Appendix A
Formal Proof Details, Function Spaces, and Convergence Structure for the KAS Sampling Theorem

A.1 Functional-Analytic and Probabilistic Setting

Let $PW_B^2 \subset L^2(\mathbb{R})$ denote the Paley–Wiener space of real-valued functions bandlimited to $[-B, B]$. For any finite observation window $[0, T]$, restriction of functions in PW_B^2 yields a bounded and equicontinuous subset of $C([0, T])$ (Wiener & Paley, 1934).

For stochastic signals, all expectations and L^2 norms are taken with respect to a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Throughout, stochastic processes are assumed to satisfy:

- zero mean,
- finite second moments,
- almost-sure membership of sample paths in $PW_B^2([0, T])$ for every finite T.

All convergence statements for stochastic signals are understood in $L^2(\mathbb{P})$ unless explicitly stated otherwise.

A.2 Continuity of RF Functionals on Bounded Sets

The RF functionals considered—average power and finite-lag correlation— are continuous on bounded subsets of $L^2([0, T])$.

Lemma A.1 (Continuity of Average Power). For $X, Y \in L^2([0, T])$,

$$\big|P_T(X) - P_T(Y)\big| \le \frac{1}{T}\|X - Y\|_{2,[0,T]}\big(\|X\|_{2,[0,T]} + \|Y\|_{2,[0,T]}\big).$$

Proof. This follows directly from polarization and the Cauchy–Schwarz inequality.

Lemma A.2 (Continuity of Lag Correlation). For fixed $\tau \ge 0$, the functional

$$R_T(\tau; X) = \frac{1}{T - \tau}\int_0^{T-\tau} X(t)X(t + \tau)\, dt$$

is continuous on bounded subsets of $L^2([0, T])$.

Proof. Apply Hölder's inequality together with boundedness on finite windows.

A.3 Sampling as an Injective Operator on PW_B^2

Let $f_s \ge 2B$ and $\Delta = 1/f_s$. The sampling operator

$$S_{f_s} : PW_B^2 \to \ell^2, \qquad X \mapsto \big(X(n\Delta)\big)_{n\in\mathbb{Z}},$$

is injective and admits a bounded left inverse R_{f_s} on restrictions to finite windows $[0, T]$.

On any finite window,

$$\|X - R_{f_s} S_{f_s} X\|_{2,[0,T]} \to 0$$

as the number of reconstruction terms increases. This justifies replacing continuous-time functionals by sample-based approximations without loss of functional information.

A.4 Riemann and Bochner Approximation of Time-Averaged Functionals

Let $g : \mathbb{R}^m \to \mathbb{R}$ be bounded and continuous. For deterministic X,

$$\frac{1}{T}\int_0^T g\big(X(t), \dots, X(t-(m-1)\Delta)\big)\,dt$$

converges uniformly to its discrete Riemann sum as $T = N\Delta \to \infty$.

For stochastic X, the integral is interpreted as a Bochner integral in $L^2(\mathbb{P})$ (Billingsley, 1999), and convergence holds in:

- $L^2(\mathbb{P})$ by dominated convergence,
- probability under finite-moment assumptions.

Almost-sure convergence is not required for any result in the paper.

A.5 Compactness of Lag-Window Domains

Under boundedness on $[0, T]$, all lag vectors

$$x_{n,m} = (x_n, \dots, x_{n-m+1})$$

lie in the compact hypercube $K = [-M, M]^m \subset \mathbb{R}^m$. This compactness enables uniform approximation by spline-KANs.

A.6 Universality of Spline-KANs on Compact Sets

By the Kolmogorov–Arnold representation theorem, any continuous $h : K \to \mathbb{R}$ admits a representation

$$h(z) = \sum_{q=1}^{2m+1} \Phi_q\left(\sum_{i=1}^{m} \Psi_{qi}(z_i)\right).$$

Cubic B-splines are dense in $C([a, b])$ under the uniform norm. Approximating each univariate component uniformly yields a spline-KAN Φ satisfying

$$\sup_{z\in K} |h(z) - \Phi(z)| < \varepsilon.$$

A.7 Cylinder Approximation of RF Functionals

On bounded subsets of $PW_B^2([0, T])$, RF functionals admit representations

$$J_\Delta^{(m)}(X) = \frac{1}{N}\sum_{n=m-1}^{N-1} g(x_{n,m}),$$

with bounded continuous g, such that

$$|J(X) - J_\Delta^{(m)}(X)| \le \varepsilon$$

uniformly in the deterministic case and in $L^2(\mathbb{P})$ for stochastic signals.

A.8 Explicit Error Decomposition

The total estimation error decomposes as

$$\varepsilon = \varepsilon_{\text{sampling}} + \varepsilon_{\text{discretization}} + \varepsilon_{\text{approx}},$$

corresponding to Nyquist truncation, Riemann/cylinder approximation, and spline-KAN approximation.

A.9 Noise Averaging Rate

Let $Y(t) = X(t) + \eta(t)$ with zero-mean noise of finite variance. Then

$$\mathbb{E}\Big[(\hat{J}(Y) - \hat{J}(X))^2\Big] = \mathcal{O}(N^{-1}),$$

independent of the spline-KAN mapping, due to boundedness and averaging (Gray, 2009).

A.10 Convergence Modes and Interpretation

- Uniform convergence applies to deterministic bounded subsets.
- $L^2(\mathbb{P})$ convergence applies to stochastic RF environments.
- No almost-sure convergence is claimed or required.

A.11 Scope and Non-Assumptions

The KAS Sampling Theorem does *not* assume Gaussianity, reconstruction, linearity, or overparameterized networks—only band limitation, boundedness on finite windows, and second-order moments.

A.12 Reproducibility Link

All constructions and bounds correspond to executable experiments in the Testing and Evidence Notebook: `https://github.com/XXJanusXX/KAS`.

A.13 Measurability of the Estimator

The estimator

$$\hat{J}(X) = \frac{1}{N} \sum_{n=m-1}^{N-1} \Phi(x_{n,m})$$

is measurable as a finite composition of measurable mappings, and hence a well-defined random variable in $L^2(\mathbb{P})$.

A.14 Non-Asymptotic Interpretation

All bounds are non-asymptotic. For any finite T and tolerance $\varepsilon > 0$, finite choices of f_s, m, N, and spline parameters suffice. No limit $T \to \infty$ is required.