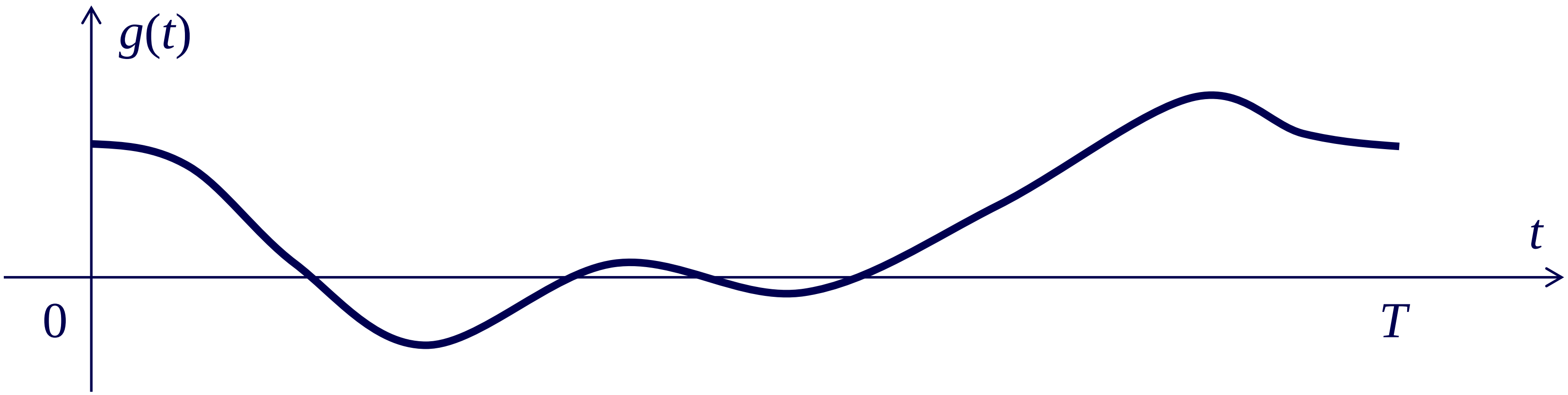
Bandlimited Signal Reconstruction From the Distribution of Unknown Sampling Locations

Animesh Kumar
Electrical Engineering
Indian Institute of Technology Bombay

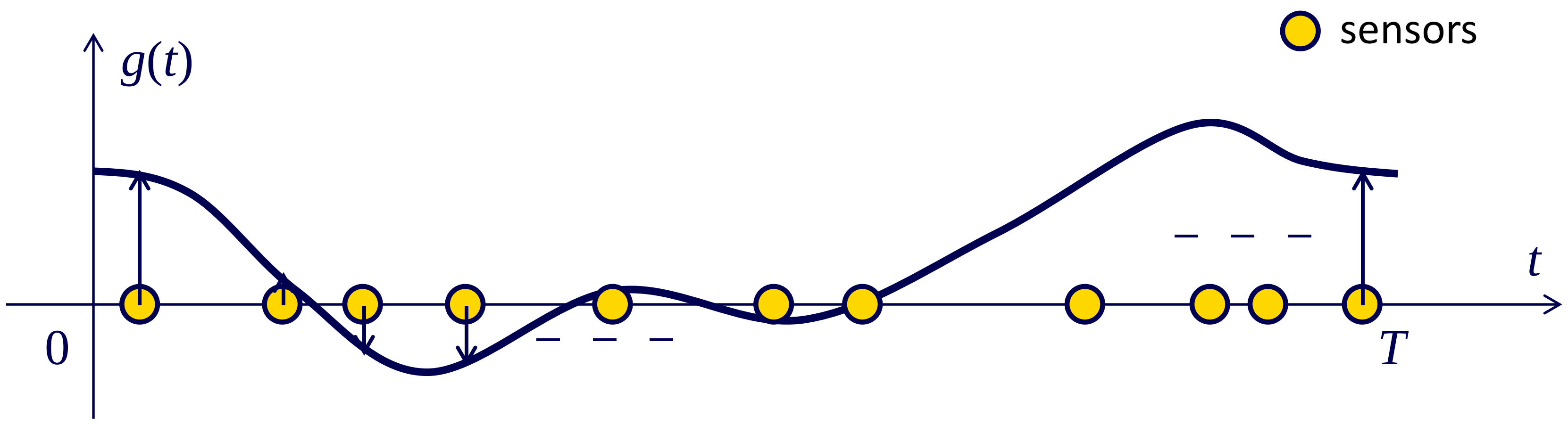
Spatial acquisition problem of interest



Consider the acquisition problem, where a smooth field in a finite interval has to be sampled or estimated

Example: acquisition of spatial fields with sensors

Sampling model

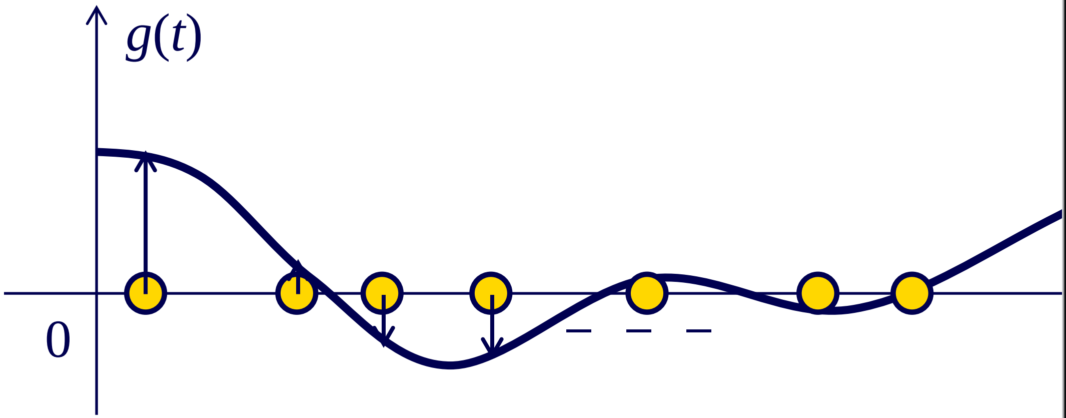


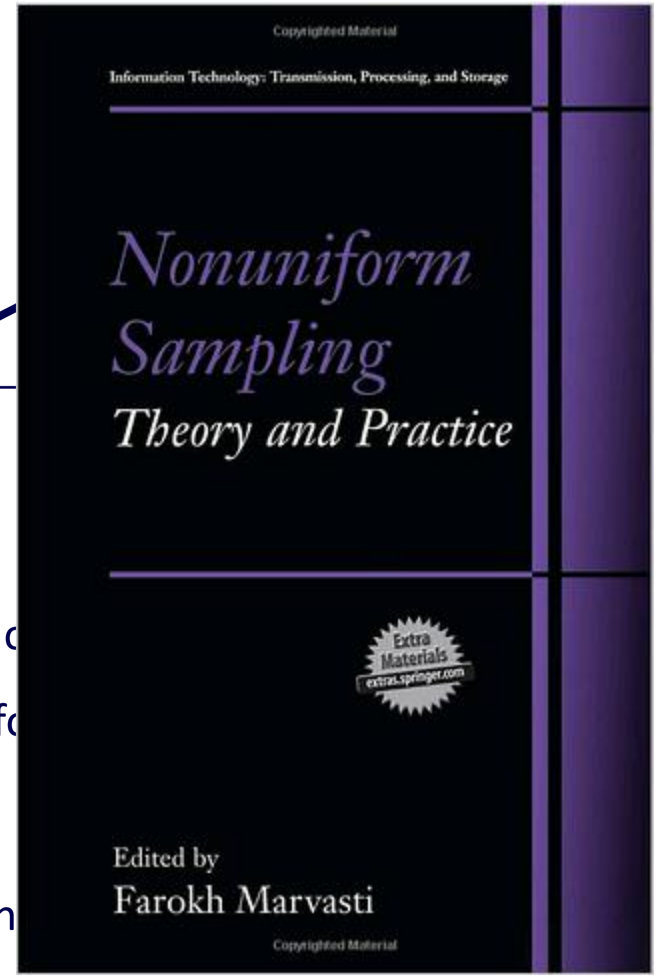
Motivated by the smart-dust paradigm, where a lot of sensors are "scattered" in a region, we consider **random** deployment of sensor for sampling the field

There are two possible scenarios:

- ◊ When the sensor locations are random but known
- ◊ When the sensor locations are **unknown** but their statistical distribution is known

Sampling model





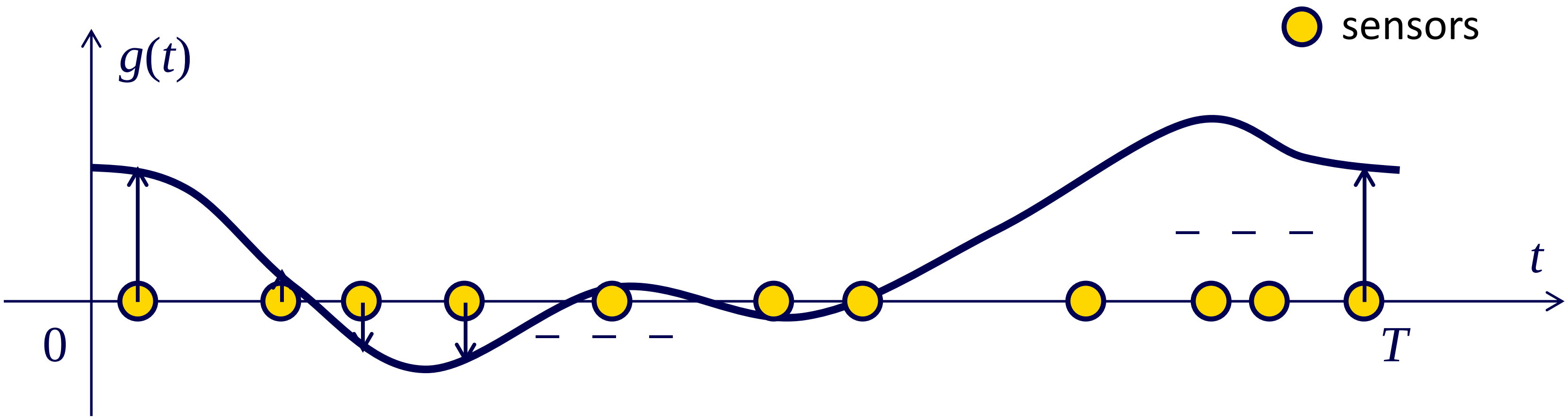
Motivated by the smart-dust paradigm, where a lot c

region, we consider **random** deployment of sensor fo

There are two possible scenarios:

- ◊ When the sensor locations are random but kn
- ◊ When the sensor locations are **unknown** but their statistical distribution is known

Sampling model



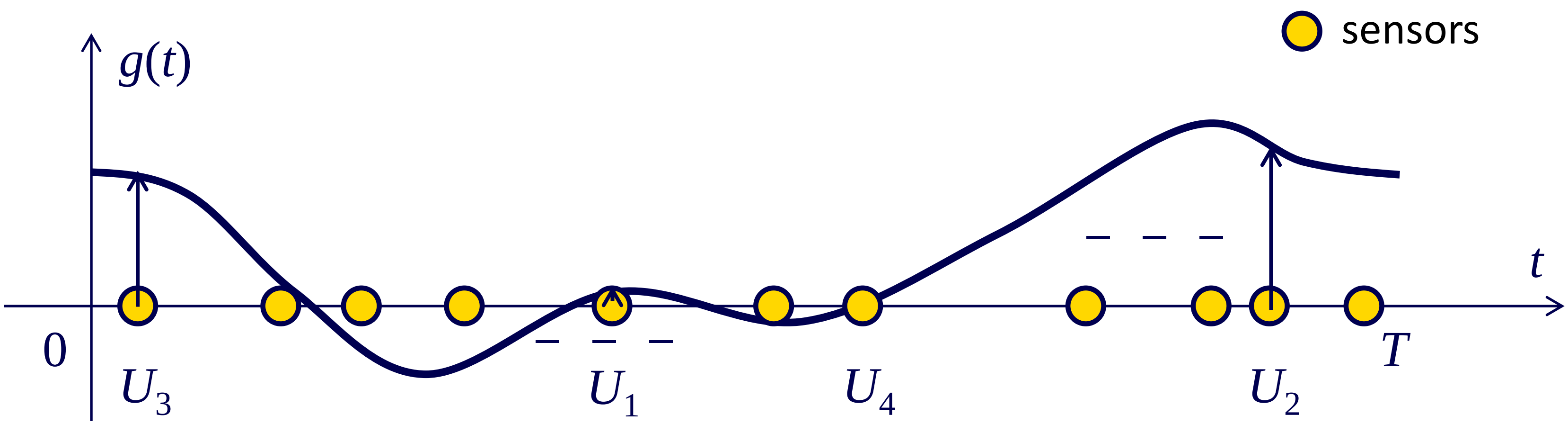
Motivated by the smart-dust paradigm, where a lot of sensors are "scattered" in a region, we consider **random** deployment of sensor for sampling the field

There are two possible scenarios:

◊ When the sensor locations are random but known

◊ When the sensor locations are **unknown** but their statistical distribution is known

Field and sensor-locations models

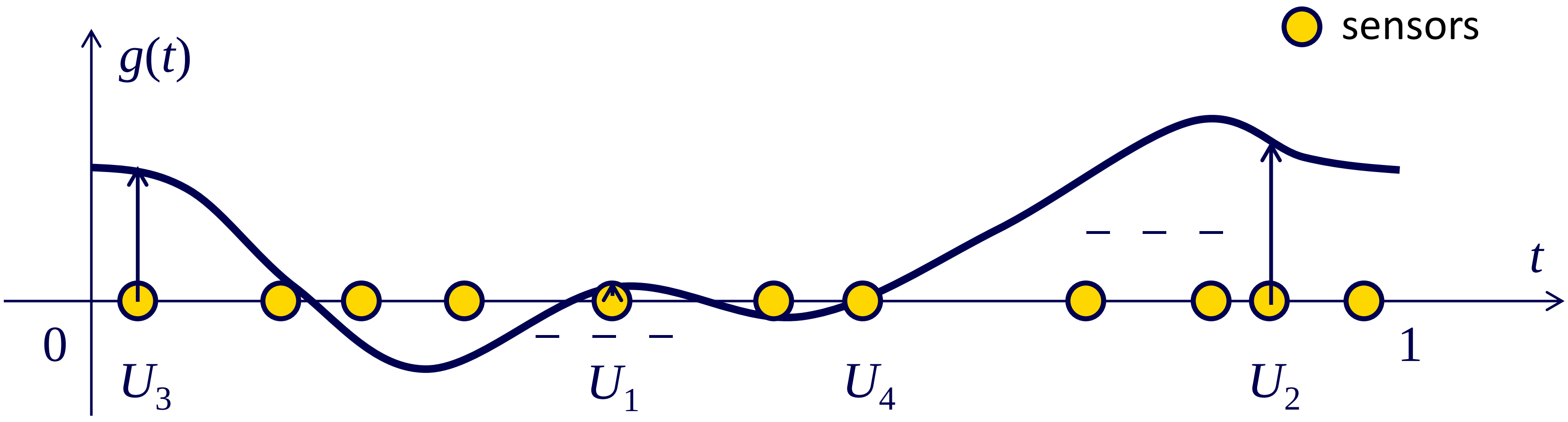


Sensor locations are **unknown** but their statistical distribution is known. For this work, $U_1^n = (U_1, U_2, \ldots, U_n)$ are i.i.d. Unif$[0,T]$

We assume that a periodic extension of the field $g(t)$ is bandlimited, that is, $g(t)$ is given by a finite number of Fourier series coefficients, (WLOG) $|g(t)| \leq 1$, and $T = 1$

$$g(t) = \sum_{k=-b}^{b} a_k \exp(j 2\pi k t)$$

Observations made and distortion criterion



$\boldsymbol{G}^T = (g(U_1), g(U_2), \ldots, g(U_n))$ is collected without the knowledge of $(U_1, U_2, \ldots, U_n)$

We wish to estimate $g(t)$ and measure the performance of estimate against the average mean-squared error, i.e., if $\hat{G}(t)$ is the estimate then

$$D := ||\widehat{G} - g||_2^2 := \int_0^1 |\widehat{G}(t) - g(t)|^2 \mathrm{d}t$$

Main results

◊ A bandlimited field *cannot be* uniquely determined with (perfect) samples obtained at statistically distributed locations, even if the number of samples is infinite

◊ If the **order** (left to right) of sample locations is known, a consistent estimate $\hat{G}(t)$ for the field of interest can be obtained

- Consistency, distortion, and weak convergence results are established for this estimate $\hat{G}(t)$. Recall that

$$D := ||\widehat{G} - g||_2^2 := \int_0^1 |\widehat{G}(t) - g(t)|^2 \mathrm{d}t$$

Related work

◊ Recovery of (narrowband) discrete-time bandlimited signals from samples taken at unknown locations **[Marziliano and Vetterli'2000]**

◊ Recovery of a bandlimited signal from a finite number of ordered nonuniform samples at unknown sampling locations **[Browning'2007]**.

◊ Estimation of periodic bandlimited signals in the presence of random sampling location under two models **[Nordio, Chiasserini, and Viterbo'2008]**

- Reconstruction of bandlimited signal affected by noise at random but known locations
- Estimation of bandlimited signal from noisy samples on a location set obtained by random perturbation of equi-spaced deterministic grid.

◊ Estimation of a bandlimited field from samples taken at i.i.d. distributed unknown locations. Asymptotic consistency (convergence in probability), mean-squared error bounds, and central-limit type weak law are the focus of this work

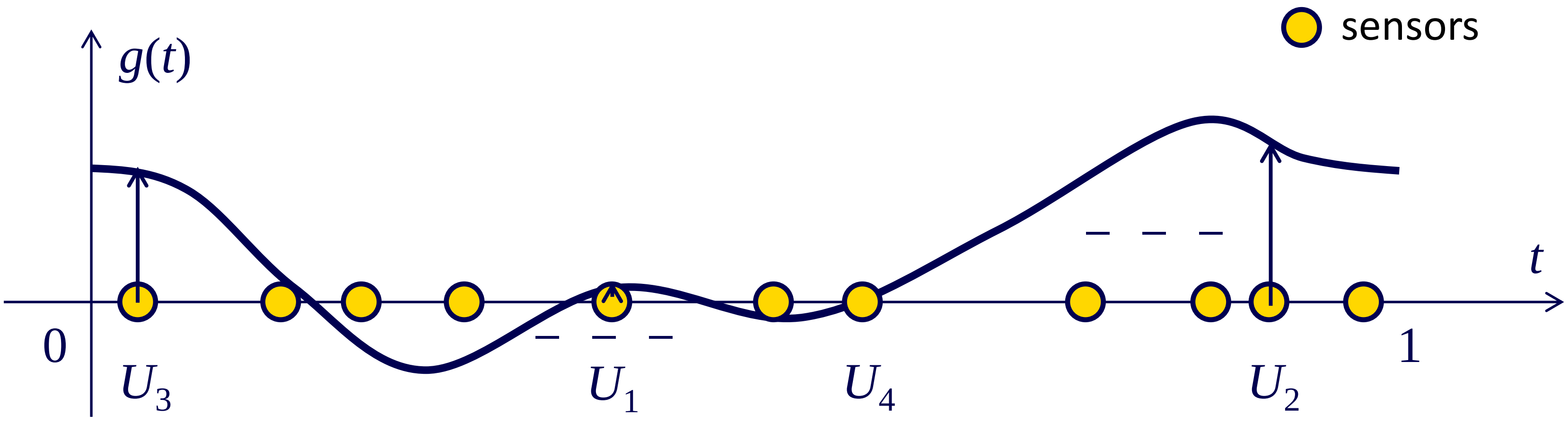
Organization

- ◊ Introduction and contributions
- ◊ Signal estimation without any knowledge of $(U_1, U_2, \ldots, U_n)$
- ◊ Signal estimation and reconstruction distortion when order of samples $(U_1, U_2, \ldots, U_n)$ is known
- ◊ Conclusions

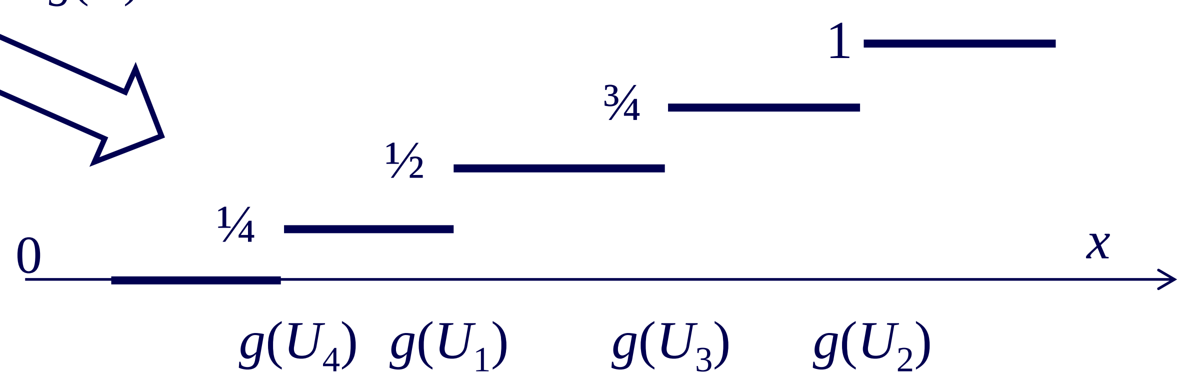
Organization

- ◊ Introduction and contributions
- ◊ Signal estimation without any knowledge of $(U_1, U_2, \ldots, U_n)$
- ◊ Signal estimation and reconstruction distortion when order of samples $(U_1, U_2, \ldots, U_n)$ is known
- ◊ Conclusions

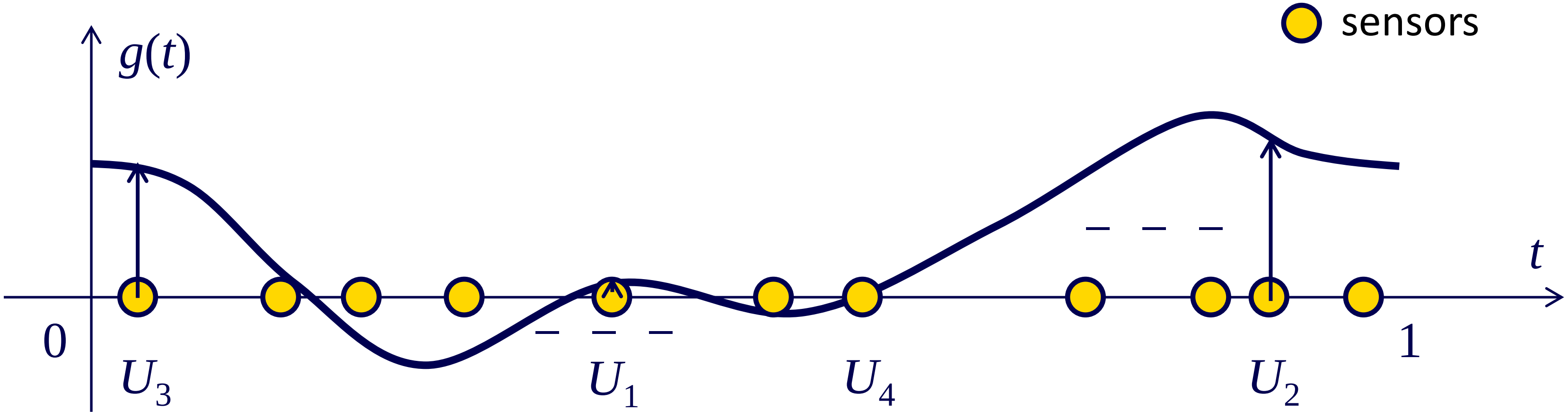
It is impossible to infer $g(t)$ from $g(U_1^\infty)$



Effectively, we are just collecting the empirical distribution or histogram of $g(U_1)$, $g(U_2)$, …, $g(U_n)$ and, in the limit of large n, the task is to estimate $g(t)$ from the distribution of $g(U)$



It is impossible to infer $g(t)$ from $g(U_1^\infty)$



Consider the statistic

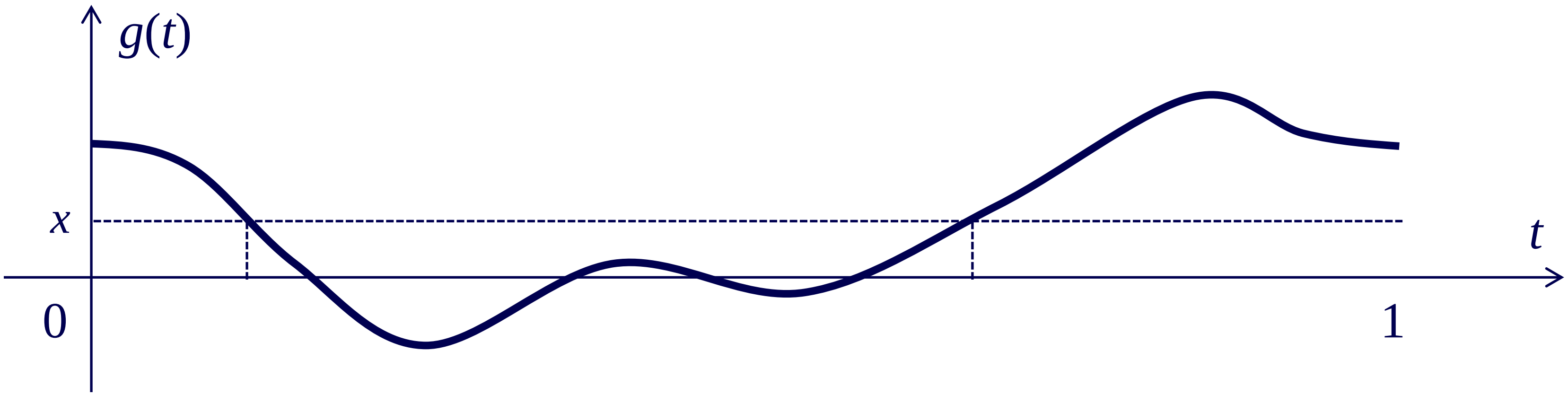
$$F_{g,n}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(g(U_i) \leq x)$$

◊ Then $F_{g,n}(x)$, x in set of reals and $g(U_1)$, $g(U_2)$, …, $g(U_n)$ are statistically equivalent

◊ By the Glivenko Cantelli theorem, $F_{g,n}(x)$ converges almost surely to Prob($g(U) \leq x$) for each x in set of real numbers **[van der Vaart'1998]**

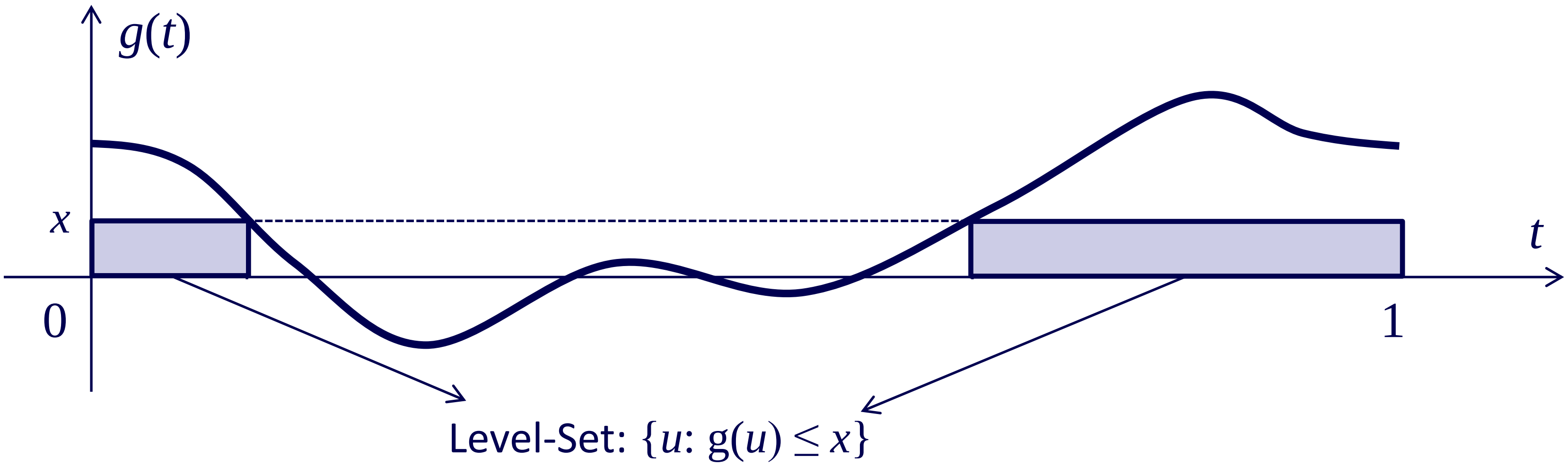
It is impossible to infer $g(t)$ from U_1^∞

So what does Prob($g(U) \leq x$), for x in set of real numbers, looks like?



It is impossible to infer $g(t)$ from U_1^∞

So what does Prob($g(U) \leq x$), for x in set of real numbers, looks like?

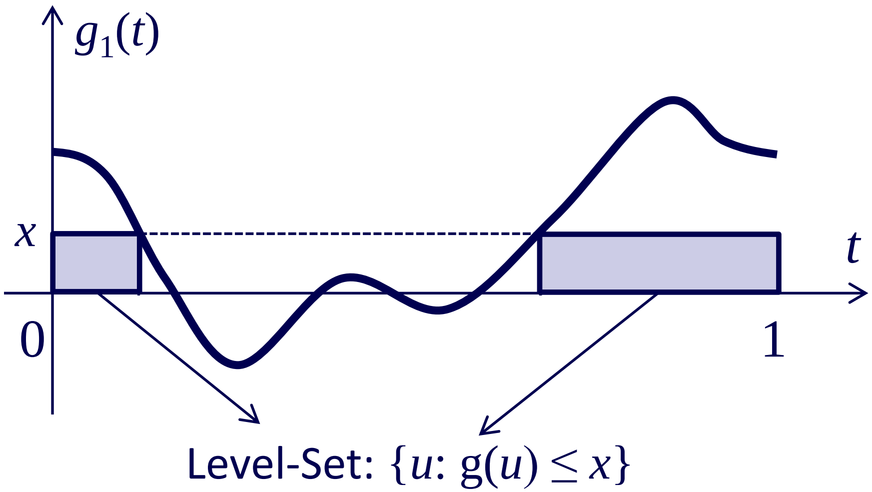


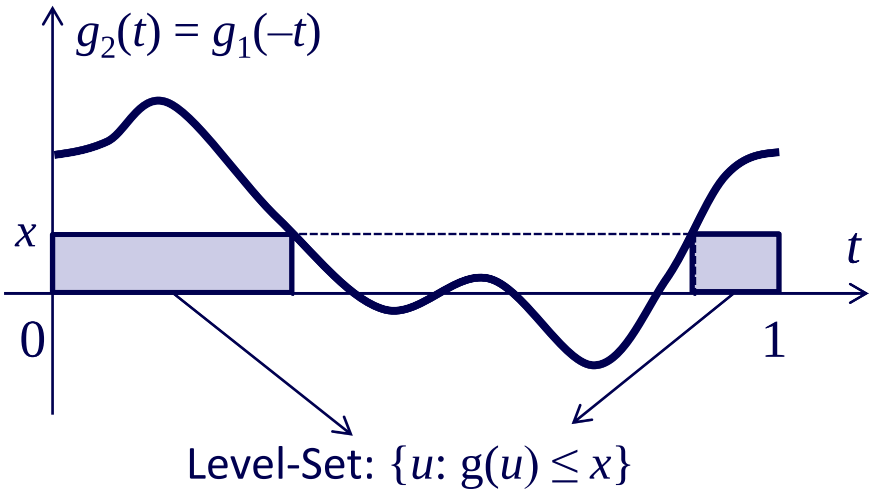
◊ Prob($g(U) \leq x$) for each x is the probability of U belonging in the level-set. Thus, it is simply the length (measure) of level-set

◊ We will now illustrate that two different fields $g_1(t) \neq g_2(t)$ can still lead to Prob($g_1(U) \leq x$) = Prob($g_2(U) \leq x$)

Graphical proof of first result

$g_1(t) \neq g_2(t)$ does not imply Prob$(g_1(U) \leq x)$ = Prob$(g_2(U) \leq x)$





◊ The length (measure) of the level-sets is the same in the two cases for every x

◊ As a recap, we showed that the Glivenko Cantelli theorem's limit, obtained from a statistical equivalent of observed samples, is the same for two different signals. Thus, the observed samples alone do not lead to a unique reconstruction of the field

Organization

- ◊ Introduction and contributions
- ◊ Signal estimation without any knowledge of $(U_1, U_2, \ldots, U_\mathrm{n})$
- ◊ Signal estimation and reconstruction distortion when order of samples $(U_1, U_2, \ldots, U_\mathrm{n})$ is known
- ◊ Conclusions

Organization

- ◊ Introduction and contributions
- ◊ Signal estimation without any knowledge of $(U_1, U_2, \ldots, U_n)$
- ◊ Signal estimation and reconstruction distortion when order of samples $(U_1, U_2, \ldots, U_n)$ is known
- ◊ Conclusions

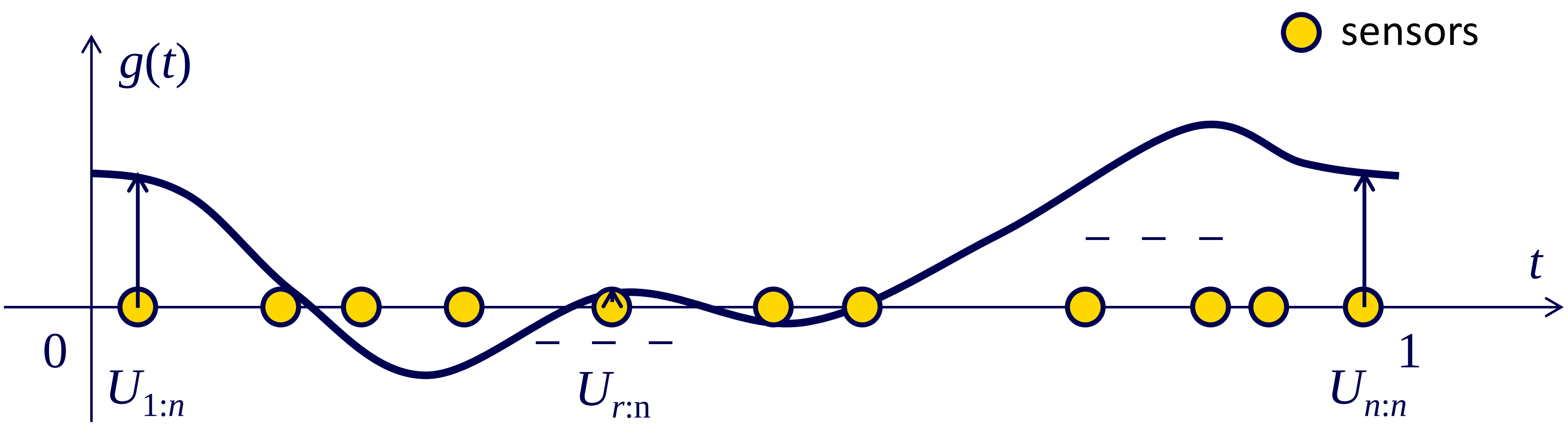
Working with ordered samples

◊ If the **order** (left to right) of sample locations is known, a consistent estimate $\hat{G}(t)$ for the field of interest can be obtained

◊ Recall that

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

◊ Thus, due to bandlimitedness, there are $(2b+1)$ parameters to be learned or estimated



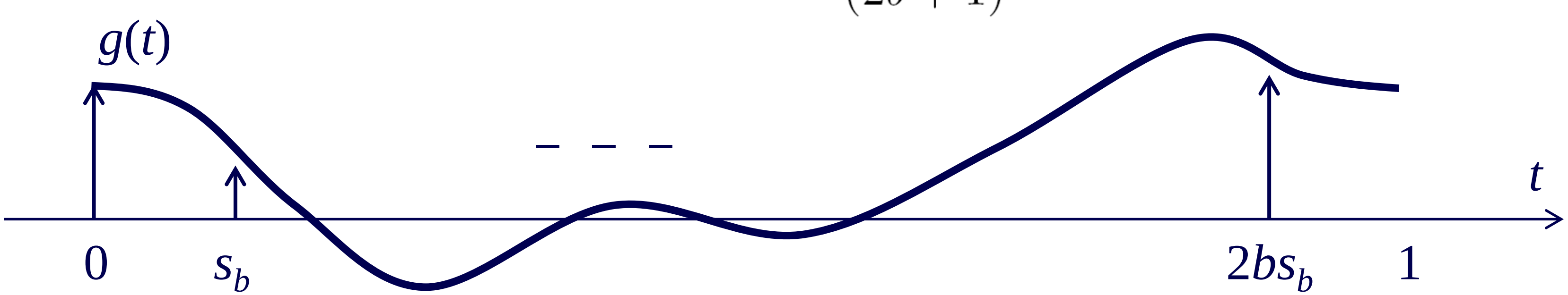
Using field samples to get the Fourier series

From $(2b+1)$ equi-spaced samples of the field, the $(2b+1)$ Fourier series coefficients (and hence the field) can be obtained as follows

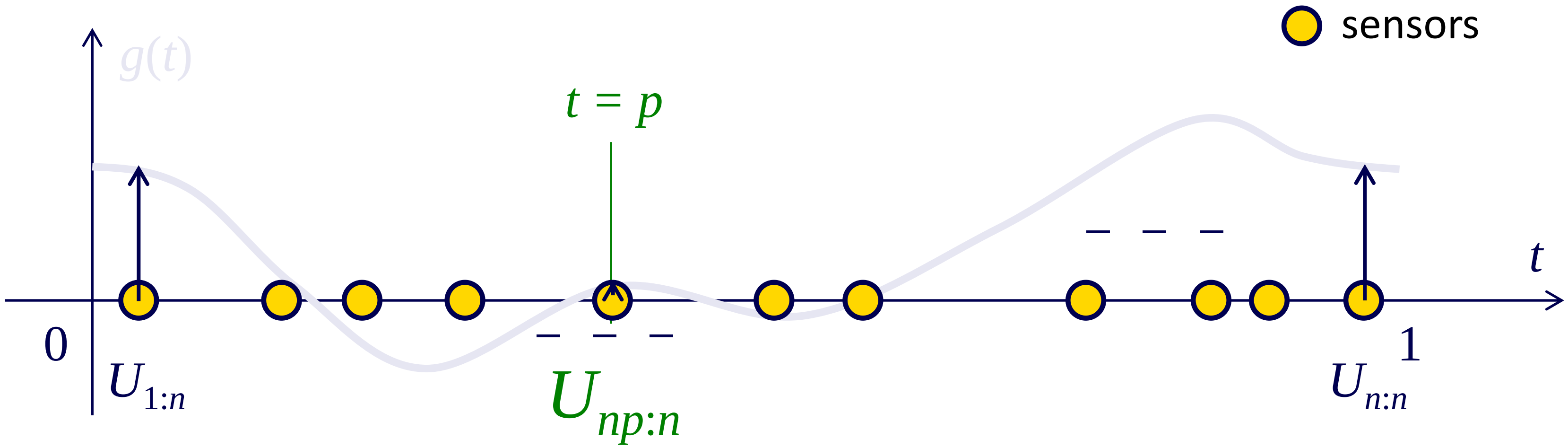
$$\begin{bmatrix} g(0) \\ g(s_b) \\ \vdots \\ g(2bs_b) \end{bmatrix} = \begin{bmatrix} 1 & \dots & 1 \\ \phi_{-b} & \dots & \phi_b \\ \vdots & & \vdots \\ (\phi_{-b})^{2b} & \dots & (\phi_b)^{2b} \end{bmatrix} \begin{bmatrix} a_{-b} \\ a_{-b+1} \\ \vdots \\ a_b \end{bmatrix}$$

where $s_b = 1/(2b+1)$ and $\phi_b = \exp(j2\pi k s_b) = \exp(j2\pi k/(2b+1))$. In matrix notation and upon inversion

$$\vec{a} = (\Phi_b)^{-1}\vec{g} = \frac{1}{(2b+1)}\Phi_b^{\dagger}\vec{g}$$



Approximation of the field samples



It is known that $U_{np:n}$ converges to p in many ways (in L^2, in almost-sure sense, and in weak-law) **[David and Nagaraja'2003]**

In the absence of field values $g(0)$, $g(s_b)$, …, $g(2bs_b)$, we use $\vec{G} = (g(U_{1:n}), g(U_{nsb:n}), \ldots, g(U_{2bsbn:n}))$, to define the Fourier series estimate and field estimate as follows

$$\vec{A} := [\widehat{A}_{-b}, \widehat{A}_{-b+1}, \ldots, \widehat{A}_b]^T := \frac{1}{(2b+1)} \Phi_b^\dagger \vec{G} \quad \text{and} \quad \widehat{G}(t) = \sum_{k=-b}^{b} \widehat{A}_k \exp(j2\pi k t)$$

Consistency of our estimate

Define $\overrightarrow{G} = (g(U_{1:n}),\ g(U_{nsb:n}),\ \ldots,\ g(U_{2bsbn:n}))$, and the Fourier series and field estimates as

$$\vec{A} = \frac{1}{(2b+1)}\Phi_b^{\dagger}\vec{G} \quad \text{and} \quad \widehat{G}(t) = \sum_{k=-b}^{b} \widehat{A}_k \exp(j2\pi kt)$$

Then

Theorem 1: $\vec{A} \xrightarrow{a.s} \vec{a}, \widehat{G}(t) \xrightarrow{a.s} g(t)$ and $\vec{A} \xrightarrow{\mathcal{L}^2} \vec{a}, \widehat{G}(t) \xrightarrow{\mathcal{L}^2} g(t)$

Key ideas:

◊ For $r = [np] + 1$, $U_{r:n} \to p$ almost surely

◊ That is $U_{[nsb]:n} \to s_b$ almost surely, $U_{[2nsb]:n} \to 2s_b$ almost surely, etc.

◊ By continuity of $g(t)$, $g(U_{[nsb]:n}) \to g(s_b)$ almost surely, $g(U_{[2nsb]:n}) \to g(2s_b)$ almost surely, etc.

◊ Finally, the estimates $\overrightarrow{A}$ and $\hat{G}(t)$ are bounded-coefficient finite linear combination of $g(U_{1:n}), g(U_{[nsb]:n}), \ldots, g(U_{[2b\ nsb]:n})$

Mean-squared error performance

If $r \approx [np]$ then the second moment of $(U_{r:n} - p)$ satisfies

$$\left.\begin{aligned} n\mathbb{E}(U_{r:n}-p)^2 &= p(1-p)\mathbb{E}(Z^2) + O(\sqrt{1/n}) \\ &\leq \frac{1}{4} + O(\sqrt{1/n}). \end{aligned}\right\}$$

[David and Nagaraja'2003]

Keep in mind that

$$\vec{A} = \frac{1}{(2b+1)}\Phi_b^{\dagger}\vec{G} \quad \text{and} \quad \widehat{G}(t) = \sum_{k=-b}^{b} \widehat{A}_k \exp(j2\pi kt)$$

Then the following mean-squared result holds for the estimate $\hat{G}(t)$

Theorem 2: $n\mathbb{E}\left[||\widehat{G} - g||_2^2\right] \leq \pi^2 b^2 (2b+1)\left[1 + O(\sqrt{1/n})\right]$

Key ideas:

◊ The matrix Φ_b has entries with magnitude $|(\phi_j)^k| = 1$. The signal's derivative $g'(t)$ is bounded. As a result, linear approximations can be used to get the above bound

◊ Observe that the mean-squared error decreases as $O(1/N)$

Weak-convergence of the estimate $\hat{G}(t)$

Fact: If $0 < p_1 < p_2 < \ldots < p_{(2b+1)} < 1$ and $(r_i/n - p_i) = o(1/\sqrt{n})$ for each i. Then,

$$\sqrt{n}[U_{r_1:n} - p_1, \ldots, U_{r_{2b+1}:n} - p_{2b+1}]^T \xrightarrow{d} \mathcal{N}\left(\vec{0}, K_U\right)$$

where $[K_U]_{j,j'} = p_j(1 - p_{j'})$ for $j \leq j'$. **[David and Nagaraja'2003]**

Once again

$$\vec{A} = \frac{1}{(2b+1)} \Phi_b^{\dagger} \vec{G} \quad \text{and} \quad \widehat{G}(t) = \sum_{k=-b}^{b} \widehat{A}_k \exp(j 2\pi k t)$$

Then the following mean-squared result holds for the estimate $\hat{G}(t)$

Theorem 2: $$\sqrt{n}(\widehat{G}(t) - g(t)) \xrightarrow{d} \mathcal{N}\left(\vec{0}, K_G(t)\right)$$

where, the variance $K_G(t)$ depends on K_U, the derivative of $g(t)$, and Φ_b

Key ideas:

◊ $(U_{1:n}, U_{sbn:n}, \ldots, U_{2bsbn:n})$ converges to a Gaussian vector

◊ Since $g(t)$ is smooth, therefore $\vec{G} = g(U_{1:n}), g(U_{sbn:n}), \ldots, g(U_{2bsbn:n})$ converges to a Gaussian vector by the Delta method **[van der Vaart'1998]**

Weak-convergence of the estimate $\hat{G}(t)$

Theorem 2: $$\sqrt{n}(\widehat{G}(t) - g(t)) \xrightarrow{d} \mathcal{N}\left(\vec{0}, K_G(t)\right)$$

where, the variance $K_G(t)$ depends on K_U, the derivative of $g(t)$, and Φ_b

Key ideas (contd.):

◊ Since the map from $\overrightarrow{G} = (g(U_{1:n}), g(U_{sbn:n}), \ldots, g(U_{2bsbn:n}))$ to $\hat{G}(t)$ is linear, therefore, Gaussian distribution is preserved

Conclusions and future work

◊ Estimation of a bandlimited field from samples taken at uniformly distributed but unknown locations was considered

◊ A bandlimited field *cannot be* uniquely determined with (perfect) samples obtained at statistically distributed locations, even if the number of samples is infinite

◊ If the **order** (left to right) of sample locations is known, a consistent estimate $\hat{G}(t)$ for the field of interest can be obtained

- The estimate $\hat{G}(t)$ converges in almost-sure sense and mean-square sense to the true field $g(t)$
- The mean-squared error between $\hat{G}(t)$ and $g(t)$ decreases as $O(1/n)$
- This leads to a central-limit type weak-law

◊ Extensions of this result to field affected by noise and multidimensional field is of immediate interest