

SAMPLING AND RECONSTRUCTING DIFFUSION FIELDS WITH LOCALIZED SOURCES

Juri Ranieri*, Amina Chebira, Yue M. Lu and Martin Vetterli†

School of Computer and Communication Sciences
 Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015, Switzerland
 E-mail: {juri.ranieri, amina.chebira, yue.lu, martin.vetterli}@epfl.ch

ABSTRACT

We study the spatiotemporal sampling of a diffusion field generated by K point sources, aiming to fully reconstruct the unknown initial field distribution from the sample measurements. The sampling operator in our problem can be described by a matrix derived from the diffusion model. We analyze the important properties of the sampling matrices, leading to precise bounds on the spatial and temporal sampling densities under which perfect field reconstruction is feasible. Moreover, our analysis indicates that it is possible to compensate linearly for insufficient spatial sampling densities by oversampling in time. Numerical simulations on initial field reconstruction under different spatiotemporal sampling densities confirm our theoretical results.

Index Terms— Diffusion equation, initial inverse problems, spatiotemporal sampling, point sources localization, compressed sensing

1. INTRODUCTION

Sensor networks are spatiotemporal sampling devices. They are often used to sense phenomena driven by well-known physical fields. In the case of diffusive fields, the dimensions – space and time – are not homogenous and increasing the spatial density is often more expensive than increasing the temporal sampling rate, as more nodes are needed in the network. In contrast, the temporal frequency can be increased until the hardware limits imposed by the sensor nodes are reached. Even if multidimensional sampling is well documented in literature, there is a lack of results concerning the trade-off between non-homogenous dimensions, while it is a fundamental aspect for these sampling scenarios.

In this paper, we consider the spatiotemporal sampling and reconstruction of a one-dimensional diffusive field $f(x, t)$, modeled by the heat equation. That is

$$\frac{\partial f(x, t)}{\partial t} = \gamma \frac{\partial^2 f(x, t)}{\partial x^2} + h(x, t), \quad t \geq 0, \quad (1)$$

where γ is the diffusion coefficient and $h(x, t)$ represents the external sources. When $t = 0$, the field $f_0(x) = f(x, 0)$ represents the *initial field distribution*. We assume K point sources appearing at $t = 0$. Generally, the sources are defined in $h(x, t)$, however, here we know when the sources appear and model them as an initial field distribution. More precisely, we define $f_0(x) =$

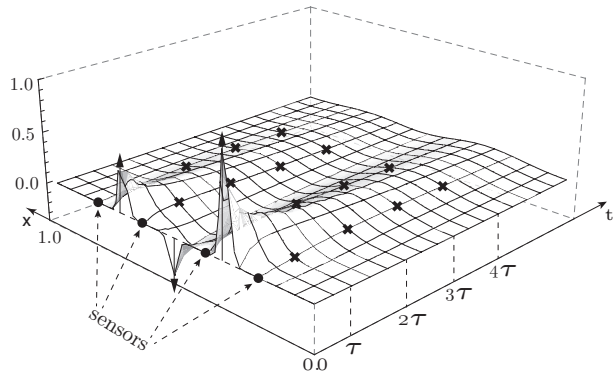


Fig. 1. A diffusion field induced by three point sources (arrows) at $t = 0$, sampled by a network of $N_s = 4$ sensors (dots ●), each one collecting $N_t = 4$ samples. The spatiotemporal sampling grid (crosses x) is the intersection between the position in space of the sensors and the temporal sampling instants ($\tau, 2\tau, 3\tau, 4\tau$).

$\sum_{m=1}^K a_m \delta(x - x_m)$, where a_m are the amplitudes of the K sources and x_m their positions. We consider a sensor network sampling $f(x, t)$ with $N_s < M$ sensors deployed uniformly in space, and each sensor collects N_t samples uniformly in time. An example of a field generated by point sources and the corresponding sampling grid is given in Fig. 1. The reconstruction scenario we consider is an *inverse initial problem*, where we aim at estimating the initial field distribution using the collected samples. In general, inverse problems of the diffusion equation are ill-conditioned [1], but the sparsity constraint on the sources helps to regularize the solution.

The spatial undersampling of the heat equation introduces another issue to the general problem of spatiotemporal sampling: if we sample the field too early (see Fig. 2(a), for $t < \tau$), some sensors may not sense the presence of the sources; conversely, if we collect samples too late (see Fig. 2(b)), the information content carried by the field is blurred by the diffusion effect.

An efficient solution to these complex issues will have an impact in different real world scenarios, such as pollution detection [2], plume source detection [3], and localization of the so-called cold and hot spots in server rooms, responsible for energy inefficiency [4]. In this paper, we tackle these challenges taking advantage of the knowledge of the physical model underlying the sampled field and the sparsity of the sources, to improve both sampling and reconstruction. In particular, we study the spatiotemporal sampling problem analyzing the meaningful properties of a matrix Ψ , which represents the sampling operator acting on the diffusive field. This operator is given as a function of one spatiotemporal parameter ρ , that represents the ex-

*This work has been partially supported by MIUR under the internalization project ACASTI and by the EDIC-EPFL fellowship.

†This work has been supported by ERC Advanced Grant - Support for Frontier Research - SPARSAM Nr : 247006.

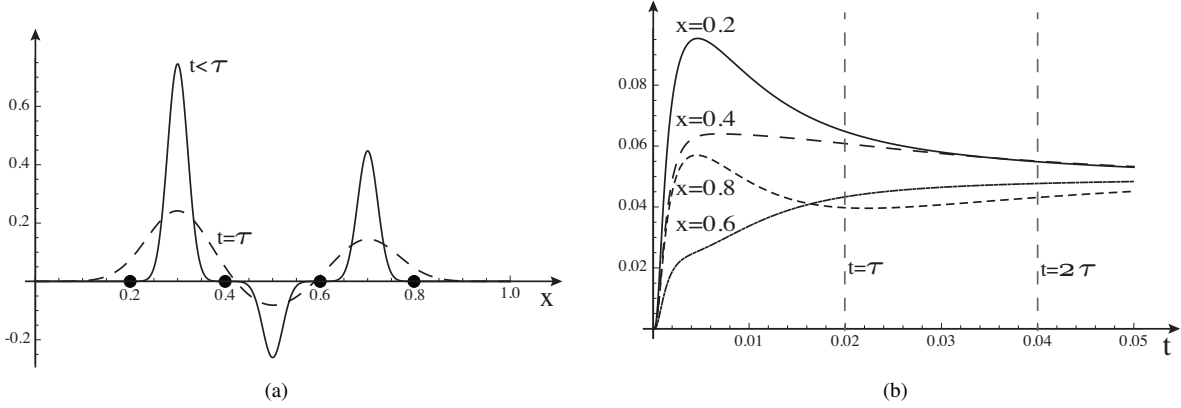


Fig. 2. The field $f(x, t)$ shown in Fig. 1 at two time instants (a), and at the four sensor positions (b). τ is the first sampling instant.

isting trade-off between the spatial and temporal density. We derive the optimal spatiotemporal sampling parameter which gives a low condition number for Ψ and an optimal sampling of all the sources in the field. We theoretically prove the optimality and give evidence of it via numerical simulations that assess the reconstruction performance. We show that it is possible to reduce the number of sensors by increasing linearly the number of samples collected in time, provided that the sensors network is dimensioned properly.

One of the earliest studies of the inverse initial diffusive problem was conducted by Fourier and Kelvin [5], who estimated the initial temperature of the earth from the current temperature distribution. More recently, many approaches have been considered, such as the hyperbolic heat equation [6], transform techniques [7], spatial super-resolution [8] and an adaptive spatiotemporal sampling scheme [9]. A similar approach to ours, applied to the wave equation, has been proposed by Fannjiang et al. [10] where the localization of remote sparse objects using electromagnetic pulses is solved by compressed sensing (CS) techniques.

The remainder of the paper is organized as follows. In Section 2, the matrix representing the sampling operator is defined and closed-form estimations its characteristics are given as a function of the spatiotemporal parameters. In Section 3, an evaluation of the reconstruction performance is conducted using an algorithm derived from CS. We conclude in Section 4.

2. SAMPLING OF THE DIFFUSION FIELD

In this section, we give a mathematical characterization of the sampling process of a diffusive field and develop it as follows: we describe the diffusive field as a linear system and construct a *sampling matrix* Ψ representing the sampling operator. The key point here is that Ψ allows us to study easily the properties of the sampler. In particular, we reach this goal by evaluating the norm of the columns, the rank and the condition number of Ψ as a function of a single spatiotemporal parameter defined by the sampling rate in space and time. From now on, \mathbf{f} denotes a vector and $\mathbf{f}(n)$ its n th element. If Ψ is a matrix, then ψ_m is its m th column and $\psi_m(n)$ is the n th element of that column. Unless stated otherwise, the norm under consideration is the ℓ_2 norm.

Two related problems of the diffusion equation are of interest in this work: the *forward problem* and the *inverse initial problem*. The former aims to find the field $f(x, t)$ for every x and t when the boundary conditions and the initial field distribution are known. We consider the solution of the forward problem to define the latter, that aims at finding the initial field $f_0(x)$ given the boundary conditions

and a set of samples from the field $f(x, t)$ at given positions and time instants. The forward problem of the heat equation can be solved as a linear system. We consider a circular domain of length $D = 1$ and $\gamma = 1$ for simplicity of notation. Note that even though the circular domain requirement seems fairly strong, it is a good approximation whenever we do not have sources appearing close to the boundaries. We determine the impulse response, called *Green's function*, as

$$G^D(x, t) = \frac{1}{2\sqrt{\pi t}} \sum_{k \in \mathbb{Z}} \exp\left\{-\frac{(x + kD)^2}{4t}\right\}, \quad (2)$$

and obtain the field $f(x, t)$ via the spatial convolution between $f_0(x)$ and $G(x, t)$. That is

$$f(x, t) = \int_0^D G^D(x - s, t) f_0(s) ds. \quad (3)$$

Note that the spatial and temporal dimensions are strictly connected in the diffusion equation. Indeed, scaling in the spatial domain can be compensated by scaling in the time domain. We see this effect in the Green's function (2), since

$$G^{\alpha D}(\alpha x, t) = \frac{1}{\alpha} G^D\left(x, \frac{t}{\alpha^2}\right). \quad (4)$$

Hereinafter, we use (2) and (3) to construct the sampling matrix.

To define the discrete inverse problem, we constrain the location x_m of the K sources to be on a uniform grid of size M , where $K \ll M$. We define two sets $\mathcal{X}_M = \{m\Delta_1 : 0 \leq m \leq M - 1\}$ and $\mathcal{Y}_{N_s} = \{n\Delta_2 : 0 \leq n \leq N_s - 1\}$, where $\Delta_1 = \frac{1}{M}$ and $\Delta_2 = \frac{1}{N_s} = \Delta_1 \frac{M}{N_s}$, to represent the spatial position of the sensors and the sources, respectively. We also define a set $\mathcal{T}_{N_t} = \{\ell\tau : 1 \leq \ell \leq N_t\}$ to define the temporal sampling grid.

Let \mathbf{f}_0 be a vector defining the amplitude a_m of the source at the m th position of the grid \mathcal{X}_M , that is $\mathbf{f}_0(m) = a_m$, and let \mathbf{f} be the vector containing all $N = N_s N_t$ samples collected by all the sensors. We want to construct a $N \times M$ matrix Φ representing the sampling operator as a discrete linear operator, $\mathbf{f} = \Phi \mathbf{f}_0$. Each row of Φ represents the sampling kernel of a sensor in a given position and at a given time instant. We organize the matrix in N_t stacked submatrices Φ^ℓ of size $N_s \times M$, where each Φ^ℓ models the sampling of the sensor network for a given time instant $\ell\tau \in \mathcal{T}_{N_t}$. We define the elements of the ℓ th submatrix $\phi_m^\ell(n)$ for $\ell = 1, \dots, N_t$ as

$$\phi_m^\ell(n) = \frac{\Delta_1}{\sqrt{4\pi\ell\tau}} \sum_{k \in \mathbb{Z}} \exp\left\{-\frac{(n\Delta_2 - k\Delta_2 N_s - m\Delta_1)^2}{4\ell\tau}\right\}. \quad (5)$$

Let $\mathbf{f}^\ell(n)$ be the sample collected by the n th sensor at the ℓ th sam-

pling instant, then we write

$$\mathbf{f}^\ell(n) = \sum_{m=0}^{M-1} \phi_m^\ell(n) \mathbf{f}_0(m), \quad (6)$$

that can be rewritten in matrix notation as $\mathbf{f}^\ell = \Phi^\ell \mathbf{f}_0$, where the vector \mathbf{f}^ℓ contains all the samples collected at $l\tau$. One can verify that the discretization of (3), considering $x = m\Delta_1 - n\Delta_2$, is equivalent to (6). We concatenate the N_t measurement vectors \mathbf{f}^ℓ and sampling operators Φ^ℓ into a single vector \mathbf{f} and single matrix Φ , respectively, to obtain $\mathbf{f} = \Phi \mathbf{f}_0$.

Recalling the spatiotemporal scaling in (4), we introduce, without loss of generality, a parameter ρ to have invariance to such scaling and to simplify the analysis of the problem. We choose

$$\rho = \frac{\tau}{\Delta_2^2}, \quad (7)$$

where Δ_2 and τ represent the spatial and temporal sampling densities, respectively. Note that since the sampling is uniform in time, the parameter τ provides us with two relevant informations: it represents the temporal sampling density and the delay between the appearance of the sources ($t = 0$), and the first sampling instant. To obtain an expression of the matrix solely as a function of ρ and to simplify the analysis of the properties of Φ , we define a new set of $N_s \times M$ matrices Ψ^ℓ , whose elements are defined as

$$\psi_m^\ell(n) = \frac{1}{\sqrt{N_t} \sqrt{2\pi\ell\rho}} \sum_{k \in \mathbb{Z}} \exp \left\{ -\frac{(n - kN_s - m \frac{N_s}{M})^2}{4\ell\rho} \right\}. \quad (8)$$

The elements of Ψ^ℓ are equal to those of Φ^ℓ up to a scaling factor unique for every submatrix Ψ^ℓ . Thus, we obtain the linear system $\mathbf{y} = \Psi \mathbf{f}_0$, where \mathbf{y} is equivalent to \mathbf{f} up to a proportional factor. In the following propositions, we give three results regarding the properties of Ψ studied as a function of ρ . All the corresponding proofs are given in [11]. We first consider $\|\psi\|^2$, since it represents the energetic relation between sources and sensors.

Proposition 1 *Let Ψ be a $N_s N_t \times M$ matrix representing the sampling process of a sensor network in a diffusion field with N_s sensors each collecting N_t samples. If ρ is bounded according to*

$$\frac{1}{8C} \leq \rho \leq \frac{(N_s - 1)^2}{N_t 72} = \rho_{MAX}, \quad (9)$$

for $0 < C < 1$, then $\|\psi_0\|^2 \approx 1$ and the maximum difference between the squared norms of any two columns is bounded according to

$$\max_{k,j} (|\|\psi_k\|^2 - \|\psi_j\|^2|) \leq N_t C \|\psi_0\|^2, \quad \forall k, j.$$

If ρ does not satisfy the lower bound in (9), the Green's function tends to be a vector with only one element significantly different from zero, and only the sources which are really close to the sensors are correctly sampled. Instead, if $\rho > \rho_{MAX}$, then the Gaussian kernel tends to be almost flat, and the sensed data contains almost no information about the position and the amplitude of the sources. We are interested in the differences between the norm of the columns, because if $\|\psi_j\|^2$ is smaller than $\|\psi_k\|^2$, the energy collected by the sensors from the j th source is smaller than the one collected by the k th one, leading to a sub-optimal estimation of the j th source in the initial field distribution. The result above guarantees that, since $\|\psi_0\|^2$ is bounded and C can be arbitrarily small, the norms of any two columns can be arbitrarily close. We extend this analysis to be more practical and establish a minimum ρ to sense properly all sources, including the ones far from the sensors.

Proposition 2 *Let us consider a sensor network composed of N_s sensors, each collecting N_t samples uniformly in time. To sense all the possible sources in the field, the sampling densities must satisfy*

$$\rho > \frac{1}{2N_t} = \rho_{MIN}. \quad (10)$$

In practice, Proposition 1 and 2 are meant to prevent the two sampling issues depicted in Fig. 2. Moreover, we wish to have a full-rank Ψ , otherwise there would be a set of initial field distributions \mathbf{f}_0 that are in the null space $\mathcal{K}(\Psi)$ and generate null measurement vectors \mathbf{f} . Thus, it would not be possible to distinguish the elements of $\mathcal{K}(\Psi)$ from the measurement vector. If ρ satisfies (9), we can prove that Ψ is full rank, but the smallest singular value is exponentially smaller for increasing values of ρ . This raises the question of the conditioning, that we improve exploiting the sparsity of the sources. Namely, if we use an algorithm that searches for a sparse solution, the optimization is concentrated only on the K columns of Ψ activated by the sources and the solution is only minimally affected by the ill-conditioning. Therefore, we introduce a method to compute the conditioning of a subset of columns of Ψ as a function of ρ and the K , which is a relevant merit figure for sparse reconstruction.

Proposition 3 *Let us consider a $N_s N_t \times M$ sampling matrix Ψ , where $N_t \geq 1$ and $N \leq M$. Let S be a set of K indices $s_k \in \{0, \dots, M - 1\}$ representing the position of the sources, let Ψ_S be the matrix composed by the columns of Ψ pointed by S and let $\text{Gr}(\Psi_S) = \Psi_S^T \Psi_S$ be the Gramian matrix of Ψ_S . If the parameter ρ satisfies (9), then the elements of $\text{Gr}(\Psi_S)$ are defined as*

$$\text{Gr}(\Psi_S)_{k,j} \approx \frac{1}{N_t} \sum_{\ell=0}^{N_t-1} \exp \left\{ -\frac{d(s_k, s_j)^2}{8\ell\rho} \left(\frac{N_s}{M} \right)^2 \right\}, \quad (11)$$

where $d(s_k, s_j)$ is the circular distance between s_k and s_j .

The approximation in (11) is given by the approximation $\|\psi_j\|^2 \approx \|\psi_k\|^2$, consequence of Proposition 1. The eigenvalues of (11), and thus the singular values of Ψ_S can be analytically expressed for low values of K . Hereinafter, we give the result for $K = 3$.

Example 1 *For $K = 3$ and $S = \{1, 2, 3\}$, Ψ_S is a $N_s \times 3$ matrix and its condition number is given by*

$$\text{cond}(\Psi_S) = \sqrt{\frac{1 + 2e^{\frac{1}{2\ell\rho}} \left(\frac{N_s}{M} \right)^2 + \sqrt{1 + 8e^{\frac{3}{4\ell\rho}} \left(\frac{N_s}{M} \right)^2}}{1 + 2e^{\frac{1}{2\ell\rho}} \left(\frac{N_s}{M} \right)^2 - \sqrt{1 + 8e^{\frac{3}{4\ell\rho}} \left(\frac{N_s}{M} \right)^2}}}. \quad (12)$$

Example 1 has a straightforward interpretation: whenever ρ satisfies (9), the condition number of a subset of columns is much smaller than that of Ψ . Moreover, we obtain a better conditioning of Ψ_S for a smaller ρ . The theoretical estimation derived in (12) is confirmed with our numerical simulation, whose results are given in Fig. 3. In this experiment we measure the condition number of Ψ_S when $S = \{1, 2, 3\}$ for different values of ρ .

As a summary, and concentrating on the effect of ρ on Ψ , we observe that ρ must satisfy (9) and (10) to have each source sampled with the same efficiency by the sensor network. Moreover, Proposition 3 together with Example 1 put in evidence that as long as the aforementioned bounds are respected, a lower ρ corresponds to a better condition number of the matrix Ψ_S . Therefore, the reconstruction of \mathbf{f}_0 is optimal whenever ρ tends to ρ_{MIN} . This is confirmed by the numerical experiments presented in the next section.

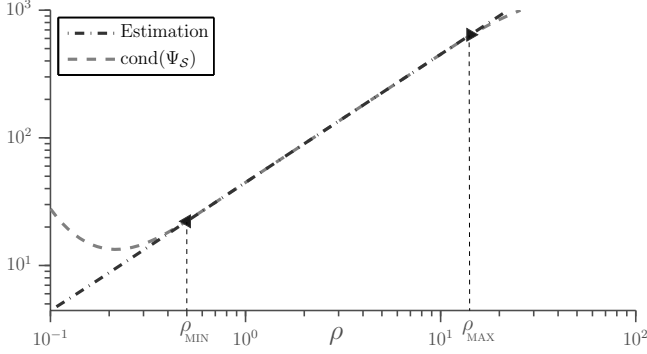


Fig. 3. Estimated and condition number of Ψ_S for $S = \{1, 2, 3\}$, that represents the worst case scenario (consecutive columns).

3. RECONSTRUCTION OF THE INITIAL FIELD

To verify our analysis of the sampling operator, we test the reconstruction performance of a CS algorithm applied to the inversion of the system (6) with an unknown sparse vector \mathbf{f}_0 and a sampling matrix Ψ . In particular, we consider Smoothed ℓ_0 (SL0) [12], that attempts to minimize directly the ℓ_0 norm of a vector subject to linear constraints, using a smooth approximation within an iterative process. We analyze the reconstruction performance fixing $M = 64$ and $K = 3$ and by varying the number of samples collected in time N_t and ρ . The merit figure is the minimum N_s necessary to reach 99% of exact reconstruction. A reconstruction is exact when the SNR of the reconstructed vector is over 30dB. The results are an average of the results obtained over 5000 sparse vectors and are depicted in Fig. 4, where we note that the parameter ρ has a strong influence on the reconstruction. Indeed, when $\rho < \rho_{\text{MIN}}$, the algorithm requires $N_s N_t > M$, meaning that undersampling is not possible, confirming Proposition 2. Namely, each sample has the information of only one source, thus we need $M = N_s$ sensors. Note that this bound is tight only for $N_t = 1$; for $N_t > 1$, ρ_{MIN} seems too conservative as it is not anymore the smallest ρ at which we obtain the best reconstruction performance. Further investigation of this matter is left for future work. On the other hand, when ρ increases, the condition number increases exponentially, and the algorithm cannot recover a valid solution, no matter how many measurements, spatial or temporal, are collected. This is an effect of the diffusion kernel, that in such conditions is too wide and each source cannot be distinguished from the other sources. Note that there is a very well-balanced trade-off between N_s and N_t . Namely, if we increase linearly the number of collected time samples N_t , we can reduce the number of sensors provided that we collect an equal amount of total samples N . As shown in Fig. 4, if we pick $\rho = \rho_{\text{MIN}}$, the possible combinations for (N_s, N_t) are (40, 1), (19, 2), (13, 3) and (10, 4).

4. CONCLUSIONS

We proposed an innovative approach to model the spatiotemporal sampling of a diffusive field generated by sparse sources. This approach allowed us to study the main properties of the sampling process through a matrix Ψ . We obtained a theoretical optimal sampling of the field by improving the conditioning of the inverse problem without any loss of information. Our experiments align with the theory and show that exact reconstruction is feasible with spatially undersampled data. Moreover, these results show we can reduce the spatial sampling density by linearly increasing the temporal one, allowing for a lower cost of the sensor network. In the future, we will

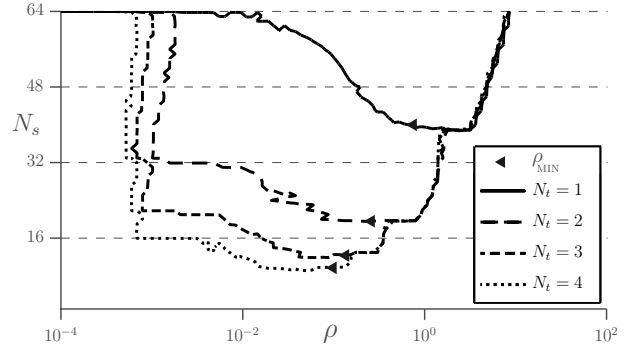


Fig. 4. Minimum number of sensors N_s for an exact reconstruction of the initial distribution as a function of N_t and ρ . If the point (ρ, N_s) is above the threshold, the probability of an exact reconstruction of the initial field is over 99%.

use our approach to study the theoretical aspects and bounds of two generalized sampling scenarios: first, where point sources appear at an unknown instant and second, where sources are not localized anymore, but sparse with regards to an orthonormal basis. In both cases, we have promising preliminary numerical results.

5. REFERENCES

- [1] C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*, Soc. for Industrial Math., 1987.
- [2] A. El Badia and T. Ha-Duong, “An inverse problem in heat equation and application to pollution problem,” *Inverse and ill-Posed Problems*, vol. 10, pp. 585 – 599, 2002.
- [3] D. M. Moreira, T. Tirabassi, and J. C. Carvalho, “Plume dispersion simulation in low wind conditions in stable and convective boundary layers,” *Atmospheric Environment*, vol. 39, no. 20, pp. 3643–3650, June 2005.
- [4] K. Chen, D. M. Ausl, C. E. Bash, and R. D. Patel, “Local temperature control in data center cooling,” Tech. Rep., HP Labs, 2006.
- [5] H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids*, Oxford University Press, USA, 2 edition, Apr. 1986.
- [6] K. Masood and F. D. Zaman, “Investigation of the initial inverse problem in the heat equation,” *J. of Heat Transfer*, vol. 126, no. 2, pp. 294–296, Apr. 2004.
- [7] G. Nakamura, S. Saitoh, and A. Syarif, “Representations of initial heat distributions by means of their heat distributions as functions of time,” *Inv. Prob.*, vol. 15, pp. 1255–1261, 1999.
- [8] Y. M. Lu and M. Vetterli, “Spatial super-resolution of a diffusion field by temporal oversampling in sensor networks,” in *IEEE Int. Conf. on Acoustics, Speech and Signal Proc.*, Taiwan, 2009, pp. 2249–2252.
- [9] Y. M. Lu and M. Vetterli, “Distributed Spatio-Temporal sampling of diffusion fields from sparse instantaneous sources,” in *Proc. 3rd Int. Workshop on Comp Adv. in Multi-Sensor Adaptive Proc.*, 2009.
- [10] A. Fannjiang, P. Yan, and T. Strohmer, “Compressed remote sensing of sparse objects,” *SIAM J. Imag. Sci.*, Accepted.
- [11] J. Ranieri, A. Chebira, Y. M. Lu, and M. Vetterli, “Sampling and reconstruction of a diffusive field: Exploiting the physical model and sparsity,” Tech. Rep., LCAV-EPFL, 2010.
- [12] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, “A fast approach for overcomplete sparse decomposition based on smoothed ℓ_0 norm,” *IEEE Trans. on Signal Proc.*, vol. 57, no. 1, pp. 289–301, 2009.