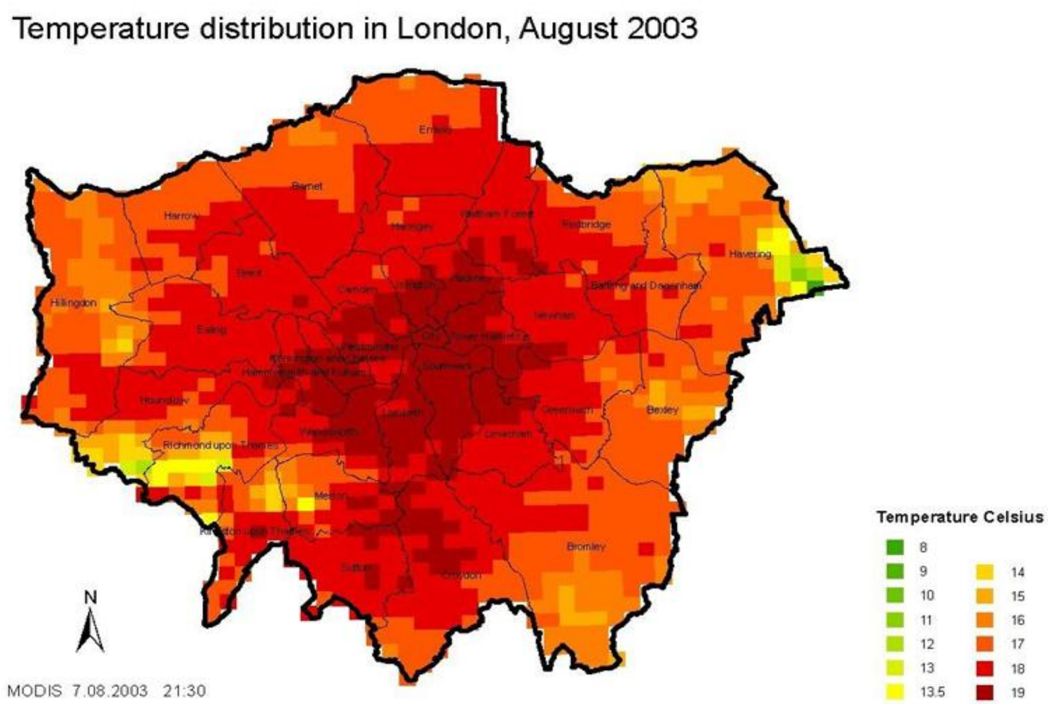
Bandlimited Field Reconstruction from Samples obtained on a Discrete Grid with Unknown Random Locations

Ankur Mallick, Animesh Kumar
Electrical Engineering
Indian Institute of Technology Bombay

Motivation – Spatial Sampling

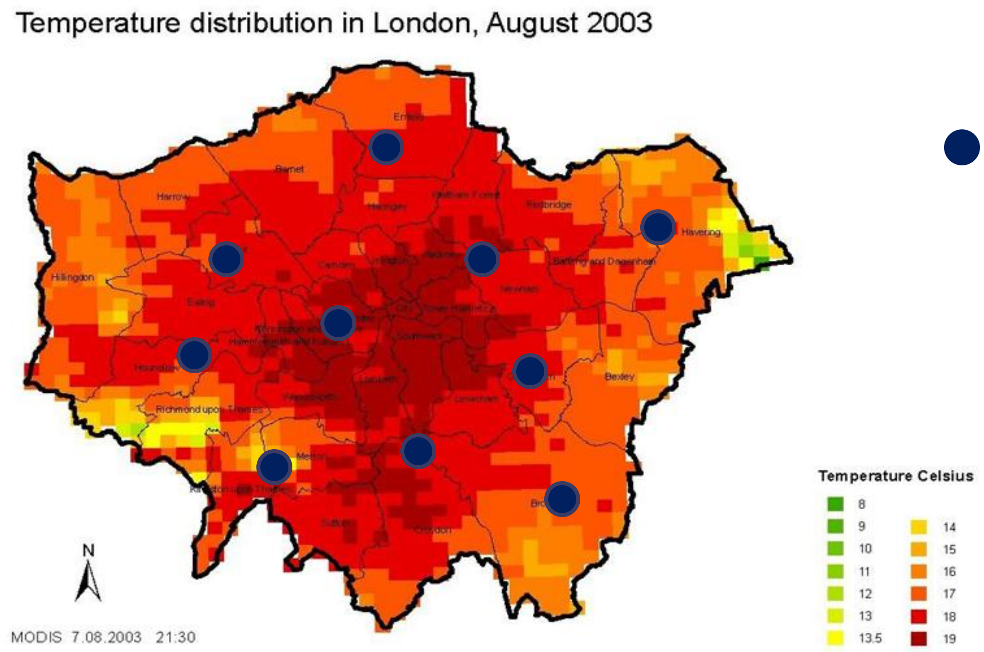
- Consider the problem of estimating a spatially varying field over a large area (For eg. Temperature)



Source: http://climatelondon.org.uk/

Standard Approach

- The usual procedure is to estimate the number of degrees of freedom of the field
- If there are 'N' degrees of freedom, 'N' samples are taken and the corresponding system of equations is solved

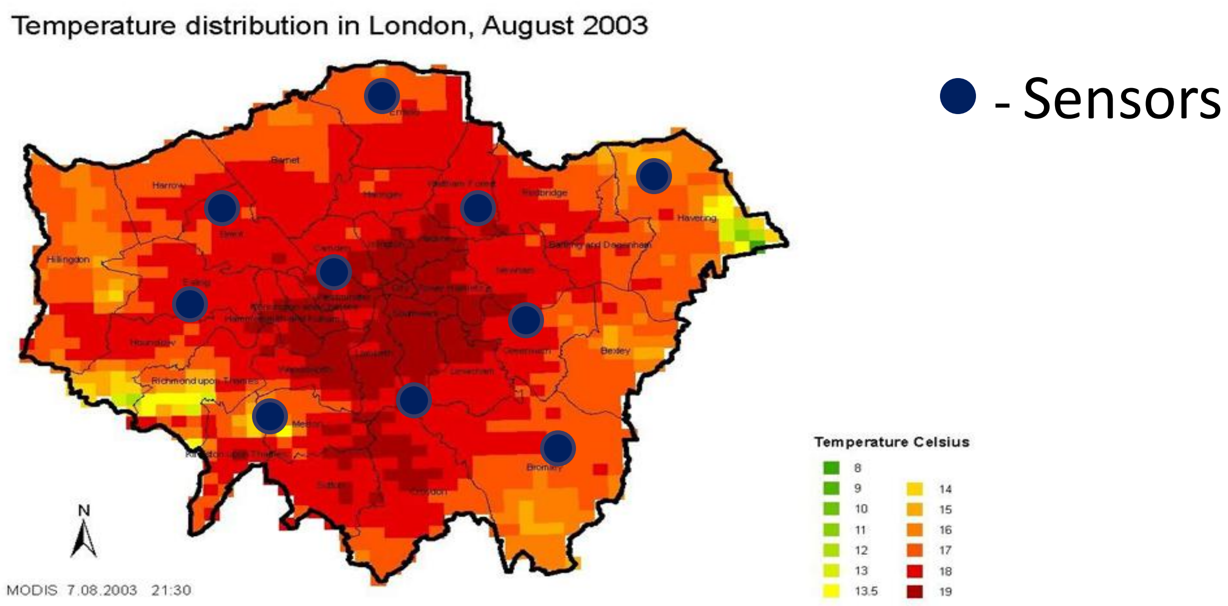


● - Sensors

Source: http://climatelondon.org.uk/

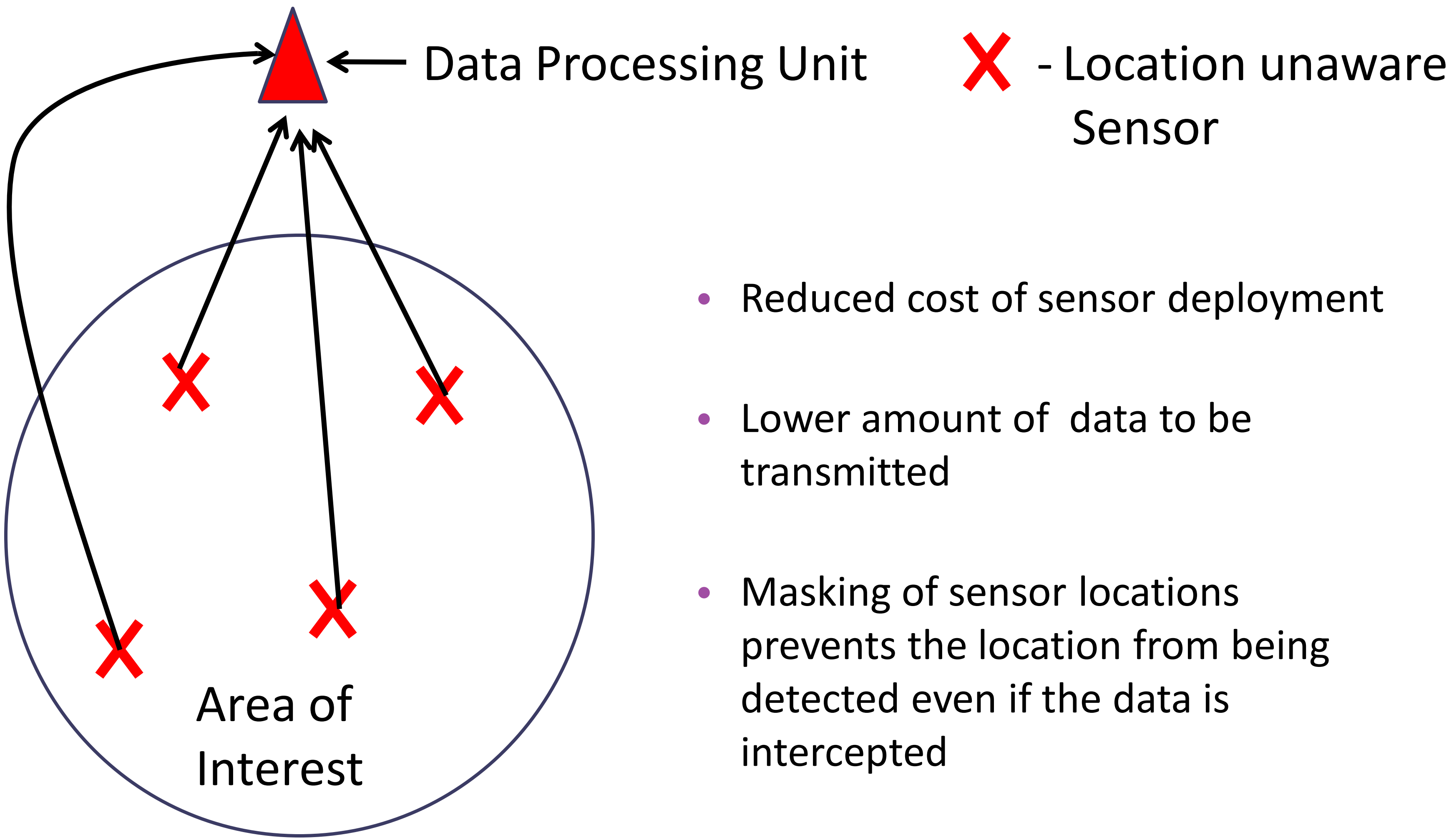
Localization of Sensors is Challenging

- Localization algorithms or GPS equipment required to estimate the coordinates of the sensors is expensive – especially if the number of sensors is large
- The location information obtained might be unreliable since sensor positions are liable to perturbations in spatial sampling



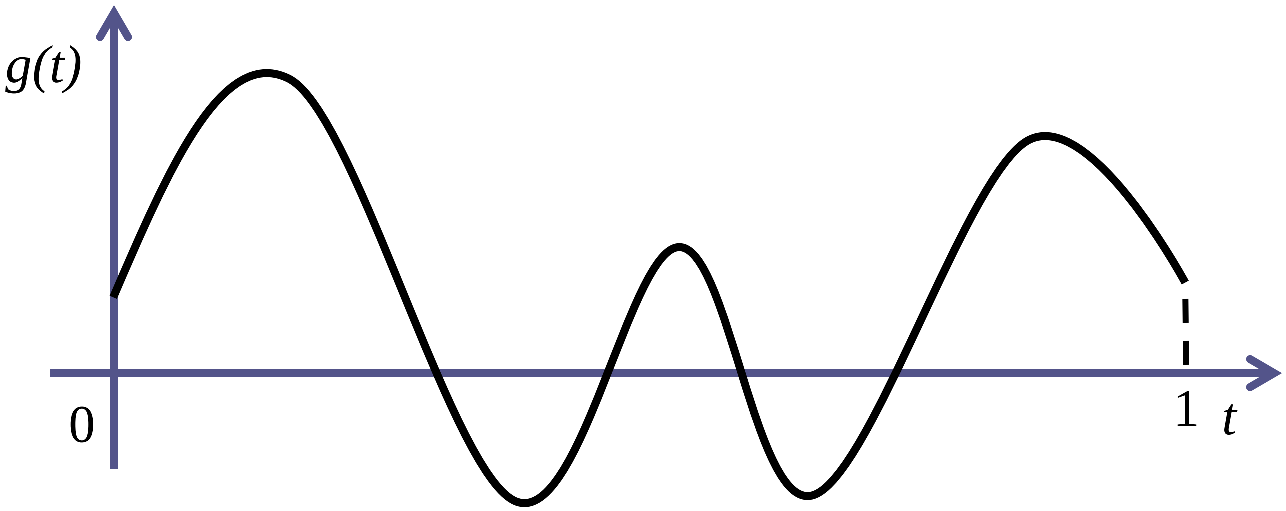
Source: http://climatelondon.org.uk/

Benefits of Location Unaware Sensors



- Reduced cost of sensor deployment
- Lower amount of data to be transmitted
- Masking of sensor locations prevents the location from being detected even if the data is intercepted

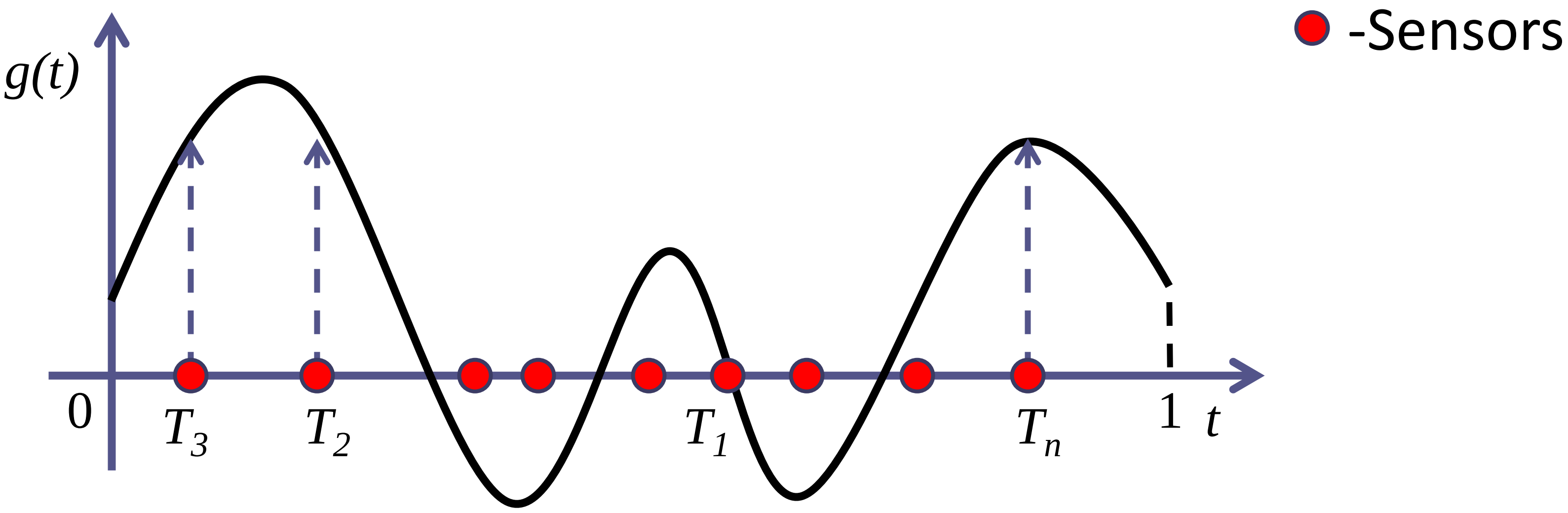
Field Model



- Consider the 1D version of the spatial sampling problem
- *g(t)* is a smooth bandlimited, periodic field (one period is shown)
- Assuming the period to be 1:

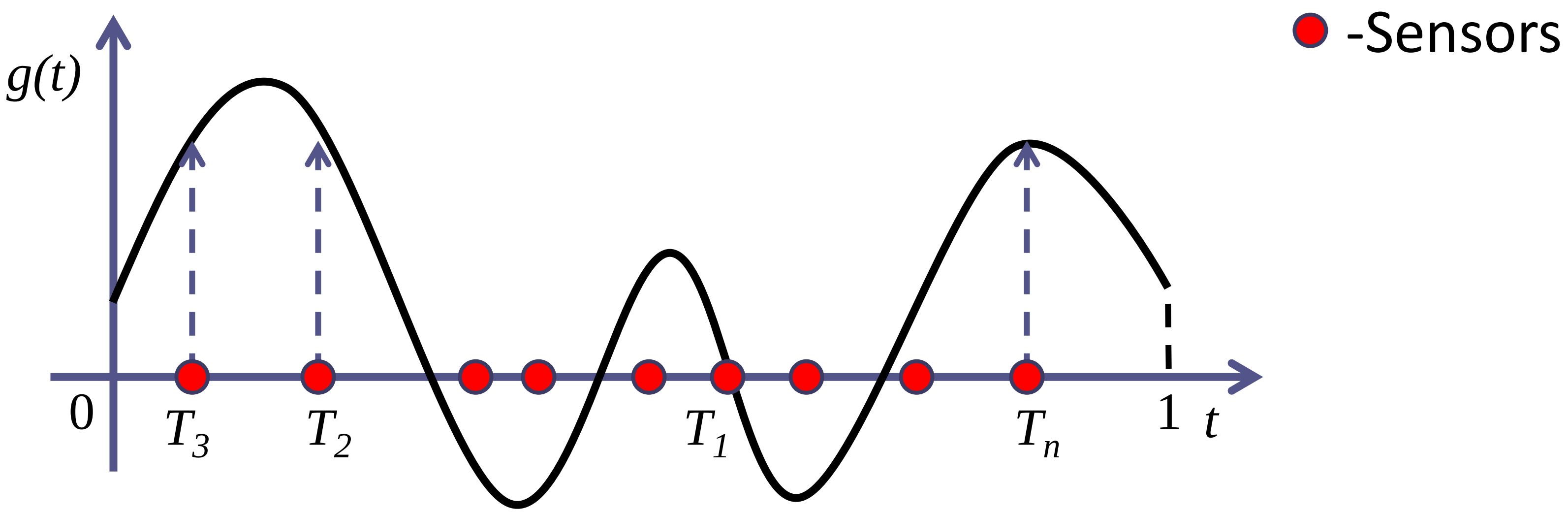
$$g(t) = \sum_{k=-b}^{b} a[k] \exp(j2\pi kt)$$

Distributed Sampling Setup



- Sensors are deployed at **unknown** locations $T_1, T_2, \ldots, T_n$ obtained according to a *random distribution*
- The ordering of the locations is also unknown
- The goal is to estimate the field using the sample values and the *distribution* on the sensor locations

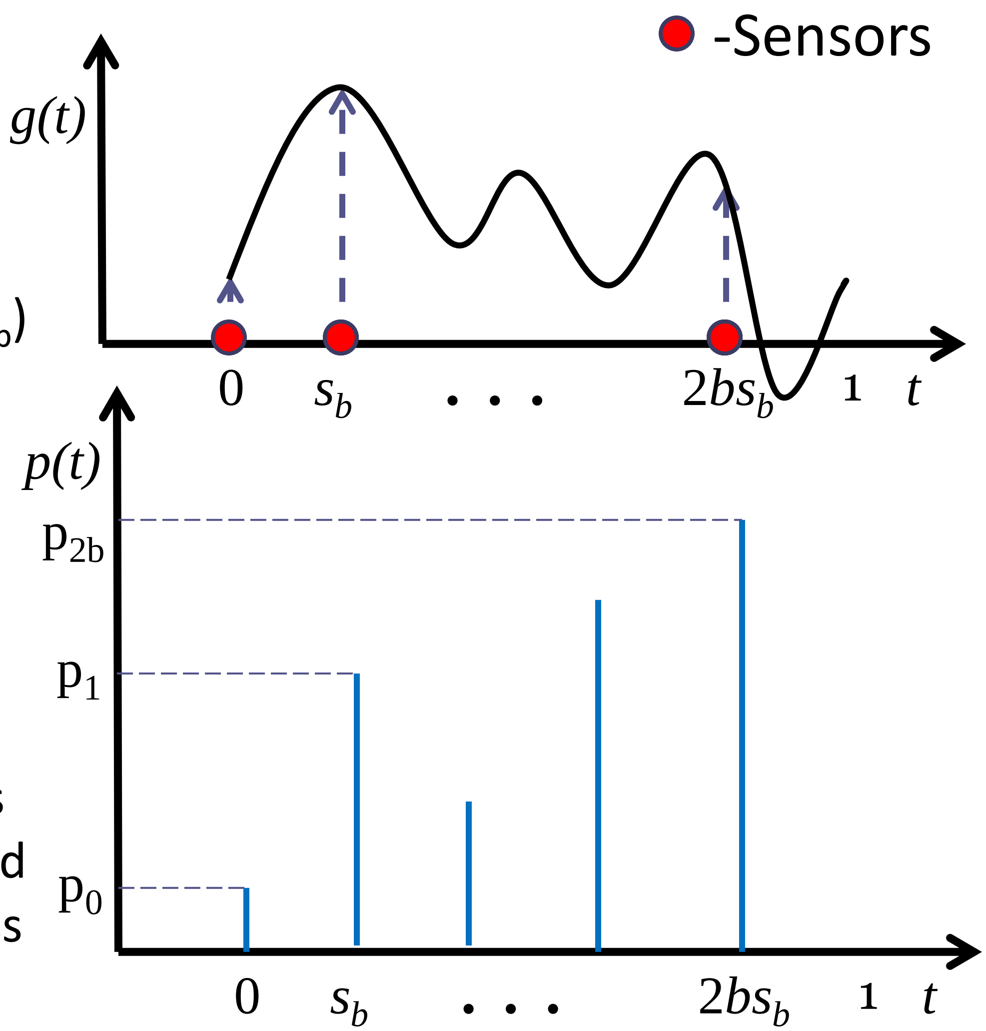
Assumption on Sensor Deployment



- The problem where sensors are deployed according to a continuous distribution is non-linear and hence difficult to solve
- We will address a simplified version of the problem where the sensors are located at a random point on a **discrete grid**

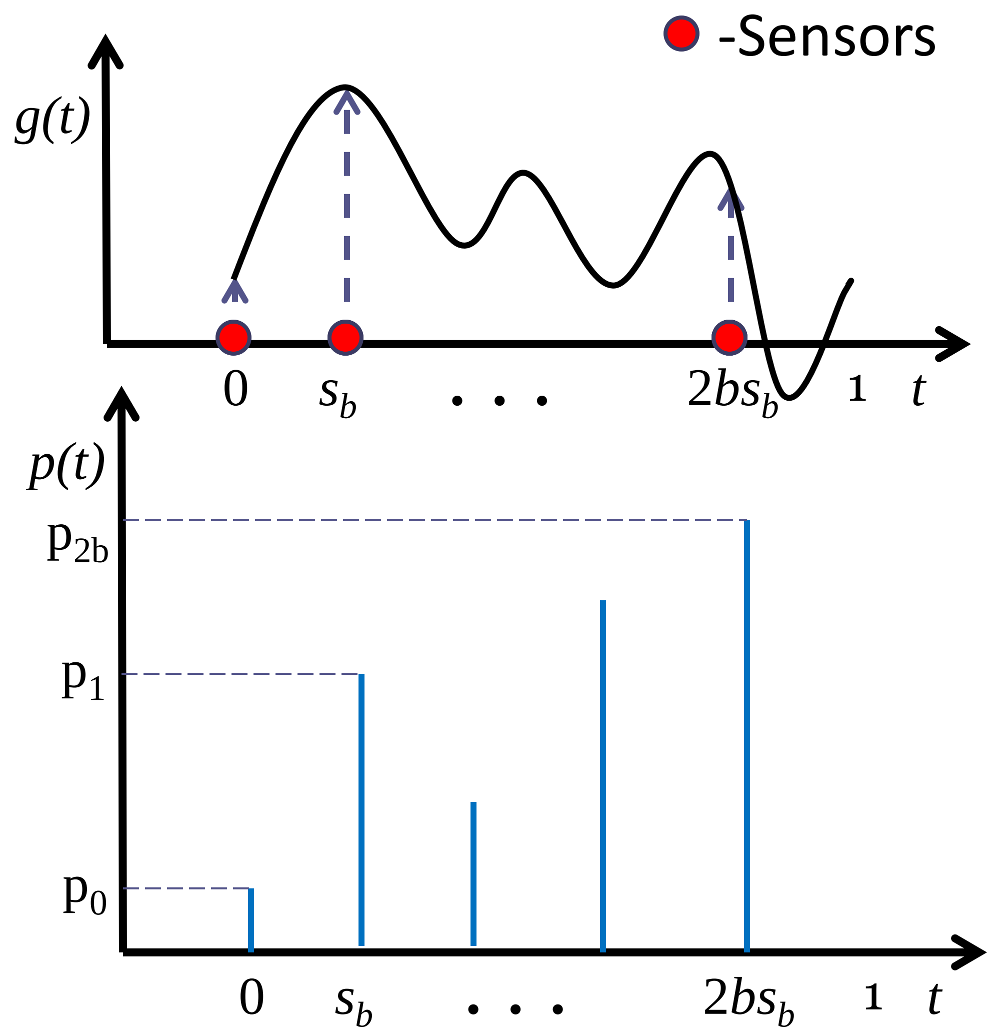
Sampling Model

- $g(t) = \sum_{k=-b}^{b} a[k] \exp(j2\pi kt)$
- $s_b = \frac{1}{(2b+1)}$ (Spacing Parameter)
- $(2b+1)$ grid points: $\{0, s_b, 2s_b, \ldots 2bs_b\}$
- Consider any sensor deployed at location T according to the distribution $p(t)$:
  $T = is_b$ w.p. p_i $(i = 0,1,\ldots,2b)$
- Sensor location, i.e. the index 'i' is unknown and oversampling is used to overcome location unawareness



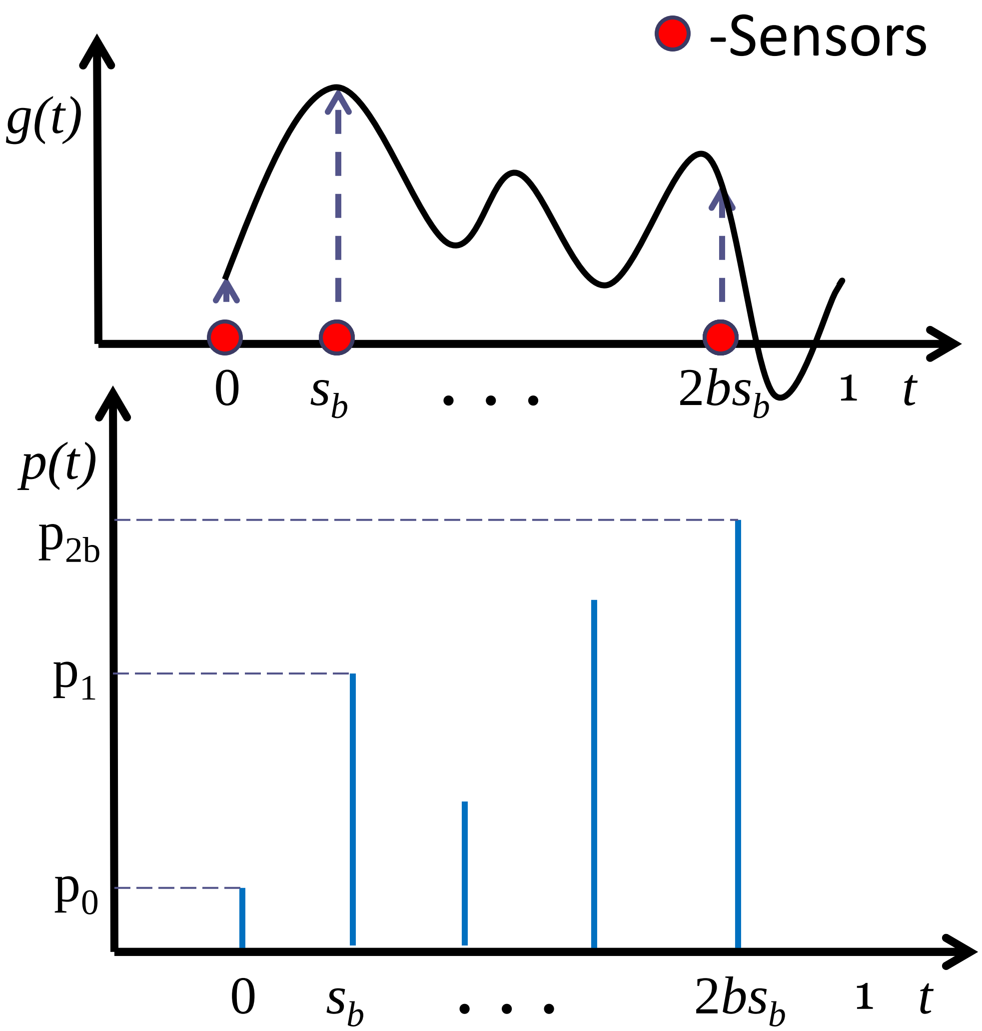
Performance Criterion

- $g(t) = \sum_{k=-b}^{b} a[k] \exp(j2\pi kt)$
- The field has $2b+1$ degrees of freedom
- Correct detection of the $2b+1$ field values, $g(is_b)$, corresponds to correct estimation of the field
- We wish to detect the field correctly with a high probability
- Hence **detection error probability** is the performance criterion to be minimized



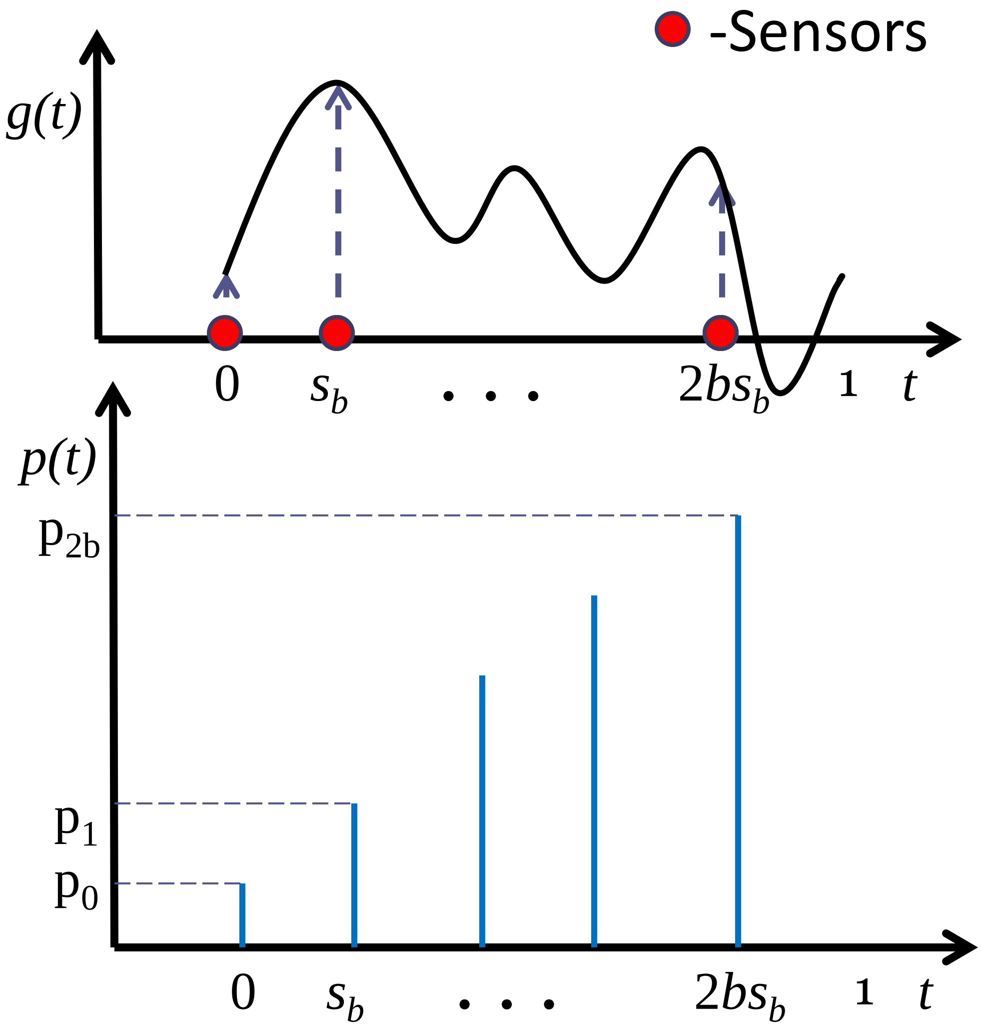
Main Result

- $g(t) = \sum_{k=-b}^{b} a[k] \exp(j2\pi kt)$
- Detection error probability depends on the distribution on the sensor locations, *p(t)*
- *p(t)* is assumed to be discrete and asymmetric
- The **main result** of our work is to find the optimal such *p(t)* that minimizes the detection error probability of any field *g(t)*

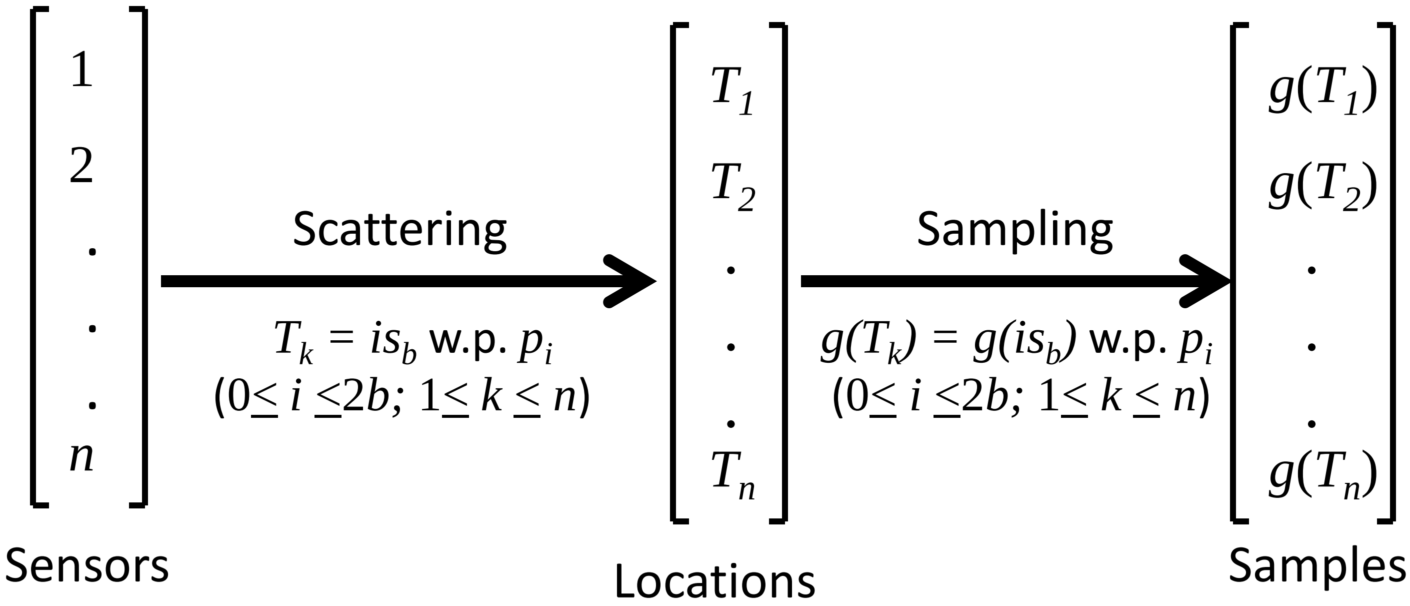


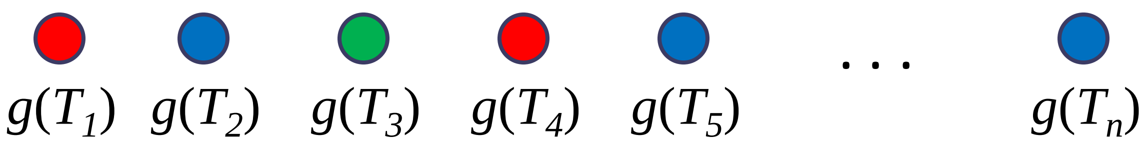
Field Detection Algorithm

- The field detection algorithm has 2 steps
- Step 1: Clustering Samples
- Step 2: Assigning Locations to Clusters
- Additional assumption: $p_0 < p_1 < p_2 < \ldots < p_{2b}$



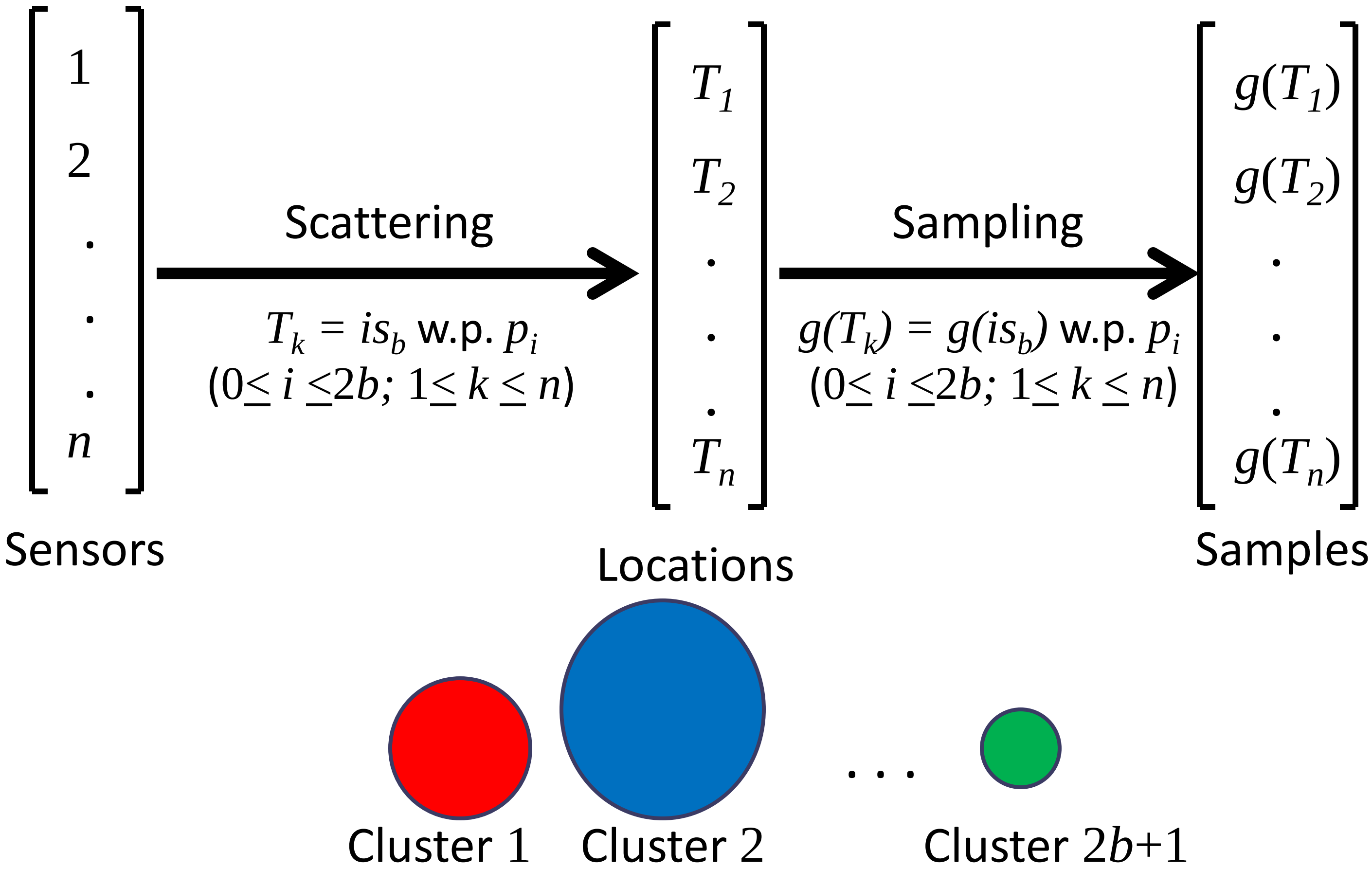
Clustering Samples





- All samples of equal value are put in the same cluster ('Value' of the cluster = Value of any sample in the cluster)
- Since there are $2b$+1 distinct sample values we form ($2b$+1) clusters

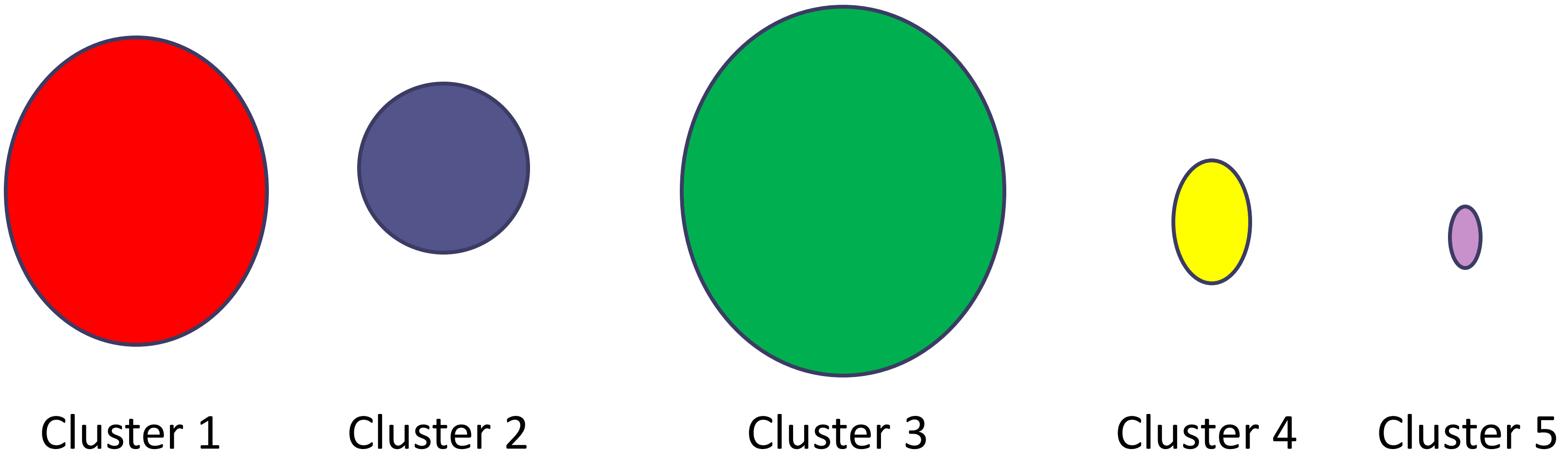
Clustering Samples



- All samples of equal value are put in the same cluster ('Value' of the cluster = Value of any sample in the cluster)
- Since there are $2b$+1 distinct sample values we form ($2b$+1) clusters

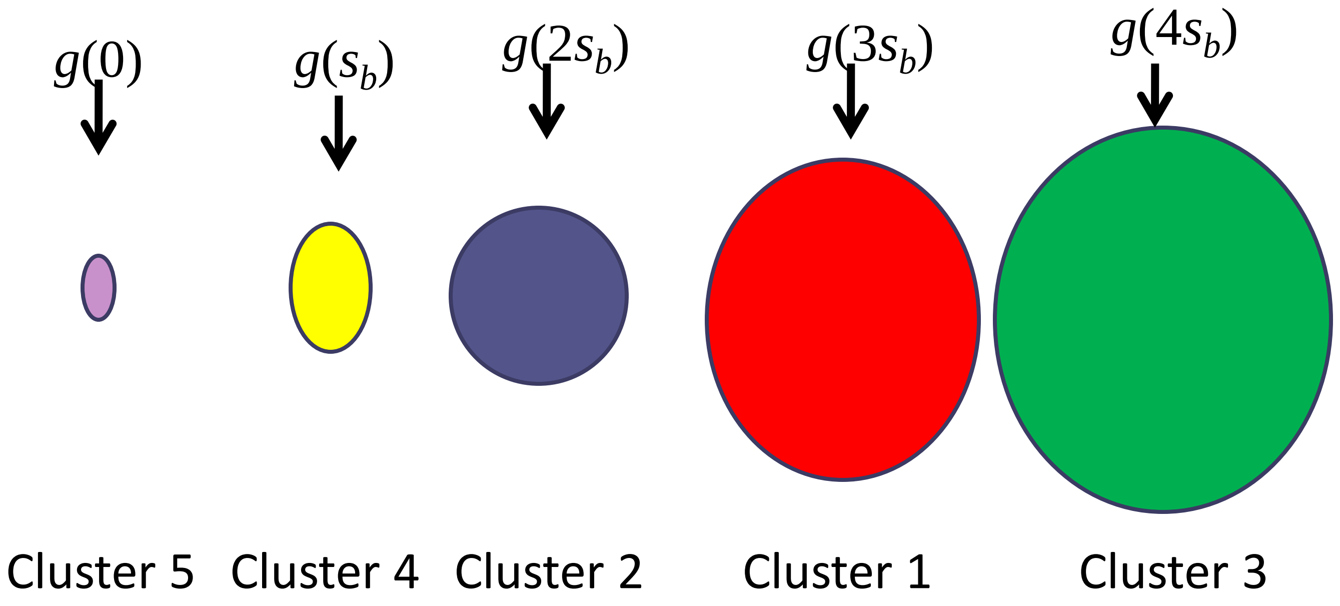
Assigning Locations to Clusters

- ‘Type’ of cluster = Number of elements in the cluster
- Clusters are sorted according to type
- ‘Value’ of cluster with smallest ‘Type’ is assigned to $g(0)$, next smallest to $g(s_b)$, and so on till $g(2bs_b)$ (since $p_0 < p_1 < \ldots < p_{2b}$)
- Consider the following illustration for the case where b=2 and so there are $2b$+1=5 clusters:



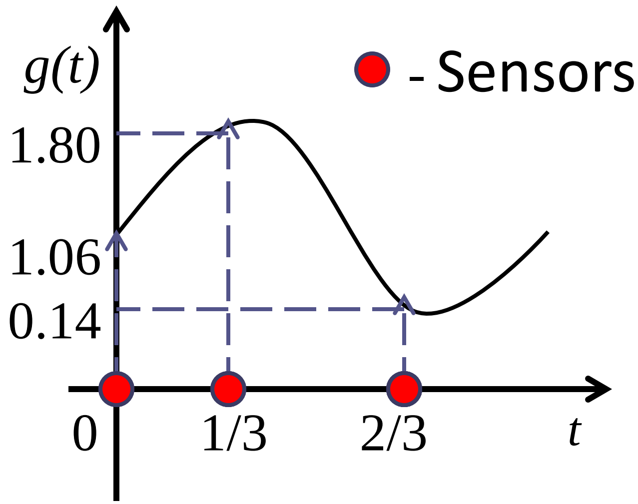
Assigning Locations to Clusters

- 'Type' of cluster = Number of elements in the cluster
- Clusters are sorted according to type
- 'Value' of cluster with smallest 'Type' is assigned to $g(0)$, next smallest to $g(s_b)$, and so on till $g(2bs_b)$ (since $p_0 < p_1 < \ldots < p_{2b}$)
- Consider the following illustration for the case where b=2 and so there are $2b$+1=5 clusters:



Illustrative Example

- Consider a field $g(t)$ as shown below with b=1, s_b=1/3 which is sampled n=10 times



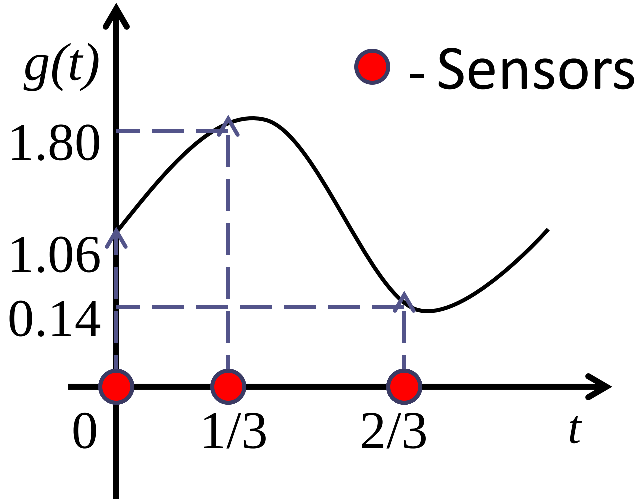
Samples= $\begin{bmatrix} 1.80 \\ 0.14 \\ 0.14 \\ 1.06 \\ 1.80 \\ 0.14 \\ 1.80 \\ 1.06 \\ 0.14 \\ 0.14 \end{bmatrix}$

Cluster	Value	Type
1	1.06	2
2	1.80	3
3	0.14	5

- Conclusion: $g(0)$=1.06, $g(1/3)$=1.80, $g(2/3)$=0.14
- Field is detected **correctly**

Illustrative Example

- Consider a field $g(t)$ as shown below with $b=1$, $s_b=1/3$ which is sampled $n=10$ times



Samples= $\begin{bmatrix} 1.06 \\ 0.14 \\ 0.14 \\ 1.06 \\ 1.80 \\ 0.14 \\ 1.80 \\ 1.06 \\ 0.14 \\ 0.14 \end{bmatrix}$

Cluster	Value	Type
1	1.80	2
2	1.06	3
3	0.14	5

- Conclusion: $g(0)=1.80$, $g(1/3)=1.06$, $g(2/3)=0.14$
- Field is detected **incorrectly**

What if the field values at 2 sensor locations are equal?

- All samples of the same value are grouped in the same cluster
- If field value is equal any 2 of the $2b+1$ grid points then all the samples from these points go into the same cluster and we will have less than $2b+1$ clusters
- If we assume the signal value to be equal at grid points '0' and 's_b' to be equal then :

$$\sum_{k=-b}^{b} a[k](exp(j2\pi k(0)) - exp(j2\pi k(s_b))) = 0$$

- To satisfy this one of the Fourier series coefficients, $a[k]$, is constrained to a fixed value
- If Fourier Series coefficients of a natural signal are instances of independent, continuous random variables then this occurs with probability zero

Detection Error Probability

- Let N_i be the 'type' of the cluster corresponding to $g(is_b)$ (i.e samples from location is_b) in a set of 'n' samples

- Our field detection algorithm is based on the assumption that $0 < N_0 < N_1 < \ldots < N_{2b}$ because $0 < p_0 < p_1 < \ldots < p_{2b}$

- Probability of detection error $(P_e) = P((0 < N_0 < N_1 < \ldots < N_{2b})^c)$

- It can be shown from the union bound that:
  $M \leq P_e \leq (2b+1)M$
  $M = max(\, P(N_0 = 0), P(N_0 \geq N_1), P(N_1 \geq N_2), \ldots, P(N_{2b-1} \geq N_{2b}))$

- It is known from Sanov's Theorem (analogous to the Chernoff Bound) that each term in M decays exponentially with an increase in 'n'

- Thus the distribution $\boldsymbol{p} = (p_0, p_1, \ldots, p_{2b})$ that minimizes M, also minimizes P_e

Deriving the Main Result

- $M = max(\, \mathrm{P}(N_0 = 0), \mathrm{P}(N_0 \geq N_1), \mathrm{P}(N_1 \geq N_2), \ldots, \mathrm{P}(N_{2b-1} \geq N_{2b}))$
- $\mathrm{P}(N_0 = 0) = (1 - p_0)^n$
- $\mathrm{P}(N_0 \geq N_1) \propto 2^{-nD^*}$ (From Sanov's Theorem)

  where $D^* = min \sum_{i=0}^{2b} \frac{N_i}{n} \log_2 \frac{N_i}{np_i}$, subject to $\sum_{i=0}^{2b} N_i = n$ and $N_1 \leq N_0$
- The other terms in M can be calculated as a function of $\boldsymbol{p}$ in similar fashion
- Minimizing M with respect to $\boldsymbol{p}$ (equivalent to minimizing P_e with respect to $\boldsymbol{p}$) gives the following distribution:

$$p_i = \frac{3(i+1)^2}{(b+1)(2b+1)(4b+3)} \text{ for } 0 \leq i \leq 2b$$

- This is the distribution that gives **minimum detection error probability** for our field detection algorithm

Simulation Setup

- Field being estimated: $g(t) = \sum_{k=-b}^{b} a[k]\exp(j2\pi kt)$ (b = 4 is assumed)
- $a[k]$'s are generated using a uniform random number generator (Table 1) with $a[-k] = (a[k])^*$ for real valued fields (conjugate symmetry)
- Number of samples collected ('n') is increased from 100 to 20,000
- The empirical detection error probability for various distributions (Table 2) on the sensor locations is simulated using 10,000 Monte-Carlo trials

Coefficient	Value
$a[0]$	1
$a[1]$	$0.9134 - j0.5469$
$a[2]$	$0.1270 - j0.2785$
$a[3]$	$0.9058 - j0.0975$
$a[4]$	$0.8147 - j0.6324$

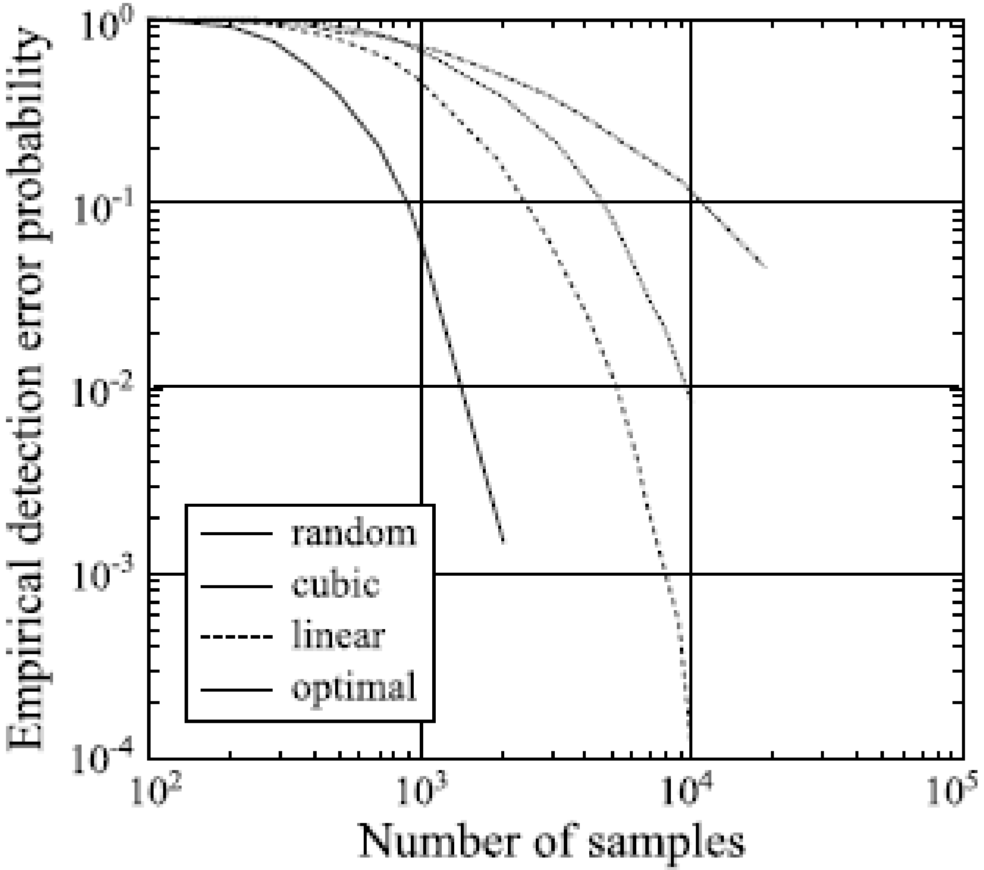
Table 1

Distribution Type	$\boldsymbol{p} = [p_0, p_1, \ldots, p_{2b}]$
Random	$\alpha_1[U(1),U(2),\ldots,U(2b+1)]$*
Linear	$\alpha_2[1,2,\ldots,2b+1]$
Cubic	$\alpha_3[1,8,\ldots,(2b+1)^3]$
Optimal	$\alpha_4[1,4,\ldots,(2b+1)^2]$
*$U(k)$'s are ordered uniform random variables	

Table 2

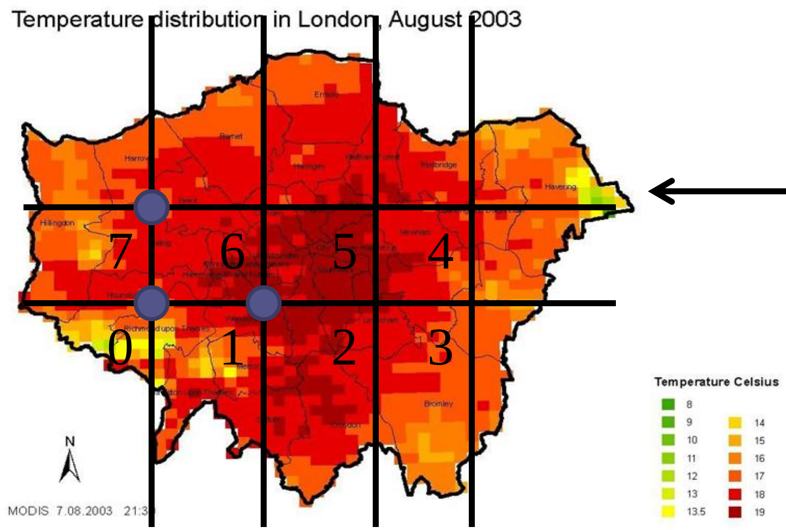
Simulation Results

- We use a log-log plot since the P_e decays exponentially with n and we are interested in modeling the exponent
- Each plot ends when the empirical detection error probability becomes zero or the maximum sample size ($n = 20000$) is reached
- It is observed that the estimated optimal distribution decays fastest and has the smallest empirical detection error probability



Extension to the 2D case

- In the 1 dimensional case the signal had $2b+1$ degrees of freedom and hence we sampled it at $2b+1$ grid points
- Similarly in the 2D case, if the signal has 'N' degrees of freedom it is sampled at 'N' grid points
- Sensors are deployed according to an asymmetric distribution and the location on the grid where the sensor lands is unknown



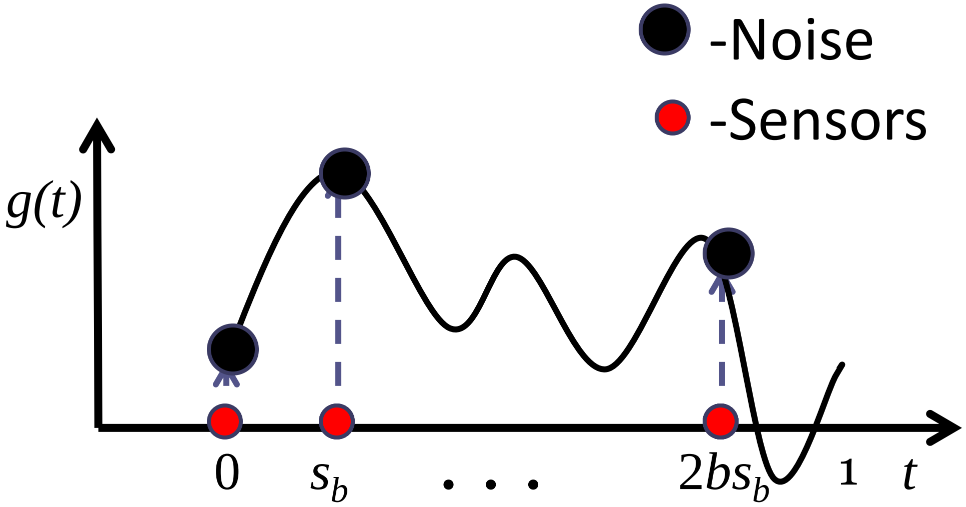
$N=8$

$p_0 < p_1 < \ldots < p_7$

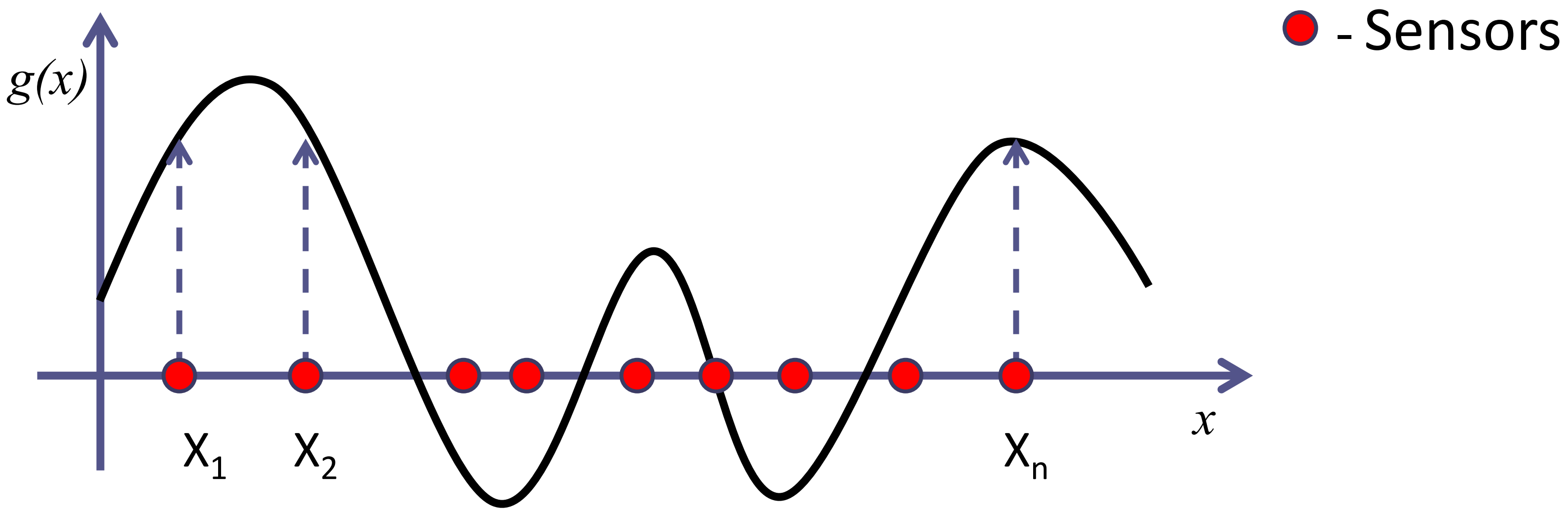
● - Sensors

Future Work

- Extending the setup to include measurement noise on the samples
- Requires application of clustering algorithms from machine learning (For eg. EM algorithm) on the noisy samples



Future Work



- Deploying sensors according to an arbitrary continuous distribution
- We are working on an algorithm to estimate the field in this case

Other Works in this area

- Animesh Kumar, "Bandlimited Spatial Field Sampling with Mobile Sensors in the Absence of Location Information." *arXiv preprint arXiv:1509.03966*(2015)
- Animesh Kumar, “On bandlimited signal reconstruction from the distribution of unknown sampling locations,” *IEEE Transactions on Signal Processing, vol. 63, no. 5, pp. 1259–1267, Mar. 2015*
- Pina Marziliano and Martin Vetterli, “Reconstruction of irregularly sampled discrete-time bandlimited signals with unknown sampling locations,” *IEEE Transactions on Signal Processing, vol. 48, no. 12, pp. 3462–3471, Dec. 2000*
- Alessandro Nordio, Carla-Fabiana Chiasserini, and Emanuele Viterbo, “Performance of linear field reconstruction techniques with noise and uncertain sensor locations,” *IEEE Transactions on Signal Processing, vol. 56, no. 8, pp. 3535–3547, Aug. 2008*
- Browning, John, "Approximating signals from nonuniform continuous time samples at unknown locations." *Signal Processing, IEEE Transactions on*55.4 (2007): 1549-1554