Bandlimited Spatiotemporal Field Sampling with Location and Time Unaware Mobile Senors

B. Tech. Project Report
Spring 2018

Submitted in the partial fulfilment of
the requirements for the degree of

Bachelor of Technology

By

Sudeep Salgia
14D070011
Department of Electrical Engineering

Under the guidance of
Prof. Animesh Kumar
Department of Electrical Engineering

Acknowledgments

I would like to thank my guide, Professor Animesh Kumar for his constant guidance and support. I am really grateful to him for introducing this problem to me, exposing me to current literature and his valuable inputs throughout the project. His able guidance helped me understand the problem and go beyond what is available to devise a solution with understanding and intuition.

I am indebted to him for his support at a time when he was very busy with multiple students, his own research and teaching responsibilities.

Abstract

Sampling of smooth spatiotemporally varying fields has been a topic that been studied extensively in literature. Classical approaches to this problem assume the precise knowledge of the sampling locations and the time instants of sampling enabling interpolation or reconstruction. Moreover, classical approaches involve sampling using sensor networks. However, with the advent of mobile sampling, it is important to develop techniques that look into the practical considerations of the reconstruction and incorporate it into the estimation methods. This report considers two such aspects, economic feasibility and vehicle modeling.

In the first part of the work, the sampling and reconstruction of a spatiotemporal bandlimited field is addressed, where the samples have been obtained by a *location and time unaware* mobile sensor. The order of the samples, however, is assumed to be known. The obtained samples have been assumed to be corrupted by measurement noise. Furthermore, the evolution of the field is modelled by a known linear constant coefficient partial differential equation. In this challenging setup, a regression style estimate has been developed for the reconstruction of the spatial field. The intersample distances and the intersample timestamp differences are assumed to be coming from independent, *unknown* renewal processes.

In the second part, the issue of independent intersample distances is addressed. The inertia of the moving vehicle constrains the independence between the intersample distances, making them correlated. This part, thus, aims at estimating spatially bandlimited fields from samples, corrupted with measurement noise, collected on sampling locations obtained from an autoregressive model on the intersample distances. The autoregressive model is used to capture the correlation between the intersample distances.

In both the cases, if n is the average number of samples of the field obtained by the sensor, then it is shown that the mean squared error decreases as $O(1/n)$.

Contents

1 Introduction

Sampling of smooth spatiotemporally varying fields is a well studied topic in the literature [1, 2]. Classical approach assumes that the field is observed at known sampling locations and known timestamps ensuring interpolation or reconstruction [1, 2]. The field is assumed to evolve according a physical law, typically governed by a suitable partial differential equation (such as the diffusion equation) [3].

Recently, mobile sensing has been proposed for sampling [4, 5, 6]. Mobile sampling comes along with its associated costs. Often that devices that are needed to be employed to have a precise knowledge of the sampling locations and the sampling instants, are expensive. As a result, in practical implementations of such a sampling technique, economic feasibility is a major concern. For cost reduction, it is desirable to work with a location unaware and preferably even time unaware, inexpensive mobile sensor for the purpose of sensing and *still* be able to reconstruct the spatiotemporal field upto a desired accuracy. Furthermore, it would be an added benefit if this could be done for *any* physical spatiotemporally varying field. This is the core motivation behind the work.

The problem is addressed in two different aspects, namely, economic feasibility and modeling of the vehicle. In the first part, the spatiotemporal variation is studied to address the economic feasibility while in second part the vehicular movement is modeled separately. To the best knowledge of the authors, there has been no work in sampling fields governed by *any* linear partial differential equation (PDE) with constant coefficients. It is to be noted here that the class of partial differential equations referred here is the one that includes only those equations that are suitable to model a physical field. Thus, whenever this work mentions a "general PDE", it is inherently assumed that the equation belongs to such a class. Also, autoregressive modeling of the intersample distances is not studied in the literature. Thus, both the aspects are relatively new.

The analysis of both the problems is done separately. However, the problem setup is very similar and in the initial part of the report no distinction is made because the topics discussed are common to both the scenarios. The only difference is that in the case of autoregressive modeling, the field is considered to be temporally fixed to simplify calculations. Thus, the reader is requested to understand the spatiotemporal field in relevant context for each of the problems.

This work assumes that such an inexpensive mobile sensor is available for sampling a spatial field of interest in a finite region. The mobile sensor is location and time unaware, which presents a challenging field reconstruction problem. Is it possible to sample and reconstruct a spatiotemporal field governed by a partial differential equation, where the samples are obtained by a location-unaware, time-unaware mobile sensor. This works answers this question in affirmative under some conditions. The results hold even when additive measurement noise is present in the samples.

Also, it can be shown that if such location unaware noisy samples are obtained locations from an autoregressive process, the field still can still be reconstructed with a vanishing error.

The field is assumed to be bandlimited. The evolution of the spatiotemporal field is characterized by a linear partial differential equation with constant coefficients, assumed to be known. Bandlimitedness implies a finite number of Fourier coefficients characterize the spatial field. The key aspects utilized in solution include oversampling, denoising, and utilization of spatial/temporal order of samples. The latter is assumed to be known. Our solution works by developing a regression style estimate for the reconstruction of Fourier coefficients of the spatiotemporal field at time $t = 0$. The work later deals with stability issues of the obtained results and proposes standard regularization methods for the same. With PDE known, the coefficients at all times can be predicted from the initial Fourier series coefficients. The sampling methods for the two cases, however, are slightly different. In the first case, the intersample spacings and the intersample time stamp differences are assumed to be from independent *unknown* renewal processes. However, in the second case, the intersample spacings are assumed to coming from an autoregressive process such that stochastic term of the same is coming from an unknown renewal process.

If n is the average number of samples of the field obtained by the mobile sensor, then our main result shows that the mean-squared error in the spatiotemporal field estimation decreases as $O(1/n)$.

For mathematical tractability, the spatiotemporal field is assumed to be one dimensional in space evolving according to a known partial differential equation. The smoothness of the field is modeled by is spatial bandlimitedness. The measurement noise is independent of the intersample spacing and intersample timestamp processes. It is also assumed to have zero mean and a finite variance. Oversampling is the key technique used to denoise and mitigate location-unawareness in this work.

1.1 Prior art

1.1.1 Reconstruction of Spatiotemporal Fields

The literature consists of two parts: (i) papers which approach source reconstruction; and (ii) papers which address this problem as a sampling/reconstruction problem. The problem of estimating sources in a diffusive field is a severely ill-conditioned problem and hence certain regularization methods need to be employed. Spatial sparsity of sources has been used to reconstruct sources[7] using maximum likelihood estimators. Lu and Vetterli have used sparsity aware super resolution techniques to obtain sources[8] and other on an adaptive spatiotemporal sampling scheme [9]. A method of localizing sources has been proposed by Ranieri et al.[3] which has been extended to real line [10]. Sampling with a mobile sensor has been a topic of interest recently [4, 5, 6]. However, mobile sensing typically assumes sampling locations to be known. When spatial field is fixed with time, sampling with a location-unaware mobile sensor has also been addressed recently [11]. This work is novel since both the sampling locations and the timestamps of the samples are unknown. It is also assumed that the field is evolving while measurements are being made by the mobile sensor.

1.1.2 Samples from an Autoregressive process

Sampling and reconstruction of discrete-time bandlimited signals from samples taken at unknown locations was first studied by Marziliano and Vetterli [12] who had addressed the problem in a discrete-time setup. Browning[13] later proposed an algorithm to recover bandlimited signals from a finite number of ordered nonuniform samples at unknown sampling locations. Nordio et al.[14] studied the estimation of periodic bandlimited signals, where sampling at unknown locations is modeled by a random perturbation of equi-spaced deterministic grid. More generally, the topic of sampling with jitter on the sampling locations [15], [16], is well known in the literature.Mallick and Kumar [17] worked on reconstruction of bandlimited fields from location-unaware sensors restricted on a discrete grid. A more generic case of sampling from unknown locations coming from a known underlying distribution was introduced recently[18]. Further, the work [11], deals with estimation of field from unknown sampling locations coming from an unknown renewal process. This work is different from all others in the sense that the sampling model incorporates the correlation between the intersample distances and thus addresses a more practical scenario. The intersample distances are *unknown* have been considered to be coming from an autoregressive model of order 1 whose stochastic part is an *unknown* renewal process.

1.2 Notation

The spatiotemporally varying field will be denoted as $g(x,t)$ and its variants. Average sampling density will be denoted by n. All vectors will be denoted in bold. The $\mathcal{L}^2$ norm of a vector $\mathbf{x}$ is denoted by $||\mathbf{x}||_2$. The expectation operator will be denoted by E[.]. The expectation is over all the random variables within the arguments. The trace of a matrix A will be denoted by $\text{tr}(A)$.

For a complex number z, $\Re(z)$ denotes the real part of z. The set of natural numbers, integers, reals and complex numbers will be denoted by $\mathbb{N}, \mathbb{Z}, \mathbb{R}$ and $\mathbb{C}$ respectively. Also, $j = \sqrt{-1}$.

1.3 Organization

The report is divided into two sections for each of the problems. Each of them has the following subsections. Subsection 2.1 describes the field model, the distortion criteria, the sampling model and the noise model. The estimation of the field is presented in Subsection 2.2. The simulations have been shown in Subsection 2.3and. Section 4 concludes the report.

2 Time varying fields with location and time unaware sensors

2.1 Field, Sampling, Noise Model and Distrortion Criteria

2.1.1 Field Model

The field is considered to be spatially smooth over a finite support, one dimensional in space and evolving with time according to a Partial Differential Equation. The PDE is linear with constant coefficients and is assumed to be known and given by

$$\sum_{i=0}^{m} p_i \frac{\partial^i}{\partial t^i} g(x,t) = \sum_{i=0}^{m'} q_i \frac{\partial^i}{\partial x^i} g(x,t) \tag{1}$$

where, $\frac{\partial^0}{\partial y^0} f(y) = f(y)$. This will be represented throughout the report in the terms of polynomials. Define two polynomials as

$$p(z) = \sum_{i=0}^{m} p_i z^i \ ; \ q(z) = \sum_{i=0}^{m'} q_i z^i \tag{2}$$

where the coefficients are same as in the differential equation. If for notational convenience, $\left(\frac{\partial}{\partial z}\right)^l = \left(\frac{\partial^l}{\partial z^l}\right)$, then the original partial differential equation can be written as

$$p\left(\frac{\partial}{\partial t}\right) g(x,t) = q\left(\frac{\partial}{\partial x}\right) g(x,t) \tag{3}$$

To incorporate the smoothness of the field, the field is assumed to be bandlimited. It is important to ensure that the field is bandlimited as it evolves with time. Intuitively speaking, this condition should hold true if we have know the field is spatially bandlimited at $t = 0$, because time evolution is unlikely to affect the spatial bandlimitedness. Formally speaking, if the degree of the polynomial p is m, then if $m-1$ partial derivatives of $g(x,t)$ along with $g(x,t)$ are spatially bandlimited then the function $g(x,t)$ will be spatially bandlimited for all $t \geq 0$ if the field evolves according to the given PDE. A detailed proof of this has been given in Appendix A. The field under consideration is assumed to be bandlimited for all $t \geq 0$. Without loss of generality, the support of the field is considered to be $[0,1]$. It can thus be represented as,

$$g(x,t) = \sum_{k=-b}^{b} a_k(t) \exp(j2\pi kx) \ ; \ a_k(t) = \int_0^1 g(x,t) \exp(-j2\pi kx) dx \tag{4}$$

WLOG field is assumed to be bounded. That is, $|g(x,t)| \leq 1 \ \forall (x,t)$.

2.1.2 Distortion Criteria

The distortion will be measured by the mean-squared error between the field and its estimate at $t = 0$. Let the estimated field be $\hat{G}(x,t)$ and its Fourier coefficients be $\hat{A}[k]$, then the distortion criterion is defined as

$$\mathcal{D}\left[\hat{G}, g\right] = \mathbb{E}\left[\int_0^1 |\hat{G}(x,t) - g(x,t)|^2\right]\Bigg|_{t=0} \tag{5}$$

$$= \mathbb{E}\left[\sum_{k=-b}^{b} |\hat{A}_k(t) - a_k(t)|^2\right]\Bigg|_{t=0}$$

$$= \sum_{k=-b}^{b} \mathbb{E}\left[|\hat{A}_k(0) - a_k(0)|^2\right] \tag{6}$$

2.1.3 Sampling Model

It is assumed that there is a mobile sensor which moves from $x = 0$ to $x = 1$ while recording samples of the field $g(x, t)$. As the sensor moves, the field also evolves. This is the key distinction from models used in the literature. Spatial samples are collected on points generated by an unknown renewal process [11]. Let $X_1, X_2, \ldots$ be the intersample distances and $N_1, N_2, \ldots$ be the corresponding intersample time intervals. These variables are assumed to be the realizations of two independent, possibly different, and unknown renewal processes. Let the intersample distributions for X and N be $f(x)$ and $g(x)$, respectively. The sampling locations S_n are given by $S_n = \sum_{i=1}^{n} X_i$. The sampling is done over $[0, 1]$, and M is the random number of samples that lie in this interval. Note that the stopping condition $S_M \leq 1$ and $S_{M+1} > 1$ results in M ; and, it means M is a well defined random variable[19]. Note that, the number of samples in the interval only depend on the spatial sampling density and not on the temporal counterpart.

For the purpose of ease of analysis and tractability, the support of the distributions of X and N are considered to be finite and inversely proportional to the sampling density. Hence, it is assumed that

$$0 < X \leq \frac{\lambda}{n}, 0 < N \leq \frac{\mu}{n} \text{ and } \mathbb{E}[X] = \mathbb{E}[N] = \frac{1}{n}, \tag{7}$$

where $1 < \lambda, \mu \ll n$ are parameters that characterize the support of the distributions. Both are finite numbers, independent of the average sampling density n. These would be important factors that govern the constant of proportionality in the expected error of the estimate. Furthermore, λ is an important factor also to determine the threshold on the minimum number of samples. Applying Wald's identity[19], on S_{M+1}, and using equation (7),

$$\mathbb{E}[M+1]\mathbb{E}[X] = \mathbb{E}[S_{M+1}] \tag{8}$$

$$(\mathbb{E}[M] + 1)\frac{1}{n} = \mathbb{E}[S_{M+1}]$$

$$\mathbb{E}[M] = n\mathbb{E}[S_{M+1}] - 1 \tag{9}$$

By definition, $S_{M+1} > 1$ and $S_M \leq 1$. Since $S_{M+1} = S_M + X_{M+1}$, therefore, $S_{M+1} \leq 1 + X_{M+1} \leq 1 + \frac{\lambda}{n}$. Use these inequalities with equation (8), to obtain,

$$n - 1 < \mathbb{E}[M] \leq n + \lambda - 1 \tag{10}$$

Also, the bound on each $X_i \leq \frac{\lambda}{n}$, along with $S_{M+1} > 1$ gives,

$$(M+1)\frac{\lambda}{n} > 1 \text{ or } M > \frac{n}{\lambda} - 1 \tag{11}$$

The results in (10) and (11), imply that expected values of M scales linearly with n thus justifying the name average sampling rate for n. The timestamps are assumed to be $T_n = \sum_{i=1}^{n} N_i$. The only assumption on the sampling model with respect to time is that it has been assumed that all samples are collected within some time T_0, i.e., $T_M \leq T_0$ and $T_{M+1} > T_0$ and it is assumed to be known. That is the time stamp at the beginning $t = 0$ and at the end of the mobile sensing experiment must be known. The value of T_0 may vary. Since $nN \leq \mu \ll n$, we expect that $T_0 \ll M$.

2.1.4 Measurement Noise Model

It will be assumed that the obtained samples have been corrupted by additive noise that is independent both of the samples and of both the renewal processes. For simplicity, the noise is considered to be varying only spatially. That is, at all time instants, the distribution of the noise remains the same, which is assumed to be *unknown* in this work. Hence, $W(x, t) \equiv W(x)$. Thus, the samples obtained would be sampled versions of $g(x, t) + W(x, t)$, where $W(x, t) \equiv W(x)$ is the noise. Also, since the measurement noise is independent, that is for any set of measurements at distinct points $s_1, s_2, s_3, \ldots s_n$, the samples $W(s_1), W(s_2), W(s_3), \ldots W(s_n)$ would be independent and identically distributed random variables. Note that the sampling instants have not been considered because of the distribution being temporally static. The only statistics known about the noise are that the noise is zero mean and has a finite variance, σ^2.

2.2 Field Estimation from Samples

This section will highlight the estimation of the field from samples whose locations and timestamps come from two *unknown* independent renewal processes.

2.2.1 Field under the PDE

A brief discussion of the field under its PDE is presented first. Using the fact that Fourier series are linear in coefficients, and combining the equations (4) and (1), and using the orthogonality of Fourier basis, we can write,

$$
\begin{aligned}
\sum_{i=0}^{m} p_i \frac{\partial^i}{\partial t^i}\left(\sum_{k=-b}^{b} a_k(t)\exp(j2\pi kx)\right) &= \sum_{i=0}^{n} q_i \frac{\partial^i}{\partial x^i}\left(\sum_{k=-b}^{b} a_k(t)\exp(j2\pi kx)\right) \\
\sum_{k=-b}^{b}\left(\sum_{i=0}^{m} p_i \frac{\partial^i a_k(t)}{\partial t^i}\right)\exp(j2\pi kx) &= \sum_{k=-b}^{b} a_k(t)\left(\sum_{i=0}^{n} q_i (j2\pi k)^i\right)\exp(j2\pi kx) \\
\sum_{k=-b}^{b}\left(\sum_{i=0}^{m} p_i \frac{\partial^i a_k(t)}{\partial t^i}\right)\exp(j2\pi kx) &= \sum_{k=-b}^{b} a_k(t) q(j2\pi k)\exp(j2\pi kx) \\
\sum_{i=0}^{m} p_i \frac{\partial^i a_k(t)}{\partial t^i} - q(j2\pi k) a_k(t) = 0 \;\; \forall k &= -b,\ldots,b
\end{aligned} \tag{12}
$$

Essentially, each $a_k(t)$ evolves by an ordinary differential equation (ODE) with constant coefficients. The solution of ODE with constant coefficients is well known (via unilateral Laplace transform [20]). For each k, this leads to the polynomial equation,

$$
\begin{aligned}
\sum_{i=0}^{m} p_i \frac{\partial^i A e^{rt}}{\partial t^i} - q(j2\pi k) A e^{rt} &= 0 \\
\left(\sum_{i=0}^{m} p_i r^i - q(j2\pi k)\right) A e^{rt} &= 0 \\
p(r) - q(j2\pi k) &= 0
\end{aligned} \tag{13}
$$

The solution for $a_k(t)$ is a of the form Ae^{rt}, where r is the root of the above polynomial and A is a constant independent of t. Let the roots of the above polynomial be $r_1(k), r_2(k), \ldots r_m(k)$. Note that the roots of the polynomial are indexed by k as well, implying there is a set of m roots for each value of k. It is essential here to realize that if the field is a physically feasible one, then $\Re(r_i(k)) \leq 0$; $i = 1, 2, \ldots, m; k = -b, \ldots, b$, that is all roots have a non positive real part. Furthermore for simplicity of analysis, all of the roots $r_1(k), r_2(k), \ldots r_m(k)$ are considered to be distinct for a given k.[1] However, it is possible that $r_i(k_1) = r_j(k_2)$ for some $i, j, k_1 \neq k_2$. The condition is similar to the one obtained in control theory, where we want the poles of the closed loop system to lie in the left half plane. Thus, we can use criteria like the Routh Hurwitz condition[21], to ensure the roots have negative real parts in our case.

In fact, the solution for $a_k(t)$ can thus be written as a linear combination of these roots. Thus,

$$
a_k(t) = \sum_{i=1}^{m} a_{ki}(0)\exp(r_i(k)t) \tag{14}
$$

The coefficients have been represented so to maintain consistency of representation of $a_k(t)$ as a function of time. Also $a_{ki}(0)$ are finite constants independent of everything else. Let

[1] If there is a repeated root r, then the solution will be of the form e^{rt} and also te^{rt}, which will make the problem complicated. Such cases can also be treated in a similar manner that has been described in the report. To be very specific, if the repeated root is 0, it can be easily taken into the given framework by combining all the repeated terms with it. This is because if $r = 0$, then $te^{rt} = t$, which diverges and hence cannot be the solution for a physical field. Such nuances have been omitted to simplify the description of the process.

$\alpha_k = \max_i |a_{ki}(0)|$. Using this representation, the partial derivatives of $g(x,t)$ wrt to space and time can be bounded as follows.

$$\begin{aligned}
\left|\frac{\partial}{\partial t}g(x,t)\right| &= \left|\frac{\partial}{\partial t}\sum_{i=1}^{m} a_{ki}(0)\exp(r_i(k)t)\exp(j2\pi kx)\right| \\
&= \left|\sum_{i=1}^{m} a_{ki}(0)r_i(k)\exp(r_i(k)t)\exp(j2\pi kx)\right| \\
&\leq \sum_{i=1}^{m}\left|a_{ki}(0)r_i(k)\exp(r_i(k)t)\exp(j2\pi kx)\right| \\
&\leq \sum_{i=1}^{m}|a_{ki}(0)||r_i(k)| \\
&\leq m\alpha_k R
\end{aligned} \tag{15}$$

$$\begin{aligned}
\left|\frac{\partial}{\partial x}g(x,t)\right| &= \left|\frac{\partial}{\partial x}\sum_{k=-b}^{b} a_k(t)\exp(j2\pi kx)\right| \\
&= \left|\sum_{k=-b}^{b} a_k(t)j2\pi k\exp(j2\pi kx)\right| \\
&\leq \sum_{i=1}^{m}\left|a_{ki}(0)j2\pi k\exp(r_i(k)t)\exp(j2\pi kx)\right| \\
&\leq \sum_{i=1}^{m}2b\pi|a_{ki}(0)| \\
&\leq m\alpha_k 2b\pi
\end{aligned} \tag{16}$$

The third step follows from triangle inequality, fourth step uses the fact that $\Re(r_i(k)) \leq 0$ and the bound on $r_i(k)$ uses the Rouche's theorem [22]. The value of R can be expressed in terms of $|p_i|$'s, which are finite and so is the upper bound. Using the above expression for $a_k(t)$ and using it in equation (4) to obtain the value at $(x,t) = (S_n, T_n)$, we get,

$$g(S_n, T_n) = \sum_{k=-b}^{b}\sum_{i=1}^{m} a_{ki}(0)\exp(r_i(k)T_n)\exp(j2\pi kS_n) \tag{17}$$

The above equation can be written in a vector notation form. Let $e_{k,i}(x,t) = \exp(r_i(k)t + j2\pi kx)$. Define,

$$\begin{aligned}
\mathbf{e}_{k,i}(x,t) &:= [e_{k,1}(x,t),\ e_{k,2}(x,t), \ldots, e_{k,m}(x,t)] \\
\mathbf{a}_k &:= [a_{k1}(0),\ a_{k2}(0),\ a_{k3}(0),\ \ldots,\ a_{km}(0)] \\
\mathbf{a} &= [\mathbf{a}_{-b}, \ldots \mathbf{a}_{-1}, \mathbf{a}_0, \mathbf{a}_1, \ldots \mathbf{a}_b]^T \\
\mathbf{e}(x,t) &= [\mathbf{e}_{-b}(x,t), \ldots, \mathbf{e}_b(x,t)]^H
\end{aligned} \tag{18}$$

Observe that $\mathbf{a}$ and $\mathbf{e}(x,t)$ are column vectors, while $\mathbf{e}_k(x,t)$ and $\mathbf{a}_k$ are row vectors. Since $\Re(r_i(k)) \leq 0$, so $|e_{k,i}(x,t)| \leq 1$. This implies,

$$\begin{aligned}
||\mathbf{e}_k(x,t)||^2 &= \sum_{i=1}^{m}|\exp(r_i(k)t + j2\pi kx)|^2 \leq \sum_{i=1}^{m}1 \leq m \\
||\mathbf{e}(x,t)||^2 &= \sum_{k=-b}^{b}||\mathbf{e}_k(x,t)||^2 \leq \sum_{k=-b}^{b}m = m(2b+1)
\end{aligned} \tag{19}$$

Therefore, on using (18), equation (17) can be rewritten as,

$$g(S_n, T_n) = \mathbf{e}^H(S_n, T_n)\ \mathbf{a} \tag{20}$$

Due to measurement-noise, the observed field values will be $g(S_i, T_i) + W(S_i, T_i), 1 \le i \le M$. Note that S_i and T_i are sampling locations and their corresponding time stamps. Define two vectors,

$$\mathbf{g} = [g_1, g_2, \ldots g_M]^T \text{ and } \mathbf{w} = [w_1, w_2, \ldots w_M]^T, \tag{21}$$

where $g_i = g(S_i, T_i)$ and $w_i = W(S_i, T_i)$. Let $\mathbf{g}_s$ be the measurement-noise affected samples obtained at $(S_1, T_1), \ldots, (S_M, T_M)$. Then, $\mathbf{g}_s = \mathbf{g} + \mathbf{w}$. Combining equation (20) and (21), we get,

$$\mathbf{g} = \begin{bmatrix} \mathbf{e}^H(S_1, T_1) \\ \mathbf{e}^H(S_2, T_2) \\ \vdots \\ \mathbf{e}^H(S_M, T_M) \end{bmatrix} \mathbf{a} = Y\mathbf{a} \text{ ; where } Y = \begin{bmatrix} \mathbf{e}^H(S_1, T_1) \\ \mathbf{e}^H(S_2, T_2) \\ \vdots \\ \mathbf{e}^H(S_M, T_M) \end{bmatrix} \tag{22}$$

2.2.2 Estimation method

The main idea behind the reconstruction of the field would be that the sampling location and time instants are "near" to the locations and time instants, had we sampled uniformly both in time and space for M points. Thus, to incorporate the same into the formulation, let

$$s_i = \frac{i}{M}, \text{ and } t_i = \frac{iT_0}{M} \text{ for } 1 \le i \le M \tag{23}$$

$$\mathbf{g}_0 = [g_{u1}, g_{u2}, \ldots g_{uM}]^T; \; Y_0 = \begin{bmatrix} \mathbf{e}^H(s_1, t_1) \\ \mathbf{e}^H(s_2, t_2) \\ \vdots \\ \mathbf{e}^H(s_M, t_M) \end{bmatrix}, \tag{24}$$

where $g_{ui} = g(s_i, t_i), i = 1, 2, \ldots, M$. This implies, $\mathbf{g}_0 = Y_0\mathbf{a}$. Note that Y_0 has Vandermonde structure[23]. Now, since we expect that the sampling locations are "near" to the grid points, we can estimate the Fourier coefficients by assuming that samples have been obtained by multiplying the Fourier coefficient vector by Y_0 instead of the unknown matrix, Y. The best least-squared estimate of the Fourier coefficients, $\hat{\mathbf{a}}$, thus would be,

$$\hat{\mathbf{a}} = \arg\min_{\mathbf{b}} = ||\mathbf{g}_s - Y_0\mathbf{b}||^2 \tag{25}$$

Since the main way to achieve to estimate the field relies on oversampling, the sampling density will be generally very large and thus, $n > m(2b + 1)$, making this problem a linear least squares estimation problem (regression). The solution to this problem is well known and uses the psuedoinverse of the matrix. Therefore,

$$\hat{\mathbf{a}} = (Y_0^H Y_0)^{-1} Y_0^H \mathbf{g}_s \tag{26}$$

$$\mathbf{a} = (Y_0^H Y_0)^{-1} Y_0^H \mathbf{g}_0 \tag{27}$$

The second equation is obtained in a similar manner. However, it is important to realize at this point that the first equation is a least-square *estimate* because of the unknown locations and noise while the second equation is an *exact* solution. The distortion given in (5) can be rewritten as,

$$\begin{aligned}
\sum_{k=-b}^{b} \mathbb{E}\left[|\hat{A}_k(0) - a_k(0)|^2\right] &= \sum_{k=-b}^{b} \mathbb{E}\left[\left|\sum_{i=1}^{m} \hat{A}_{ki}(0) - \sum_{i=1}^{m} a_{ki}(0)\right|^2\right] \\
&= \sum_{k=-b}^{b} \mathbb{E}\left[\left|\sum_{i=1}^{m} \left(\hat{A}_{ki}(0) - a_{ki}(0)\right)\right|^2\right] \\
&\leq \sum_{k=-b}^{b} \mathbb{E}\left[m \sum_{i=1}^{m} \left|\hat{A}_{ki}(0) - a_{ki}(0)\right|^2\right] \\
&= m \sum_{k=-b}^{b} \mathbb{E}\left[||\hat{\mathbf{a}}_k - \mathbf{a}_k||^2\right] \\
&\leq m \sum_{k=-b}^{b} \mathbb{E}\left[||\hat{\mathbf{a}} - \mathbf{a}||^2\right] \\
&= m(2b+1)\mathbb{E}\left[||\hat{\mathbf{a}} - \mathbf{a}||^2\right] \qquad (28)
\end{aligned}$$

The third step uses the inequality $\left(\sum_{i=1}^{n} a_i\right)^2 \leq n \sum_{i=1}^{n} a_i^2$ for $a_1, a_2, \ldots, a_n \in \mathbb{R}$. The following is the main result of this work.

Theorem 1. *Let* $\hat{\mathbf{a}}$ *and* $\mathbf{a}$ *be as defined in equation* (26)*. Under the sampling model discussed and the corruption by the measurement noise, the following result holds*

$$\mathbb{E}\left[||\hat{\mathbf{a}} - \mathbf{a}||^2\right] \leq \frac{C'}{n}$$

where n *is the average sampling density and* C' *is a positive constant independent of* n*. It depends on the bandwidth,* b *of the signal, the support parameters of the renewal processes,* λ *and* μ*, the coefficients of the PDE and the noise variance,* σ^2*. The proportionality constant,* C' *is directly proportional to* b*,* λ*,* μ *and* σ^2*. The dependence on the coefficients of the PDE is in a very non linear way through the roots of the equations whose almost all coefficients are determined by these values. Correspondingly, the distortion error can be bounded as* $\frac{m(2b+1)C'}{n}$

Proof. Thus, the value of $\mathbb{E}\left[\left|\hat{\mathbf{a}} - \mathbf{a}\right|^2\right]$ will be upper bounded to obtain a bound on distortion. Let $A = (Y_0^H Y_0)^{-1} Y_0^H$.

$$\begin{aligned}
\mathbb{E}\left[||\hat{\mathbf{a}} - \mathbf{a}||^2\right] &= \mathbb{E}\left[||(Y_0^H Y_0)^{-1} Y_0^H \mathbf{g}_s - \mathbf{a}||^2\right] \\
&= \mathbb{E}\left[||(Y_0^H Y_0)^{-1} Y_0^H (\mathbf{g} + \mathbf{w}) - (Y_0^H Y_0)^{-1} Y_0^H \mathbf{g}_0||^2\right] \\
&= \mathbb{E}\left[||A(\mathbf{g} + \mathbf{w} - \mathbf{g}_0)||^2\right] \\
&\leq 2\mathbb{E}\left[||A(\mathbf{g} - \mathbf{g}_0)||^2\right] + 2\mathbb{E}\left[||A\mathbf{w}||^2\right] \\
&\leq 2\mathbb{E}\left[\lambda_{\max}^A ||\mathbf{g} - \mathbf{g}_0||^2\right] + 2\mathbb{E}\left[||A\mathbf{w}||^2\right] \qquad (29)
\end{aligned}$$

where second step follows from equation (26) and definition of $\mathbf{g}_s$, the fourth step from Cauchy Scwharz inequality and $\lambda_{\max}^A$ is the largest eigen value of $A^H A$. Both the terms, along with the bound on $\lambda_{\max}^A$, in the last step will be analyzed separately to obtain the bound on the error. Also consider,

$$\begin{aligned}
||\mathbf{g} - \mathbf{g}_0||^2 &= \sum_{i=1}^{M} |g(S_i, T_i) - g(s_i, t_i)|^2 \\
&\leq \sum_{i=1}^{M} \left(\max_{\mathbf{x}, t} ||\nabla g||_2\right) \left\{|S_i - s_i|^2 + |T_i - t_i|^2\right\} \\
&\leq C_0 \left\{\sum_{i=1}^{M} \left|S_i - \frac{i}{M}\right|^2 + \sum_{i=1}^{M} \left|T_i - \frac{iT_0}{M}\right|^2\right\} \qquad (30)
\end{aligned}$$

The second step follows from the fact that for a smooth function $h(\mathbf{x})$, for $\mathbf{x} \in \mathbb{R}^n$, $|h(\mathbf{x}_1) - h(\mathbf{x}_2)| \leq \left(\max_{\mathbf{x}} ||\nabla h||\right) ||\mathbf{x}_1 - \mathbf{x}_2||$. The third step uses the fact that $||\nabla g||_2$ is upper bounded. This follows from the fact that $||\nabla g||_2^2 = \left(\frac{\partial}{\partial x} g(x,t)\right)^2 + \left(\frac{\partial}{\partial t} g(x,t)\right)^2$ along with the bounds on partial derivatives. From (15) and (16), we can write,

$$||\nabla g||_2^2 = \left(\frac{\partial}{\partial x} g(x,t)\right)^2 + \left(\frac{\partial}{\partial t} g(x,t)\right)^2 \leq (m\alpha_k R)^2 + (m\alpha_k R 2b\pi)^2 \leq C_0^2 \tag{31}$$

for some $C_0 \in \mathbb{R}$. Now consider,

$$\begin{aligned}
\lambda_{\max}^A &\overset{(a)}{\leq} \text{tr}(A^H A) \\
&= \text{tr}(AA^H) \\
&= \text{tr}((Y_0^H Y_0)^{-1} Y_0^H Y_0 (Y_0^H Y_0)^{-1}) \\
&= \text{tr}((Y_0^H Y_0)^{-1})
\end{aligned} \tag{32}$$

(a) follows from the fact that trace of a matrix is the sum of its eigen values and since $A^H A$ is symmetric, all its eigen values will be non negative therefore, the sum will be greater than the largest eigen value. For the second term in the equation (29), the structure of the noise model can be exploited to simplify the expression. Note that using the assumptions on the noise model, $\mathbb{E}[\mathbf{w}] = 0$ and $\mathbb{E}[\mathbf{w}\mathbf{w}^T] = \sigma^2 I$, where I is the identity matrix.

$$\begin{aligned}
\mathbb{E}\left[||A\mathbf{w}||^2\right] &= \mathbb{E}\left[\mathbf{w}^T(A^H A\mathbf{w})\right] \\
&\overset{(a)}{=} \mathbb{E}\left[\text{tr}\left(\mathbf{w}^T(A^H A\mathbf{w})\right)\right] \\
&\overset{(b)}{=} \mathbb{E}\left[\text{tr}\left((A^H A\mathbf{w})\mathbf{w}^T\right)\right] \\
&\overset{(c)}{=} \text{tr}\left(\mathbb{E}\left[A^H A\mathbf{w}\mathbf{w}^T\right]\right) \\
&\overset{(d)}{=} \text{tr}\left(\mathbb{E}\left[A^H A\right]\mathbb{E}\left[\mathbf{w}\mathbf{w}^T\right]\right) \\
&= \text{tr}\left(\mathbb{E}\left[A^H A\right]\sigma^2 I\right) \\
&= \mathbb{E}\left[\text{tr}(A^H A\sigma^2 I)\right] \\
&= \mathbb{E}\left[\sigma^2 \text{tr}(A^H A)\right] \\
&= \sigma^2 \mathbb{E}\left[\text{tr}((Y_0^H Y_0)^{-1})\right]
\end{aligned} \tag{33}$$

where,
(a) uses the fact that $||A\mathbf{w}||^2$ is scalar hence, it equals its trace,
(b) follows from $\text{tr}(AB) = \text{tr}(BA)$
(c) uses linearity of expectation and the trace operator
(d) is a result of independence of noise and sampling

Thus, from equation (29), (32) and (33), it is clear that characterizing the bound on $\text{tr}((Y_0^H Y_0)^{-1})$ is required and will also suffice for the purpose.

Let $\lambda_1, \lambda_2, \ldots, \lambda_{m(2b+1)}$ be eigen values of $Y_0^H Y_0$. Since the matrix is symmetric, $\lambda_i \geq 0$ $i = 1, 2, \ldots, m(2b+1)$. Using the property of eigen values, the eigen values of $(Y_0^H Y_0)^{-1}$, will be $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots, \frac{1}{\lambda_{m(2b+1)}}$. Let $\lambda_{\max}$ and $\lambda_{\min}$ be the maximum and minimum eigen values of $Y_0^H Y_0$. Therefore, $\frac{1}{\lambda_{\min}}$ and $\frac{1}{\lambda_{\max}}$ will be the maximum and minimum eigen values of $(Y_0^H Y_0)^{-1}$. Applying the Polya-Szego's inequality[24] on the sequence formed by eigenvalues of $Y_0^H Y_0$ and by those of $(Y_0^H Y_0)^{-1}$, we can write

$$\frac{\sum_{i=1}^{m(2b+1)} \lambda_i \sum_{i=1}^{m(2b+1)} (1/\lambda_i)}{(\sum_{i=1}^{m(2b+1)} \sqrt{\lambda_i . 1/\lambda_i})^2} \leq \frac{1}{4}\left(\frac{\lambda_{\max}}{\lambda_{\min}} + \frac{\lambda_{\min}}{\lambda_{\max}}\right)^2 \tag{34}$$

Noting that $\text{tr}(Y_0^H Y_0) = \sum_{i=1}^{m(2b+1)} \lambda_i, \text{tr}((Y_0^H Y_0)^{-1}) = \sum_{i=1}^{m(2b+1)} (1/\lambda_i)$ and κ is the condition number of the matrix $Y_0^H Y_0$, it can be written as,

$$\begin{aligned}
\text{tr}(Y_0^H Y_0)\ \text{tr}((Y_0^H Y_0)^{-1}) &\leq \frac{m^2(2b+1)^2}{4}\left(\kappa + \frac{1}{\kappa}\right)^2 \\
\text{tr}((Y_0^H Y_0)^{-1}) &\leq \frac{m^2(2b+1)^2}{4\ \text{tr}(Y_0^H Y_0)}\left(\kappa + \frac{1}{\kappa}\right)^2 \qquad (35)
\end{aligned}$$

Consider,

$$\begin{aligned}
\text{tr}(Y_0^H Y_0) &= \text{tr}\left([\mathbf{e}(s_1,t_1), \mathbf{e}(s_2,t_2), \ldots, \mathbf{e}(s_M,t_M)] \begin{bmatrix} \mathbf{e}^H(s_1,t_1) \\ \mathbf{e}^H(s_2,t_2) \\ \vdots \\ \mathbf{e}^H(s_M,t_M) \end{bmatrix} \right) \\
&= ||\mathbf{e}(s_1,t_1)||^2 + ||\mathbf{e}(s_2,t_2)||^2 + \cdots + ||\mathbf{e}(s_M,t_M)||^2 \\
&\overset{(a)}{=} \sum_{i=1}^{M} \sum_{k=-b}^{b} ||\mathbf{e}_k(s_i,t_i)||^2 \\
&\overset{(b)}{=} \sum_{i=1}^{M} \sum_{k=-b}^{b} \sum_{j=1}^{m} |e_k^j(s_i,t_i)|^2 \\
&= \sum_{i=1}^{M} \sum_{k=-b}^{b} \sum_{j=1}^{m} \exp(2\Re(r_j(k))t_i) \\
&= \sum_{j=1}^{m} \sum_{k=-b}^{b} \sum_{i=1}^{M} \exp\left(2\Re(r_j(k))\frac{iT_0}{M}\right) \qquad (36)
\end{aligned}$$

Both (a) and (b) follow directly from equation (18). For a given value of j and k, the sum $\sum_{i=1}^{M} \exp\left(2\Re(r_j(k))\frac{iT_0}{M}\right)$ can be considered as a scalar multiple of Riemann sum approximation of the integral $\int_0^{T_0} \exp(2\Re(r_j(k))t)dt$ with partitions chosen uniformly over the interval. The only difference here that the terms in the sum are not multiplied by the interval difference, which in this case is the same for all intervals and hence can be taken out as a scalar. It is interesting to note here that the value of this integral can be used as a bound on the value of the sum because of that fact that $\exp(2\Re(r_j(k))t)$ is decreasing since $\Re(r_j(k)) \leq 0\ \forall j,k$. The bound can be obtained as,

$$\begin{aligned}
\int_0^{T_0} \exp(2\Re(r_j(k))t)dt &= \sum_{i=1}^{M} \int_{\frac{(i-1)T_0}{M}}^{\frac{iT_0}{M}} \exp(2\Re(r_j(k))t)dt \\
&\overset{(a)}{\leq} \sum_{i=1}^{M} \int_{\frac{(i-1)T_0}{M}}^{\frac{iT_0}{M}} \exp\left(2\Re(r_j(k))\frac{(i-1)T_0}{M}\right) dt \\
&= \sum_{i=1}^{M} \exp\left(2\Re(r_j(k))\frac{iT_0}{M}\right) \exp\left(-2\Re(r_j(k))\frac{T_0}{M}\right) \int_{\frac{(i-1)T_0}{M}}^{\frac{T_0}{M}} dt \\
&= \exp\left(-2\Re(r_j(k))\frac{T_0}{M}\right) \frac{T_0}{M} \sum_{i=1}^{M} \exp\left(2\Re(r_j(k))\frac{iT_0}{M}\right) \\
&\leq \exp(-2\Re(r_j(k)))\frac{T_0}{M} \sum_{i=1}^{M} \exp\left(2\Re(r_j(k))\frac{iT_0}{M}\right) \qquad (37)
\end{aligned}$$

where (a) follows from the decreasing nature of $\exp(2\Re(r_j(k))t)$ and the last step uses that fact that $T_0 \leq M$. This implies,

$$\exp\left(2\Re(r_j(k))\frac{iT_0}{M}\right) \geq M\exp(2\Re(r_j(k)))\int_0^{T_0}\exp(2\Re(r_j(k))t)dt = MC_{jk} \tag{38}$$

where $C_{jk} = \exp(2\Re(r_j(k)))\int_0^{T_0}\exp(2\Re(r_j(k))t)dt$ is a finite constant for each j,k. Using equation (38) in equation (36), we get

$$\text{tr}(Y_0^H Y_0) \geq \sum_{j=1}^{m}\sum_{k=-b}^{b} MC_{jk} = MC_3 \tag{39}$$

where C_3 is a finite deterministic constant given by $C_3 = \sum_{j=1}^{m}\sum_{k=-b}^{b} C_{jk}$. Clearly, it is independent of n. Since Y_0 is a random Vandermonde matrix, using the result in [25] for large random Vandermonde matrices with complex entries lying in the unit circle, we can also say that the condition number of every realization of Y_0 is independent of its dimensions, more specifically, M and is upper bounded by a finite constant that does not scale with M. Hence, κ, the condition number of $Y_0^H Y_0$ is also upper bounded by a finite constant, $C_k > 0$, independent of M and hence n. Since $\kappa \geq 1$ and $\kappa \leq C_k, \exists\ K > 0$, such that the term $\left(\kappa + \frac{1}{\kappa}\right)^2 \leq K$. Combining this results with ones obtained in (35) and (39),

$$\text{tr}((Y_0^H Y_0)^{-1}) \leq \frac{m^2(2b+1)^2}{4MC_3}K = \frac{C_t}{M} \tag{40}$$

From (11), it is noted that $M > \frac{n-\lambda}{\lambda}$ or $\frac{1}{M} < \frac{\lambda}{n-\lambda}$. This implies, $\mathbb{E}\left[\frac{1}{M}\right] < \left[\frac{\lambda}{n-\lambda}\right]$. Therefore,

$$\mathbb{E}\left[\text{tr}((Y_0^H Y_0)^{-1})\right] \leq \mathbb{E}\left[\frac{C_t}{M}\right] \leq \frac{C_t\lambda}{n-\lambda} \tag{41}$$

Substituting the results obtained above in equations (32) and (33) and combining them with the equations (29) and (30), we get,

$$\mathbb{E}\left[||\hat{\mathbf{a}} - \mathbf{a}||^2\right] \leq 2\mathbb{E}\left[\frac{C_t}{M}\left\{\sum_{i=1}^{M}\left|S_i - \frac{i}{M}\right|^2 + \sum_{i=1}^{M}\left|T_i - \frac{i}{M}\right|^2\right\}\right] + \frac{2C_t\lambda}{n-\lambda} \tag{42}$$

From Appendix B, it is noted that,

$$\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|S_i - \frac{i}{M}\right|^2\right] \leq \frac{C_S}{n} \text{ and } \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|T_i - \frac{iT_0}{M}\right|^2\right] \leq \frac{C_T}{n}$$

. Therefore, we can conclude that

$$\mathbb{E}\left[||\hat{\mathbf{a}} - \mathbf{a}||^2\right] \leq \frac{C_tC_S}{n} + \frac{C_tC_T}{n} + \frac{C_t\lambda}{n-\lambda} \leq \frac{C'}{n} \tag{43}$$

This completes the main result. □

2.3 Simulations

This section presents the results of simulations and explanations to the same. The simulations have been presented in Fig. 2 The simulations help ease out verification of the results obtained of different PDEs and analyze the effect of sampling density.

Firstly, for the purpose of analysis, a field $g(x,t)$ with $b = 3$ is considered and its Fourier coefficients have been generated using independent trials of a Uniform distribution over $[-1,1]$ for all real and imaginary parts separately. It is ensured that the field is real by using conjugate symmetry. Finally, the field is scaled to have $|g(x)| \leq 1$. Three differential equations have been considered for the purpose and Fourier coefficients have been reused. The simulations have been carried out using the following PDEs

$$\frac{\partial^2}{\partial t^2} g(x,t) + 3\frac{\partial}{\partial t} g(x,t) = 0.01 \left(\frac{\partial^2}{\partial x^2} g(x,t) - 0.125 \frac{\partial^4}{\partial x^4} g(x,t) \right) \tag{44}$$

$$\frac{\partial^2}{\partial t^2} g(x,t) + 3\frac{\partial}{\partial t} g(x,t) = 0.01 \frac{\partial^2}{\partial x^2} g(x,t) \tag{45}$$

$$\frac{\partial}{\partial t} g(x,t) = 0.01 \frac{\partial^2}{\partial x^2} g(x,t) \tag{46}$$

The corresponding polynomials are (i) $p_1(z) = z^2 + 3z, q_1(z) = 0.01(z^2 - 0.0125z^4)$, (ii) $p_2(z) = z^2 + 3z, q_2(z) = 0.01z^2$, and (iii) $p_3(z) = z, q_3(z) = 0.01z^2$. Note that the last one is the diffusion equation. The same two sets of Fourier coefficients were used in first two equations. Note that the others are conjugate of these to ensure real field.

$$\begin{gathered} a_1[0] = 0.3002;\ a_2[0] = 0.2445; \\ a_1[1] = -0.0413 + j0.0216;\ a_2[1] = -0.0357 + j0.0478; \\ a_1[2] = 0.0871 + j0.0343;\ a_2[2] = 0.0978 + j0.0729; \\ a_1[3] = -0.1679 - j0.0586;\ a_2[3] = -0.1796 - j0.0756; \end{gathered} \tag{47}$$

The Fourier coefficients used in the last equation are

$$\begin{aligned} a[0] &= 0.11 \\ a[1] &= 0.023 - j0.076 \\ a[2] &= 0.0669 + j0.0551 \\ a[3] &= 0.2 + j0.0821 \end{aligned} \tag{48}$$

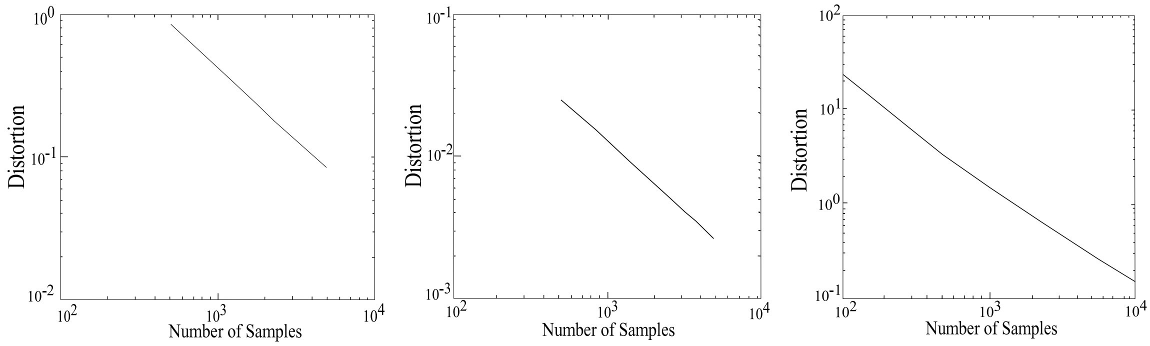

Figure 1: The three different figures show the variation of error with the number of samples for different PDEs. The first one is the corresponding to first PDE in equation (44), the middle one corresponds to the second equation and the last one to the third equation in (44), which is the standard diffusion equation. The variation in the distortion is clearly of $O(1/n)$ as denoted by the slopes in the plots. However, the error is slightly large because of the large condition number of the matrix Y_0 giving issues regarding numerical stability.

The Figure 2 shows the mean square error of the estimate for different PDEs. The plots are shown for the different PDEs as they have been listed in the equation (44). The slopes of the lines obtained are $-1.0019, -1.0110$ and -1.0086 which confirms the $O(1/n)$ decrease.

3 Samples from an Autoregressive Process

3.1 Field Model, Distortion Metric, Sampling Model and Measurement Noise Model

The section describes the models which have been used for analysis in this work. Firstly the field model is described followed by the distortion metric, the sampling model and finally the measurement noise model.

3.1.1 Field Model

The Field Model is very similar to the one presented in the previous case. However, the only difference is that a temporally fixed field is considered to increase tractability. The field is given by $g(x)$ and has a Fourier series expansion given by

$$g(x) = \sum_{k=-b}^{b} a[k] \exp(j2\pi x) \quad ; \quad a[k] = \int_0^1 g(x) \exp(-j2\pi x) dx \tag{49}$$

where b is a known positive integer. Also,

$$|g'(x)| \leq 2b\pi ||g||_\infty \leq 2b\pi \tag{50}$$

directly follows from the Bernstein's inequality.[26]

3.1.2 Distortion Metric

The distortion metric will be the same as in the previous case just incorporating the temporally fixed nature of the field. If $\hat{G}(x)$ is an estimate of the field, then the distortion, $\mathcal{D}(\widehat{G}, g)$ is defined as

$$\mathcal{D}(\widehat{G}, g) := \mathbb{E}\left[\int_0^1 |\widehat{G}(x) - g(x)|^2 dx\right] = \sum_{k=-b}^{b} \mathbb{E}\left[|\hat{A}[k] - a[k]|^2\right] \tag{51}$$

where, $\hat{A}[k] = \displaystyle\int_0^1 \widehat{G}(x) \exp(-j2\pi x) dx$

3.1.3 Sampling Model

Let $\{X_i\}_{i=1}^{M}$ be the set of intersample distances where X_i is the distance between i^{th} and $(i-1)^{\text{th}}$ sample and X_1 is the distance of the first sample from $x = 0$. The intersample distances have been modeled as an autoregressive process of order 1 using a parameter ρ, $X_i = \rho X_{i-1} + Y_i \; \forall \; i \geq 2$ and $X_1 = Y_1$. The ρ models the dependence of the current intersample distance on the previous one and Y_i corresponds to the stochastic term. Y_i's are realized using an unknown renewal process. That is to say, Y_i's are independent and identically distributed random variables realized from an *unknown* common distribution Y, such that $Y > 0$. Using these intersampling distances, the sampling locations, S_n are given by $S_n = \sum_{i=1}^{n} X_i$. The sampling is done over an interval $[0, 1]$ and M is the random number of samples that lie in the interval i.e. it is defined such that, $S_M \leq 1$ and $S_{M+1} > 1$. Thus M is a well defined measurable random variable[19].

For the purpose of ease of analysis and tractability, the support of the distribution of Y is considered to be finite and inversely proportional to the sampling density. Hence, it is assumed that

$$0 < Y \leq \frac{\lambda}{n} \text{ and } \mathbb{E}[Y] = \frac{1}{n}, \tag{52}$$

where $\lambda > 1$ is a parameter that characterizes the support of the distribution. It is a finite number and is independent of n. This would be a crucial factor that governs the constant of proportionality in the expected mean squared error in the estimate of the field. Furthermore, it is also an important factor that determines the threshold on the minimum number of samples. Note that,

$$X_1 = Y_1; \quad X_2 = \rho X_1 + Y_2 = \rho Y_1 + Y_2; \quad X_3 = \rho X_2 + Y_3 = \rho^2 Y_1 + \rho Y_2 + Y_3 \tag{53}$$

which can be generalized as $X_n = \sum_{r=1}^{n} \rho^{n-r} Y_r$. This can be used to find a closed form expression of S_n which can be written as $S_n = \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} \sum_{r=1}^{i} \rho^{i-r} Y_r$. Therefore,

$$S_n = \sum_{i=1}^{n} \sum_{r=1}^{i} \rho^{i-r} Y_r = \sum_{i=1}^{n} \sum_{r=0}^{n-i} \rho^{r} Y_i = \frac{1}{1-\rho} \sum_{i=1}^{n} (1 - \rho^{n-i+1}) Y_i = \frac{1}{1-\rho} \sum_{i=1}^{n} c_{i,n} Y_i \tag{54}$$

where, $c_{i,n} = 1 - \rho^{n-i+1}$, for $i = 1, 2, 3 \ldots n$. For $0 \leq \rho < 1$,

$$\min_i c_{i,n} = 1 - \rho \text{ and } \max_i c_{i,n} \leq 1 \text{ for all } n. \tag{55}$$

. We know that by definition, $S_{M+1} > 1$. Substiuting S_{M+1} from (54),

$$1 < \frac{1}{1-\rho} \sum_{i=1}^{M+1} c_{i,M+1} Y_i < \frac{(\max_k c_{k,M+1})}{1-\rho} \sum_{i=1}^{M+1} Y_i \leq \frac{1}{1-\rho} \sum_{i=1}^{M+1} Y_i \leq \frac{\lambda(M+1)}{n(1-\rho)} \tag{56}$$

The last step follows from (52). Therefore,

$$M > \frac{n(1-\rho)}{\lambda} - 1 \tag{57}$$

This gives a lower bound on the value of M. To find the bounds on the expected number of samples, $\mathbb{E}[M]$, the following lemma is noted.

Lemma: For the sampling model described as above with the intersample distances coming from an autoregressive model, the average number of samples taken in over the interval obeys the following bounds

$$n(1-\rho) - 1 \leq \mathbb{E}[M] \leq n + \frac{\lambda}{1-\rho} - 1$$

Proof: For the proof of the lemma, we first need to consider the upper bound on the value of $\mathbb{E}[S_{M+1}]$,

$$\mathbb{E}[S_{M+1}] = \mathbb{E}\left[\frac{1}{1-\rho} \sum_{i=1}^{M+1} c_{i,M+1} Y_i\right] = \frac{1}{1-\rho} \mathbb{E}\left[\sum_{i=1}^{M+1} c_{i,M+1} Y_i\right] \leq \frac{(\max_k c_{k,M+1})}{1-\rho} \mathbb{E}\left[\sum_{i=1}^{M+1} Y_i\right] \tag{58}$$

Now using Wald's identity[19], one can write,

$$\mathbb{E}\left[\sum_{i=1}^{M+1} Y_i\right] = \mathbb{E}[Y]\mathbb{E}[M+1] = \mathbb{E}[Y](\mathbb{E}[M]+1) \tag{59}$$

It is important to note that Wald's identity is applicable on this expression and not directly on S_{M+1} because Y_i's are a set of independent and identically distributed random variables while X_i's are not. Using this with (58) and the bounds obtained in (55), we can write, $\mathbb{E}[S_{M+1}] \leq \frac{1}{1-\rho}\left(\frac{1}{n}\right)(\mathbb{E}[M]+1)$. Since, by definition, $S_{M+1} > 1$, therefore, $\mathbb{E}[S_{M+1}] > 1$. Combine this with the result in (58) and (59) to get,

$$\mathbb{E}[M] > n(1-\rho) - 1 \tag{60}$$

Similarly we can consider a lower bound on $\mathbb{E}[S_{M+1}]$,

$$\mathbb{E}[S_{M+1}] \geq \frac{1}{1-\rho}(\min_k c_{k,M+1})\mathbb{E}\left[\sum_{i=1}^{M+1} Y_i\right] = \frac{1-\rho}{1-\rho}\mathbb{E}[Y].(\mathbb{E}[M]+1) = \frac{1}{n}(\mathbb{E}[M]+1) \tag{61}$$

Thus to upper bound the value of $\mathbb{E}[M]$, the following equation is considered,

$$\begin{aligned} S_{M+1} - S_M &= \frac{1}{1-\rho}\sum_{i=1}^{M+1}(1-\rho^{M-i+2})Y_i - \frac{1}{1-\rho}\sum_{i=1}^{M}(1-\rho^{M-i+1})Y_i \\ &= \frac{1}{1-\rho}\sum_{i=1}^{M+1}(\rho^{M-i+1}-\rho^{M-i})Y_i = \sum_{i=1}^{M+1}\rho^{M-i+1}Y_i \end{aligned}$$

Since, $S_M \leq 1$ and $Y \leq \frac{\lambda}{n}$, we can write,

$$S_{M+1} = S_M + \sum_{i=1}^{M+1}\rho^{M-i+1}Y_i \leq 1 + \sum_{i=1}^{M+1}\rho^{M-i+1}\frac{\lambda}{n} \leq 1 + \frac{\lambda}{n}\sum_{i=0}^{\infty}\rho^i = 1 + \frac{\lambda}{n(1-\rho)} \tag{62}$$

This implies,

$$\mathbb{E}[S_{M+1}] \leq 1 + \frac{\lambda}{n(1-\rho)} \tag{63}$$

Combine equation (63) with equations (61) and (60) to obtain,

$$n(1-\rho) - 1 \leq \mathbb{E}[M] \leq n + \frac{\lambda}{1-\rho} - 1 \tag{64}$$

This completes the proof of the lemma. Since the expected number of samples is of the order of n, therefore, n is termed the sampling density. However, the results are governed by the *effective* sampling density which is given by $n(1-\rho)$. The difference becomes relevant at values of ρ close to 1 and finitely large n. More detailed explanation about this has been given in Section III, in the light of obtained results.

3.1.4 Measurement Noise model

The noise model is same as in the previous case.

3.2 Field Estimation from the Obtained Samples

The primary idea in the reconstruction of the field would be the estimation of the Fourier coefficients of the field using the noisy samples that have been obtained at unknown sampling locations where the intersample distances have been modeled using an autoregressive model. The approach is similar to[11], however, there would be a difference in the analysis arising because of the correlated intersample distances. A Riemann sum kind of approximation using the obtained samples will be used to get estimates of the Fourier coefficients of the signals. Define the estimate of the Fourier coefficients as

$$\hat{A}_{\text{gen}}[k] = \frac{1}{M}\sum_{i=1}^{M}\left[g(S_i) + W(S_i)\right]\exp\left(-\frac{j2\pi ki}{M}\right) \tag{65}$$

The motivation to use this as the estimate is the same as the one in the previous paper. The difference is in how this estimate behaves. This is because in [11], the sample locations were considered to be "near" the grid points as the sampling was on locations obtained from i.i.d. intersample distances. Thus, because of independence, each point was individually likely to be close to the grid points, unaffected by others. However, in this case, the intersample distances are correlated and hence the error in one will propagate to all the further ones breaking the premise of these locations being "near" to the grid points. Despite this error propagation, it will be shown that the bound on error is still $O(1/n)$ which is non-trivial. Thus, even though the work is inspired from the approach taken in [11], the analysis in the given setup is far more challenging. With our estimate under this scenario, the bound on the mean square error is analysed and the following theorem is noted.

Theorem 2. *Define the estimate of the Fourier coefficients $\hat{A}_{gen}[k]$ as in (65). Then in the scenario that the intersample distances are obtained from an autoregressive model, the expected mean squared error between the estimated and the actual Fourier coefficients is upper bounded as*

$$\mathbb{E}\left[\left|\hat{A}_{gen}[k] - a[k]\right|^2\right] \leq \frac{C - C'\rho^n}{n}, \tag{66}$$

where C, C' are finite positive constants independent of n. The constants depend on λ, a finite constant independent of n that characterises the support of the distribution of Y, the bandwidth b of the field, the 'correlation coefficient' ρ and the variance of the measurement noise, σ^2.

Proof. The proof will be coupled step approach by using a Riemann sum approximation of the integral in (49) and bounding the error between the Riemann sum and the estimate of the Fourier coefficients given in (65). The detailed calculations will be separately shown in the Appendices while using the main results to maintain the flow of the proof. Define,

$$\hat{A}[k] := \frac{1}{M}\sum_{i=1}^{M} g(S_i)\exp\left(-\frac{j2\pi ki}{M}\right) \ ; \ \hat{W}[k] := \frac{1}{M}\sum_{i=1}^{M} W(S_i)\exp\left(-\frac{j2\pi ki}{M}\right)$$

Therefore, $\hat{A}_{\text{gen}}[k] = \hat{A}[k] + \hat{W}[k]$. The above terms are signal and noise components in the estimates and they have been separated to ease out calculations. The integral in (49), defining $a[k]$, can be approximated using an M-point Riemann sum as,

$$A_R[k] = \frac{1}{M}\sum_{i=1}^{M} g\left(\frac{i}{M}\right)\exp\left(-\frac{j2\pi ki}{M}\right) \tag{67}$$

It is important to note here that is this sum is actually random because of M and thus when used to calculate the distortion, the expected value of the sum should be taken to average over different estimates of the fields.

$$\begin{aligned}
\mathcal{D}(\widehat{G}, g) &= \mathbb{E}\left[\left|\hat{A}_{\text{gen}}[k] - a[k]\right|^2\right] \\
&= \mathbb{E}\left[\left|\hat{A}[k] + \hat{W}[k] - a[k]\right|^2\right] \\
&\leq 2\mathbb{E}\left[\left|\hat{A}[k] - a[k]\right|^2\right] + 2\mathbb{E}\left[\left|\hat{W}[k]\right|^2\right] \\
&= 2\mathbb{E}\left[\left|\hat{A}[k] - A_R[k] + A_R[k] - a[k]\right|^2\right] + 2\mathbb{E}\left[\left|\hat{W}[k]\right|^2\right] \\
&\leq 4\mathbb{E}\left[\left|\hat{A}[k] - A_R[k]\right|^2\right] + 4\mathbb{E}\left[\left|A_R[k] - a[k]\right|^2\right] + 2\mathbb{E}\left[\left|\hat{W}[k]\right|^2\right]
\end{aligned} \tag{68}$$

where the second and the fourth step follow from the inequality $\left|\sum_{i=1}^{n} a_i\right|^2 \leq n\sum_{i=1}^{n}|a_i|^2$, for any numbers $a_1, a_2, \ldots, a_n \in \mathbb{C}$. The three terms as obtained in the final step of (68) obtained will be analyzed separately and solved one by one. The three terms are considered in their specific order.

$$\begin{aligned}
\left|\hat{A}[k] - A_R[k]\right| &= \left|\frac{1}{M}\sum_{i=1}^{M} g(S_i)\exp\left(-\frac{j2\pi ki}{M}\right) - \frac{1}{M}\sum_{i=1}^{M} g\left(\frac{i}{M}\right)\exp\left(-\frac{j2\pi ki}{M}\right)\right| \\
&= \frac{1}{M}\left|\sum_{i=1}^{M}\left[g(S_i) - g\left(\frac{i}{M}\right)\right]\exp\left(-\frac{j2\pi ki}{M}\right)\right| \\
&\leq \frac{1}{M}\sum_{i=1}^{M}\left|\left[g(S_i) - g\left(\frac{i}{M}\right)\right]\exp\left(-\frac{j2\pi ki}{M}\right)\right| \\
&= \frac{1}{M}\sum_{i=1}^{M}\left|g(S_i) - g\left(\frac{i}{M}\right)\right|
\end{aligned} \tag{69}$$

The third step follows from the triangle inequality. Squaring the expression obtained in the above equation,

$$
\begin{aligned}
\left|\hat{A}[k] - A_R[k]\right|^2 &= \frac{1}{M^2}\left\{\sum_{i=1}^{M}\left|g(S_i) - g\left(\frac{i}{M}\right)\right|\right\}^2 \\
&\le \frac{M}{M^2}\sum_{i=1}^{M}\left|g(S_i) - g\left(\frac{i}{M}\right)\right|^2 \\
&= \frac{1}{M}\sum_{i=1}^{M}\left|g(S_i) - g\left(\frac{i}{M}\right)\right|^2 \\
&\le \frac{1}{M}\sum_{i=1}^{M}||g'||_\infty^2\left|S_i - \frac{i}{M}\right|^2
\end{aligned} \tag{70}
$$

The second step follows from the inequality stated before, i.e., $\left(\sum_{i=1}^n a_i\right)^2 \le n\sum_{i=1}^n a_i^2$, for any numbers $a_1, a_2 \ldots a_n \in \mathbb{R}$. The last step uses the smoothness property of the field $g(x)$. For any smooth field $g(x)$ over a domain $\mathcal{D} \subset \mathbb{R}$ and any $x_1, x_2 \in \mathcal{D}$, $|g(x_1) - g(x_2)| \le ||g'|||_\infty|x_1 - x_2|$. This follows from the Lagrange's mean value theorem. To get the first term in (68), we need to take the expectations on either side in (70). Therefore,

$$
\mathbb{E}\left[\left|\hat{A}[k] - A_R[k]\right|^2\right] = ||g'||_\infty^2\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|S_i - \frac{i}{M}\right|^2\right] \tag{71}
$$

The expectation in the right hand side has been calculated in detail in Appendix C and from the result obtained there,

$$
\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|S_i - \frac{i}{M}\right|^2\right] \le \frac{C_0(1 - C_1\rho^n)}{n} \tag{72}
$$

which is independent of the distribution of the renewal process and only depends on the support parameter λ and C_0, C_1 are constants independent of n. The important part here is to note that the bound is guaranteed as the average sampling density, n, becomes large and more specifically if it is at least $\frac{\lambda}{1-\rho}\left(1 - \frac{2}{\ln\rho}\right)$. Note that since $0 < \rho < 1$, therefore, $\ln\rho < 0$. Thus, the bound is always a positive number. In fact, since $\ln x$ is an increasing function of x, so is this threshold value an increasing function of ρ. It is important to note that this is a sufficient condition and not a necessary condition. Also, this value is relatively small for mobile sensing setups. For examples, for $\rho = 0.9$ and $\lambda = 2$, this value is roughly about 400.

The other two terms obtained in equation (68) are exactly the same as the ones obtained in [11] and reproducing the bounds obtained there (Appendix B and equation (15)), we can write,

$$
\mathbb{E}\left[\left|A_R[k] - a[k]\right|^2\right] \le \mathbb{E}\left[\frac{16b^2\pi^2}{M^2}\right] \tag{73}
$$

$$
\mathbb{E}\left[\left|\hat{W}[k]\right|^2\right] \le \mathbb{E}\left[\frac{\sigma^2}{M}\right] \tag{74}
$$

Combining these results with the one obtained in equation (57), we get

$$
\mathbb{E}\left[\left|A_R[k] - a[k]\right|^2\right] \le \frac{16b^2\pi^2\lambda^2}{(n(1-\rho) - \lambda)^2} \tag{75}
$$

$$
\mathbb{E}\left[\left|\hat{W}[k]\right|^2\right] \le \frac{\sigma^2\lambda}{n(1-\rho) - \lambda} \tag{76}
$$

Putting together the results obtained in (50), (68), (72), (75),

$$
\begin{aligned}
\mathbb{E}\left[\left|\hat{A}_{\text{gen}}[k]-a[k]\right|^2\right] &\leq 4||g'||_\infty^2 \frac{C_0(1-C_1\rho^n)}{n} + \frac{64b^2\pi^2\lambda^2}{(n(1-\rho)-\lambda)^2} + \frac{2\sigma^2\lambda}{n(1-\rho)-\lambda} \\
&\leq 4(2b\pi)^2 \frac{C_0(1-C_1\rho^n)}{n} + \frac{64b^2\pi^2\lambda^2}{(n(1-\rho)-\lambda)^2} + \frac{2\sigma^2\lambda}{n(1-\rho)-\lambda} \\
&\leq \frac{C-C'\rho^n}{n} \qquad (77)
\end{aligned}
$$

as n becomes large. C, C' are positive constants independent of the sampling density n. These are mainly functions of the bandwidth parameter b, the support parameter λ, the 'correlation coefficient' ρ and the measurement noise σ^2.

This completes the proof of the main result in the theorem.

□

The dependence on the coefficient ρ that characterizes the autoregressive model is rather interesting and deserves special attention. On a broad scale, due to the autoregressive model, two things have significantly changed from the result shown in [11]. Firstly, it has resulted in a lower bound on the average sampling density n, that is sufficient to ensure the bound on the mean squared error. Even though it is not a necessary condition, the bound fails when $n \ll \frac{\lambda}{1-\rho}\left(1-\frac{2}{\ln\rho}\right)$. That is, if sampling density is made too small, the aberration is visible and the error does not go down as $1/n$. This can be seen in the simulations section of the paper. Moreover, this threshold gets larger as ρ gets closer to 1 and in fact is unbounded and shoots off to infinity when ρ is in the neighborhood of 1. Another factor is the presence of the term of $1-\rho^n$ in the numerator. This does not affect the asymptotic bounds on the error because as n becomes larger, ρ^n becomes smaller since $0 \leq \rho 1$.

Another thing that is worth noting that even though the stochastic process is such that n is expected sampling density, the *effective* sampling density is lesser due to the autoregressive model. Even though n has been termed as sampling density for the ease of understanding, it is just that the sampling density is of the order of n. The *actual* or *effective* sampling density is $n(1-\rho)$. Asymptotically, it cannot be made this is same as n, however, for finitely large n, a clear difference is noted when ρ is in the neighborhood of 1. Consider the case of $n = 10000$, which is a reasonably large number of samples and one expects almost perfect reconstruction at this sampling density. If ρ is close to 1, say, 0.99, the value of *effective* sampling density becomes about 100 which is not sufficient for a good reconstruction. Thus, because of this *effective* sampling densities, the error begins to decrease as $1/n$ from much larger values of n when ρ is close to 1 as opposed to when it is not.

3.3 Simulations

This section presents the results of simulations. The simulations have been presented in Fig. 2. The simulations have a large scope as they help confirm the results obtained, help in examining the effect of the 'correlation coefficient', ρ on the results and help in analyzing the effect of different renewal processes that characterize the stochastic term of the autoregressive model on the estimation error.

Firstly, for the purpose of simulations, a field $g(x)$ with $b = 3$ is considered and its Fourier coefficients have been generated using independent trials of a Uniform distribution over $[-1, 1]$ for all real and imaginary parts separately. To ensure that the field is real, conjugate symmetry is employed, i.e., $a[k] = \bar{a}[-k]$ and $a[0]$ is ensured that it is real. Finally, the field is scaled to have $|g(x)| \leq 1$. The following coefficients were obtained.

$$
\begin{aligned}
a[0] = 0.3002;\ a[1] = -0.04131 + j0.0216; \\
a[2] = 0.0871 + j0.0343;\ a[3] = -0.1679 - j0.0586; \qquad (78)
\end{aligned}
$$

The distortion was estimated using Monte Carlo simulation with 10000 trials.

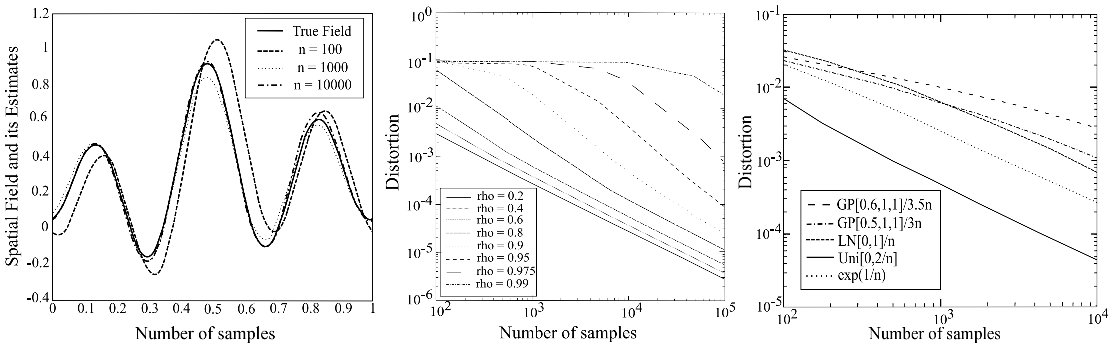

Figure 2: Fig. 2. (a) The convergence of random realizations of $\widehat{G}(x)$ to $g(x)$ is illustrated for the field whose Fourier coefficients have been mentioned above. The field estimate for $n = 10000$ almost converges to the true value of the field (and is not visible in the plot). (b) The distortion scales as $O(1/n)$ with additive Gaussian Noise of variance 0.125 and Uniform$[0, 2/n]$ renewal process distribution. Note the variation with ρ. For small values of ρ, a slope of -1 is distinctly noted. However, as the value of ρ increase, we see that the slope does not initially follow the trend. This can be accounted for in two ways. One being that at large values of ρ, the threshold is not being satisfied at these sampling densities and the other is the *effective* sampling rate. Note that how the slopes are eventually getting close to -1 as the sampling density is increased. It be shown that at very large sampling densities the plot will behave as expected.

Field estimates were obtained in various cases using different distributions for the renewal process and the measurement noise distributions. Uniform distribution between a and b is denoted as Uniform$[a, b]$. The $\exp(\mu)$ denotes an exponential distribution with mean $1/\mu$. $\mathcal{N}(\mu, s)$ denotes a Gaussian distribution with mean μ and variance s. The distribution LN$[m, s]$ represents a log-normal distribution where m is the mean and s is the variance of the underlying Gaussian.

In Fig.2(a) the measurement noise was generated using $\mathcal{N}(0, 0.125)$. The autoregressive model that was employed had $\rho = 0.5$ and the stochastic part was generated using Uni$[0, 2/n]$. Therefore, $\lambda = 2$. The method for reconstruction in (65) is agnostic to this distribution though. The figure has random realizations of the estimated field $\hat{G}(x)$ along with the true field, $g(x)$. It can be clearly seen that as the value of n increases the estimated field gets closer to the actual field.

The Fig.2(b) has the plots of the mean squared error between the field and its estimate. For the plots, the noise model and stochastic part of the autoregressive model is same as in above paragraph. However, the plots have been plotted for different values of ρ in the range 0.2 to 0.99. Note that the illustrated plots are log-log plots.

The Fig.2(c) has plots for different renewal processes for the stochastic part of the autoregressive process. For heavy-tailed, infinite support distributions like the General Pareto distribution, the error does not decrease as $1/n$. For the other finite support distributions, the error follows the $O(1/n)$ decrease.

4 Conclusion

The problem of sampling spatially bandlimited fields was studied with a practical viewpoint. Firstly, the sampling of spatially bandlimited field evolving according to the constant coefficient linear partial differential equation using a mobile sensor was studied. The field was estimated using the noisy samples obtained at unknown locations and time instants obtained from two independent and unknown renewal processes and it was shown that the mean squared error between the estimated field and the true field decreased as $O(1/n)$, where n was the average sampling density. The main idea that was leveraged was the fact that the locations of the samples got closer to the ones corresponding to uniform sampling as the sampling density increased and thus oversampling was used to reduce the error.

Secondly, the problem of sampling with unknown sampling locations obtained from an autoregressive model on the intersample distances was studied. The field was estimated using the noisy samples and the mean squared error between the estimated and true field was shown to decrease as $O(1/n)$, where n was the average sampling density.

The most important part was that two interrelated problems which arise in the case of practical modeling of sampling were studied separately. However, the results obtained are such that they can be combined together to form a single result that states that if a spatiotemporally varying field is sampled with location and time unaware sensors such that intersample distances come from an autoregressive process, it is possible to reconstruct the field with an error of $O(1/n)$, where n is the average sampling density. This is enabled by the separability of the spatial and the temporal part in the analysis of the field.

Appendix A

This appendix mainly deals with proving the bandlimitedness of the field for all $t \geq 0$ given that the field was spatially bandlimited at $t = 0$. Spatial bandlimitednes, as defined previously, refers to the fact that the Fourier series of the field over its spatial support has finite number of terms.The main result shown is that for a field evolving according to the equation (3), if it is known that the field and its $m-1$, temporal derivatives are spatially bandlimited at $t = 0$, then the field will always remain bandlimited. Here m is the degree of the polynomial p. A major part of this proof will be similar to the approach in Section 2.2 and has been reproduced here for ease of understanding. Since the field is assumed to have a finite support, the field can be written as

$$g(x,t) = \sum_{k=\infty}^{\infty} a_k(t) \exp(j2\pi kx)) \ ; \ a_k(t) = \int_{-\infty}^{\infty} g(x,t) \exp(-j2\pi kx) dx \tag{79}$$

Substituting this in the equation (1), and using (2) along with the orthogonality property for the Fourier basis we can write,

$$\begin{aligned}
\sum_{i=0}^{m} p_i \frac{\partial^i}{\partial t^i} \left(\sum_{k=-\infty}^{\infty} a_k(t) \exp(j2\pi kx) \right) &= \sum_{i=0}^{n} q_i \frac{\partial^i}{\partial x^i} \left(\sum_{k=-\infty}^{\infty} a_k(t) \exp(j2\pi kx) \right) \\
\implies \sum_{k=-\infty}^{\infty} \left(\sum_{i=0}^{m} p_i \frac{\partial^i a_k(t)}{\partial t^i} \right) \exp(j2\pi kx) &= \sum_{k=-\infty}^{\infty} a_k(t) \left(\sum_{i=0}^{n} q_i (j2\pi k)^i \right) \exp(j2\pi kx) \\
\overset{(a)}{\implies} \sum_{k=-\infty}^{\infty} \left(\sum_{i=0}^{m} p_i \frac{\partial^i a_k(t)}{\partial t^i} \right) \exp(j2\pi kx) &= \sum_{k=-\infty}^{\infty} a_k(t) q(j2\pi k) \exp(j2\pi kx) \\
\overset{(b)}{\implies} \sum_{i=0}^{m} p_i \frac{\partial^i a_k(t)}{\partial t^i} - q(j2\pi k) a_k(t) &= 0 \ \ \forall k \in \mathbb{Z}
\end{aligned} \tag{80}$$

Essentially, each $a_k(t)$ evolves by an ordinary differential equation (ODE) with constant coefficients. The solution of ODE with constant coefficients is well known (via unilateral Laplace transform [20]). For each k, this leads to the polynomial equation,

$$\begin{aligned}
\sum_{i=0}^{m} p_i \frac{\partial^i A e^{rt}}{\partial t^i} - q(j2\pi k) A e^{rt} &= 0 \\
\implies \left(\sum_{i=0}^{m} p_i r^i - q(j2\pi k) \right) A e^{rt} &= 0 \\
\implies p(r) - q(j2\pi k) &= 0
\end{aligned} \tag{81}$$

The solution for $a_k(t)$ is a of the form Ae^{rt}, where r is the root of the above polynomial and A is a constant independent of t. Let the roots of the above polynomial be $r_1(k), r_2(k), \ldots r_m(k)$. Note that the roots of the polynomial are indexed by k as well, implying there is a set of m roots for each value of k. Furthermore for simplicity of analysis, all of the roots $r_1(k), r_2(k), \ldots r_m(k)$ are considered to be distinct for a given k as in Section 2.2. However, it is possible that $r_i(k_1) = r_j(k_2)$ for some $i, j, k_1 \neq k_2$. In fact, the solution for $a_k(t)$ can thus be written as a linear combination of these roots. Thus $a_k(t) = \sum_{i=1}^{m} a_{ki}(0) \exp(r_i(k)t)$. The coefficients have been chosen to maintain consistency of represent as a function of time. Now, we know that the field is bandlimited and so are its $m-1$ partial derivatives at $t = 0$. This can be written for $i = 0, 1, 2 \ldots m-1$

$$\left. \frac{\partial^i a_k(t)}{\partial t^i} \right|_{t=0} = \begin{cases} c_{ki} & \text{if } |k| \leq b \\ 0 & \text{otherwise} \end{cases} \tag{82}$$

where c_{ki}'s are real constants. Since $a_k(t) = \sum_{i=1}^{m} a_{ki}(0) \exp(r_i(k)t)$, therefore,

$$\frac{\partial^j a_k(t)}{\partial t^j} = \sum_{i=1}^{m} a_{ki}(0) r_i^j(k) \exp(r_i(k)t) \ \forall \ j \geq 0 \tag{83}$$

Consider k' in the range $|k| > b$. Then for k' and $\forall\ i = 0, 1, 2 \ldots m$, we have $\left.\frac{\partial^i a_{k'}(t)}{\partial t^i}\right|_{t=0} = 0$. Then for k' using equation (83), we can combine all the equations for all i and the resulting expression can be written in matrix form as,

$$\begin{bmatrix} 1 & 1 & \ldots & 1 \\ r_1(k') & r_2(k') & \ldots & r_m(k') \\ r_1^2(k') & r_2^2(k') & \ddots & r_m^2(k') \\ \vdots & \vdots & \ddots & \vdots \\ r_1^m(k') & r_2^m(k') & \ldots & r_m^m(k') \end{bmatrix} \begin{bmatrix} a_{k'1} \\ a_{k'2} \\ \vdots \\ a_{k'm} \end{bmatrix} = 0 \tag{84}$$

Since the roots are assumed to be distinct, therefore the matrix on the left is a Vandermonde matrix and is always invertible[23]. This means that $a_{k'i} = 0\ \forall i = 1, 2, 3 \ldots m$ and all $|k'| > b$. This implies that the field is bandlimited, ie, $a_k(t) \equiv 0$ for all $|k| > b$.

Appendix B

This section primarily deals with establishing upper bound on the terms $\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|S_i - \frac{i}{M}\right|^2\right]$ and $\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|T_i - \frac{iT_0}{M}\right|^2\right]$ in the case of independent samples. The renewal based sampling model for the spatial terms is the same that has been considered in [11]. Moreover, the bound on the term $\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|S_i - \frac{i}{M}\right|^2\right]$ has been elaborately derived there([11], Appendix A). Using that we have the bound,

$$\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|S_i - \frac{i}{M}\right|^2\right] \le \frac{C_S}{n} \tag{85}$$

for some $C_S > 0$ and independent of n. The proof for the other term follows in a similar manner. Define,

$$\begin{aligned} y_m &:= \mathbb{E}\left[\left(N_1 - \frac{T_0}{M}\right)^2 \middle| M = m\right] \\ z_m &:= \mathbb{E}\left[\left(N_1 - \frac{T_0}{M}\right)\left(N_2 - \frac{T_0}{M}\right) \middle| M = m\right] \end{aligned} \tag{86}$$

From equation (31) in [11], we have

$$\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|T_i - \frac{iT_0}{M}\right|^2 \middle| M = m\right] = \frac{m+1}{2}y_m + \frac{m^2-1}{3}z_m \tag{87}$$

Consider the expression,

$$\begin{aligned} \mathbb{E}\big[(T_M - T_0)^2 | M = m\big] &= \mathbb{E}\left[\left\{\sum_{i=1}^{M}\left(N_i - \frac{T_0}{M}\right)\right\}^2 \middle| M = m\right] \\ &= m y_m + m(m-1) z_m \end{aligned} \tag{88}$$

where the second step is obtained from evaluating and rearranging the expression along the exchangability of N_i's. Since we know that, $T_M \le T_0$ and $T_{M+1} > T_0$, define

$$J_M = T_0 - T_M \implies J_M < T_{M+1} - T_M \le \frac{\mu}{n} \tag{89}$$

Also $J_M^2 = (T_M - T_0)^2 \implies \mathbb{E}\big[(T_M - T_0)^2 | M = m\big] = \mathbb{E}[J_M^2 | M = m] \le \frac{\mu^2}{n^2}$. Using the above two results, we can conclude that, $m y_m + m(m-1) z_m = \mathbb{E}[J_M^2 | M = m]$. Combining this with (87), we can write,

$$\begin{aligned} \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|T_i - \frac{iT_0}{M}\right|^2 \middle| M = m\right] &= \frac{m+1}{2}y_m + \frac{m^2-1}{3m(m-1)}\big(-m y_m + \mathbb{E}[J_M^2 | M = m]\big) \\ &= \frac{m+1}{2}y_m + \frac{m+1}{3m}\mathbb{E}[J_M^2 | M = m] \\ &\le \frac{m+1}{2}y_m + \frac{2}{3}\frac{\mu^2}{n^2} \end{aligned} \tag{90}$$

This is exactly the same result as obtained for $\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|S_i - \frac{i}{M}\right|^2\right]$ in [11]. Using the same steps for the expression in S_i, we can conclude,

$$\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|T_i - \frac{iT_0}{M}\right|^2 \middle| M = m\right] \le \frac{C_T}{n} \tag{91}$$

An important thing to note here is that even if $J_M \leq \frac{K\mu}{n}$ for some positive constant K, the result will hold. Interestingly, K, can be $O(\sqrt{n})$, and still the result will hold. The idea is that the difference between T_0 and T_M should of $O(1/\sqrt{n})$, that is T_0 cannot be simply any number larger than T_M. It has to be a reasonably accurate estimation of the time taken. It is important to know this to help decide the construction of Y_0 because the entire idea is based on the assumption that the samples are "near"the grid points. But to determine the grid points, we must have the knowledge of the support of the function which has to be finite.

Appendix C

This section will mainly be elaborating the bound on $\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|S_i - \frac{i}{M}\right|^2\right]$ for the case when the samples come from an autoregressive process. To begin with consider the remainder term R_M, which is the distance remaining after M samples. Therefore, $R_M = 1 - S_M$. An upper bound on this will be established as this as this will be of use later.

$$R_M = 1 - S_M < S_{M+1} - S_M = \sum_{i=1}^{M+1} \rho^{M-i+1} Y_i \leq \sum_{i=0}^{\infty} \rho^i \frac{\lambda}{n} = \frac{\lambda}{n(1-\rho)}$$

Thus, we have,

$$R_M \leq \frac{\lambda}{n(1-\rho)} \tag{92}$$

The main expresstion that this Appendix will be dealing with is the upper bound on the mean square error between the sampling locations and the uniform grid locations. The expression is given by

$$\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|S_i - \frac{i}{M}\right|^2\right] = \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|\sum_{k=1}^{i} X_k - \frac{i}{M}\right|^2\right] = \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left\{\sum_{k=1}^{i}\left(X_k - \frac{1}{M}\right)\right\}^2\right]$$

To ease out the calculations, a conditional version of this expectation will be dealt with that is, for $M = m$ which is given by $\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left\{\sum_{k=1}^{i}\left(X_k - \frac{1}{M}\right)\right\}^2 \middle| M = m\right]$ and then later the assumption will be relaxed for the final results. Define the following,

$$b(i) := \mathbb{E}\left[\left(X_i - \frac{1}{m}\right) \middle| M = m\right] = \sum_{r=1}^{i} \rho^{i-r}\mathbb{E}[Y_r | M = m] - \frac{1}{m} \tag{93}$$

$$c(i) := \mathrm{Var}\left[\left(X_i - \frac{1}{m}\right) \middle| M = m\right] = \mathrm{Var}[X_i | M = m] \tag{94}$$

Since we know that for a random variable X, $\mathbb{E}[X^2] = \mathrm{Var}[X] + \mathbb{E}[X]^2$. Therefore, we can write,

$$\mathbb{E}\left[\left(X_i - \frac{1}{m}\right)^2 \middle| M = m\right] = c(i) + b^2(i) \tag{95}$$

Consider,

$$
\begin{aligned}
\mathbb{E}[X_iX_j|M=m] &= \mathbb{E}\left[\left(\sum_{k=1}^{i}\rho^{i-k}Y_k\right)\left(\sum_{l=1}^{j}\rho^{j-l}Y_l\right)\middle|M=m\right] \\
&= \mathbb{E}\left[\left(\sum_{k=1}^{i}\rho^{i-k}\sum_{l=1}^{j}\rho^{j-l}Y_kY_l\right)\middle|M=m\right] \\
&= \sum_{k=1}^{i}\sum_{l=1}^{j}\rho^{i-k}\rho^{j-l}\mathbb{E}[Y_kY_l|M=m] \\
&= \sum_{l=1}^{\min\{i,j\}}\rho^{i+j-2l}\mathbb{E}[Y_l^2|M=m] + 2\sum_{k=1}^{i}\sum_{l=k+1}^{j}\rho^{i-k+j-l}\ \mathbb{E}[Y_k|M=m]\mathbb{E}[Y_l|M=m] \\
&= \sum_{l=1}^{\min\{i,j\}}\rho^{i+j-2l}\mathrm{Var}[Y_l|M=m] + \sum_{k=1}^{i}\sum_{l=1}^{j}\rho^{i-k+j-l}\mathbb{E}[Y_k|M=m]\mathbb{E}[Y_l|M=m] \\
&= \rho^{|i-j|}\sum_{l=1}^{a}\rho^{2(a-l)}\mathrm{Var}[Y_1|M=m] + \left(b(i)+\frac{1}{m}\right)\left(b(j)+\frac{1}{m}\right) \\
&= \rho^{|i-j|}c(a) + \left(b(i)+\frac{1}{m}\right)\left(b(j)+\frac{1}{m}\right) \qquad (96)
\end{aligned}
$$

where $a=\min\{i,j\}$ and $b(i)$ and $b(j)$ have been defined in (93).

$$
\begin{aligned}
&\mathbb{E}\left[\left(X_i-\frac{1}{m}\right)\left(X_j-\frac{1}{m}\right)\middle|M=m\right] \\
&= \mathbb{E}[X_iX_j|M=m] - \frac{1}{m}(\mathbb{E}[X_i|M=m]+\mathbb{E}[X_j|M=m]) + \frac{1}{m^2} \\
&= \rho^{|i-j|}c(a) + \left(b(i)+\frac{1}{m}\right)\left(b(j)+\frac{1}{m}\right) - \frac{1}{m}\left(b(i)+\frac{1}{m}+b(j)+\frac{1}{m}\right) + \frac{1}{m^2} \\
&= \rho^{|i-j|}c(\min\{i,j\}) + b(i)b(j) \qquad (97)
\end{aligned}
$$

Define for convenience of notation,

$$
T_m = \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|S_i-\frac{i}{M}\right|^2\right]\ ;\ Z_k = \left(X_k-\frac{1}{m}\right) \qquad (98)
$$

Therefore we can write,

$$
\begin{aligned}
T_m &= \mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left\{\sum_{k=1}^{i}\left(X_k-\frac{1}{M}\right)\right\}^2\middle|M=m\right] = \mathbb{E}\left[\frac{1}{m}\sum_{i=1}^{m}\left\{\sum_{k=1}^{i}\left(X_k-\frac{1}{m}\right)\right\}^2\middle|M=m\right] \\
&= \mathbb{E}\left[\frac{1}{m}\sum_{i=1}^{m}\left\{\sum_{k=1}^{i}Z_k\right\}^2\middle|M=m\right] = \mathbb{E}\left[\frac{1}{m}\sum_{i=1}^{m}(Z_1+Z_2+\dots Z_i)^2\middle|M=m\right] \\
&= \mathbb{E}\Bigg[\frac{1}{m}\big(mZ_1^2+(m-1)Z_2^2+\cdots+Z_m^2+2\{(m-1)Z_1Z_2+(m-2)Z_1Z_3+\cdots+Z_1Z_m \\
&\quad +(m-2)Z_2Z_3+(m-3)Z_2Z_4+\dots Z_2Z_4+\cdots+Z_{m-1}Z_m\}\big)\Bigg|M=m\Bigg] \\
&= \mathbb{E}\Bigg[\frac{1}{m}\big(2\{mZ_1^2+(m-1)Z_2^2+\cdots+Z_m^2+(m-1)Z_1Z_2+(m-2)Z_1Z_3+\cdots+Z_1Z_m+ \\
&\quad (m-2)Z_2Z_3+(m-3)Z_2Z_4+\cdots+Z_{m-1}Z_m\}-\{mZ_1^2+(m-1)Z_2^2+\cdots+Z_m^2\}\big)\Bigg|M=m\Bigg]
\end{aligned}
$$

The similar terms can be grouped together to form a simpler expression like,

$$\begin{aligned}T_m = \mathbb{E}\Bigg[\frac{1}{m}\big(2\{mZ_1^2 + (m-1)Z_1Z_2 + (m-2)Z_1Z_3 + \cdots + Z_1Z_m +\\ (m-1)Z_2^2 + (m-2)Z_2Z_3 + (m-3)Z_2Z_4 + \ldots Z_2Z_m + \ldots\\ +2Z_{m-1}^2 + \cdots + Z_{m-1}Z_m + Z_m^2\} - \{mZ_1^2 + (m-1)Z_2^2 + \cdots + Z_m^2\}\big)\Bigg| M = m\Bigg]\end{aligned}$$

Using the expressions from (93) and (97) and noting that $Z_i = X_i - \frac{1}{m}$, the above expression can be rewritten as,

$$\begin{aligned}T_m = \frac{2}{m}\Bigg\{\sum_{r=0}^{m-1}[(m-r)(\rho^r c(1) + b(1)b(r+1))] + \sum_{r=0}^{m-1}[(m-1-r)(\rho^r c(2) + b(2)b(r+2))] + \ldots\Bigg\}\\ -\frac{1}{m}\Bigg\{m[c(1)+b^2(1)] + (m-1)[c(2)+b^2(2)] + \ldots\Bigg\}\end{aligned}$$

This can be condensed to be written as,

$$T_m = \frac{2}{m}\sum_{i=1}^{m}\sum_{r=0}^{m-i}[(m-r-i+1)(\rho^r c(i) + b(i)b(r+i))] - \frac{1}{m}\sum_{i=1}^{m}(m-i+1)[c(i)+b^2(i)] \quad (99)$$

Inspired by the above equation (99), define the following,

$$\begin{aligned}k_m &:= 2\sum_{i=1}^{m}\sum_{r=0}^{m-i}\rho^r c(i) + b(i)b(r+i)\\ l_m &:= \sum_{i=1}^{m}c(i) + b^2(i) = \sum_{i=1}^{m}\mathbb{E}\left[\left(X_i - \frac{1}{m}\right)^2\Bigg| M = m\right]\end{aligned} \quad (100)$$

The last equality follows from (93). The following inequalities are noted.

$$\begin{aligned}2\sum_{i=1}^{m}\sum_{r=0}^{m-i}[(m-r-i+1)(\rho^r c(i) + b(i)b(r+i))] &\leq 2\sum_{i=1}^{m}\sum_{r=0}^{m-i}m(\rho^r c(i) + b(i)b(r+i)) = mk_m\\ \sum_{i=1}^{m}(m-i+1)[c(i)+b^2(i)] &\geq \sum_{i=1}^{m}c(i) + b^2(i) = l_m\end{aligned} \quad (101)$$

Combining the expressions obtained in equations (99), (100) and (101), one can conclude $T_m \leq \frac{1}{m}(mk_m - l_m)$. Again consider the remainder term, $R_M = 1 - S_M$. Squaring both sides and taking expectations, we get,

$$\mathbb{E}[(S_M - 1)^2 | M = m] = \mathbb{E}[R_M^2 | M = m] \quad (102)$$

For the left hand side term in the above equation,

$$\begin{aligned}\mathbb{E}[(S_M - 1)^2 | M = m] &= \mathbb{E}\left[\left\{\sum_{k=1}^{m}\left(X_k - \frac{1}{m}\right)\right\}^2\Bigg| M = m\right]\\ &= \mathbb{E}\left[\left(\sum_{k=1}^{m}Z_k\right)^2\Bigg| M = m\right]\\ &= \mathbb{E}\left[Z_1^2 + Z_2^2 + \cdots + Z_m^2 + 2\{Z_1Z_2 + Z_1Z_3 + \cdots + Z_{m-1}Z_m\}\Big| M = m\right]\\ &= \mathbb{E}\left[2\{Z_1^2 + Z_1Z_2 + \cdots + Z_{m-1}Z_m\} - \{Z_1^2 + Z_2^2 \cdots + Z_m^2\}\Big| M = m\right]\\ &= 2\sum_{i=1}^{m}\sum_{r=0}^{m-i}(\rho^r c(i) + b(i)b(r+i)) - \sum_{i=1}^{m}(c(i) + b^2(i))\\ &= k_m - l_m\end{aligned} \quad (103)$$

For the term on the right hand side in (102),

$$R_M \leq \frac{\lambda}{n(1-\rho)} \implies \mathbb{E}[R_M^2 | M = m] \leq \frac{\lambda^2}{n^2(1-\rho)^2} \tag{104}$$

The equations (102), (103) and (104) can be combined to give,

$$k_m - l_m \leq \frac{\lambda^2}{n^2(1-\rho)^2} \implies k_m \leq \frac{\lambda^2}{n^2(1-\rho)^2} + l_m \tag{105}$$

Using the fact that $T_m \leq \frac{1}{m}(mk_m - l_m)$ and above inequality, the upper bound on T_m now becomes,

$$T_m \leq \frac{1}{m}(mk_m - l_m) = k_m - \frac{l_m}{m} \leq \frac{\lambda^2}{n^2(1-\rho)^2} + l_m - \frac{l_m}{m} = \frac{\lambda^2}{n^2(1-\rho)^2} + l_m \left(1 - \frac{1}{m}\right) \tag{106}$$

Finally, taking off the conditional expectation and using expressions from (98), (100) and (106), one can write,

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|S_i - \frac{i}{M}\right|^2\right] &\leq \frac{\lambda^2}{n^2(1-\rho)^2} + \mathbb{E}\left[\left(\frac{M-1}{M}\right)\left\{\sum_{i=1}^{M}\left(X_i - \frac{1}{M}\right)^2\right\}\right] \\
&\overset{(a)}{\leq} \frac{\lambda^2}{n^2(1-\rho)^2} + \mathbb{E}\left[\left\{\sum_{i=1}^{M}\left(X_i - \frac{1}{M}\right)^2\right\}\right] \\
&\overset{(b)}{\leq} \frac{\lambda^2}{n^2(1-\rho)^2} + \mathbb{E}\left[2\sum_{i=1}^{M}\left(X_i^2 + \frac{1}{M^2}\right)\right] \\
&\overset{(c)}{=} \frac{\lambda^2}{n^2(1-\rho)^2} + \mathbb{E}\left[2\sum_{i=1}^{M}X_i^2\right] + \mathbb{E}\left[\frac{2}{M}\right] \\
&\overset{(d)}{\leq} \frac{\lambda^2}{n^2(1-\rho)^2} + 2\mathbb{E}\left[\sum_{i=1}^{M}X_i^2\right] + \frac{2\lambda}{n(1-\rho)-\lambda}
\end{aligned} \tag{107}$$

where (a) follows from the fact that $\frac{M-1}{M} < 1$, (b) is a direct application of Cauchy Schwarz, (c) follows from linearity of expectation and (d) uses the following result obtained from (57) which can be restated as $\frac{1}{M} < \frac{\lambda}{n(1-\rho)-\lambda}$. Hence,

$$\mathbb{E}\left[\frac{1}{M}\right] < \frac{\lambda}{n(1-\rho)-\lambda}$$

The proof will be complete if the term $\mathbb{E}\left[\sum_{i=1}^{M}X_i^2\right]$ is upper bounded by a term of order $1/n$. For that purpose, consider

$$X_i = \sum_{r=1}^{i}\rho^{i-r}Y_r \leq \sum_{r=1}^{i}\rho^{i-r}\frac{\lambda}{n} = \frac{\lambda}{n}\left(\frac{1-\rho^i}{1-\rho}\right)$$

This implies, $X_i^2 \leq \frac{\lambda^2}{n^2}\left(\frac{1-\rho^i}{1-\rho}\right)^2$. Taking expectations,

$$\mathbb{E}\left[\sum_{i=1}^{M}X_i^2\right] \leq \mathbb{E}\left[\sum_{i=1}^{M}\frac{\lambda^2}{n^2}\left(\frac{1-\rho^i}{1-\rho}\right)^2\right] = \frac{\lambda^2}{(n(1-\rho))^2}\mathbb{E}\left[\sum_{i=1}^{M}(1-\rho^i)^2\right] \tag{108}$$

Since, $0 < \rho < 1$, therefore, $(1-\rho^i)^2 < (1-\rho^i)\ \forall\ i \geq 1$ and $\rho^i \geq \rho^j\ \forall\ j \geq i$. Hence,

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{M} X_i^2\right] &\leq \frac{\lambda^2}{(n(1-\rho))^2}\mathbb{E}\left[\sum_{i=1}^{M}(1-\rho^i)^2\right] \\
&\leq \frac{\lambda^2}{(n(1-\rho))^2}\mathbb{E}\left[\sum_{i=1}^{M}(1-\rho^i)\right] = \frac{\lambda^2}{(n(1-\rho))^2}\left(\mathbb{E}[M] - \mathbb{E}\left[\sum_{i=1}^{M}\rho^i\right]\right) \\
&\leq \frac{\lambda^2}{(n(1-\rho))^2}\left(\mathbb{E}[M] - \mathbb{E}\left[\sum_{i=1}^{M}\rho^M\right]\right) = \frac{\lambda^2}{(n(1-\rho))^2}\left(\mathbb{E}[M] - \mathbb{E}\left[M\rho^M\right]\right) \qquad (109)
\end{aligned}$$

To establish bounds on $\mathbb{E}\left[M\rho^M\right]$, consider the function $f : \mathbb{R} \to \mathbb{R}$, given by $f(x) = x\rho^x$, where $0 < \rho < 1$ is a finite constant independent of x. The function is differentiable and its second derivative is given by $f^{''}(x) = \ln\rho(2 + x\ln\rho)\rho^x$. This is positive as long as $x > -\frac{2}{\ln\rho}$. Thus the in region defined by $x > -\frac{2}{\ln\rho}$, the function is convex. Since the function is convex in that interval, we can apply the Jensen's inequality. Jensen's inequality states for any function $h(x)$ that is convex over an interval $\mathcal{I} \subset \mathcal{D}$ where $\mathcal{D}$ is the domain of $h(x)$, and a random variable X whose support is a subset of $\mathcal{I}$, the inequality $\mathbb{E}[h(X)] \geq h(\mathbb{E}[X])$ holds true.

We know that f is convex for $x > -\frac{2}{\ln\rho}$ and from (57), we know that $M > \frac{n(1-\rho)}{\lambda} - 1$. Therefore if $M > -\frac{2}{\ln\rho}$, then using Jensen's inequality, we can say that $\mathbb{E}\left[M\rho^M\right] \geq \mathbb{E}[M]\rho^{\mathbb{E}[M]}$. The condition on M will be always true if $\dfrac{n(1-\rho)}{\lambda} - 1 > -\dfrac{2}{\ln\rho}$. That is if

$$n > \frac{\lambda}{1-\rho}\left(1 - \frac{2}{\ln\rho}\right). \qquad (110)$$

Therefore, if n is large enough, we can write from (109),

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{M} X_i^2\right] &\leq \frac{\lambda^2}{(n(1-\rho))^2}\left(\mathbb{E}[M] - \mathbb{E}\left[M\rho^M\right]\right) \\
&\leq \frac{\lambda^2}{(n(1-\rho))^2}\left(\mathbb{E}[M] - \mathbb{E}[M]\rho^{\mathbb{E}[M]}\right) = \frac{\lambda^2}{(n(1-\rho))^2}\left(\mathbb{E}[M](1-\rho^{\mathbb{E}[M]})\right) \\
&\overset{(a)}{\leq} \frac{\lambda^2}{(n(1-\rho))^2}\left(n + \frac{(1-\rho)}{\lambda}\right)\left(1 - \rho^{n+\frac{1}{\lambda}}\right) \qquad (111)
\end{aligned}$$

where (a) follows from (64) and decreasing nature of ρ^x. From (64), we have $\mathbb{E}[M] \leq n + \frac{(1-\rho)}{\lambda} - 1 \leq n + \frac{(1-\rho)}{\lambda}$ and thus $1 - \rho^{\mathbb{E}[M]} \leq 1 - \rho^{n+\frac{(1-\rho)}{\lambda}} \leq 1 - \rho^{n+\frac{1}{\lambda}}$

Using (107), (109), (110) and (111), we can write

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{M}\sum_{i=1}^{M}\left|S_i - \frac{i}{M}\right|^2\right] &\leq \frac{\lambda^2}{n^2(1-\rho)^2} + 2\frac{\lambda^2}{(n(1-\rho))^2}\left(n + \frac{(1-\rho)}{\lambda}\right)\left(1-\rho^{n+\frac{1}{\lambda}}\right) + \frac{2\lambda}{n(1-\rho)-\lambda} \\
&\leq \frac{2\lambda^2(1-\rho^{n+\frac{1}{\lambda}})}{(1-\rho)^2}\frac{1}{n} + \frac{2\lambda}{n(1-\rho)-\lambda} + \frac{\lambda^2}{n^2(1-\rho)^2}\left(1 + \frac{2}{\lambda}\right) \\
&\leq \frac{C_0(1 - C_1\rho^n)}{n} \qquad (112)
\end{aligned}$$

for some positive constants C_0, C_1 independent of n. This proves the upper bound. It is very essential to note that this bound holds surely under the condition that $n > \dfrac{\lambda}{1-\rho}\left(1 - \dfrac{2}{\ln\rho}\right)$. Also, that is condition is a sufficient one, but not necessary. This completes the proof.

References

[1] R. Marks, *Introduction to Shannon Sampling and Interpolation Theory*, Springer-Verlag, New York, 1991.

[2] R. Marks, *Advanced Topics in Shannon Sampling and Interpolation Theory*, Springer-Verlag, New York, 1993.

[3] Y. M. Lu J. Ranieri, A. Chebira and M. Vetterli, "Sampling and reconstructing diffusion fields with localized sources," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4016–4019, 2011.

[4] Jayakrishnan Unnikrishnan and Martin Vetterli, "Sampling and reconstruction of spatial fields using mobile sensors," *IEEE Trans. Signal Proc.*, vol. 61, no. 9, pp. 2328–2340, May 2013.

[5] Jayakrishnan Unnikrishnan and Martin Vetterli, "Sampling high-dimensional bandlimited fields on low-dimensional manifolds," *IEEE Trans. Info. Theory*, vol. 59, no. 4, pp. 2103–2127, 2013.

[6] Jayakrishnan Unnikrishnan and Martin Vetterli, "On optimal sampling trajectories for mobile sensing," in *Proceedings of the 10th International Conference on Sampling Theory and Applications*, 2013, pp. 352–355.

[7] A. Nehorai, B. Porat, and E. Paldi, "Detection and localization of vapor-emitting sources," *IEEE Trans. Signal Process.*, vol. 43, pp. 243–253, 1995.

[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.*, 2009, pp. 2249–2252.

[9] Y M Lu and M Vetterli, "Distributed spatio-temporal sampling of diffusion fields from sparse instantaneous sources," *Proc. 3rd Int. Workshop on Comp Adv. in Multi-Sensor Adaptive Proc*, 2009.

[10] A Chebira J Ranieri, I Dokmanic and M Vetterli, "Sampling and reconstruction of time-varying atmospheric emissions," *IEEE Int. Conf. on Acoustics,Speech and Signal Proc.*, pp. 3673–3676, 2012.

[11] A. Kumar, "On bandlimited field estimation from samples recorded by a location-unaware mobile sensor," *IEEE Trans. Inf. Theory*, vol. 63, pp. 2188–2200, Apr. 2017.

[12] P. Marziliano and M. Vetterli, "Reconstruction of irregularly sampled discrete-time bandlimited signals with unknown sampling locations," *IEEE Trans. Signal Process.*, vol. 48, pp. 3462–3471, Dec. 2000.

[13] J. Browning, "Approximating signals from nonuniform continuous time samples at unknown locations," *IEEE Trans. Signal Process.*, vol. 55, pp. 1549–1554, Apr. 2007.

[14] C. F. Chiasserini A. Nordio and E. Viterbo, "Performance of linear field reconstruction techniques with noise and uncertain sensor locations," *IEEE Trans. Signal Process.*, vol. 56, pp. 3535–3547, Aug. 2008.

[15] A. Papoulis, "Error analysis in sampling theory," *Proc. IEEE*, vol. 54, pp. 947–955, Jul. 1966.

[16] A. I. Zayed, *Advances in Shannon's Sampling Theory*, CRC Press, Boca Raton, FL, USA, 1993.

[17] A. Mallick and A. Kumar, "Bandlimited field reconstruction from samples obtained on a discrete grid with unknown random locations," *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, pp. 4014–4018, Mar. 2016.

[18] A. Kumar, "On bandlimited signal reconstruction from the distribution of unknown sampling locations," *IEEE Trans. Signal Process.*, vol. 63, pp. 1259–1267, Mar. 2015.

[19] R. Durrett, *Probability: Theory and Examples*, Duxbury Press, Belmont, CA, USA, 2nd edition, 1996.

[20] Alan V. Oppenheim, Alan S. Willsky, and S. Hamid Nawab, *Signals and Systems*, Prentice Hall, USA, 2nd edition, 1996.

[21] Norman S. Nise, *Control Systems Engineering*, John Wiley & Sons, Inc., New York, NY, USA, 3rd edition, 2000.

[22] John Conway, *Functions of One Complex Variable*, Springer-Verlag New York, 1978.

[23] W. Gautschi, "On inverses of vandermonde and confluent vandermonde matrices," *Numerische Mathematik*, vol. 4, no. 1, pp. 117–123, 1962.

[24] G. Polya and G. Szego, "Aufgaben und lehrs atze aus der analysis," vol. 1, pp. 57 and 213–214, Berlin 1925.

[25] A. Moitra, "Super-resolution, extremal functions and the condition number of vandermonde matrices," *Proceedings of the 47th Annual ACM Symposium on Theory of Computing (STOC)*, Portland, 2015.

[26] J. E. Littlewood G. H. Hardy and G. Polya, *Inequalities*, Cambridge Univ. Press, London, U.K., 1959.