

The hidden cost of correlated noise

When more samples do not mean more information

Angelo Maria Sabatini

2026-06-19

Short-range dependent stochastic processes

The random-walk model introduced in the previous post of this series ([here](#)) is often unrealistic for slowly varying quantities. A random walk possesses infinite memory; once a perturbation affects the system, its impact is never fully forgotten. Consequently, the variance grows indefinitely, and the process continuously accumulates uncertainty. Many real physical systems, however, gradually lose memory of past disturbances. Mechanical damping, thermal relaxation, fluid dynamics, electronic feedback, and various dissipative phenomena typically produce stochastic dynamics whose memory fades over time.

The simplest mathematical framework to capture this fading memory is the first-order Gauss–Markov (GM1) process. This model effectively bridges the gap between pure randomness and deterministic decay. Like any first-order Markov process, the GM1 process depends on the past only through its current state, making the future conditionally independent of the remaining history. This property yields a highly tractable linear stochastic differential equation:

$$\dot{x}(t) = -bx(t) + w(t), \tag{1}$$

where $b > 0$ is a damping rate and $w(t)$ is continuous-time white Gaussian noise with strength q :

$$R_{WW}(\tau) = \mathbb{E}[w(t)w(t + \tau)] = q\delta(\tau).$$

In probability theory and statistical physics, this process is commonly known as the Ornstein–Uhlenbeck process, originally introduced to model Brownian motion with friction. In engineering and signal-processing applications, the equivalent formulation is often referred to as a first-order Gauss–Markov process.

The term $-bx(t)$ in Equation 1 continuously pulls the process back toward a mean state, formalizing the concept of thermal relaxation or mechanical damping introduced above. Meanwhile, the stochastic forcing injects random fluctuations into the dynamics. The variance does not grow indefinitely but converges to a finite, stationary value σ_x^2 , thereby providing a simple model for systems with short-range dependence.

The corresponding autocorrelation function is

$$R_{XX}(\tau) = \sigma_X^2 \exp(-b|\tau|) = \frac{q}{2b} \exp(-b|\tau|), \quad (2)$$

showing that the process forgets its past over the characteristic correlation time

$$\tau_c = \frac{1}{b}.$$

Using the correlation time τ_c , Equation 2 can be rewritten as

$$R_{XX}(\tau) = \frac{q\tau_c}{2} \exp\left(-\frac{|\tau|}{\tau_c}\right), \quad (3)$$

For a sampling time ΔT the corresponding exact discrete-time model becomes

$$x_{k+1} = A x_k + \Delta w_k,$$

with

$$A = \exp\left(-\frac{\Delta T}{\tau_c}\right),$$

where Δw_k represents the stochastic increment resulting from integrating white noise over the sampling interval ΔT .

The corresponding discrete-time process-noise variance is

$$Q = \frac{q\tau_c}{2} \left[1 - \exp\left(-\frac{2\Delta T}{\tau_c}\right)\right].$$

Two important limiting regimes immediately emerge from these expressions. For very small sampling intervals, $A \approx 1 - \Delta T/\tau_c \approx 1$ and $Q \approx q \Delta T$, and the process locally behaves like the random walk introduced in the previous post. Conversely, for sampling intervals much longer than the correlation time, $A \approx 0$ and $Q \approx q\tau_c/2$, meaning that consecutive samples progressively lose statistical dependence and the process approaches a memoryless regime.

💡 Kalman estimator

The exact discrete-time representation is particularly useful because it can be incorporated directly into a Kalman-filter (KF) model. The state is not observed directly. Instead, measurements are assumed to be affected by additive Gaussian noise:

$$z_k = x_k + v_k, \quad v_k \sim \mathcal{N}(0, R).$$

The resulting state-space model is therefore

$$\begin{cases} x_{k+1} = Ax_k + \Delta w_k, \\ z_k = x_k + v_k, \end{cases}$$

where A and Q are given by the exact discretization above.

At steady state, the updated covariance satisfies the algebraic Riccati equation

$$P_\infty = A^2 P_\infty + Q - \frac{(A^2 P_\infty + Q)^2}{A^2 P_\infty + Q + R}, \quad (4)$$

whose positive solution is

$$P_\infty = \frac{-(Q + R - A^2 R) + \sqrt{(Q + R - A^2 R)^2 + 4A^2 R Q}}{2A^2}. \quad (5)$$

The corresponding steady-state Kalman gain is

$$K_\infty = \frac{A^2 P_\infty + Q}{A^2 P_\infty + Q + R}. \quad (6)$$

After the initial transient, the KF behaves as a stationary recursive estimator with constant asymptotic gain.

From a discrete-time perspective, the sampled GM1 process is equivalent to a first-order autoregressive process, usually denoted AR(1). This equivalence connects the continuous-time physical interpretation of the GM1 process with the classical theory of correlated discrete-time stochastic sequences. Because the discretization is exact, the discrete-time autocorrelation sequence is simply obtained by sampling the continuous-time autocorrelation function of Equation 2:

$$r_{XX}[h] = R_{XX}(h\Delta T) = \frac{q\tau_c}{2} \exp\left(-\frac{\Delta T}{\tau_c}|h|\right),$$

showing that adjacent samples become strongly correlated when the sampling interval ΔT is much shorter than the correlation time τ_c .

It is noteworthy that, for a sequence of n observations generated by a GM1 process, the variance of the sample mean is

$$\text{Var}(\bar{X}_n) = \frac{q\tau_c}{2n} \left[1 + 2 \sum_{h=1}^{n-1} \left(1 - \frac{h}{n}\right) \phi^h \right],$$

where $\phi = \exp(-\Delta T/\tau_c)$. This variance is always larger than in the limiting case $\phi = 0$, corresponding to statistically independent observations. In that case, the classical result is recovered:

$$\text{Var}(\bar{X}_n) = \frac{q\tau_c}{2n}$$

In practical terms, increasing the sampling frequency within the correlation-time range of the process provides far less independent information than expected under independence assump-

tions.

Simulation

The discrete-time GM1 model introduced above can be incorporated directly into a scalar Kalman filter. In this simulation, the filter is deliberately matched to the process that generated the data: the same values of A , Q , H , and R are used both in the simulation model and in the estimator.

This represents the simplest possible scenario. Any discrepancy between the true state and the estimate therefore arises solely from the stochastic uncertainty associated with the process and measurement noise, rather than from model mismatch.

Figure 1 shows the estimation error together with the approximate 95% confidence bounds predicted by the KF. After the initial transient, the error behaves as expected for a correctly specified model: it fluctuates around zero and mostly remains within the predicted uncertainty bounds.

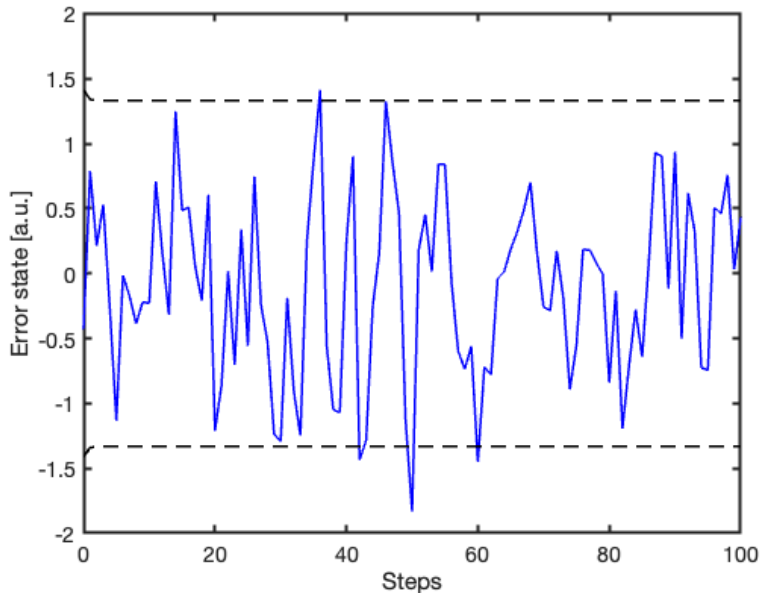


Figure 1: Estimation error and approximate 95% confidence bounds predicted by the KF ($\pm 2\sqrt{P_k}$). After the initial transient, the error remains mostly confined within the uncertainty interval associated with the estimated covariance.

Figure 2 shows the corresponding Kalman gain. After a short transient, the gain reaches a steady-state value, reflecting the stationary nature of the underlying process and the steady balance between prediction and measurement uncertainty.

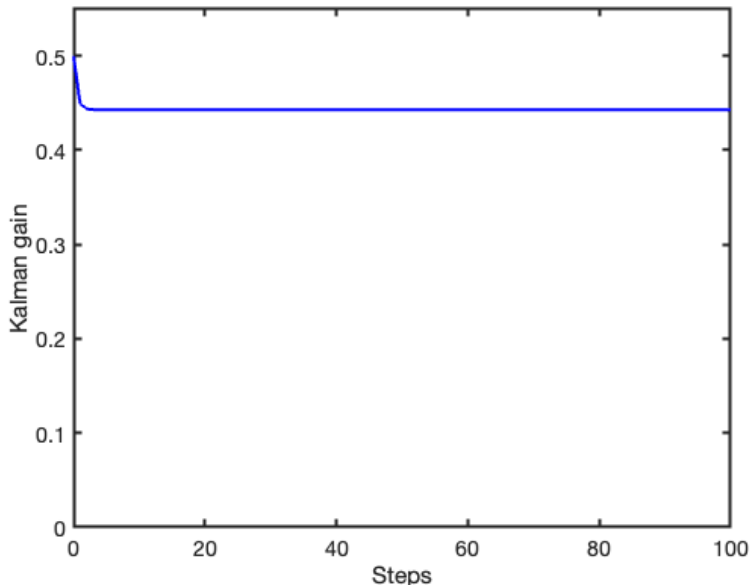


Figure 2: Evolution of the Kalman gain.

For the parameters used in this simulation ($b = 0.5 \text{ s}^{-1}$, $\Delta t = 1 \text{ s}$, $\sigma_X = 1 \text{ a.u.}$, and $R = 1 \text{ (a.u.)}^2$), the scalar Riccati equation (Equation 4) predicts a steady-state covariance (Equation 5) and Kalman gain (Equation 6) with the same numerical value, although the associated physical units differ:

$$P_\infty = K_\infty \approx 0.443.$$

This result is in excellent agreement with the simulated gain trajectory shown in Figure 2. The equality $P_\infty = K_\infty$ follows here from the scalar structure of the problem with $H = 1$ and $R = 1$.

Probing the environment

The GM1 process introduced above should not be regarded merely as a convenient mathematical model. In many sensing applications, temporally correlated fluctuations are not simply a nuisance: they are the signature of the physical medium being observed.

A barometric pressure sensor, for example, does not only measure the electronic noise of its readout chain. It also responds to slow pressure fluctuations, thermal effects, air motion, and environmental dynamics. Similarly, an ultrasonic time-of-flight system probes a propagation medium whose temperature, humidity, turbulence, and local gradients affect the measured delay.

In this sense, many sensors act as environmental probes. Their output reflects not only the sensor itself, but also the dynamics of the surrounding medium, which often evolves over finite correlation times.

The key point is that correlated noise is not always “bad noise”. Sometimes it is the phenomenon of interest. Rather than averaging it away blindly, it is often more informative to characterize its temporal structure. From a measurement perspective, a related question is how much inde-

pendent information is actually gained by sampling faster: is oversampling really a good recipe to improve signal-to-noise ratio?

Allan deviation as a multiscale diagnostic tool

Allan deviation provides exactly the multiscale view we need to reveal how variability changes across different averaging times. Originally introduced in the context of oscillator stability analysis, it is now routinely used to characterize temporally correlated stochastic processes across many areas of measurement science and signal processing.

The following simulation reproduces a long-duration acquisition from a stationary GM1 process with correlation time $\tau_c = 1$ s. The process is sampled every $T_s = 20$ ms, corresponding to an oversampling ratio of 50 samples per correlation time. The total observation time is approximately three hours (about 175 minutes), yielding a record of 2^{19} samples. A white Gaussian sequence with the same stationary variance is generated as a reference. This allows the Allan-deviation analysis to isolate the effects of temporal correlation from those associated solely with signal variance.

The GM1 process can be interpreted as the output of a first-order stochastic shaping filter driven by white Gaussian noise. Figure 3 compares a short segment of the reference white-noise sequence with the corresponding correlated GM1 output generated by the discrete-time model. Although the two processes have approximately the same variance, their temporal structure is markedly different. The GM1 process appears visibly smoother because its dynamics suppress rapid fluctuations and introduce finite temporal correlation. This apparent smoothness should not be confused with a larger amount of independent information. Samples acquired within a correlation time remain strongly redundant, even when the signal appears less noisy.

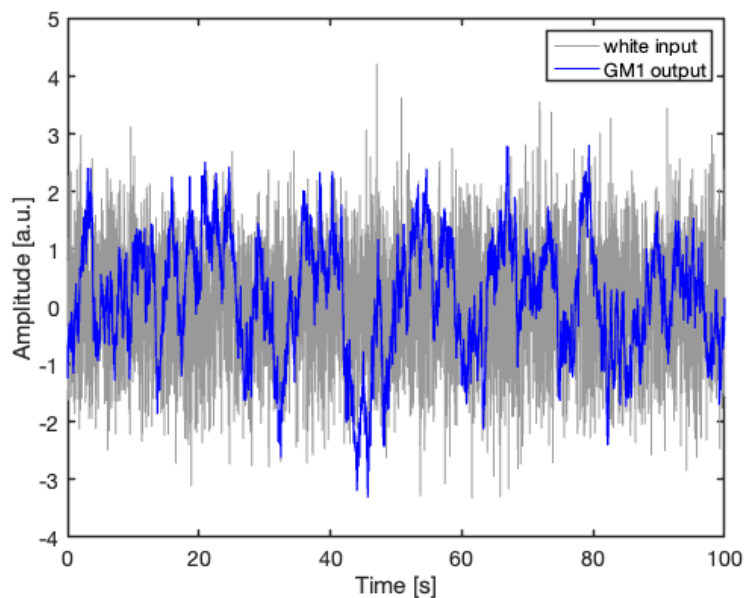


Figure 3: Short segment of the simulated white-noise input and corresponding GM1 output.

The Allan-deviation curves obtained from the simulated sequences are shown in Figure 4.

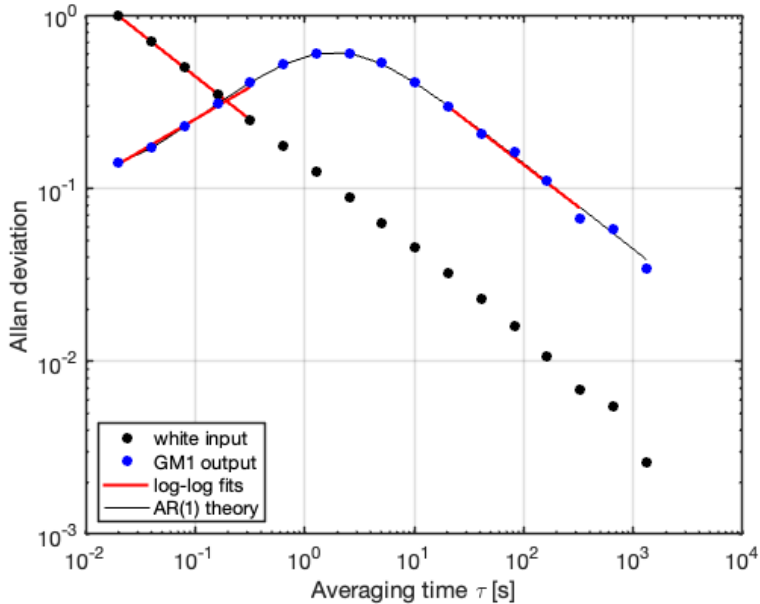


Figure 4: Allan-deviation comparison between white noise and a GM1 process with correlation time $\tau_c = 1$ s. Red segments indicate local log–log regressions used for slope estimation, while the solid black curve represents the theoretical AR(1) prediction.

Several features of temporal correlation become immediately apparent.

For the white-noise input, the Allan deviation follows the familiar power law

$$\sigma_A(\tau) \propto \tau^{-1/2},$$

indicating that longer averaging times continuously increase the amount of statistically independent information available through averaging. The corresponding log–log regression yields a slope of -0.500 (95% CI $[-0.502, -0.497]$), in excellent agreement with the theoretical value $-1/2$.

The GM1 process behaves very differently. At short averaging times, its Allan deviation lies below the white-noise reference because the first-order dynamics suppress rapid fluctuations and smooth the process locally. However, as the averaging time approaches the correlation-time region, the effects of temporal memory progressively dominate the averaging process.

In this regime, consecutive observations no longer behave as statistically independent measurements. The Allan-deviation curve departs from the white-noise scaling law and exhibits a markedly different slope structure. The short-time log–log regression yields a positive slope of 0.371 (95% CI $[0.296, 0.446]$), directly reflecting the fact that local oversampling is collecting strongly correlated information rather than statistically independent samples. In other words, additional samples are being acquired faster than new independent information is being generated by the underlying process.

Only for averaging times significantly longer than the correlation time does the process progressively recover a white-noise-like averaging behavior. In this asymptotic regime, the long-time

slope approaches the classical $-1/2$ value expected for effectively decorrelated observations. The corresponding regression yields a slope of -0.489 (95% CI $[-0.582, -0.395]$).

The theoretical Allan-deviation curve derived from the equivalent AR(1) representation closely matches the simulated GM1 behavior across all averaging scales, confirming the exact equivalence between the sampled Gauss–Markov process and its discrete-time autoregressive representation.

i A closer look at Allan averaging

The Allan variance can be interpreted in a particularly illuminating way when the underlying process possesses temporal correlation.

For a sequence divided into adjacent blocks of length n , define the local averages

$$\bar{x}_{1,n} = \frac{1}{n} \sum_{k=1}^n x_k,$$

and

$$\bar{x}_{2,n} = \frac{1}{n} \sum_{k=n+1}^{2n} x_k.$$

The Allan variance is constructed from the squared difference between consecutive local averages:

$$\sigma_A^2(n) = \frac{1}{2} \mathbb{E} [(\bar{X}_{2,n} - \bar{X}_{1,n})^2].$$

Using the elementary identity

$$\text{Var}(A - B) = \text{Var}(A) + \text{Var}(B) - 2\text{Cov}(A, B),$$

and exploiting stationarity,

$$\text{Var}(\bar{X}_{1,n}) = \text{Var}(\bar{X}_{2,n}),$$

one immediately obtains

$$\sigma_A^2(n) = \text{Var}(\bar{X}_n) - \text{Cov}(\bar{X}_{1,n}, \bar{X}_{2,n}). \quad (7)$$

This expression provides a remarkably intuitive interpretation of Allan variance. The first term, $\text{Var}(\bar{X}_n)$, describes the uncertainty reduction produced by averaging. The second term, $\text{Cov}(\bar{X}_{1,n}, \bar{X}_{2,n})$, measures the temporal redundancy that survives between adjacent local averages.

For white noise, adjacent averages are statistically independent and the covariance term vanishes. Allan averaging therefore behaves exactly as expected from the familiar $1/\sqrt{n}$ law. For correlated processes such as GM1 or AR(1), however, adjacent averages remain statistically dependent over time scales comparable with the correlation time.

This expression reveals that Allan variance is not simply the variance of a sample mean. A correlated process may even exhibit a smaller Allan deviation than white noise over some averaging ranges. This does not imply that more independent information is being collected.

Rather, it reflects the fact that temporal correlation smooths short-term fluctuations while simultaneously increasing redundancy between neighboring samples.

The key question is therefore not how small the Allan deviation becomes at a particular averaging time, but how efficiently additional averaging continues to reduce uncertainty.

Concluding remarks

At first sight, averaging appears to be a straightforward recipe for reducing uncertainty: collect more samples, average them, and the noise decreases. This intuition is perfectly valid when observations are statistically independent. However, many real-world processes possess temporal memory, and successive samples may carry strongly redundant information.

The first-order Gauss–Markov process provides a simple yet powerful framework for describing this behavior. Through its finite correlation time, it bridges the gap between purely random fluctuations and deterministic dynamics, while remaining mathematically tractable and directly compatible with Kalman filtering.

The analysis based on Allan deviation highlights an important consequence of temporal correlation. Two processes may exhibit the same variance while carrying very different amounts of independent information across time. In such situations, increasing the sampling frequency does not necessarily yield the statistical benefits predicted by classical independence assumptions.

Understanding how uncertainty, correlation, and information evolve across time scales is therefore just as important as understanding the magnitude of the fluctuations themselves. In many measurement problems, the key question is not simply how noisy a signal is, but how much new information each additional sample actually provides.