In many applications expectation values are calculated by partitioning a single experimental time series into an ensemble of data segments of equal length. Such single trajectory ensemble (STE) is a counterpart to a multiple trajectory ensemble (MTE) used whenever independent measurements or realizations of a stochastic process are available. The equivalence of STE and MTE for stationary systems was postulated by Wang and Uhlenbeck in their classic paper on Brownian motion (Rev. Mod. Phys. 17, 323 (1945)) but surprisingly has not yet been proved. Using the stationary and ergodic paradigm of statistical physics – the Ornstein-Uhlenbeck (OU) Langevin equation, we revisit Wang and Uhlenbeck’s postulate. In particular, we find that the variance of the solution of this equation is different for these two ensembles. While the variance calculated using the MTE quantifies the spreading of independent trajectories originating from the same initial point, the variance for STE measures the spreading of two correlated random walkers. Thus, STE and MTE refer to two completely different dynamical processes. Guided by this interpretation, we introduce a novel algorithm of partitioning a single trajectory into a phenomenological ensemble, which we name a threshold trajectory ensemble (TTE), that for an ergodic system is equivalent to MTE. We find that in the cohort of healthy volunteers, the ratio of STE and TTE asymptotic variances of stage 4 sleep electroencephalogram is equal to 1.96 ± 0.04 which is in agreement with the theoretically predicted value of 2.

PACS numbers: 05.40.-a, 05.10.Gg, 05.45.Tp, 87.10.Mn
In Fig. 1 we present a solution to the Ornstein-Uhlenbeck (OU) Langevin equation:

\[ \frac{dX(t)}{dt} = -\lambda X(t) + \eta(t) \]  

(1)

where \( \lambda \) is the dissipation rate. In the above equation, a zero-centered Gaussian random force \( \eta(t) \) is delta correlated in time:

\[ \langle \eta(t) \eta(t + \tau) \rangle = \sigma^2 \delta(\tau) \]  

(2)

and the angular brackets denote an average over an ensemble of realizations of the random force. \( X(t) \) is also a Gaussian random process with a spectrum:

\[ G(f) = \frac{2\sigma^2}{\lambda^2 + 4\pi^2 f^2}. \]  

(3)

Consequently, from the convolution theorem one obtains the autocorrelation function of \( X(t) \):

\[ \rho(t) = e^{-\lambda t}. \]  

(4)

\( X(t) \) may be expressed as the formal solution to the first-order, linear, stochastic differential equation:

\[ X(t) = e^{-\lambda t} \left[ X(0) + \int_0^t \eta(t')e^{\lambda t'} dt' \right]. \]  

(5)

In Fig. 1 we present a solution \( X(t) \) generated by numerical integration of Eq. (1) with a constant time step \( \Delta t = 1 \) (\( \lambda = 0.025 \) and \( \sigma_\eta = 7.8 \)).

According to Wang and Uhlenbeck’s prescription, a long time series generated by the OU Langevin equation can be partitioned to yield a STE. We know that the solution to the OU Langevin equation is ergodic so that averages obtained using STE and MTE should coincide. Variance is the most commonly used measure of time series variability. Therefore, let us perform computer simulations to calculate this metric for both ensembles. In Fig. 2, \( \sigma^2 \) of the solution \( X(t) \) of the Langevin equation is plotted as a function of the length \( \tau \) of the data window. The MTE variance \( \sigma^2 \) is denoted by open squares and was calculated using \( n_M = 3000 \) trajectories of length \( N_M = 1000 \) originating at zero. The STE variance \( \sigma^2_X \), represented by open circles, was computed by partitioning a single trajectory of length \( N_S = n_M N_M \) into segments of length \( \tau \) (we used a sliding window algorithm). It is obvious that the two ways of calculating the variance are not equivalent.

Figure 1: An example of the solution \( X(t) \) of the OU Langevin equation for \( \lambda = 0.025 \) and \( \sigma_\eta = 7.8 \). The horizontal bars in the upper part of the figure represent those segments of the displayed trajectory whose left endpoints are equal, with a predetermined accuracy, to \( X_L \). In other words, the left endpoints are the intersection of the trajectory with the chosen threshold which is marked in the graph by the horizontal gridline. Such segments of length \( \tau \) are used to construct a threshold trajectory ensemble discussed later in the text.

To clarify the difference between the two phenomenological ensembles depicted in Fig. 2 we derive the analytical expressions for the variance of the solution \( X(t) \) for both ensembles. Let us consider first the MTE of an infinite number of trajectories all starting from zero. Using Eq. (5) we obtain

\[ \sigma^2_M(t) \equiv \langle X^2(t) \rangle_M - \langle X(t) \rangle_M^2 = \frac{\sigma_\eta^2}{2\lambda} [1 - e^{-2\lambda t}]. \]  

Please note that STE averaging, by its very nature, in-
volves relative displacements: \( Z(t, \tau) = X(t + \tau) - X(t) \).
Using Eq. (5), we can write \( X(t + \tau) \) as

\[
X(t + \tau) = e^{-\lambda \tau} \left[ X(t) + \int_0^\tau \eta(t + t') e^{\lambda t'} dt' \right] (7)
\]

and express \( Z(t, \tau) \) in the following form

\[
Z(t, \tau) = e^{-\lambda \tau} \int_0^\tau \eta(t + t') e^{\lambda t'} dt' + X(t) \left[ e^{-\lambda \tau} - 1 \right]. (8)
\]

Recall that the random force \( \eta \) is stationary and, therefore, the time translation of \( \eta \) in the integrand in Eq. (8) does not affect the statistical properties of \( Z(t, \tau) \). Without loss of generality, we may assume that a trajectory \( X(t) \) starts at zero and then \( Z(t, \tau) \) may be written in a particularly illuminating form

\[
Z(t, \tau) = X(\tau) + X(t) \left[ e^{-\lambda \tau} - 1 \right]. (9)
\]

For a fixed value of \( t \) (fixed left endpoint of the interval), the relative displacement \( Z(t, \tau) \) is a function of segment length \( \tau \). In particular, the first term on the r.h.s. of Eq. (9) is stochastic while the second one is purely deterministic. Consequently, the covariation of these two terms vanishes. Thus, the partitioning of a single trajectory is equivalent to building up deterministic trends. \( Z(t, \tau) \) being the sum of Gaussian variables is a Gaussian variable itself with zero mean value \( \langle E[Z(t, \tau)] = 0 \rangle \) and the following variance

\[
E[Z^2(t, \tau)] = \sigma^2_M(\tau) + \left[ e^{-\lambda \tau} - 1 \right]^2 \sigma^2_M(t). (10)
\]

The STE variance \( \sigma^2_S(\tau) \) is just \( E[Z^2(t, \tau)] \) time averaged along the trajectory of length \( T \)

\[
\sigma^2_S(\tau) = \frac{1}{T - \tau} \int_0^{T-\tau} E[Z^2(t, \tau)] dt (11)
\]

\[
= \sigma^2_M(\tau) + \frac{1}{2\lambda} \left[ e^{-2\lambda(T-\tau)} - 1 \right].
\]

Taking into account that \( \lambda T \gg 1 \), we obtain the following approximation

\[
\sigma^2_S(\tau) \approx \frac{\sigma^2_M(\tau)}{\lambda} (1 - e^{-\lambda \tau}). (12)
\]

Thus, the variance for the single trajectory ensemble is given by the formula Eq. (12) for the MTE, albeit with the effective dissipation rate \( \lambda_{eff} = \lambda/2 \) which is half that of the MTE. Consequently, the ratio of the asymptotic variances for STE and MTE is

\[
\frac{\sigma^2_S(\infty)}{\sigma^2_M(\infty)} = 2. (13)
\]

Both curves \( \sigma^2_S(\tau) \) and \( \sigma^2_M(\tau) \) are plotted in Fig. 2 and are in agreement with the relevant numerical calculations.

In hindsight, the observed disagreement between the two ways of calculating the variance is less surprising than it ought to have been. \( \sigma_M \) is the measure of spreading of statistically independent trajectories that start at \( X(0) \). On the other hand, the endpoints of intervals used to calculate the relative displacements \( Z(t, \tau) \) may be interpreted as the final positions of two correlated random walkers who both start at \( X(0) \) and whose correlation function \( \rho(t) \) has the exponential time dependence given by Eq. (14). \( \sigma_S(\tau) \) quantifies the spread of the distance between such walkers after time \( \tau \) and consequently refers to a completely different dynamical process. The formal proof of this interpretation is given elsewhere [7].

Variance is certainly the most prevalent measure of time series variability. Moreover, it is often a critical part of fractal scaling detection algorithms, such as detrended fluctuation analysis (DFA) [8, 9]. In light of the difference between the STE and MTE variances, one can easily envision the situation when simultaneous application of both ensembles appears rational, but ultimately leads to systematic errors. For example, one may perform the measurements on a cohort of subjects to determine the variability of a physiological quantity. However, when the variability determined for a given patient is compared with that of the cohort, one may be inclined to improve the statistics by averaging over a single trajectory ensemble. We know that for the OU Langevin model, this approach leads to the gross overestimation of the asymptotic variance. Thus, the question arises as to whether it is possible to partition a single trajectory in such a way that the resulting ensemble is equivalent to MTE and the improvement in statistics is accomplished.

The solution to this problem presents itself as soon as we realize that in STE both endpoints of intervals contribute to spreading, whereas in MTE all left endpoints are the same initial condition. It is this difference between the two ensembles that explains why the asymptotic variance for STE is exactly twice that for MTE. The horizontal bars in the upper part of Fig. 1 represent those segments of the displayed trajectory whose left endpoints are equal, with a predetermined accuracy \( \epsilon \), to the \( X_L \). In other words, the left endpoints are the intersections of the trajectory with the chosen threshold which is marked in the graph by the horizontal gridline. Such segments of length \( \tau \) are used to construct a threshold trajectory ensemble (TTE). The filled squares in Fig. 2 correspond to the variance \( \sigma^2_T \) for such an ensemble. The data segments with \( X_L = 0 \) (\( \epsilon = 0.87 \)) which corresponds to 2.5% of the standard deviation of the trajectory of length \( N_S = n_M N_M \) shown in Fig. 1 were selected. There were approximately 60000 segments that satisfied the imposed criteria and, despite the relatively small size of the TTE, the agreement with the MTE is apparent.
The nonequilibrium statistical mechanics of open and closed systems (VCH New York, 1990).

Let us finish with a cautionary note. The ergodic hypothesis dates back to the very beginning of statistical physics. The recent studies [11–17] have once again brought ergodicity from the backstage into the limelight. The growing list of non-ergodic systems should warn against indiscriminate application of single trajectory (single particle) ensembles.

**References**

[1] L. Boltzmann, *Lectures on Gas Theory* (Dover, 1995), first published in 1895, translated by S.G. Brush.

[2] G. Birkoff, *PNAS* **17**, 656 (1931).

[3] A. Kinchin, *Mathematical Foundations of Statistical Mechanics* (Dover, New York, 1949), translated by G. Gamow.

[4] M. Wang and G. Uhlenbeck, Reviews of Modern Physics **17**, 323 (1945).

[5] M. Ignaccolo, M. Latka, W. Jernajczyk, P. Grigolini, and B. J. West, *Phys Rev E* **81**, 031909 (2010).

[6] K. Lindenberg and B. J. West, *The nonequilibrium statistical mechanics of open and closed systems* (VCH New York, 1990).

[7] M. Ignaccolo, M. Latka, and B. J. West, *Europhys Lett* (2010).

[8] C. K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, *Phys Rev E* **49**, 1685 (1994).

[9] A. Bashan, R. Bartsch, J. Kantelhardt, and S. Havlin, *Physica A* **387**, 5080 (2008).

[10] E. Niedermeyer and F. Da Silva, *Electroencephalography: basic principles, clinical applications, and related fields* (Lippincott Williams & Wilkins, 2004).

[11] N. Masuda and K. Aihaara, *Neural Comput* **15**, 1341 (2003), URL *http://dx.doi.org/10.1162/089976603321780308*.

[12] X. Brokmann, J.-P. Hermier, G. Messin, P. Desbiolles, J.-P. Bouchaud, and M. Dahan, *Phys Rev Lett* **90**, 120601 (2003).

[13] G. Margolin and E. Barkai, *Phys Rev Lett* **94**, 080601 (2005).

[14] J. Tang and R. A. Marcus, *J Chem Phys* **123**, 204511 (2005), URL *http://dx.doi.org/10.1063/1.2128409*.

[15] G. Margolin, V. Protasenko, M. Kuno, and E. Barkai, *J Phys Chem B* **110**, 19053 (2006), URL *http://dx.doi.org/10.1021/jp061487m*.

[16] J. McCauley, *Physica A* **387**, 5518 (2008).

[17] A. Lubelski, I. M. Sokolov, and J. Klafter, *Phys. Rev. Lett.* **100**, 250602 (2008).