arXiv:2606.03942v3 [eess.SP] 4 Oct 2026

Sampling with Autoregressive Clock Jitter: Infinite-Dimensional Stability Failure and Finite-Dimensional Conditioning

Daniele Gerosa[a,*], Thomas Eriksson[a]

[a] *Communication, Antennas and Optical Networks, Electrical Engineering Dept., Chalmers University of Technology, Chalmersplatsen 1, Göteborg, 41296, Sweden*

Abstract

Motivated by recent developments in stochastic modeling of clock jitter in Analog-to-Digital Converters (ADCs) as autoregressive processes of order one (AR(1)), we study the density and stability properties of AR(1)-jittered sampling sets for Paley-Wiener signals. While the stability of sampling sets under i.i.d. perturbations of the integer lattice is classical, no analogous analysis was previously available for the correlated jitter observed in practice. We show that, despite having the correct asymptotic density both on average and almost surely, such sets almost surely fail to be stable sampling sets. We complement this negative result with a finite-dimensional analysis, providing explicit non-asymptotic conditions on the number of samples $2N + 1$, the jitter standard deviation σ_ξ and the confidence level $1 - \eta$ under which the smallest singular value of the $(2N+1) \times (2N+1)$ jittered sinc matrix exceeds

*© 2026. This manuscript version is made available under the CC-BY-NC-ND 4.0 license https://creativecommons.org/licenses/by-nc-nd/4.0/

*Corresponding author

Email addresses: danieleg@chalmers.se (Daniele Gerosa), thomase@chalmers.se (Thomas Eriksson)

a prescribed threshold ϵ with probability at least $1-\eta$, so that the matrix is well-conditioned with high probability.

Keywords: autoregressive sampling, jitter, stable sampling

---

1. Introduction

In signal processing, an analog waveform is typically modeled as a function of a continuous time parameter $t \in \mathbb{R}$. Such functions, also called *analog signals*, can be deterministic or random; in the latter case, we usually speak of continuous-time stochastic processes. Analog-to-digital converters (ADCs) are hardware systems that convert an analog signal into a digital representation, and one of the basic steps in this operation is *sampling*. Samples may be described as instantaneous point evaluations of s at t_n, i.e. $s(t_n)$, or as averages over the sampling interval $T_s > 0$, i.e. $\frac{1}{T_s}\int_{nT_s}^{(n+1)T_s} s(t)\,dt$ [1]. In this work, sampling will always be understood as point evaluation.

An analog signal s whose frequency spectrum is bounded can be fully reconstructed from its samples using the classical Shannon-Whittaker-Nyquist-Kotelnikov theorem (or some extension thereof), provided that 1. all the infinitely many sample values $\{s(t_n)\}_{n\in\mathbb{Z}}$ are accessible, 2. the sampling interval T_s is sufficiently small, and 3. the sampling instants t_n are either (a) exact multiples of T_s, or (b) form an irregular sampling set satisfying suitable stability conditions; density conditions provide important necessary constraints, see e.g. Chapter 10 in [2]. The theory concerning the reconstruction of bandlimited signals from samples taken on irregular time grids is by now mathematically mature, having been developed from multiple perspectives [2, 3, 4, 5, 6, 7, 8, 9]. As an example, the following theorem provides a coun-

terpart to the Shannon-Whittaker-Nyquist-Kotelnikov theorem for irregular sampling grids:

Theorem 1 (10.9-10.11 in [2])**.** *Let* $\{t_n\}_{n\in\mathbb{Z}} \subseteq \mathbb{R}$ *satisfy Kadec's "one quarter" condition*

$$|n - t_n| \leq D < 1/4 \qquad \text{[K]}$$

for all $n \in \mathbb{Z}$*, and let* $s : \mathbb{R} \to \mathbb{R}$ *such that supp*$(\mathcal{F}(s)) \subseteq [-1/2, 1/2]$*. Then*

$$s(t) = \sum_{n\in\mathbb{Z}} s(t_n) \frac{L(t)}{L'(t_n)(t - t_n)}$$

in norm and globally uniformly over $\mathbb{R}$*, with*

$$L(t) = (t - t_0) \prod_{n=1}^{\infty} \left(1 - \frac{t}{t_n}\right) \left(1 - \frac{t}{t_{-n}}\right). \qquad (1)$$

Although the product representation in (1) is generally not explicit, it reduces to closed-form expressions in some special cases [2]. An extensive compendium of sampling theorems can be found in [10], chapters 2 and 3.

In concrete applications, hardware practitioners still tend to rely on the standard sinc interpolation series, mostly for reasons of complexity and computational cost [11]. Furthermore, in practice the actual sampling instants t_n are unknown and must be estimated. There is also significant evidence, both from measurements and from the properties of oscillators, that the sampling instants t_n are random in nature [12] and should therefore be modeled as a stochastic process. Building on [13, 14], the authors of [15] recently proposed to model the correlated component of the stochastic clock jitter caused by oscillator imperfections in ADCs using a two-sided discrete autoregressive

process of order 1 (AR(1)), so that $t_n = T_s n + \xi_n$, where

$$\xi_n = \sum_{k=0}^{\infty} \varphi^k \epsilon_{n-k} \tag{2}$$

and $\{\epsilon_n\}_{n\in\mathbb{Z}}$ is a sequence of independent and identically distributed Gaussian random variables, $\epsilon_n \sim \mathcal{N}(0, \sigma_\epsilon^2)$ for all $n \in \mathbb{Z}$, with $0 < \varphi < 1$ and φ typically close to 1. The process (2) has power spectral density [16]

$$\mathcal{S}_\xi(e^{i\omega}) = \frac{\sigma_\epsilon^2}{1 + \varphi^2 - 2\varphi\cos(\omega)}, \quad \omega \in [-\pi, \pi],$$

so that requiring $\varphi \approx 1$ correctly encodes the fact that this jitter is "slowly varying" and that much of its energy is therefore concentrated at low frequencies.

Since $\sum_{k\geq 0} \operatorname{var}(\varphi^k \epsilon_{n-k}) = \sigma_\epsilon^2 \sum_{k\geq 0} \varphi^{2k} < \infty$, the Khintchine-Kolmogorov convergence theorem ensures that the infinite sum in (2) converges almost surely for every n, and thus the ξ_n are well-defined. Lévy's continuity theorem then yields $\xi_n \sim \mathcal{N}(0, \sigma_\epsilon^2/(1-\varphi^2))$ for all $n \in \mathbb{Z}$. We remark that, since the ξ_n are Gaussian and hence unbounded, the jittered lattice $t_n = T_s n + \xi_n$ almost surely fails to satisfy Kadec's condition [K], so that Theorem 1 does not apply to the model (2).

1.1. Basic definitions

We recall here basic definitions.

Definition 1 (Fourier transform)**.** The *Fourier transform* $\mathcal{F}(s)(\omega)$ of a function $s : \mathbb{R} \to \mathbb{C}$ is formally defined as

$$\mathcal{F}(s)(\omega) \coloneqq \int_{-\infty}^{\infty} s(t) e^{-2\pi i \omega t}\, dt.$$

$\mathcal{F}$ defines a bounded linear operator $L^1(\mathbb{R}) \to L^\infty(\mathbb{R})$ that can be extended to all of $L^2(\mathbb{R})$ by standard density arguments. If $B \subseteq \mathbb{R}$ is a union of finitely many measurable sets of finite and positive Lebesgue measure, we define

Definition 2 (Paley-Wiener space)**.** The *Paley-Wiener space* is the set

$$\mathrm{PW}_B \coloneqq \{s \in C(\mathbb{R}) \cap L^2(\mathbb{R}) \,:\, \mathrm{supp}(\mathcal{F}(s)) \subseteq \mathrm{cl}(B)\}.$$

In sampling theory, one basic question is whether a signal can be recovered from its samples. For applications, however, uniqueness alone may not be enough: the recovery should also be stable with respect to measurement, numerical, or modeling errors. In other words, a signal with large L^2-energy should not be almost "invisible" on the sampling set. This robustness requirement is encoded in the notion of *stable sampling*.

Definition 3 (Stable sampling)**.** A set $\{t_n\} \subseteq \mathbb{R}$ is of *stable sampling* for PW_B if

$$s \in \mathrm{PW}_B \Longrightarrow \|s\|_{L^2} \leq K\|s(t_n)\|_{\ell^2}$$

with K independent of s.

Sampling stability for Paley-Wiener functions is connected to the notion of *density*, at least through a theorem of H. J. Landau [17], which states that the density (3) of a sampling set must be at least the Lebesgue measure of B. In this note we adopt the following classical definition of density.

Definition 4 (Sampling set density)**.** The density $D(t_n)$ of a sampling set $\{t_n\} \subseteq \mathbb{R}$ is defined as

$$D(t_n) \coloneqq \lim_{r\to\infty} \frac{\mathrm{card}(\{t_n \,:\, t_n \in [-r, r]\})}{2r} \tag{3}$$

whenever the latter limit exists.

1.2. Symbols and notations

We adopt the notation of using uppercase boldface letters $\mathbf{A}$ for matrices and lowercase boldface letters $\mathbf{a}$ for column vectors. Due to the structural problem setup, we indicate with $\mathbf{a}_N$ vectors belonging to $\mathbb{R}^{2N+1}$ indexed from $-N$ to N. Similarly for $(2N+1)\times(2N+1)$ matrices, that will be indicated by $\mathbf{A}_N$ or $\mathbf{A}^{(N)}$ if ambiguities may arise. We use $\sigma_j(\mathbf{A}^{(N)})$ for the $j-$th singular value of $\mathbf{A}^{(N)}$, and $\odot$ for the Hadamard component-wise product. The operator norm $\|\mathbf{A}^{(N)}\|_{\mathrm{op}}$ of a matrix equals its largest singular value $\sigma_{\max}(\mathbf{A}^{(N)})$.

1.3. Motivations

The study of the properties of systems of complex exponentials $\{e^{i\lambda_n x} : \lambda_n \in \Lambda\}$, with $\Lambda \subseteq \mathbb{R}$, is classical and the corresponding theory is well developed, cf. [18]. Roughly speaking, "good" properties of the $\{e^{i\lambda_n x}\}$ imply "good" properties of the set Λ seen as a sampling set for Paley-Wiener functions, and vice versa [2]. Much less is known when the λ_n are assumed to be random variables, although a fair amount has been established in the i.i.d. case [8, 9, 19, 7]; when correlation enters the picture, the results are quite sparse. We are not aware of results addressing sampling stability for correlated random perturbations of the integer lattice; the present note is a modest attempt to fill this gap, taking as a starting point the autoregressive jitter model of [15]. The mathematical machinery used in [8, 9, 19, 7] does not generalize well to the correlated scenario: independence is used there to

reduce the analysis to product measures and to invoke Borel-Cantelli-type arguments on independent events, neither of which is available for a correlated jitter process. We therefore rely on the ergodicity of the shift acting on the AR(1) path, which supplies the required almost sure statements without any independence assumption. Concretely, we investigate whether the autoregressively jittered lattice $t_n = n + \xi_n$ retains the stable sampling properties of Definition 3 of the integer lattice.

1.4. Contributions

The contributions of this manuscript may be summarized as follows:

- We prove that the autoregressive jittered lattice $t_n = n + \xi_n$, with ξ_n defined in (2), has the correct critical density (3) on average (Lemma 2.1) and almost surely (Lemma 2.2);
- We prove that $\{t_n\}$ is *not* of stable sampling for the Paley-Wiener space (Theorem 2);
- We prove that the corresponding sinc matrices are well-conditioned with high probability, both for the full sampling matrix (Proposition 3) and for its first-order Taylor surrogate (Proposition 4);
- We complement these results with a Monte Carlo study (Section 4) of the empirical distributions of the smallest singular values and condition numbers, and we numerically quantify the conservativeness of the bounds obtained above.

2. Critical density and global stability failure

As observed by Bass and Gröchenig in [8], random sampling sets may, on average, attain the critical (lower Beurling) density required by Landau's results and still fail to be stable. We show in this section that this is precisely what happens for the AR(1)-jittered lattice.

Lemma 2.1. *Let $\{\eta_n\}_{n\in\mathbb{Z}}$ be a sequence of jointly Gaussian real-valued random variables with 0 mean and variance σ_η^2. Define $\theta_n \coloneqq n + \eta_n$. Then*

$$\frac{\mathbb{E}[\mathrm{card}(\{\theta_n \,:\, \theta_n \in [-r,r]\})]}{2r} = 1 + O(r^{-1}).$$

Proof. Consider the nonnegative random variables $\mathbf{1}_{\{|\theta_n|\le r\}}$. By the Tonelli theorem $\mathbb{E}[\sum_{n\in\mathbb{Z}} \mathbf{1}_{\{|\theta_n|\le r\}}] = \sum_{n\in\mathbb{Z}} \mathbb{E}[\mathbf{1}_{\{|\theta_n|\le r\}}]$ and

$$\begin{aligned}
\sum_{n\in\mathbb{Z}} \mathbb{E}[\mathbf{1}_{\{|\theta_n|\le r\}}] &= \frac{1}{\sqrt{2\pi}\sigma_\eta} \sum_{n\in\mathbb{Z}} \int_{-r}^{r} e^{-\frac{(s-n)^2}{2\sigma_\eta^2}}\, ds \\
&= \frac{1}{\sqrt{2\pi}\sigma_\eta} \int_{-r}^{r} \sum_{n\in\mathbb{Z}} e^{-\frac{(s-n)^2}{2\sigma_\eta^2}}\, ds \qquad (4) \\
&= \frac{1}{\sqrt{2\pi}\sigma_\eta} \int_{-r}^{r} e^{-\frac{s^2}{2\sigma_\eta^2}} \boxed{\sum_{n\in\mathbb{Z}} e^{\frac{-n^2+2sn}{2\sigma_\eta^2}}}\, ds.
\end{aligned}$$

The boxed term is the Jacobi's theta function $\vartheta(z;\tau)$ with $\tau = i/(2\pi\sigma_\eta^2)$ and $z = -is/(2\sigma_\eta^2)$. The whole integrand becomes then equal to

$$\begin{aligned}
&= e^{-s^2/(2\sigma_\eta^2)} \vartheta\left(-\frac{is}{2\sigma_\eta^2}; \frac{i}{2\pi\sigma_\eta^2}\right) \\
&= e^{-s^2/(2\sigma_\eta^2)} \sqrt{2\pi\sigma_\eta^2} e^{s^2/(2\sigma_\eta^2)} \vartheta(\pi s; 2\pi i\sigma_\eta^2) \\
&= \sqrt{2\pi}\sigma_\eta \vartheta(\pi s; 2\pi i\sigma_\eta^2)
\end{aligned}$$

by using modular properties of the Theta function. (4) becomes then

$$\begin{aligned}\int_{-r}^{r} \vartheta(\pi s; 2\pi i\sigma_\eta^2)\, ds &= \int_{-r}^{r} \left[1 + 2\sum_{n=1}^{\infty} e^{-2\pi^2\sigma_\eta^2 n^2} \cos(2\pi n s)\right] ds \\ &= 2r + 4\sum_{n=1}^{\infty} e^{-2\pi^2\sigma_\eta^2 n^2} \frac{\sin(2\pi n r)}{2\pi n}\end{aligned}$$

from which we have that, as $r \to \infty$,

$$\frac{\mathbb{E}[\mathrm{card}(\{\theta_n \,:\, \theta_n \in [-r, r]\})]}{2r} = 1 + O(r^{-1}).$$

□

So far we have not exploited the autoregressive structure of our jitter. Indeed, Lemma 2.1 assumes nothing more than Gaussianity of the η_n. In the next lemma we show that, under a mild summability assumption on the autocorrelation function, the convergence to the correct density can be upgraded from expectation to almost sure convergence.

Lemma 2.2. *Let $\{\xi_n\}_{n\in\mathbb{Z}}$ be a centered stationary Gaussian process with variance σ_ξ^2, and let*

$$\rho(k) \coloneqq \mathrm{Corr}(\xi_n, \xi_{n+k}), \qquad k \in \mathbb{Z},$$

denote its autocorrelation function. Assume that

$$\sum_{k\in\mathbb{Z}} |\rho(k)| < \infty.$$

Define $\theta_n \coloneqq n + \xi_n$. Then

$$S(r) \coloneqq \frac{1}{2r} \sum_{n\in\mathbb{Z}} \mathbf{1}_{\{|\theta_n| \le r\}} \to 1$$

a.s. as $r \to \infty$.

Proof. We study $\mathrm{var}(S(r))$ for a fixed r. We have

$$\begin{aligned}\mathrm{var}(S(r)) &= \frac{1}{4r^2}\left[\sum_{n\in\mathbb{Z}}\mathrm{var}\big(\mathbf{1}_{\{|\theta_n|\le r\}}\big) + 2\sum_{m<n}\mathrm{cov}\big(\mathbf{1}_{\{|\theta_m|\le r\}},\mathbf{1}_{\{|\theta_n|\le r\}}\big)\right]\\ &= \frac{1}{4r^2}\left[\sum_{n\in\mathbb{Z}}p_{n,r}(1-p_{n,r}) + 2\sum_{m<n}\mathrm{cov}\big(\mathbf{1}_{\{|\theta_m|\le r\}},\mathbf{1}_{\{|\theta_n|\le r\}}\big)\right].\end{aligned} \tag{5}$$

with

$$p_{n,r} = \frac{1}{\sqrt{2\pi}\sigma_\xi}\int_{-r}^{r} e^{-\frac{(s-n)^2}{2\sigma_\xi^2}}\,ds.$$

To deal with the term $\sum_n p_{n,r}(1-p_{n,r})$ we modify the argument in (4):

$$\begin{aligned}p_{n,r}(1-p_{n,r}) &= \frac{1}{2\pi\sigma_\xi^2}\int_{-r}^{r}\int_{\mathbb{R}\setminus[-r,r]} e^{-\frac{(s-n)^2}{2\sigma_\xi^2}}\,e^{-\frac{(t-n)^2}{2\sigma_\xi^2}}\,ds\,dt\\ &= \frac{1}{2\pi\sigma_\xi^2}\int_{-r}^{r}\int_{\mathbb{R}\setminus[-r,r]} e^{-\frac{(s-t)^2}{4\sigma_\xi^2}}\,e^{-\frac{((t+s)/2-n)^2}{\sigma_\xi^2}}\,ds\,dt.\end{aligned}$$

and

$$\sum_{n\in\mathbb{Z}} e^{-((t+s)/2-n)^2/\sigma_\xi^2} \le C(\sigma_\xi)$$

because the infinite sum defines a continuous and 1-periodic function. Therefore

$$\sum_n p_{n,r}(1-p_{n,r}) \le C_1(\sigma_\xi)\int_{-r}^{r}\int_{\mathbb{R}\setminus[-r,r]} e^{-\frac{(s-t)^2}{4\sigma_\xi^2}}\,ds\,dt. \tag{6}$$

To bound the right-hand side of (6) it is enough to observe that

$$\int_r^\infty e^{-(s-t)^2/(4\sigma_\xi^2)}\,ds \lesssim e^{-(r-t)^2/(4\sigma_\xi^2)}$$

(and similarly for $\int_{-\infty}^{-r}$) as a consequence of elementary bounds of the Error function, and hence (6) is bounded from above by a constant independent of r.

We need now to estimate the series of covariances in (5), and to do so we make use of Gebelein [20] as well as the AM-GM inequalities:

$$\begin{aligned}|\operatorname{cov}(\mathbf{1}_{\{|\theta_m|\le r\}}, \mathbf{1}_{\{|\theta_n|\le r\}})| &\le |\rho(n-m)|\sqrt{\operatorname{var}(\mathbf{1}_{\{|\theta_m|\le r\}})\operatorname{var}(\mathbf{1}_{\{|\theta_n|\le r\}})} \\ &\le |\rho(n-m)|\frac{\operatorname{var}(\mathbf{1}_{\{|\theta_m|\le r\}}) + \operatorname{var}(\mathbf{1}_{\{|\theta_n|\le r\}})}{2}.\end{aligned}$$

Thus

$$\sum_{m<n} |\rho(n-m)|\operatorname{var}(\mathbf{1}_{\{|\theta_m|\le r\}}) = \sum_{m\in\mathbb{Z}} \operatorname{var}(\mathbf{1}_{\{|\theta_m|\le r\}}) \sum_{h=1}^{\infty} |\rho(h)| \le C_2(\sigma_\xi, \rho),$$

and similarly for the series with $\operatorname{var}(\mathbf{1}_{\{|\theta_n|\le r\}})$. We can finally conclude that

$$\operatorname{var}(S(r)) \le C_3(\sigma_\xi, \rho)/r^2, \tag{7}$$

which together with Lemma 2.1 implies that $S(r) \to 1$ in mean square, and hence in probability. Convergence almost everywhere requires a bit more work: we first show that $S(n) \to 1$ almost everywhere along integers, and then extend the result to arbitrary sequences $r_k \to \infty$.

As a consequence of Chebyshev's inequality and (7), for a fixed $\epsilon > 0$ we have that

$$\mathbb{P}[\{|S(n) - \mathbb{E}[S(n)]| > \epsilon\}] \lesssim \frac{1}{\epsilon^2 n^2}$$

and thus

$$\sum_{n\ge 1} \mathbb{P}[\{|S(n) - \mathbb{E}[S(n)]| > \epsilon\}] < \infty.$$

The Borel-Cantelli lemma then gives

$$\mathbb{P}\left[\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{|S(k) - \mathbb{E}[S(k)]| > \epsilon\}\right] = 0, \tag{8}$$

which implies $S(n) - \mathbb{E}[S(n)] \to 0$ almost surely. Combined with Lemma 2.1, this yields $S(n) \to 1$ almost surely.

To conclude the argument, pick a sequence $r_k \to \infty$ and define the two integers $m_k^- = \lfloor r_k \rfloor - 1$ and $m_k^+ = \lfloor r_k \rfloor + 1$, so that $m_k^- \leq r_k \leq m_k^+$ for all k (the m's depend, of course, on the sequence r_k). Then

$$S(r_k) \leq \frac{1}{2r_k} \sum_{n \in \mathbb{Z}} \mathbf{1}_{\{|\theta_n| \leq m_k^+\}} = \frac{m_k^+}{r_k} S(m_k^+)$$

and similarly

$$S(r_k) \geq \frac{m_k^-}{r_k} S(m_k^-).$$

Since by construction both m_k^-/r_k and m_k^+/r_k tend to 1 and $S(m_k^-)$, $S(m_k^+)$ converge to 1 a.s., then so does $S(r_k)$. □

Remark 1. Strictly speaking, the sum $\sum_{n \in \mathbb{Z}} \mathbf{1}_{\{|\theta_n| \leq r\}}$, which appears in the proof of Lemma 2.1 and equals $2rS(r)$ in Lemma 2.2, and $\mathrm{card}(\{\theta_n : \theta_n \in [-r, r]\})$ are not exactly the same quantity; however, since joint Gaussianity implies that $\bigcap_{n \neq m} \{\theta_n \neq \theta_m\}$ has probability 1, the two agree almost surely for every $r > 0$.

We are now in position to prove the main negative result of this section. In the following theorem we adapt the construction of Proposition 2.2 in [8] to show that the stable sampling condition introduced above is violated by the autoregressively jittered lattice. The bandlimited functions used in the counterexample are precisely those of [8]; the proof strategy, however, must be adapted to account for the fact that the random variables ξ_n are no longer independent.

Theorem 2. *Let $\theta_n(\omega) \coloneqq n + \xi_n(\omega)$, where $\{\xi_n\}_{n \in \mathbb{Z}}$ is the two-sided Gaussian AR(1) process defined in* (2)*, with $0 < \varphi < 1$. Then with probability* 1 *there*

exists a sequence of nonzero bandlimited signals $\{s_k^\omega\}_{k\in\mathbb{N}} \subseteq \mathrm{PW}_{[-1/2,1/2]}$ *such that*

$$\sum_{n\in\mathbb{Z}} |s_k^\omega(\theta_n(\omega))|^2 \leq \frac{\|s_k^\omega\|_2^2}{k} \quad \forall\, k.$$

Consequently, the jittered lattice is almost surely not a stable sampling set for $\mathrm{PW}_{[-1/2,1/2]}$.

Proof. We work in the critical normalization $B = [-1/2, 1/2]$. The process $\{\xi_n\}_{n\in\mathbb{Z}}$ is realized as a measurable map $\boldsymbol{\xi} : (\Omega', \mathbb{P}) \to \mathbb{R}^{\mathbb{Z}}$ on an underlying probability space $(\Omega', \mathbb{P})$, and we denote by $\mu \coloneqq \mathbb{P} \circ \boldsymbol{\xi}^{-1}$ its law on $\mathbb{R}^{\mathbb{Z}}$. All the events introduced below are sets in $\mathbb{R}^{\mathbb{Z}}$; for such an event E we write $\mathbb{P}(\boldsymbol{\xi} \in E) = \mu(E)$, and the finite-dimensional marginals of μ are the corresponding Gaussian laws of the relevant coordinates of $\boldsymbol{\xi}$.

The function $S(x) \coloneqq \sin(\pi x/2)\, \Psi(x)$, with Ψ as in the proof of Proposition 2.2 in [8], is a nonzero element of $\mathrm{PW}_{[-1/2,1/2]}$ satisfying $S(2j) = 0$ for all $j \in \mathbb{Z}$, with $|S(x)| \leq a/(1 + |x|^2)$ for some positive constant a and all $x \in \mathbb{R}$. For each fixed k, we introduce the two events

$$A_{\delta_k, N_k} \coloneqq \{x \in \mathbb{R}^{\mathbb{Z}} : |x_{2j}| \leq \delta_k,\ |x_{2j+1} - 1| \leq \delta_k,\ |j| \leq N_k\}$$

and

$$B_{N_k} \coloneqq \{x \in \mathbb{R}^{\mathbb{Z}} : |x_j| \leq |j|/4,\ j \leq -(2N_k + 1) \text{ or } j \geq 2N_k + 2\}.$$

Since S is bounded, so is S', by e.g. Theorem 11.1.2 in [21]. We then choose $\delta_k^2 = \|S\|_2^2/(8N_k k\|S'\|_\infty^2) = C_1/(kN_k)$. Using the autoregressive structure of

ξ_n, the random vector $(\xi_{-2N_k}, \dots, \xi_{2N_k+1})$ has precision matrix [22]

$$\mathbf{\Sigma}^{-1}_{4N_k+2} = \frac{1}{\sigma_\epsilon^2} \begin{pmatrix} 1 & -\varphi & 0 & \dots & 0 \\ -\varphi & 1+\varphi^2 & -\varphi & \dots & 0 \\ \vdots & \ddots & \ddots & \dots & \vdots \\ 0 & 0 & -\varphi & 1+\varphi^2 & -\varphi \\ 0 & 0 & 0 & -\varphi & 1 \end{pmatrix} \tag{9}$$

and $\det\left(\mathbf{\Sigma}^{-1}_{4N_k+2}\right) = (1-\varphi^2)/\sigma_\epsilon^{2(4N_k+2)}$; equivalently, the marginal of μ on the coordinates $-2N_k, \dots, 2N_k+1$ has density

$$f_{4N_k+2}(\mathbf{x}) = \frac{\sqrt{1-\varphi^2}}{(\sqrt{2\pi}\sigma_\epsilon)^{4N_k+2}} \exp\left(-\mathbf{x}^t \mathbf{\Sigma}^{-1}_{4N_k+2}\mathbf{x}/2\right).$$

By using (9) we can write explicitly the quadratic form in the exponential, which gives us

$$\begin{aligned} \sigma_\epsilon^2 \mathbf{x}^t \mathbf{\Sigma}^{-1}_{4N_k+2}\mathbf{x} &= x_1^2 + x_{4N_k+2}^2 + (1+\varphi^2)\sum_{i=2}^{4N_k+1} x_i^2 - 2\varphi \sum_{i=1}^{4N_k+1} x_i x_{i+1} \\ &\le \left(\delta_k^2 + (1+\delta_k)^2 + (1+\varphi^2)(2N_k(1+\delta_k)^2 + 2N_k\delta_k^2)\right. \\ &\quad \left. + 2\varphi(4N_k+1)\delta_k(1+\delta_k)\right) = \ell \end{aligned}$$

if $\mathbf{x} \in (0,1,\dots,0,1) + [-\delta_k, \delta_k]^{4N_k+2}$. Therefore

$$\begin{aligned} \mu(A_{\delta_k,N_k}) &\ge C_2(\varphi)\left(\frac{\sqrt{2}\delta_k}{\sqrt{\pi}\sigma_\epsilon}\right)^{4N_k+2} e^{-\frac{\ell}{2\sigma_\epsilon^2}} \\ &= C_2(\varphi)\left(\frac{C_3}{\sqrt{N_k}}\right)^{4N_k+2} e^{-\frac{1}{2\sigma_\epsilon^2}\left(O(N_k)+O\left(N_k^{1/2}\right)+O(1)+O\left(N_k^{-1/2}\right)+O\left(N_k^{-1}\right)\right)}. \end{aligned}$$

In logarithmic scale, the latter becomes

$$\log(\mu(A_{\delta_k,N_k})) \ge C_4 - C_5 N_k \log(N_k) - C_6 N_k$$

for large N_k, where the constants C_i are strictly positive and depend only on φ, σ_ϵ, and k. On the other hand, by the union bound,

$$\begin{aligned}\mu(B_{N_k}^c) &\leq \sum_{j\leq -(2N_k+1)} \mu(|x_j| > |j|/4) + \sum_{j\geq 2N_k+2} \mu(|x_j| > |j|/4) \\ &\lesssim \sum_{j\geq 2N_k+1} \exp(-C_7 j^2) \lesssim C_8 e^{-C_9 N_k^2}\end{aligned}$$

for some other strictly positive constants C_7, C_8, C_9. Combining the two,

$$\mu(A_{\delta_k,N_k} \cap B_{N_k}) \geq \mu(A_{\delta_k,N_k}) - \mu(B_{N_k}^c), \tag{10}$$

and by the asymptotics in N_k the right-hand side is eventually strictly positive. By choosing N_k large enough we can also ensure that

$$\sum_{|j|\geq 2N_k} \frac{a^2}{(1+(3|j|/4)^2)^2} \leq \frac{\|S\|_2^2}{4k}. \tag{11}$$

We henceforth fix N_k satisfying both (11) and (10). Consider now the shift operator $T : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}^{\mathbb{Z}}$ defined by $(Tx)_n = x_{n+1}$. The stationary Gaussian AR(1) process with $0 < \varphi < 1$ is mixing by Theorem 10 in [23]; equivalently, its law $\mu = \mathbb{P} \circ \boldsymbol{\xi}^{-1}$ is mixing, hence ergodic, with respect to T. It follows that the system $(\mathbb{R}^{\mathbb{Z}}, \mu, U)$ with $U \coloneqq T^2$ is also mixing, and thus ergodic. Set $E_k \coloneqq A_{\delta_k,N_k} \cap B_{N_k} \subseteq \mathbb{R}^{\mathbb{Z}}$, so that $\mu(E_k) > 0$ by (10), and define

$$F_k \coloneqq \bigcup_{m\geq 1} U^{-m} E_k.$$

By the characterization of ergodicity (cf. Theorem 1.5.(iii) in [24])[1] we have $\mu(F_k) = 1$, and hence $\mathbb{P}(\Omega_k) = 1$ for $\Omega_k \coloneqq \boldsymbol{\xi}^{-1}(F_k)$. Therefore, for every

[1]Theorem 1.5.(iii) in [24] requires the maps T and U to be measure-preserving, i.e. $\mu(A) = \mu \circ T^{-1}(A)$ for all $A \in \mathcal{B}(\mathbb{R}^{\mathbb{Z}})$. This is true in our case but it requires a bit

$\omega \in \Omega_k$ there exists at least one $\overline{m}$ with $U^{\overline{m}}(\boldsymbol{\xi}(\omega)) \in E_k$, i.e. $T^{2\overline{m}}(\boldsymbol{\xi}(\omega)) \in A_{\delta_k,N_k} \cap B_{N_k}$. Define $s_k^\omega(x) \coloneqq S(x - 2\overline{m})$. Since $(T^{2\overline{m}}\boldsymbol{\xi}(\omega))_j = \xi_{j+2\overline{m}}(\omega)$, the membership $T^{2\overline{m}}(\boldsymbol{\xi}(\omega)) \in B_{N_k}$ means that for indices with $j - 2\overline{m} \leq -(2N_k+1)$ or $j-2\overline{m} \geq 2N_k+2$ one has $|\xi_j(\omega)| \leq |j-2\overline{m}|/4$, so $|\theta_j(\omega)-2\overline{m}| \geq 3|j-2\overline{m}|/4$, and

$$\sum_{\substack{j-2\overline{m}\leq -(2N_k+1)\\ \text{or } j-2\overline{m}\geq 2N_k+2}} |s_k^\omega(\theta_j(\omega))|^2 \leq \sum_{|\ell|\geq 2N_k+1} \frac{a^2}{(1+(3|\ell|/4)^2)^2} \leq \frac{\|S\|_2^2}{4k},$$

where the last inequality follows from (11) since $\{|\ell| \geq 2N_k + 1\} \subset \{|\ell| \geq 2N_k\}$. Likewise, the membership $T^{2\overline{m}}(\boldsymbol{\xi}(\omega)) \in A_{\delta_k,N_k}$ means that for indices with $j-2\overline{m} \in \{-2N_k, \ldots, 2N_k+1\}$ the point $\theta_j(\omega)-2\overline{m}$ lies within distance δ_k of an even integer, where S vanishes; hence

$$\sum_{j-2\overline{m}\in\{-2N_k,\ldots,2N_k+1\}} |s_k^\omega(\theta_j(\omega))|^2 \leq (4N_k+2)\|S'\|_\infty^2\delta_k^2 = \left(1+\frac{1}{2N_k}\right)\frac{\|S\|_2^2}{2k}.$$

Combining the two estimates and using $\|s_k^\omega\|_2 = \|S\|_2$,

$$\sum_{j\in\mathbb{Z}} |s_k^\omega(\theta_j(\omega))|^2 \leq \left(\frac{3}{4}+\frac{1}{4N_k}\right)\frac{\|S\|_2^2}{k} \leq \frac{\|s_k^\omega\|_2^2}{k},$$

the last bound holding for all $N_k \geq 1$. Taking $\overline{\Omega} = \bigcap_{k\geq 1}\Omega_k$ yields the desired conclusion, since $\mathbb{P}(\overline{\Omega}) = 1$. □

Remark 2. The hypothesis on ξ_n in Theorem 2 can be relaxed: the conclusion remains valid for any stationary mixing process for which, for every k, there

---

of a tedious proof, that we sketch here: for $I = (n_1, \ldots, n_d) \in \mathbb{Z}^d$ with $n_1 < \cdots < n_d$ consider the canonical projections $\pi_I : \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}^{|I|} \cong \mathbb{R}^d$ and the sets $\mathcal{C}_d = \{\pi_I^{-1}(A) : A \in \mathcal{B}(\mathbb{R}^d), |I| = d\}$. By definition $\sigma(\mathcal{C}_1) = \mathcal{B}(\mathbb{R}^{\mathbb{Z}})$ and moreover $\mathcal{C} \coloneqq \bigcup_{d\geq 1}\mathcal{C}_d$ is a $\pi-$system, thus it is enough to prove $\mu(C) = \mu \circ T^{-1}(C)$ for all $C \in \mathcal{C}$ by Dynkin's lemma (1.6(a) in [25]). The conclusion follows from the strict stationarity of $\boldsymbol{\xi}$.

exist N_k and $\delta_k(N_k)$ as defined in the previous proof such that

$$\mathbb{P}(A_{\delta_k,N_k}) - \mathbb{P}(B^c_{N_k}) > 0,$$

in the sense of (10).

The previous proof shows that for each $\omega \in \overline{\Omega}$ there exists a sequence of pathological bandlimited signals whose energy is concentrated precisely at the locations where the jitter $\boldsymbol{\xi}(\omega)$ exhibits the bad pattern E_k. This is an ad hoc deterministic construction, designed to disprove a worst-case property, and it crucially relies on aligning the signal with a specific shift $2\overline{m}$ of $\boldsymbol{\xi}$.

A natural question, left open here, is whether the AR(1)-jittered lattice stably samples a *random* bandlimited signal with probability one - for instance a stationary bandlimited Gaussian process, for which the relevant comparison is between time-averaged signal power and the corresponding average sample power. The mechanism of Theorem 2 does not apply in that setting, since a stationary signal cannot align with any prescribed shift of the jitter path.

3. Finite-dimensional analysis

The negative result of Theorem 2 concerns the full Paley-Wiener space and the infinite jittered lattice. In practice, however, one works with finitely many samples and a finite-dimensional approximation of the signal, and the question of stability must be reformulated accordingly. The goal of this section is to develop such a finite-dimensional counterpart and to identify regimes in which the associated sampling matrix is *well-conditioned* with high probability. We remark that, unlike Theorem 2, the results of this section

require only that $\boldsymbol{\xi}_N$ be a centered Gaussian vector with $\mathrm{var}(\xi_n) = \sigma_\xi^2$; the autoregressive structure enters through the explicit form (20) of the Cholesky factor $\mathbf{L}_N$, but not through the resulting bounds.

If $s \in \mathrm{PW}_{[-1/2,1/2]}$, the Shannon-Whittaker-Nyquist-Kotelnikov theorem gives

$$s(t) = \sum_{n\in\mathbb{Z}} s(n)\,\mathrm{sinc}(t-n),$$

and in applications one typically resorts to the approximation

$$s(t) \approx \sum_{|n|\le N} s(n)\,\mathrm{sinc}(t-n)$$

for N sufficiently large. In this sense, s can be regarded as living in $\mathrm{span}\{\mathrm{sinc}(\cdot - n)\}_{|n|\le N}$. The vector of jittered samples $\widetilde{\mathbf{s}}_N \coloneqq (s(-N+\xi_{-N}), \dots, s(N+\xi_N))^t$ is related to the non-jittered one $\mathbf{s}_N \coloneqq (s(-N), \dots, s(N))^t$ through the random $(2N+1)\times(2N+1)$ sampling matrix

$$\mathbf{A}^{(N)} \coloneqq \begin{pmatrix} \mathrm{sinc}(\xi_{-N}) & \mathrm{sinc}(-1+\xi_{-N}) & \dots & \mathrm{sinc}(-2N+\xi_{-N}) \\ \mathrm{sinc}(1+\xi_{-N+1}) & \mathrm{sinc}(\xi_{-N+1}) & \dots & \mathrm{sinc}(-2N+1+\xi_{-N+1}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{sinc}(2N+\xi_N) & \mathrm{sinc}(2N-1+\xi_N) & \dots & \mathrm{sinc}(\xi_N) \end{pmatrix}, \tag{12}$$

so that $\widetilde{\mathbf{s}}_N = \mathbf{A}^{(N)}\mathbf{s}_N$. Then

$$\sigma_{\min}(\mathbf{A}^{(N)})\|\mathbf{s}_N\|_{\ell^2} = \sigma_{\min}(\mathbf{A}^{(N)})\|s\|_{L^2} \le \|\mathbf{A}^{(N)}\mathbf{s}_N\|_{\ell^2} = \|\widetilde{\mathbf{s}}_N\|_{\ell^2},$$

where the first identity follows from Parseval. Rearranging yields $\|s\|_{L^2} \le \sigma_{\min}(\mathbf{A}^{(N)})^{-1}\|\widetilde{\mathbf{s}}_N\|_{\ell^2}$ (cf. Definition 3), so the stability of finite-dimensional reconstruction reduces to the conditioning of $\mathbf{A}^{(N)}$. Note that, in general, $\mathbf{A}^{(N)}$ is neither symmetric nor Toeplitz. Observe also that $\boldsymbol{\xi}_N \mapsto \det(\mathbf{A}^{(N)}(\boldsymbol{\xi}_N))$

is analytic and not identically zero, and thus $\mathbb{P}(\{\det(\mathbf{A}^{(N)}) = 0\}) = 0$ for all N, making the mere invertibility a trivial matter. We prove the following.

Proposition 3. *Let $\mathbf{A}^{(N)}$ be the sensing matrix defined in* (12)*, where $\boldsymbol{\xi}_N = (\xi_{-N}, \dots, \xi_N)$ is a centered Gaussian random vector with $\mathrm{var}(\xi_n) = \sigma_\xi^2 \leq 10^{-2}$ for all $|n| \leq N$ and arbitrary correlation structure. Let $\eta \in (0,1)$, $\epsilon \in (0, 2 - e^{16\pi/1000})$ and set $c \coloneqq \log(2-\epsilon)/\pi$. Then $\sigma_{\min}(\mathbf{A}^{(N)}) \geq \epsilon$ with probability at least $1-\eta$ if*

$$N \leq \frac{1}{2}\left\lceil \frac{\log(1-\eta)}{\log\left(1 - \sqrt{\frac{8}{\pi}}\,\frac{e^{-c^2/(2\sigma_\xi^2)}}{\frac{c}{\sigma_\xi} + \sqrt{\frac{c^2}{\sigma_\xi^2}+2}}\right)} - 1 \right\rceil. \tag{13}$$

Proof. Observe preliminarily that the assumptions on ϵ and σ_ξ automatically imply that

$$\sqrt{\frac{8}{\pi}}\, e^{-c^2/(2\sigma_\xi^2)} < \frac{c}{\sigma_\xi} + \sqrt{\frac{c^2}{\sigma_\xi^2}+2}. \tag{14}$$

By Weyl's inequality for singular values,

$$\begin{aligned} \sigma_{\min}(\mathbf{A}^{(N)}) &\geq \sigma_{\min}(\mathbf{I}_N) - \|\mathbf{A}^{(N)} - \mathbf{I}_N\|_{\mathrm{op}} \\ &\geq 1 - \|\mathbf{A}^{(N)} - \mathbf{I}_N\|_{\mathrm{op}}, \end{aligned} \tag{15}$$

where $\mathbf{I}_N$ denotes the $(2N+1)\times(2N+1)$ identity matrix. We now study $\|\mathbf{A}^{(N)} - \mathbf{I}_N\|_{\mathrm{op}}$. Introduce the auxiliary matrices

$$\mathbf{T}^{(N)} \coloneqq \begin{pmatrix} 0 & -1 & -2 & \dots & -2N \\ 1 & 0 & -1 & \dots & -2N+1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 2N & 2N-1 & 2N-2 & \dots & 0 \end{pmatrix}, \quad \mathbf{D}_{\boldsymbol{\xi}_N^r} \coloneqq \mathrm{diag}(\xi_{-N}^r, \dots, \xi_N^r),$$

and observe that

$$\mathbf{A}^{(N)} - \mathbf{I}_N = \sum_{r \geq 1} \frac{1}{r!} \mathbf{D}_{\boldsymbol{\xi}_N^r} \mathrm{sinc}^{(r)} \circ \mathbf{T}^{(N)},$$

where $\mathrm{sinc}^{(r)}$ is the r-th derivative of the sinc function and $\circ$ denotes componentwise application of $\mathrm{sinc}^{(r)}$ to the entries of $\mathbf{T}^{(N)}$. Setting $\delta_N \coloneqq \max_{|n| \leq N} |\xi_n|$, we obtain

$$\|\mathbf{A}^{(N)} - \mathbf{I}_N\|_{\mathrm{op}} \leq \sum_{r \geq 1} \frac{\delta_N^r \|\mathrm{sinc}^{(r)} \circ \mathbf{T}^{(N)}\|_{\mathrm{op}}}{r!} \leq \sum_{r \geq 1} \frac{\delta_N^r \pi^r}{r!} = e^{\delta_N \pi} - 1,$$

where the last inequality uses the fact that the (infinite) Toeplitz matrix $\mathrm{sinc}^{(r)} \circ \mathbf{T}^{(\infty)}$ has Fourier symbol $(2\pi i \omega)^r$ on $[-1/2, 1/2]$, and that

$$\|\mathrm{sinc}^{(r)} \circ \mathbf{T}^{(N)}\|_{\mathrm{op}} \leq \|\mathrm{sinc}^{(r)} \circ \mathbf{T}^{(\infty)}\|_{\mathrm{op}} = \max_{\omega \in [-1/2, 1/2]} |2\pi\omega|^r = \pi^r;$$

see Theorem 1.1 in [26]. Substituting back into (15) gives $\sigma_{\min}(\mathbf{A}^{(N)}) \geq 2 - e^{\delta_N \pi}$. It remains to determine when $2 - e^{\delta_N \pi} \geq \epsilon > 0$ with high probability, which is equivalent to requiring

$$\mathbb{P}[|\xi_{-N}| \leq c,\ |\xi_{-N+1}| \leq c,\ \ldots\ , |\xi_N| \leq c] \geq 1 - \eta \tag{16}$$

for η small.

By the classical Šidák theorem[2] [28] we obtain

$$\mathbb{P}[|\xi_{-N}| \leq c,\ |\xi_{-N+1}| \leq c,\ \ldots,\ |\xi_N| \leq c] \geq \prod_{n=-N}^{N} \mathbb{P}[|\xi_n| \leq c] = \left[\operatorname{erf}\left(\frac{c}{\sigma_\xi \sqrt{2}}\right)\right]^{2N+1},$$

and Komatsu's inequality [29] finally yields

$$\left[\operatorname{erf}\left(\frac{c}{\sigma_\xi \sqrt{2}}\right)\right]^{2N+1} \geq \left[1 - \sqrt{\frac{8}{\pi}} \frac{e^{-c^2/(2\sigma_\xi^2)}}{\frac{c}{\sigma_\xi} + \sqrt{\frac{c^2}{\sigma_\xi^2} + 2}}\right]^{2N+1}. \tag{17}$$

For fixed c and σ_ξ, the right-hand side of (17) tends to 0 as $N \to \infty$, so the lower bound is asymptotically useless. In the finite-dimensional setting, however, it allows us to estimate how large the sensing matrix $\mathbf{A}^{(N)}$ may be "allowed to be" in order to still have $\sigma_{\min}(\mathbf{A}^{(N)}) \geq \epsilon$ with probability at least $1 - \eta$. A direct calculation from (17) gives

$$N \leq \frac{1}{2}\left\lceil \frac{\log(1-\eta)}{\log\left(1 - \sqrt{\frac{8}{\pi}} \frac{e^{-c^2/(2\sigma_\xi^2)}}{\frac{c}{\sigma_\xi} + \sqrt{\frac{c^2}{\sigma_\xi^2} + 2}}\right)} - 1 \right\rceil.$$

□

[2] It is unclear to the authors whether the problem of finding a non-trivial non-asymptotic lower bound for the left-hand side of (16) has been satisfactorily solved in the literature when the ξ_n are Gaussian and correlated. Without the absolute values, the "excess" probability caused by the correlations was estimated by Li and Shao in [27], but they themselves remark that "[...] it may be possible to have a comparisons of the two sided probabilities [with absolute values], [...] however, at this time, we are unable to find a right formulation [...]".

3.1. First-order Taylor expansion analysis

Proposition 3 is based on a uniform tail bound on $\max_{|n|\le N}|\xi_n|$ and is consequently loose: it captures the worst single coordinate of $\boldsymbol{\xi}$ rather than the collective effect on the operator norm. In the small-jitter regime $\sigma_\xi \ll 1$, typical of hardware applications, a sharper analysis is possible via first-order linearization. The sensing matrix $\mathbf{A}^{(N)}$ may then be replaced with the first-order Taylor surrogate

$$\mathbf{B}^{(N)} \coloneqq \mathbf{I}_N + \mathbf{D}_{\boldsymbol{\xi}_N^1}\mathbf{S}_N, \tag{18}$$

where $(\mathbf{S}_N)_{mn} = (-1)^{m-n}/(m-n)$ for $m \neq n$ and $(\mathbf{S}_N)_{mm} = 0$, obtained by evaluating the first derivative of sinc at integer arguments. In signal-processing practice, $\mathbf{B}^{(N)}$ often entirely replaces $\mathbf{A}^{(N)}$.

Proposition 4. *Let $\mathbf{B}^{(N)}$ be the first-order Taylor surrogate defined in* (18)*, where $\boldsymbol{\xi}_N = (\xi_{-N},\dots,\xi_N)$ is a centered Gaussian random vector with $\operatorname{var}(\xi_n) = \sigma_\xi^2$ for all $|n| \le N$ and arbitrary correlation structure. Let $\epsilon,\eta \in (0,1)$. Then $\sigma_{\min}(\mathbf{B}^{(N)}) \ge \epsilon$ with probability at least $1-\eta$ if*

$$N \le \frac{1}{4}\left[\eta \exp\!\big((1-\epsilon)^2/(2\sigma_\xi^2\pi^2)\big) - 2\right], \tag{19}$$

which should be compared to condition (13)*.*

Proof. The matrix $\mathbf{S}_N$ is Toeplitz and skew-symmetric with symbol $2\pi i\omega$ (cf. the proof of Proposition 3), and therefore $\|\mathbf{S}_N\|_{\mathrm{op}} \le \pi$. What we would like, however, is a direct bound on $\|\mathbf{D}_{\boldsymbol{\xi}_N^1}\mathbf{S}_N\|_{\mathrm{op}}$. We can write

$$\mathbf{D}_{\boldsymbol{\xi}_N^1}\mathbf{S}_N = \sum_{|n|\le N} g_n \operatorname{diag}((\mathbf{L}_N)_{:n})\,\mathbf{S}_N =: \sum_{|n|\le N} g_n \widetilde{\mathbf{S}}_{N,n},$$

where $(\mathbf{L}_N)_{:n}$ is the n-th column of the Cholesky decorrelation matrix $\mathbf{L}_N$, $\mathbf{L}_N\mathbf{L}_N^t = \boldsymbol{\Sigma}_N$, and $(g_{-N},\dots,g_N) =: \mathbf{g}_N \sim \mathcal{N}(\mathbf{0},\mathbf{I}_N)$. In the AR(1) case

(2), whose stationary variance is $\sigma_\xi^2 = \sigma_\epsilon^2/(1-\varphi^2)$, the matrix $\mathbf{L}_N$ may be obtained explicitly by "unrolling" the recurrence:

$$\mathbf{L}_N = \begin{pmatrix} \sigma_\xi & 0 & 0 & \cdots & 0 \\ \sigma_\xi\varphi & \sigma_\xi\sqrt{1-\varphi^2} & 0 & \cdots & 0 \\ \sigma_\xi\varphi^2 & \sigma_\xi\varphi\sqrt{1-\varphi^2} & \sigma_\xi\sqrt{1-\varphi^2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_\xi\varphi^{2N} & \sigma_\xi\varphi^{2N-1}\sqrt{1-\varphi^2} & \sigma_\xi\varphi^{2N-2}\sqrt{1-\varphi^2} & \cdots & \sigma_\xi\sqrt{1-\varphi^2} \end{pmatrix}. \tag{20}$$

The argument below, however, only uses the identity $\sum_n (\mathbf{L}_N)_{mn}^2 = (\mathbf{\Sigma}_N)_{mm} = \sigma_\xi^2$, which holds for any Cholesky factor of a covariance matrix with constant diagonal. To apply Theorem 4.1.1 in [30], which addresses random series with matrix coefficients, we need to estimate

$$v(\mathbf{D}_{\boldsymbol{\xi}_N^1}\mathbf{S}_N) \coloneqq \max\Big\{ \big\|\textstyle\sum_n \widetilde{\mathbf{S}}_{N,n}\widetilde{\mathbf{S}}_{N,n}^t\big\|_{\mathrm{op}},\ \big\|\textstyle\sum_n \widetilde{\mathbf{S}}_{N,n}^t\widetilde{\mathbf{S}}_{N,n}\big\|_{\mathrm{op}} \Big\}.$$

For the first term, we have $\sum_n \widetilde{\mathbf{S}}_{N,n}^t\widetilde{\mathbf{S}}_{N,n} = -\sigma_\xi^2\mathbf{S}_N^2$, because $\sum_n(\mathbf{L}_N)_{mn}^2 = \sigma_\xi^2$ for all m; hence $\|\sum_n \widetilde{\mathbf{S}}_{N,n}^t\widetilde{\mathbf{S}}_{N,n}\|_{\mathrm{op}} \le \sigma_\xi^2\pi^2$. For the second term, we observe that

$$\begin{aligned} \sum_{|n|\le N} \widetilde{\mathbf{S}}_{N,n}\widetilde{\mathbf{S}}_{N,n}^t &= -\sum_{|n|\le N} \mathrm{diag}((\mathbf{L}_N)_{:n})\mathbf{S}_N^2\mathrm{diag}((\mathbf{L}_N)_{:n}) \\ &= -\underbrace{\left[\sum_{|n|\le N}(\mathbf{L}_N)_{:n}(\mathbf{L}_N)_{:n}^t\right]}_{=:\widetilde{\mathbf{L}}_N} \odot \mathbf{S}_N^2. \end{aligned}$$

By construction $(\widetilde{\mathbf{L}}_N)_{nn} = \sum_{j\le n}(\mathbf{L}_N)_{nj}^2 = \sigma_\xi^2$, and the matrix $\widetilde{\mathbf{L}}_N$ is positive semidefinite, being a sum of rank-one positive semidefinite matrices. Therefore

$$\|\widetilde{\mathbf{L}}_N \odot \mathbf{S}_N^2\|_{\mathrm{op}} \le \max_n(\widetilde{\mathbf{L}}_N)_{nn}\,\|\mathbf{S}_N^2\|_{\mathrm{op}} \le \sigma_\xi^2\pi^2,$$

where the first inequality is Theorem 5.5.18 in [31]. We conclude that $v(\mathbf{D}_{\boldsymbol{\xi}_N^1}\mathbf{S}_N) \leq \sigma_\xi^2\pi^2$. Formula (4.1.6) in Theorem 4.1.1 of [30], together with the previous calculations, gives

$$\mathbb{P}\left[\|\mathbf{D}_{\boldsymbol{\xi}_N^1}\mathbf{S}_N\|_{\text{op}} \leq 1-\epsilon\right] \geq 1 - 2(2N+1)\exp\left(-(1-\epsilon)^2/(2\sigma_\xi^2\pi^2)\right), \quad (21)$$

and since $\sigma_{\min}(\mathbf{B}^{(N)}) \geq 1 - \|\mathbf{D}_{\boldsymbol{\xi}_N^1}\mathbf{S}_N\|_{\text{op}}$ by Weyl's inequality, the right-hand side of (21) is $\geq 1-\eta$ provided that

$$N \leq \frac{1}{4}\left[\eta\exp\left((1-\epsilon)^2/(2\sigma_\xi^2\pi^2)\right) - 2\right].$$

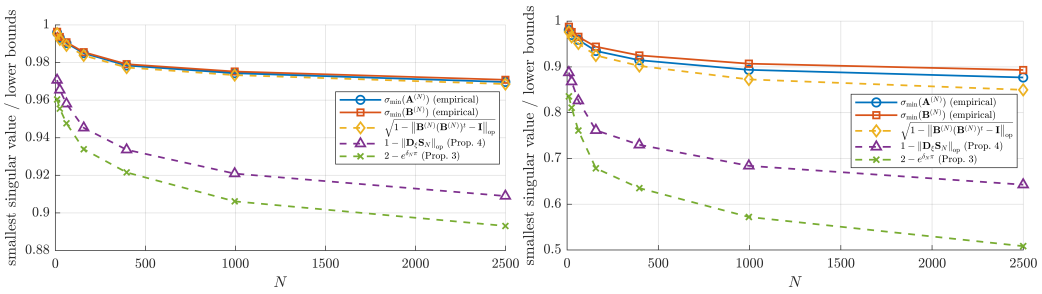
□

Remark 3. The same techniques used in the proofs of Propositions 3 and 4 yield probabilistic upper bounds for $\sigma_{\max}(\mathbf{A}^{(N)})$ and $\sigma_{\max}(\mathbf{B}^{(N)})$, namely $\sigma_{\max}(\mathbf{A}^{(N)}) \leq e^{\delta_N\pi}$ and $\sigma_{\max}(\mathbf{B}^{(N)}) \leq 1 + \|\mathbf{D}_{\boldsymbol{\xi}_N^1}\mathbf{S}_N\|_{\text{op}}$. Consequently, on the events considered in Propositions 3 and 4 respectively,

$$1 \leq \kappa(\mathbf{A}^{(N)}) = \frac{\sigma_{\max}(\mathbf{A}^{(N)})}{\sigma_{\min}(\mathbf{A}^{(N)})} \leq \frac{e^{\delta_N\pi}}{2-e^{\delta_N\pi}}, \qquad 1 \leq \kappa(\mathbf{B}^{(N)}) \leq \frac{2-\epsilon}{\epsilon},$$

where the lower bounds are immediate from $\sigma_{\min} \leq \sigma_{\max}$. These provide information not only on the invertibility of $\mathbf{A}^{(N)}$ and $\mathbf{B}^{(N)}$, but also on their condition numbers, and hence on their numerical stability. In particular, the regions depicted in Figure 2 are also regions of high-probability control on $\kappa(\mathbf{A}^{(N)})$ and $\kappa(\mathbf{B}^{(N)})$.

The figures 1 show that the bounds $\sigma_{\min}(\mathbf{A}^{(N)}) \geq 2 - e^{\delta_N\pi}$ and $\sigma_{\min}(\mathbf{B}^{(N)}) \geq 1 - \|\mathbf{D}_{\boldsymbol{\xi}_N^1}\mathbf{S}_N\|_{\text{op}}$ are rather loose. The same figures numerically suggest that a sharper lower bound for $\sigma_{\min}(\mathbf{B}^{(N)})$ is $\sqrt{1 - \|\mathbf{I}_N - \mathbf{B}^{(N)}(\mathbf{B}^{(N)})^t\|_{\text{op}}}$.



(a) $\sigma_\xi = 0.01$, $\varphi = 0.99$ (b) $\sigma_\xi = 0.04$, $\varphi = 0.99$

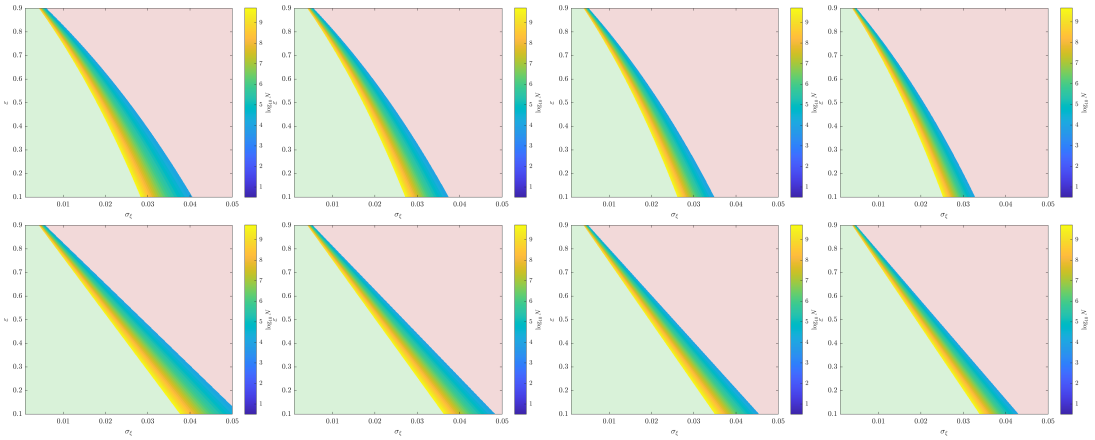
Figure 1: The figures above show that the obtained bounds are loose.

An analytic justification proceeds by exploiting the skew-symmetry $\mathbf{S}_N^t = -\mathbf{S}_N$, which gives

$$\mathbf{B}^{(N)}(\mathbf{B}^{(N)})^t - \mathbf{I}_N = [\mathbf{D}_{\boldsymbol{\xi}_N^1}, \mathbf{S}_N] - \mathbf{D}_{\boldsymbol{\xi}_N^1}\mathbf{S}_N^2\mathbf{D}_{\boldsymbol{\xi}_N^1},$$

where $[\cdot,\cdot]$ denotes the commutator. The commutator is again a Gaussian series in the variables $\{g_n\}$ of Section 3.1, with coefficient matrices $\mathbf{C}_n \coloneqq \mathbf{S}_N \operatorname{diag}((\mathbf{L}_N)_{:n}) - \operatorname{diag}((\mathbf{L}_N)_{:n})\,\mathbf{S}_N$, and is therefore amenable to the same Theorem 4.1.1 of [30] that we used in the proof of Proposition 4. The entries of $\mathbf{C}_n$ involve differences of entries of $\mathbf{L}_N$, which encode the AR(1) coherence between adjacent jitter samples and are small when $\varphi \approx 1$; these cancellations heuristically explain the empirically observed sharpness of the Gram-matrix bound in the highly correlated regime relevant to the ADC application. Translating them into an operator-norm bound with an explicit φ-dependent improvement over Proposition 4, however, appears to require non-standard Schur-multiplier estimates that lie outside the scope of this note.

The figures below show a qualitative illustration of (13) and (19) for different values of σ_ξ and ϵ.



(a) $\eta = 10^{-2}$ (b) $\eta = 10^{-3}$ (c) $\eta = 10^{-4}$ (d) $\eta = 10^{-5}$

Figure 2: Contour plots of the right-hand side of (13) (first row) and (19) (second row) as functions of σ_ξ and ϵ. In the pink region the bound falls below 10^3, while in the light green region it exceeds 10^{10}. We remark that, in ADC applications, σ_ξ represents the average deviation from the ideal sampling instants nT_s, so that $\sigma_\xi \ll T_s$ and σ_ξ/T_s is typically of the order of 1% or less. Higher values of σ_ξ may correspond to harsher scenarios, but more likely reflect poor hardware designs.

Remark 4. Real-world systems do not allow for infinite numerical precision, and typically some form of quantization follows the sampling operation in ADCs. If Q is an infinite uniform quantizer and $Q(\mathbf{M})$ is its component-wise application to all the coefficients of a matrix $\mathbf{M}$, it is straightforward to extend both Propositions 3 and 4 to $Q(\mathbf{M})$ by means of a (deterministic) bound on $\|Q(\mathbf{M}) - \mathbf{M}\|_{\text{op}}$ that follows from finite-dimensional norm equivalence and quantizer resolution. If the quantizer is finite within a finite range $[-R, R]$, the bound on $\|Q(\mathbf{M}) - \mathbf{M}\|_{\text{op}}$ will be "probabilistic", i.e. for example subjected to the probability of $|\xi_i| \leq R$ for all $i = -N, \ldots, N$.

4. Numerical Experiments

In this section we complement the analytical results of Section 3 with Monte Carlo simulations. Our aims are threefold: to describe the empirical distributions of $\sigma_{\min}$ and of the condition number for both $\mathbf{A}^{(N)}$ and $\mathbf{B}^{(N)}$, to assess how these distributions depend on the autoregressive parameter φ, and to quantify the conservativeness of the bounds obtained in Propositions 3 and 4. The first two aims are addressed in Section 4.1 while the third one in Section 4.2.

4.1. Singular values and condition number distributions

In Figure 3 we display the empirical cumulative distribution function (ECDF) of $\sigma_{\min}(\mathbf{A}^{(N)})$, $\sigma_{\min}(\mathbf{B}^{(N)})$, $\kappa(\mathbf{A}^{(N)})$ and $\kappa(\mathbf{B}^{(N)})$ respectively with $N = 1000$, derived by simulating 1000 matrix instances for five different correlation levels $\varphi \in \{0.1, 0.25, 0.5, 0.9, 0.99\}$ and $\sigma_\xi = 0.04$. The plots show that both random matrices $\mathbf{A}^{(1000)}$ and $\mathbf{B}^{(1000)}$ are very well-conditioned with high probability in all the correlation regimes tested (including the high-correlation regime, $\varphi \approx 1$, which is of particular interest to ADC applications), consistently with Propositions 3 and 4. The distributions relative to $\mathbf{A}^{(1000)}$ and $\mathbf{B}^{(1000)}$ are moreover very similar, which supports the use of the first-order surrogate (18) in place of the full sensing matrix. Finally, the higher the correlation, the larger $\sigma_{\min}(\mathbf{A}^{(N)})$ and $\sigma_{\min}(\mathbf{B}^{(N)})$, with little-to-no overlapping between the ECDF curves; we remark that this dependence on φ is not captured by the bounds of Section 3, which only involve the marginal variance σ_ξ^2.

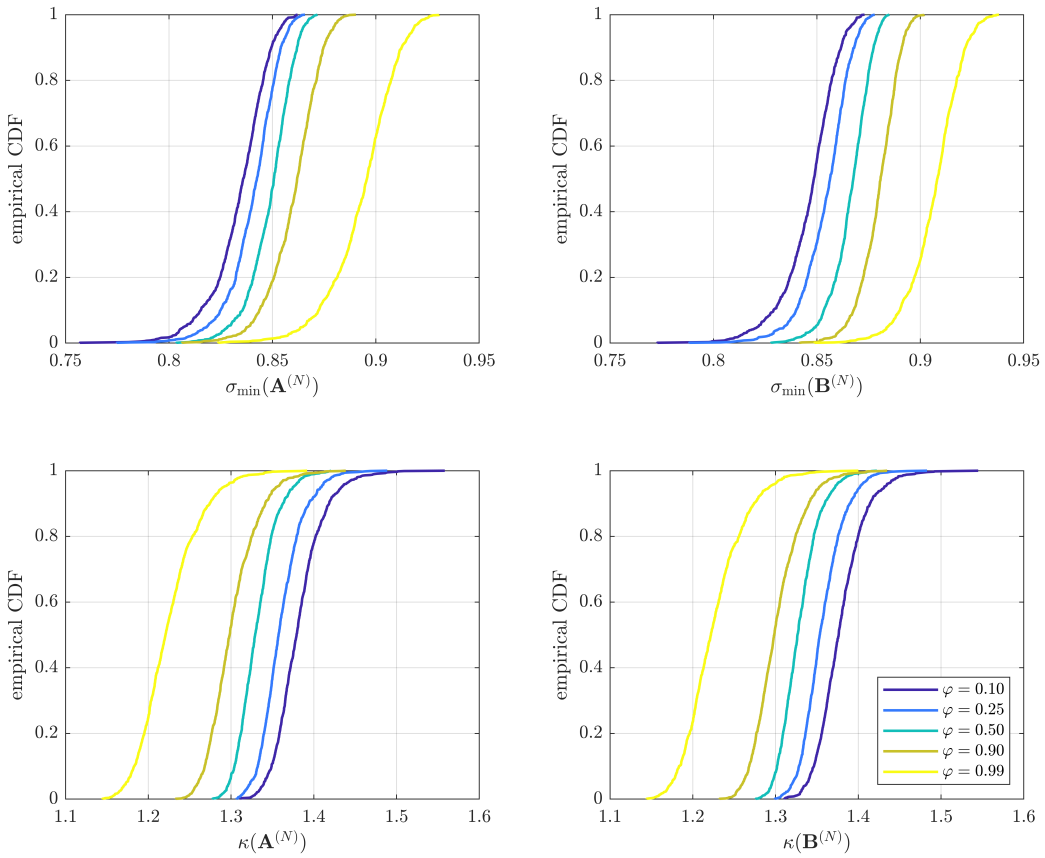


Figure 3: Empirical cumulative distribution functions obtained via Monte Carlo.

4.2. Lower bounds quality

As Figure 1 already suggests, the lower bounds used in Propositions 3 and 4 are somewhat crude, at least for "large" values of the process standard deviation σ_ξ, and analytically deriving *how* loose they are is beyond the scope of this work. Such bounds are tractable proxies that made it possible to rigorously demonstrate the well-conditioning of the jittered matrix with high probability. This self-contained subsection is meant to numerically assess the gap between such proxies and the true underlying quantities, as a function

of φ and the matrix size N.

Figure 4 displays the results of this assessment. We ran two separate Monte Carlo simulations in which we fixed the jitter process standard deviation to $\sigma_\xi = 0.01$ (top) and $\sigma_\xi = 0.04$ (bottom), and swept over the matrix dimensions and the geometric factor φ. Notice that for an AR(1) process $\sigma_\epsilon^2 = (1-\varphi^2)\sigma_\xi^2$, so that when $\varphi \to 1$ the innovation noise variance shrinks to 0; in this case the process degenerates into the uncorrelated case $\xi_i \approx \xi_{-N}$ for all i. In the heatmaps we plotted the mean relative gap (MRG)

$$\frac{1}{N_{\text{MC iters}}} \sum_{n=1}^{N_{\text{MC iters}}} \frac{\sigma_{\min} - s}{\sigma_{\min}}$$

with $N_{\text{MC iters}} = 250$, where $s = 2 - e^{\delta_N \pi}$ is compared against $\sigma_{\min} = \sigma_{\min}(\mathbf{A}^{(N)})$, while $s = 1 - \|\mathbf{D}_\xi \mathbf{S}_N\|_{\text{op}}$ and $s = \sqrt{1 - \|\mathbf{I}_N - \mathbf{B}^{(N)}(\mathbf{B}^{(N)})^t\|_{\text{op}}}$ are compared against $\sigma_{\min} = \sigma_{\min}(\mathbf{B}^{(N)})$.

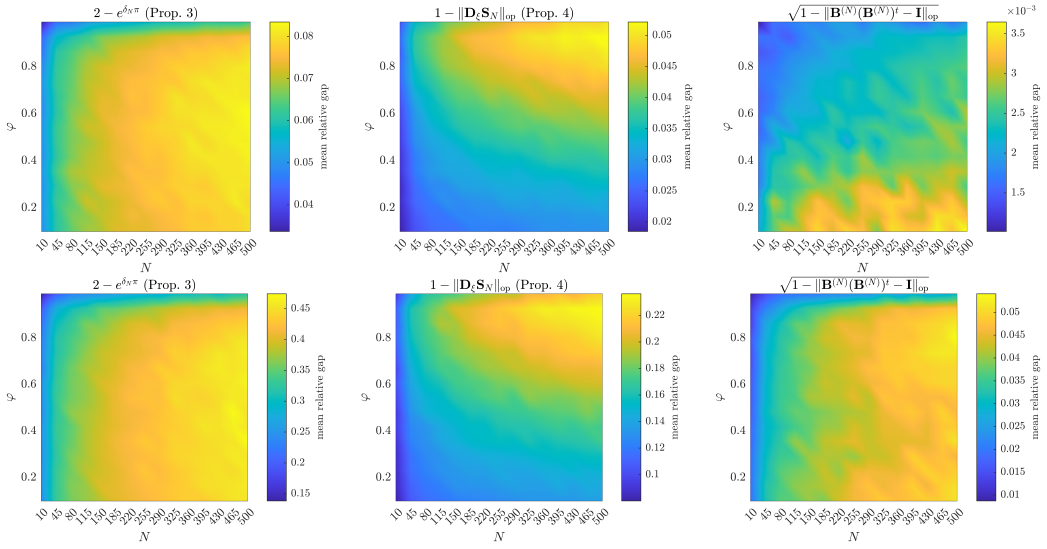


Figure 4: MRG of the different lower bounds obtained in Section 3.

Figure 4 confirms the preliminary conclusions drew from Figure 1, includ-

ing the degradation of the lower bounds performance as the process standard deviation increases.

4.3. Jitter cumulative correlation length and relation to the upper bounds in Propositions 3 and 4

The cumulative correlation length τ of a sequence indexed by $\mathbb{Z}$ is defined as the bilateral sum of its correlation coefficients, and it measures the degree of self-similarity of such a sequence; or, in other words, it provides an interval length within which all samples may be approximately considered as "highly correlated". For the process (2) it is easy to see that $\tau(\varphi) = (1+\varphi)/(1-\varphi)$. In this section we test whether the results from Section 3 remain informative when compared against the correlation length, i.e. whether N can be large enough to span one or many correlation lengths. The quantity of interest will be the ratio of the upper bounds in Propositions 3 and 4 to τ. Notice that in our case N/τ is a strictly decreasing function of φ over $[0, 1)$. Tables 1 and 2 show that the ratio N/τ exceeds 100 throughout the tested range of φ in many setups of interest, and by several orders of magnitude in most of them.

Table 1: (Rounded) Values of φ at which $N/\tau = 1$ and $N/\tau = 100$ (reported as $\cdot/\cdot$), for the bound of Proposition 3. An entry at the right endpoint of the tested range indicates that the ratio exceeds the corresponding threshold throughout $\varphi \in [0, 1)$ (and analogously with entries at the left endpoint).

ϵ	η	$\sigma_\xi = 0.005$	$\sigma_\xi = 0.01$	$\sigma_\xi = 0.025$	$\sigma_\xi = 0.04$
0.05	10^{-2}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9957
	10^{-3}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.9995 / 0.9575
	10^{-4}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.9957 / 0.6426
	10^{-5}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.9570 / 0
0.25	10^{-2}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.9965 / 0.7064
	10^{-3}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.9658 / 0
	10^{-4}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.6821 / 0
	10^{-5}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0 / 0
0.75	10^{-2}	0.9999 / 0.9999	0.9999 / 0.9999	0 / 0	$N \leq 0$
	10^{-3}	0.9999 / 0.9999	0.9999 / 0.9999	$N \leq 0$	$N \leq 0$
	10^{-4}	0.9999 / 0.9999	0.9999 / 0.9999	$N \leq 0$	$N \leq 0$
	10^{-5}	0.9999 / 0.9999	0.9999 / 0.9999	$N \leq 0$	$N \leq 0$

Table 2: Same as in table 1, but for Proposition 4.

ϵ	η	$\sigma_\xi = 0.005$	$\sigma_\xi = 0.01$	$\sigma_\xi = 0.025$	$\sigma_\xi = 0.04$
0.05	10^{-2}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999
	10^{-3}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999
	10^{-4}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999
	10^{-5}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999
0.25	10^{-2}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9985
	10^{-3}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9853
	10^{-4}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.9985 / 0.8628
	10^{-5}	0.9999 / 0.9999	0.9999 / 0.9999	0.9999 / 0.9999	0.9853 / 0.1500
0.75	10^{-2}	0.9999 / 0.9999	0.9999 / 0.9999	$N \leq 0$	$N \leq 0$
	10^{-3}	0.9999 / 0.9999	0.9999 / 0.9999	$N \leq 0$	$N \leq 0$
	10^{-4}	0.9999 / 0.9999	0.9999 / 0.9999	$N \leq 0$	$N \leq 0$
	10^{-5}	0.9999 / 0.9999	0.9999 / 0.9999	$N \leq 0$	$N \leq 0$

5. Conclusions and future work

This note examined the sampling properties of the AR(1)-jittered lattice $t_n = n + \xi_n$ at two scales. At the infinite-dimensional scale, Lemmas 2.1 and 2.2 show that the jittered set attains the critical Nyquist density both in expectation and almost surely, yet Theorem 2 shows that, with probability one, it fails to be a stable sampling set for $\mathrm{PW}_{[-1/2,1/2]}$. The mechanism behind this failure, adapted from [8], exploits the ergodicity of the shift on the AR(1) path to locate, almost surely, an arbitrarily long block on which the jitter aligns with the zeros of a fixed bandlimited function. At the finite-dimensional scale relevant to ADC applications, Propositions 3 and 4 give explicit, non-asymptotic conditions on the sample size N under which the jittered sinc sampling matrix, and its first-order Taylor surrogate, respectively, is well-conditioned with high probability. The Monte Carlo experiments of Section 4 corroborate these guarantees: for $N = 1000$ and $\sigma_\xi = 0.04$, the smallest singular values of both matrices remain above 0.8 and the condition numbers below 1.5 across all tested correlation levels $\varphi \in [0.1, 0.99]$, with conditioning improving monotonically in φ; this dependence on φ is not captured by the bounds, which involve only the marginal variance σ_ξ^2.

Future work will focus on understanding whether ergodicity can again be exploited to show that the autoregressive jittered sinc family also violates the upper frame (Bessel) bound, and whether the autoregressive structure can be leveraged to sharpen the conclusions of Propositions 3 and 4.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors occasionally used Anthropic's Claude as interlocutor for discussions of mathematical arguments, exposition, and English grammar. All mathematical content, proofs, results, code, and conclusions are the authors' own, as is responsibility for any errors.

References

[1] M. Vetterli, J. Kovačević, V. K. Goyal, Foundations of Signal Processing, Cambridge University Press, 2016.

[2] J. R. Higgins, Sampling Theory in Fourier and Signal Analysis - Foundations, Clarendon Press, 1996.

[3] K. Gröchenig, Reconstruction algorithms in irregular sampling, Mathematics of Computation 59 (199) (1992) 181–194.

[4] E. Margolis, Y. C. Eldar, Nonuniform sampling of periodic bandlimited signals, IEEE Transactions on Signal Processing 56 (7) (2008) 2728–2745.

[5] T. Strohmer, J. Tanner, Fast reconstruction methods for bandlimited functions from periodic nonuniform sampling, SIAM Journal on Numerical Analysis 44 (3) (2006) 1073–1094.

[6] J. L. Yen, On nonuniform sampling of bandwidth-limited signals, IRE Transactions on Circuit Theory 3 (1956) 251–257.

[7] K. Seip, A. M. Ulanovskii, Random exponential frames, Journal of the London Mathematical Society 53 (1996) 560–568.

[8] R. F. Bass, K. Gröchenig, Random sampling of bandlimited functions, Israel Journal of Mathematics 177 (2010) 1–28.

[9] R. F. Bass, K. Gröchenig, Relevant sampling of bandlimited functions, Illinois Journal of Mathematics 57 (1) (2013) 43–58.

[10] A. I. Zayed, Advances in Shannon's Sampling Theory, CRC Press, 1993.

[11] S. Tertinek, C. Vogel, Reconstruction of nonuniformly sampled bandlimited signals using a differentiator-multiplier cascade, IEEE Transactions on Circuits and Systems I: Regular Papers 55 (8) (2008) 2273–2286.

[12] A. Zanchi, C. Samori, Analysis and characterization of the effects of clock jitter in A/D converters for subsampling, IEEE Transactions on Circuits and Systems I: Regular Papers 55 (2) (2008) 522–534.

[13] Z. J. Towfic, S.-K. Ting, A. H. Sayed, Clock jitter compensation in high-rate ADC circuits, IEEE Transactions on Signal Processing 60 (11) (2012) 5738–5753.

[14] Y.-S. Chang, C.-L. Lin, W.-S. Wang, C.-C. Lee, C.-Y. Shih, An analytical approach for quantifying clock jitter effects in continuous-time sigma-delta modulators, IEEE Transactions on Circuits and Systems I: Regular Papers 53 (9) (2006) 1861–1868.

[15] D. Gerosa, R. Hou, V. Björk, U. Gustavsson, T. Eriksson, Autoregres-

sive stochastic clock jitter compensation in analog-to-digital converters, IEEE Transactions on Signal ProcessingTo appear (2026).

[16] S. Kay, Fundamentals of Statistical Signal Processing - Estimation Theory, Pearson, 1993.

[17] H. J. Landau, Sampling, data transmission, and the Nyquist rate, Proceedings of the IEEE 55 (10) (1967) 1701–1706.

[18] R. E. A. C. Paley, N. Wiener, Fourier Transforms in the Complex Domain, American Mathematical Society, New York, 1934.

[19] G. Chistyakov, Y. Lyubarskii, Random perturbations of exponential Riesz bases in $L^2(-\pi, \pi)$, Annales de l'Institut Fourier 47 (1) (1997) 201–255.

[20] H. Gebelein, Das statistische problem der korrelation als variations- und eigenwertproblem und sein zusammenhang mit der ausgleichsrechnung, Zeitschrift für Angewandte Mathematik und Mechanik 21 (1941) 364–379.

[21] R. P. Boas, Entire Functions, Academic Press Inc., 1954.

[22] G. Fikioris, Spectral properties of Kac–Murdock–Szegö matrices with a complex parameter, Linear Algebra and its Applications 553 (2018) 183–210.

[23] M. Krishnapur, Gaussian processes, Lecture Notes, Indian Institute of Science, `https://math.iisc.ac.in/~manju/GP20/gp_2024.pdf` (2024).

[24] P. Walters, An Introduction to Ergodic Theory, Springer-Verlag, 2000.

[25] D. Williams, Probability with Martingales, Cambridge: Cambridge University Press, 1991.

[26] A. Böttcher, S. M. Grudsky, Toeplitz Matrices, Asymptotic Linear Algebra and Functional Analysis, Hindustan Book Agency, 2000.

[27] W. V. Li, Q.-M. Shao, A normal comparison inequality and its applications, Probability Theory and Related Fields 122 (2002) 494–508.

[28] Z. Šidák, Rectangular confidence regions for the means of multivariate normal distributions, Journal of the American Statistical Association 62 (318) (1967) 626–633.

[29] L. Dümbgen, Bounding standard Gaussian tail probabilities, arXiv:1012.2063, `https://arxiv.org/abs/1012.2063` (2010).

[30] J. A. Tropp, An Introduction to Matrix Concentration Inequalities, Foundations and Trends in Machine Learning, 2014, `https://tropp.caltech.edu/books/Tro14-Introduction-Matrix-FnTML-rev.pdf`.

[31] R. A. Horn, C. R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991.