," *J. Amer. Statist. Assoc.*, vol. 113, no. 523, pp. 1156–1171, Jun. 2018.

[24] A. M. Zoubir and D. R. Iskander, *Bootstrap Techniques for Signal Processing*. Cambridge University Press, Jan. 2001.

[25] M. Gölz, V. Koivunen, and A. M. Zoubir, "Nonparametric detection using empirical distributions and bootstrapping," in *Proc. 25th Eur. Signal Process. Conf.*, IEEE, Aug. 2017.

[26] T. T. Cai, W. Sun, and Y. Xia, "LAWS: A locally adaptive weighting and screening approach to spatial multiple testing," *J. Amer. Statist. Assoc.*, pp. 1–14, Jan. 2021.

[27] S. Chen and S. Kasiviswanathan, "Contextual online false discovery rate control," in *Proc. 23rd Int. Conf. Artif. Intell. Statist.*, ser. Proceedings of Machine Learning Research, vol. 108, PMLR, 26–28 Aug 2020, pp. 952–961.

[28] T. Hastie, J. Friedman, and R. Tibshirani, "Kernel methods," in *The Elements of Statistical Learning*, Springer New York, 2001, pp. 165–192.

[29] E. A. Nadaraya, "On estimating regression," *Theory Probab. Appl.*, vol. 9, no. 1, pp. 141–142, Jan. 1964.

[30] G. S. Watson, "Smooth regression analysis," *Sankhyā Ser. A*, vol. 26, no. 4, pp. 359–372, 1964.

[31] X. Cai and G. Giannakis, "A two-dimensional channel simulation model for shadowing processes," *IEEE Trans. Veh. Technol.*, vol. 52, no. 6, pp. 1558–1567, Nov. 2003.