# Data Assimilation Using Noisy Time-Averaged Measurements

Jordan Blocher
Department of Mathematics
University of Nevada-Reno
Reno, NV 89557, USA
jblocher@nevada.unr.edu

Vincent R. Martinez
Department of Mathematics
Tulane University
New Orleans, LA 70118, USA
vmartin6@tulane.edu

Eric Olson
Department of Mathematics
University of Nevada-Reno
Reno, NV 89557, USA
ejolson@unr.edu

November 24, 2017

**Abstract**

We study the synchronization of chaotic systems when the coupling between them contains both time averages and stochastic noise. Our model dynamics—inspired by the partial differential equations which govern the atmosphere—are given by the Lorenz equations which are a system of three ordinary differential equations in the variables $X$, $Y$ and $Z$. Our theoretical results show that coupling two copies of the Lorenz equations using a feedback control which consists of time averages of the $X$ variable leads to exact synchronization provided the time-averaging window is known and sufficiently small. In the presence of noise the convergence is to within a factor of the variance of the noise. We also consider the case when the time-averaging window is not known and show that it is possible to tune the feedback control to recover the size of the time-averaging window. Further numerical computations show that synchronization is more accurate and occurs under much less stringent conditions than our theory requires.

*Dedicated to Edriss S. Titi*
*on the occasion of his 60th birthday*
*with great respect and admiration*

# 1 Introduction

A nudging algorithm similar to the one studied here, but without random noise or blur due to time averaging, was introduced in [1] as a data assimilation algorithm and then analyzed for the model problem of the two-dimensional incompressible Navier–Stokes equations. A related numerical study was performed in [10], see also [11]. Farhat, Lunasin and Titi [7] show that such algorithms work when observational data is collected only for horizontal velocity measurements. Such results also hold in the context of Bénard convection [8]. Noiseless measurements that include time blur were studied in [15] in the context of the surface quasi-geostrophic equation. In this work we focus on the interaction between time blur and noise.

The effects of noisy measurements of the two-dimensional Navier–Stokes equations without time blur were treated in [2], which led to a stochastic differential equation. Nudging algorithms based on discrete-in-time observations of the two-dimensional Navier–Stokes equations were analyzed in [13] and with errors by Foias, Mondaini and Titi [9], see also [4]. None of these works explore the interaction between time blur and noise and we remark that such interaction results in stress terms which lead to a non-closed system of differential inequalities for the model problem of the two-dimensional Navier–Stokes equations. Fortunately, such stress terms do not occur in the treatment of the Lorenz system considered here.

Synchronization between copies of the Lorenz system was first observed by Pecora and Carroll in [21]. The Lorenz system was considered in the context of data assimilation by Hayden [12], see also [13], and to study the effects of noisy observational measurements by Law, Shukla and Stuart in [16]. The present work considers noisy measurements that have been further contaminated by a moving time average with respect to an averaging window of size $\delta$. This type of moving average is inspired by real-world scientific instrumentation that approximates measurements at any instant in time by averages taken over a very small time interval.

The system of ordinary differential equations

$$\begin{cases} \dot{X} = -\sigma X + \sigma Y \\ \dot{Y} = -\sigma X - Y - XZ \\ \dot{Z} = -bZ + XY - b(r+\sigma) \end{cases} \tag{1.1}$$

was proposed by Lorenz in 1963 as a simplified model of atmospheric convection [18]. Here $\sigma$ is the Prandtl number, $r$ is the Rayleigh number, and $b$ is a geometric factor. We will use the standard parameter values $\sigma = 10$, $b = 8/3$ and $r = 28$. With these values the Lorenz system is known to posses

a chaotic global attractor $\mathcal{A}$, see Tucker [23].

Throughout this paper we shall refer to the solution $U = (X, Y, Z)$ as the free-running solution and assume it lies on the global attractor. Thus, $U$ reflects the long term behavior of the Lorenz systems and obeys the *a priori* bounds given in Doering and Gibbon [6], see also [5], stated here as

**Theorem 1.1.** *Let $U$ be a trajectory with $U_0 \in \mathcal{A}$. Then $\|U(t)\|^2 \leq K$ for all $t \in \boldsymbol{R}$ where*

$$K = \frac{b^2(r+\sigma)^2}{4(b-1)} \approx 1540.27. \tag{1.2}$$

As the previous studies consider coupling through the $X$ variable as a model for data assimilation, we, therefore, suppose that measurements of the free-running solution are given by time averages of $X$. Although the result for a time-averaged measurement could not physically be available until the end of the averaging window, theoretically such a measurement most accurately represents the state of the system at a time corresponding to the midpoint of the averaging window. Thus, we might best approximate $X$ by a moving time average as

$$X(t) \approx \frac{1}{\delta} \int_{t-\delta/2}^{t+\delta/2} X(s)\, ds. \tag{1.3}$$

While this is a reasonable approximation, it is necessary for the feedback control to use exactly the same approximation in order to obtain the important property that in the noise-free case if both copies of the Lorenz system are given the same initial conditions then they remain identical for all time. However, for reasons of causality and mathematical well-posedness it is difficult to use the approximation given by (1.3) in a feedback control.

Rather than giving up the possibility of exact synchronization in the noise-free case, we obtain better results by taking

$$X(t) \approx \frac{1}{\delta} \int_{t-\delta}^{t} X(s)\, ds$$

so that the same approximation can be used in the feedback term for the nudging algorithm. We, therefore, suppose that our blurred-in-time noisy measurements of the free-running solution $U$ are given by

$$\mathcal{M}(t) = \frac{1}{\delta} \int_{t-\delta}^{t} \big(X(s)\, ds + \varepsilon\, dW(s)\big) = \frac{1}{\delta} \int_{t-\delta}^{t} X(s) ds + \xi(t),$$

where $W(t)$ is a standard one-dimensional Brownian motion with underlying probability space $\Omega$ and $\xi(t) = \varepsilon\big(W(t) - W(t-\delta)\big)/\delta$. Note for every $\omega \in \Omega$ that $\xi$ is continuous and at every time $t$ that

$$\mathbf{E}\big[\xi(t)\big] = 0 \qquad \text{and} \qquad \mathbf{E}\big[\xi(t)^2\big] = \varepsilon^2/\delta.$$

Thus, $\mathcal{M}(t)$ represents an approximation of $X(t)$ given by incomplete noisy measurements of the phase space of $U$ that have been blurred in time by means of a moving average.

Under ideal conditions the exact size $\delta$ of the averaging window of the measuring equipment would be known. In practice, only an approximation might be known which we denote here by $\delta_*$. We now use the observations $\mathcal{M}(t)$ to couple a second copy of the Lorenz system to the first by means of the nudging algorithm

$$\begin{aligned} \dot{x} &= \begin{cases} -\sigma x + \sigma y & \text{for } t < \delta_* \\ -\sigma x + \sigma y - \dfrac{\mu}{\delta_*} \displaystyle\int_{t-\delta_*}^{t} x + \mu \mathcal{M}(t) & \text{for } t > \delta_* \end{cases} \\ \dot{y} &= -\sigma x - y - xz \\ \dot{z} &= -bz + xy - b(r+\sigma) \end{aligned} \tag{1.4}$$

with arbitrary initial condition $u_0 = (x_0, y_0, z_0)$. Here $\mu$ represents a tunable relaxation parameter with dimensions of inverse time. We shall refer to $u = (x, y, z)$ as the driven solution and insist that it be continuous. Note, however, that the above piecewise definition implies $\dot{x}$ will be discontinuous unless by chance

$$\frac{1}{\delta} \int_{\delta_* - \delta}^{\delta_*} X(s) ds + \xi(\delta_*) = \frac{1}{\delta_*} \int_0^{\delta_*} x(s) ds.$$

In the absence of a time blur (1.4) reduces to the stochastic differential equations studied in [16], see also [2] and [3]. In the present case, the addition of a time blur leads to a random integro-differential equation. In particular $u$ depends on $\omega \in \Omega$ and when it is important to explicitly denote this we shall refer to the solution of (1.4) as $u(\omega; t)$. Our aim is to find analytic bounds on $\mathbf{E}\big[\|U - u\|^2\big]$ in terms of $\delta$, $\delta_*$, $\varepsilon$ and $\mu$ and to then check those bounds through numeric simulation.

Along these lines, our main analytical result may be stated as

**Theorem 1.2.** *Let $\mu \geq K$ where $K$ is the bound given in (1.2) and suppose*

$$\max(\delta, \delta_*) < \frac{1 - e^{-1}}{16(\sigma + \mu)^2}.$$

*Then*

$$\limsup_{t\to\infty} \mathbf{E}\big[\|U-u\|^2\big] \le 4e\Big(\frac{\delta+\delta_*}{\delta}\Big)\mu^2\varepsilon^2 + 16e\mu K\Big(\frac{\delta_*-\delta}{\delta}\Big)^2.$$

The upper bound in the above theorem consists of two terms: the first term is proportional to $\varepsilon^2$ the variance of the error in the observations; the second term is proportional to the square of the relative error in $\delta_*$ our estimate of the length of the averaging window. Note that as $\varepsilon \to 0$ the first term vanishes. Similarly, if $\delta_* \to \delta$ then the second term vanishes while the first becomes independent of $\delta$. When $\delta_* = \delta$ we obtain a bound depending only on the variance of the noise which is comparable to the bound

$$\limsup_{t\to\infty} \mathbf{E}\big[\|U-u\|^2\big] \le \frac{\mu^2\varepsilon^2}{\sigma+\mu+1-\sqrt{(\sigma+\mu-1)^2+K}}$$

found in [3] and [16] for the stochastic case when there are no time averages. When $\varepsilon = 0$ there is no randomness and we obtain

$$\limsup_{t\to\infty} \|U-u\|^2 \le 16e\mu K\Big(\frac{\delta_*-\delta}{\delta}\Big)^2.$$

When $\varepsilon > 0$ it would be nice to obtain similar bounds on $\|U-u\|$; however, no matter how small $\varepsilon$ is chosen, there is a positive probability for every $h > 0$, $M > 0$ and $T > 0$ that

$$|\xi(t)| \ge M \qquad \text{for} \qquad t \in (T, T+h). \tag{1.5}$$

As these probabilities are identical and independent when $(T_1-\delta, T_1+h)$ and $(T_2-\delta, T_2+h)$ are non-intersecting, it follows for almost every $\omega \in \Omega$ that there exists $T_n \to \infty$ as $n \to \infty$ such that $|\xi(t)| \ge M$ for $t \in (T_n, T_n+h)$. Consequently,

$$\limsup_{t\to\infty} \big[\|U-u\|^2\big] = \infty \qquad \text{for almost every} \qquad \omega \in \Omega.$$

At the same time, given any bound of the form

$$\limsup_{t\to\infty} \mathbf{E}\big[\|U-u\|^2\big] < B_1.$$

Chebyshev's inequality implies there exists $T > 0$ such that

$$P\big\{\, \|U-u\|^2 \le B_2 \,\big\} \ge 1 - B_1/B_2, \tag{1.6}$$

for every $B_2 > B_1$ and $t > T$. Thus, the error in $\|U - u\|^2$ has a 90 percent chance of being bounded by $B_2 = 10B_1$ at any fixed point in time.

The outline of this paper is as follows. In section 2 we present a proof of the existence and uniqueness of our solution. In addition, we present a generalized integro-differential Gronwall inequality that will be used throughout the paper. In section 3 we prove Theorem 1.2 and a number of corollaries while section 4 presents our numerical results. The paper ends with some concluding remarks in section 5.

## 2 Preliminaries

In this section we prove a local existence theorem which is then used to show that the nudging equations (1.4) are globally well posed. We end with an integro-Gronwall inequality that will later be used to bound the difference between the driven solution and the free-running solution.

Defining

$$F = \begin{bmatrix} -\sigma x + \sigma y \\ -\sigma x - y - xz \\ -b(z + r + \sigma) + xy \end{bmatrix}, \quad f = F + \begin{bmatrix} \mu\mathcal{M} \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad g = \begin{bmatrix} -\mu x \\ 0 \\ 0 \end{bmatrix}$$

allows (1.4) to be written as the general integro-differential equation

$$u'(t) = \begin{cases} F\big(t, u(t)\big) & \text{for } t < \delta_* \\ f\big(t, u(t)\big) + \dfrac{1}{\delta_*} \displaystyle\int_{t-\delta_*}^{t} g\big(s, u(s)\big)\, ds & \text{for } t > \delta_*, \end{cases} \tag{2.1}$$

where $u \in \mathbf{R}^n$ is a continuous function and

$$F, f \text{ and } g\colon [0, \infty) \times \mathbf{R}^n \to \mathbf{R}^n$$

are continuous in the first variable and uniformly Lipschitz in the second. This equation, similar to the Volterra equations studied by Levin in [17], is locally well posed as shown by

**Lemma 2.1.** *Suppose $u\colon [0, T] \to \mathbf{R}^n$ satisfies equations (2.1), then there exists $h > 0$ such that $u$ can be uniquely extended to a function on $[0, T + h]$ which also satisfies equations (2.1).*

*Proof.* If $T < \delta_*$, standard existence and uniqueness results for ordinary differential equations show there exists $h > 0$ such that the solution $u$ can

be extended to the interval $[0, T+h)$. If $T \geq \delta_*$, the result follows from a straight-forward fixed point argument included here for completeness.

Define

$$B = \max\big\{\, \|u(t)\| : t \in [0,T] \,\big\} + 1$$

and

$$M_f = \max\big\{\, \|f(t,v)\| : t \in [0,T+1] \text{ and } \|v\| \leq B \,\big\}.$$

Let $\gamma_f$ be the Lipschitz constant such that

$$\big\|f(t,v_1) - f(t,v_2)\big\| \leq \gamma_f \|v_1 - v_2\| \quad \text{for} \quad \|v_1\|, \|v_2\| \leq B.$$

Corresponding to the function $g$ similarly define $M_g$ and $\gamma_g$. Let

$$h = \min\left(1, \frac{1}{M_f + M_g}, \frac{1}{2(\gamma_f + \gamma_g)}\right).$$

and define

$$\mathcal{X} = \left\{ v \in \mathcal{C}([0,T+h], \mathbf{R}^n) : \begin{array}{l} v(t) = u(t) \text{ for } t \leq T, \text{ and} \\ \|v(t)\| \leq B \text{ for } t \in [T, T+h] \end{array} \right\}.$$

Consider the mapping $\mathcal{J}\colon \mathcal{X} \to \mathcal{C}([0,T+h], \mathbf{R}^n)$ given when $t \leq T$ by $\mathcal{J}(v)(t) = u(t)$ and when $t > T$ by

$$\mathcal{J}(v)(t) = u(T) + \int_T^t f\big(\tau, v(\tau)\big)\, d\tau + \int_T^t \frac{1}{\delta_*} \int_{\tau-\delta_*}^{\tau} g(s, v(s))\, ds\, d\tau.$$

We claim that $\mathcal{J}\colon \mathcal{X} \to \mathcal{X}$ and that $\mathcal{J}$ is a contraction. We estimate

$$\begin{aligned} \|\mathcal{J}(v)(t)\| &\leq \|u(T)\| + \int_T^t \|f\big(\tau, v(\tau)\big)\|\, d\tau \\ &\qquad + \int_T^t \frac{1}{\delta_*} \int_{\tau-\delta_*}^{\tau} \|g(s, v(s))\|\, ds\, d\tau \\ &\leq (B-1) + hM_f + hM_g \leq B. \end{aligned}$$

Therefore, $\mathcal{J}\colon \mathcal{X} \to \mathcal{X}$. Moreover,

$$\begin{aligned} \|\mathcal{J}(v_1)(t) - \mathcal{J}(v_2)(t)\| &\leq \int_T^t \|f\big(\tau, v_1(\tau)\big) - f(\tau, v_2(\tau)\|\, d\tau \\ &\qquad + \int_T^t \frac{1}{\delta_*} \int_{\tau-\delta_*}^{\tau} \|g(s, v_1(s)) - g(s, v_2(s))\|\, ds\, d\tau \\ &\leq h(\gamma_f + \gamma_g) \max\{\, \|v_1(s) - v_2(s)\| : s \in [T, T+h] \,\} \\ &\leq (1/2) \max\{\, \|v_1(s) - v_2(s)\| : s \in [T, T+h] \,\}, \end{aligned}$$

which shows that $\mathcal{J}$ is a contraction. Now, by the contraction mapping theorem, there is a unique fixed point $v \in \mathcal{X}$ such that $\mathcal{J}(v) = v$. Since this fixed point clearly satisfies (2.1), the lemma is proved. □

Now, we use the general local existence result in Lemma 2.1 and the specific properties of the Lorenz equations to show that the solution $u$ to (1.4) exists for all time. We prove

**Theorem 2.1.** *Given $\omega \in \Omega$ and $T > 0$ there exists $\gamma > 0$ and $R > 0$ such that $\|u(\omega; t)\|^2 \leq Re^{\gamma t}$ for $t \in [0, T]$.*

*Proof.* For each $\omega \in \Omega$ fixed, $\xi$ is continuous and consequently bounded on the interval $[\delta_*, T]$. Since $U$ lies on the global attractor, then $X$ is also bounded. It follows that there exists $M_\xi > 0$ such that

$$\frac{3}{\sigma}\Big\{\frac{\mu^2}{\delta}\int_{t-\delta}^{t} X^2 + \mu^2\xi^2\Big\} + (r+\sigma)^2 \leq M_\xi \quad \text{for} \quad t \in [\delta_*, T]. \tag{2.2}$$

Until $\delta_*$ the solution $u$ is governed by the standard Lorenz system and therefore exists up to that time. Applying Lemma 2.1 yields $h > 0$ such that $u$ is continuous on the interval $[0, \delta_* + h]$. We further assume, by taking $h$ smaller if necessary, that

$$h\left(\frac{6\mu^2}{\sigma} + M_\xi\right) \leq \frac{1}{2}.$$

Now set $V = x^2 + y^2 + z^2$ and choose $\gamma > 0$ and $R > 0$ so large that

$$e^{-(1+\gamma)h} \leq 1/2 \qquad \text{and} \qquad V(t) \leq Re^{\gamma t} \quad \text{for} \quad t \in [0, \delta_* + h].$$

We proceed by employing some specific properties of the Lorenz equations to show that $V(t) \leq Re^{\gamma t}$ for $t \in [0, T]$.

Assume $t > \delta_*$ and respectively multiply the equations in (1.4) by $x$, $y$ and $z$ to obtain

$$\begin{aligned}
\frac{1}{2}\frac{dx^2}{dt} &= -\sigma x^2 + \sigma xy + x\frac{\mu}{\delta}\int_{t-\delta}^{t} X - x\frac{\mu}{\delta_*}\int_{t-\delta_*}^{t} x + \mu x\xi \\
\frac{1}{2}\frac{dy^2}{dt} &= -y^2 - \sigma xy - xyz \\
\frac{1}{2}\frac{dz^2}{dt} &= -bz^2 + xyz - b(r+\sigma)z.
\end{aligned}$$

Young's and Cauchy's inequalities yield

$$\begin{aligned}\frac{1}{2}\dot{V} + V &\leq x\frac{\mu}{\delta}\int_{t-\delta}^{t} X - x\frac{\mu}{\delta_*}\int_{t-\delta_*}^{t} x + \mu x\xi - b(r+\sigma)z \\ &\leq \frac{\sigma}{2}x^2 + \frac{3}{2\sigma\delta}\Big\{\frac{\mu^2}{\delta}\int_{t-\delta}^{t} X^2 + \frac{\mu^2}{\delta_*}\int_{t-\delta_*}^{t} x^2 + \mu^2\xi^2\Big\} + \frac{b}{2}z^2 + \frac{1}{2}(r+\sigma)^2.\end{aligned}$$

Consequently,

$$\dot{V} + V \leq \frac{3\mu^2}{\sigma\delta_*}\int_{t-\delta_*}^{t} V + M_\xi, \tag{2.3}$$

where $M_\xi$ is the bound given in (2.2).

Let $t_n = \delta_* + nh$. By definition $V(t) \leq Re^{\gamma t}$ holds for $t \in [0, t_1]$. We proceed by induction. Suppose, $V(t) \leq Re^{\gamma t}$ holds for $t \in [0, t_n]$. We claim this same inequality holds on $[0, t_{n+1}]$.

Now, either $V(t) \leq 2Re^{\gamma t}$ for $t \in (t_n, t_{n+1}]$ or there exists $t_* \in (t_n, t_{n+1}]$ such that $V(t_*) = 2Re^{\gamma t_*}$ and $V(t) \leq 2Re^{\gamma t_*}$ for $t \in (t_n, t_*]$. We claim that second case cannot happen. Suppose for the sake of contradiction that such $t^*$ exists and take $t \in (t_n, t_*]$. Since $t - h > \delta_*$ then multiplying (2.3) by $e^t$ and integrating from $t-h$ to $t$ yields

$$V(t) \leq V(t-h)e^{-h} + \frac{3\mu^2}{\sigma\delta_*}\int_{t-h}^{t} e^{\rho-t}\int_{\rho-\delta_*}^{\rho} V(s)ds\,d\rho + hM_\xi.$$

By assumption $V(s) \leq 2Re^{\gamma s}$ for $s \leq t_*$. Therefore,

$$\int_{t-h}^{t} e^{\rho-t}\int_{\rho-\delta_*}^{\rho} V(s)ds\,d\rho \leq \int_{t-h}^{t} e^{\rho-t}\int_{\rho-\delta_*}^{\rho} 2Re^{\gamma s}ds\,d\rho \leq 2\delta_* hRe^{\gamma t}.$$

Since $t - h \leq t_n$ then $V(t-h) \leq Re^{\gamma(t-h)}$. It follows that

$$V(t) \leq Re^{\gamma t}\Big\{e^{-(1+\gamma)h} + h\Big(\frac{6\mu^2}{\sigma} + M_\xi\Big)\Big\} \leq Re^{\gamma t}. \tag{2.4}$$

Taking $t = t_*$ now gives a contradiction.

Consequently $V(t) \leq 2Re^{\gamma t}$ for all $t \in (t_n, t_{n+1}]$. In this case (2.4) again holds and further implies that $V(t) \leq Re^{\gamma t}$. This completes the induction and the proof. □

We now prove the following integro-differential Gronwall inequality.

**Lemma 2.2.** *Suppose $\mathcal{V}$ is continuous and for $t \geq h$ that*

$$\frac{d\mathcal{V}(t)}{dt} + \mathcal{V}(t) \leq A \int_{t-h}^{t} \mathcal{V}(s)\ ds + B, \tag{2.5}$$

*where $A > 0$, $B > 0$ and*

$$0 < h < \min\left\{\frac{1 - e^{-1}}{2A}, 1\right\},$$

*Then* $\limsup_{t\to\infty} \mathcal{V}(t) < eB$.

*Proof.* By hypothesis, there exists $\gamma \in (0, 1)$ such that

$$h < \frac{e^{-\gamma} - e^{-1}}{2A}.$$

Therefore,

$$e^{-(1-\gamma)} + 2Ahe^{\gamma h} < e^{-(1-\gamma)} + 2A\frac{e^{-\gamma} - e^{-1}}{2A}e^{\gamma} = 1 \tag{2.6}$$

and

$$2e^{-1} + 2Ah(1 - e^{-1}) - e^{-2} < e^{-1} + e^{-\gamma}(1 - e^{-1}) = \beta < 1. \tag{2.7}$$

Moreover,

$$\beta + (e^{-1} - e^{-2})(\beta^{-1} - 1) \leq 1. \tag{2.8}$$

Inequality (2.8) follows by taking $\phi(\beta) = \beta + (e^{-1} - e^{-2})(\beta^{-1} - 1)$ and differentiating to get $\phi'(\beta) = 1 - (e^{-1} - e^{-2})\beta^{-2}$. Therefore $\phi$ is increasing when $\beta > \sqrt{e^{-1} - e^{-2}}$. Since $\gamma \in (0, 1)$ implies $\beta > 2e^{-1} - e^{-2} > \sqrt{e^{-1} - e^{-2}}$ then $\phi(\beta) \nearrow 1$ as $\beta \nearrow 1$.

Let $m$ be the smallest integer such that $1 + h \leq m$ and $\mathcal{B} = e\beta B$. Since $\mathcal{V}$ is continuous there is a constant $R$ such that $\mathcal{V}(s) \leq Re^{-\gamma s} + \mathcal{B}$ for all $s \in [0, m]$. For induction, suppose $\mathcal{V}(s) \leq Re^{-\gamma s} + \mathcal{B}$ for all $s \in [0, n]$. We claim that $\mathcal{V}(s) < 2(Re^{-\gamma s} + \mathcal{B})$ for all $s \in [n, n + 1]$. If not, then there is some $t \in (n, n + 1]$ such that $\mathcal{V}(t) = 2(Re^{-\gamma t} + \mathcal{B})$ and $\mathcal{V}(s) \leq 2(Re^{-\gamma s} + \mathcal{B})$ for $s \in [n, t]$. Multiplying (2.5) by $e^t$ and integrating from $t - 1$ to $t$ we obtain

$$\mathcal{V}(t) \leq \mathcal{V}(t - 1)e^{-1} + e^{-t}A \int_{t-1}^{t} e^{\rho} \int_{\rho-h}^{\rho} \mathcal{V}(s)\, ds\, d\rho + (1 - e^{-1})B.$$

Since

$$\begin{aligned}\int_{t-1}^{t} e^{\rho} \int_{\rho-h}^{\rho} \mathcal{V}(s)\, ds\, d\rho &\le \int_{t-1}^{t} e^{\rho} \int_{\rho-h}^{\rho} 2(Re^{-\gamma s} + \mathcal{B})\, ds\, d\rho \\ &\le 2h \int_{t-1}^{t} e^{\rho}(Re^{-\gamma(\rho-h)} + \mathcal{B})\, d\rho \\ &\le 2Rhe^{\gamma h} \int_{t-1}^{t} e^{(1-\gamma)\rho}\, d\rho + 2\mathcal{B}h \int_{t-1}^{t} e^{\rho}\, d\rho \\ &\le 2Rhe^{\gamma h} e^{(1-\gamma)t} + 2\mathcal{B}h(e^{t} - e^{t-1}),\end{aligned}$$

then using (2.6) and (2.8) we obtain

$$\begin{aligned}\mathcal{V}(t) &\le (Re^{-\gamma(t-1)} + \mathcal{B})e^{-1} \\ &\qquad + A\big(2Rhe^{\gamma h}e^{-\gamma t} + 2\mathcal{B}h(1 - e^{-1})\big) + (1 - e^{-1})B \\ &= Re^{-\gamma t}\big(e^{-(1-\gamma)} + 2Ahe^{\gamma h}\big) \\ &\qquad + \mathcal{B}\big(2e^{-1} + 2Ah(1 - e^{-1}) - e^{-2}\big) \\ &\qquad + \mathcal{B}\big((e^{-1} - e^{-2})(\beta^{-1} - 1)\big) \\ &\le Re^{-\gamma t} + \mathcal{B}\big(\beta + (e^{-1} - e^{-2})(\beta^{-1} - 1)\big). \\ &\le Re^{-\gamma t} + \mathcal{B},\end{aligned} \tag{2.9}$$

which contradicts $\mathcal{V}(t) \ge 2(Re^{-\gamma t} + \mathcal{B})$. Therefore $\mathcal{V}(s) \le 2(Re^{-\gamma s} + \mathcal{B})$ for $s \in [n, n+1]$ and it immediately follows from (2.9) and induction that $\mathcal{V}(t) \le Re^{-\gamma t} + \mathcal{B}$ for $t > 0$. Consequently $\limsup_{t\to\infty} \mathcal{V}(t) \le \mathcal{B} < eB$. □

We remark that the proof of Lemma 2.2 is still valid when $B = 0$ except for the strict inequality in the last step. Consequently we also obtain

**Corollary 2.1.** *Suppose the hypothesis of Lemma 2.2 are satisfied except that $B = 0$. Then $\mathcal{V}(t) \to 0$ as $t \to \infty$.*

# 3 Analytical Results

In this section we prove our main analytical result stated in the introduction as Theorem 1.2 followed by a number of immediate corollaries. Recall that $\delta$ is the length of the unknown averaging window in the observations of the free running solution and $\delta_*$ is an approximation of $\delta$ that will be used in the feedback term of the driven system.

We first write differential equations governing the evolution of $\Delta X$, $\Delta Y$ and $\Delta Z$ where

$$\Delta X = X - x, \qquad \Delta Y = Y - y \qquad \text{and} \qquad \Delta Z = Z - z.$$

Upon setting

$$\Delta \bar{X} = \frac{1}{\delta_*} \int_{t-\delta_*}^{t} \Delta X, \quad \mathcal{E}_1 = \frac{1}{\delta} \int_{t-\delta_*}^{t-\delta} X \quad \text{and} \quad \mathcal{E}_2 = \Big(\frac{1}{\delta_*} - \frac{1}{\delta}\Big) \int_{t-\delta_*}^{t} X$$

we obtain

$$\begin{aligned}
\Delta \dot{X} &= -\sigma \Delta X + \sigma \Delta Y - \mu \Delta \bar{X} + \mu \mathcal{E}_1 + \mu \mathcal{E}_2 + \mu \xi \\
\Delta \dot{Y} &= -\sigma \Delta X - \Delta Y - \Delta X Z - x \Delta Z \\
\Delta \dot{Z} &= -b \Delta Z + \Delta X Y + x \Delta Y.
\end{aligned}$$

The terms $\big|\Delta X(t) - \Delta \bar{X}(t)\big|$ and $\big|\Delta X(t) - \Delta X(t-\delta)\big|$ appear in our analysis and will be bounded in terms of the integral of $|\Delta \dot{X}|$ using

**Lemma 3.1.** *The inequality*

$$\begin{aligned}
\Big( \int_{t-h}^{t} |\Delta \dot{X}| \Big)^2 \le 4(h + \delta_*)(\sigma + \mu)^2 \int_{t-(h+\delta_*)}^{t} V \\
+ 4h\mu^2 \int_{t-h}^{t} |\xi|^2 + 16 h^2 \mu^2 K \Big( \frac{\delta - \delta_*}{\delta} \Big)^2.
\end{aligned}$$

*holds with* $V = \Delta X^2 + \Delta Y^2 + \Delta Z^2$.

*Proof.* Estimate

$$|\Delta \dot{X}| \le \sigma |\Delta X| + \sigma |\Delta Y| + \mu |\Delta \bar{X}| + \mu |\mathcal{E}_1| + \mu |\mathcal{E}_2| + \mu |\xi|$$

Since

$$|\mathcal{E}_1| + |\mathcal{E}_2| \le \frac{1}{\delta} \Big| \int_{t-\delta_*}^{t-\delta} \sqrt{K} \Big| + \Big| \frac{1}{\delta_*} - \frac{1}{\delta} \Big| \int_{t-\delta_*}^{t} \sqrt{K} \le 2 \Big| \frac{\delta - \delta_*}{\delta} \Big| \sqrt{K} \tag{3.1}$$

and

$$\begin{aligned}
\int_{t-h}^{t} |\Delta \bar{X}(\rho)| d\rho &\le \int_{t-h}^{t} \frac{1}{\delta_*} \int_{\rho-\delta_*}^{\rho} |\Delta X(s)| ds \, d\rho \\
&= \frac{1}{\delta_*} \int_{t-\delta_*}^{t} \int_{s-h}^{s} |\Delta X(\rho)| d\rho \, ds \le \int_{t-(h+\delta_*)}^{t} |\Delta X|,
\end{aligned}$$

it follows that

$$\int_{t-h}^{t} \left|\Delta\dot{X}\right| \leq I_1 + I_2 + I_3 + I_4$$

where

$$I_1 = (\sigma+\mu)\int_{t-(h+\delta_*)}^{t} |\Delta X|, \qquad I_2 = \sigma\int_{t-h}^{t} |\Delta Y|,$$

$$I_3 = 2h\mu\Big|\frac{\delta_*-\delta}{\delta}\Big|\sqrt{K} \qquad \text{and} \qquad I_4 = \mu\int_{t-h}^{t} |\xi|.$$

By Cauchy's inequality,

$$I_1^2 \leq (h+\delta_*)(\sigma+\mu)^2\int_{t-(h+\delta_*)}^{t} |\Delta X|^2, \qquad I_2^2 \leq h\sigma^2\int_{t-h}^{t} |\Delta Y|^2$$

$$I_3^2 \leq 4h^2\mu^2K\Big(\frac{\delta-\delta_*}{\delta}\Big)^2 \qquad \text{and} \qquad I_4^2 \leq h\mu^2\int_{t-h}^{t} |\xi|^2.$$

Consequently,

$$\begin{aligned}\Big(\int_{t-h}^{t}|\Delta\dot{X}|\Big)^2 &\leq 4I_1^2 + 4I_2^2 + 4I_3^2 + 4I_4^2\\ &\leq 4(h+\delta_*)(\sigma+\mu)^2\int_{t-(h+\delta_*)}^{t} V\\ &\quad + 16h^2\mu^2K\Big(\frac{\delta-\delta_*}{\delta}\Big)^2 + 4h\mu^2\int_{t-h}^{t}|\xi|^2,\end{aligned}$$

as was to be shown. □

We are now ready for the proof our our main analytical result.

*Proof of Theorem 1.2.* Multiply the equations for $\Delta\dot{X}$, $\Delta\dot{Y}$ and $\Delta\dot{Z}$ respectively by $\Delta X$, $\Delta Y$ and $\Delta Z$ to obtain

$$\begin{aligned}\Delta X\Delta\dot{X} &= -(\sigma+\mu)\Delta X^2 + \sigma\Delta X\Delta Y + \mu\Delta X(\Delta X - \Delta\bar{X})\\ &\quad + \mu\Delta X\mathcal{E}_1 + \mu\Delta X\mathcal{E}_2 + \mu\Delta X\xi\\ \Delta Y\Delta\dot{Y} &= -\Delta Y^2 - \sigma\Delta X\Delta Y - \Delta X\Delta YZ - x\Delta Y\Delta Z\\ \Delta Z\Delta\dot{Z} &= -b\Delta Z^2 + \Delta XY\Delta Z + x\Delta Y\Delta Z\end{aligned}$$

Consequently,

$$\begin{aligned}\frac{1}{2}\dot{V} + (\sigma+\mu)\Delta X^2 + \Delta Y^2 + b\Delta Z^2 \\ = -\Delta X\Delta Y Z + \Delta X Y \Delta Z + \mu\Delta X(\mathcal{E}_1+\mathcal{E}_2) \\ + \mu\Delta X(\Delta X - \Delta\bar{X}) + \mu\Delta X\xi.\end{aligned}$$

By Young's inequality,

$$\begin{aligned}|\Delta X\Delta Y Z| + |\Delta X Y\Delta Z| &\le \frac{1}{2}\Delta Y^2 + \frac{1}{2}\Delta X^2 Z^2 + \frac{b}{2}\Delta Z^2 + \frac{1}{2b}\Delta X^2 Y^2 \\ &\le \frac{1}{2}\Delta Y^2 + \frac{b}{2}\Delta Z^2 + \frac{K}{2}\Delta X^2.\end{aligned}$$

Similarly, using (3.1) we have

$$\mu\big|\Delta X(\mathcal{E}_1+\mathcal{E}_2)\big| \le \frac{\mu}{4}\Delta X^2 + 4\mu K\Big(\frac{\delta-\delta_*}{\delta}\Big)^2.$$

On the other hand, observe that

$$\mu\xi(t)\Delta X(t) = \mu\xi(t)\big(\Delta X(t) - \Delta X(t-\delta)\big) + \mu\xi(t)\Delta X(t-\delta)$$

and further that

$$\begin{aligned}&\mu\big|\xi(t)(\Delta X(t) - \Delta X(t-\delta))\big| \\ &\qquad \le (\delta+\delta_*)\mu^2\xi(t)^2 + \frac{1}{4(\delta+\delta_*)}\big|\Delta X(t) - \Delta X(t-\delta)\big|^2.\end{aligned}$$

Then, upon taking $h=\delta$ in Lemma 3.1, it follows that

$$\begin{aligned}\frac{1}{4(\delta+\delta_*)}\big|\Delta X(t) - \Delta X(t-\delta)\big|^2 &\le \frac{1}{4(\delta+\delta_*)}\Big(\int_{t-\delta}^t |\Delta\dot{X}|\Big)^2 \\ &\le (\sigma+\mu)^2\int_{t-(\delta+\delta_*)}^t V + \mu^2\int_{t-\delta}^t \xi^2 + 4\delta\mu^2 K\Big(\frac{\delta-\delta_*}{\delta}\Big)^2.\end{aligned}$$

Lastly, observe that

$$\mu\big|\Delta X(\Delta X - \Delta\bar{X})\big| \le \frac{\mu}{4}\Delta X^2 + \mu(\Delta X - \Delta\bar{X})^2.$$

Since

$$\begin{aligned}|\Delta X - \Delta\bar{X}| &= \frac{1}{\delta_*}\Big|\int_{t-\delta_*}^t \big(\Delta X(t) - \Delta X(s)\big)\,ds\Big| \\ &\le \frac{1}{\delta_*}\int_{t-\delta_*}^t\int_s^t |\Delta\dot{X}(\rho)|\,d\rho\,ds \le \int_{t-\delta_*}^t |\Delta\dot{X}|,\end{aligned}$$

then taking $h = \delta_*$ in Lemma 3.1 yields

$$\begin{aligned}\mu|\Delta X - \Delta\bar{X}|^2 &\le \mu\Big(\int_{t-\delta_*}^{t} |\Delta\dot{X}|\Big)^2 \\ &\le 8\delta_*\mu(\sigma+\mu)^2\int_{t-2\delta_*}^{t} V + 4\delta_*\mu^3\int_{t-\delta_*}^{t}\xi^2 + 16{\delta_*}^2\mu^3 K\Big(\frac{\delta-\delta_*}{\delta}\Big)^2.\end{aligned}$$

We now combine the above estimates to obtain

$$\begin{aligned}&\frac{1}{2}\dot{V} + \Big(\sigma + \frac{\mu}{2} - \frac{K}{2}\Big)\Delta X^2 + \frac{1}{2}\Delta Y^2 + \frac{b}{2}\Delta Z^2 \\ &\quad\le (\sigma+\mu)^2\int_{t-(\delta+\delta_*)}^{t} V + 8\delta_*\mu(\sigma+\mu)^2\int_{t-2\delta_*}^{t} V \\ &\qquad + \mu^2\int_{t-\delta}^{t}\xi^2 + 4\delta_*\mu^3\int_{t-\delta_*}^{t}\xi^2 + (\delta+\delta_*)\mu^2\xi^2 \\ &\qquad + 4\mu K\big(1+\delta\mu+4\delta_*^2\mu^2\big)\Big(\frac{\delta-\delta_*}{\delta}\Big)^2 + \mu\xi(t)\Delta X(t-\delta).\end{aligned}$$

By hypothesis

$$\sigma + \frac{\mu}{2} - \frac{K}{2} \ge \frac{1}{2} \qquad \text{and} \qquad \max(\delta,\delta_*) \le \frac{1}{8\mu}.$$

Set $\mathcal{V} = \mathbf{E}[V]$ and recall that $\mathbf{E}[\xi^2] = \varepsilon^2/\delta$. Also, since Brownian motions have independent increments, we have that $x(t-\delta)$ is independent from $\xi(t) = \varepsilon\big(W(t) - W(t-\delta)\big)/\delta$. Consequently,

$$\mathbf{E}[\xi(t)\Delta X(t-\delta)] = \mathbf{E}[\xi(t)]\mathbf{E}[\Delta X(t-\delta)] = 0.$$

Therefore,

$$\dot{\mathcal{V}} + \mathcal{V} \le A\int_{t-h}^{t}\mathcal{V} + B \tag{3.2}$$

with $h = 2\max(\delta_*,\delta)$ where

$$A = 4(\sigma+\mu)^2 \qquad \text{and} \qquad B = 4\Big(\frac{\delta+\delta_*}{\delta}\Big)\mu^2\varepsilon^2 + 16\mu K\Big(\frac{\delta-\delta_*}{\delta}\Big)^2.$$

We now apply Lemma 2.2 to obtain for $\mu \ge K$ and

$$\max(\delta,\delta_*) < \frac{1-e^{-1}}{16(\sigma+\mu)^2}$$

that

$$\limsup_{t\to\infty}\mathbf{E}[V] \le 4e\Big(\frac{\delta+\delta_*}{\delta}\Big)\mu^2\varepsilon^2 + 16e\mu K\Big(\frac{\delta-\delta_*}{\delta}\Big)^2,$$

which is the desired result. □

Before turning to our numeric results, we discuss what Theorem 1.2 implies when the averaging window $\delta$ is known, when the observations are noise free and when both conditions hold. As there is nothing in the proof of Theorem 1.2 to prevent taking $\delta_* = \delta$ we immediately obtain

**Corollary 3.1.** *Under the assumptions of Theorem 1.2, if the length of the time averaging window is known, then $\delta_* = \delta$ and*

$$\limsup_{t\to\infty} \mathbf{E}\big[\|U-u\|^2\big] \le 8e\mu^2\varepsilon^2.$$

Similarly, at no point in the proof of Theorem 1.2 is it assumed that $\varepsilon > 0$. Therefore, the proof applies when $\varepsilon = 0$. In this case there is no randomness present and $\mathcal{V} = V$ in (3.2). Consequently, we obtain

**Corollary 3.2.** *Under the assumptions of Theorem 1.2, if there is no noise, then $\varepsilon = 0$ and*

$$\limsup_{t\to\infty} \|U-u\|^2 \le 16e\mu K\Big(\frac{\delta-\delta_*}{\delta}\Big)^2.$$

In the noiseless and errorless case we obtain, in light of Corollary 2.1, the exact convergence of the solution driven by blurred-in-time observational measurements to the free running solution. This we state as

**Corollary 3.3.** *Under the hypothesis of Theorem 1.2, if $\delta_* = \delta$ and $\varepsilon = 0$ then $\|U-u\| \to 0$ as $t \to \infty$.*

Note that exact convergence in the absence of errors and noise is a property important for any method used to solve a data assimilation problem.

# 4 Numerical Results

The behavior of the nudging algorithm (1.4) studied analytically in the previous sections can also be simulated numerically. In doing so we test not only the sharpness of our analysis, but also the practicality of the algorithm. As with our analysis, our main goal is to understand a feedback controller designed to handle observations which have been blurred in time by means of a moving time average. We are also interested in the effects of stochastic noise in the underlying measurement and consider the question whether time averaging might in practice actually improve accuracy by filtering the noise. While the answer to this last question is that time blurring does not help

filter the noise, on a more positive note, we find that it does not impact the efficacy of the synchronization either.

We compute both the free-running solution and the driven solution using Euler's explicit method using double-precision floating-point arithmetic with a step size of $\Delta t = 2^{-p}$ for some $p \in \mathbf{N}$. Let $t_n = n\Delta t$ and suppose $\delta = m\Delta t$ for some $m \in \mathbf{N}$. Approximate the free-running solution as $U_n \approx U(t_n)$ where

$$U_{n+1} = U_n + hF(t_n, U_n), \qquad U_0 = (-5.5751, -3.24345, -11.7028)$$

and $F$ is the function defined at the beginning of Section 2.

Blurred-in-time noisy observations of the current state of the free running solution $U_n = (X_n, Y_n, Z_n)$ are given by $\mathcal{M}_n \approx \mathcal{M}(t_n)$ where

$$\mathcal{M}_n = \frac{1}{m} \sum_{k=0}^{m-1} X_{n-k} + \xi_n. \tag{4.1}$$

The noise term is given by $\xi_n = \varepsilon(W_n - W_{n-m})/\delta$ where the Brownian motion $W_n \approx W(t_n)$ is computed using the Lévy–Ciesielski algorithm as described by Björn Böttcher in Schilling and Partzsch [22]. This algorithm was implemented using 1024-bit Marsaglia pseudo-random number generators [19] in such a way that for a fixed random number seed the same Brownian path could be sampled at any resolution on a dyadic grid. This results in a natural correspondence between $\omega \in \Omega$ and the random number seeds. Additional details are given in [3].

Although the sample paths of our Brownian motions are independent of $\Delta t$, the free-running solution $U$ strongly depends on $\Delta t$ as a result of the sensitive dependence of the Lorenz system on initial conditions and the resulting deterministic chaos. Therefore, we fix $\Delta t = 1/2048$ to ensure the same reference solution $U$ appears in all our simulations. Note that initial condition used for $U$ was obtained as the final state of a long-time initializing computation of the Lorenz equations. Therefore $U_0$ lies very close to the global attractor and we suppose that the resulting solution $U$ satisfies the conditions of Theorem 1.1 for at least $t \geq 0$. A sharper bound may be obtained *a posteriori* as the maximum value of the numerical solution. In particular, the specific solution considered here satisfies

$$\|U(t)\|^2 \leq 1247.0911 \qquad \text{for} \qquad t \in [0, 500], \tag{4.2}$$

which is about 20 percent smaller than the theoretical bound $K$.

For $\delta_* > 0$ arbitrary, define $m_* = \lceil \delta_*/\Delta t \rceil$ to be the least integer greater than or equal to $\delta_*/\Delta t$ and let

$$G_n \approx \frac{1}{\delta_*} \int_{t_n - \delta_*}^{t_n} g\big(s, u(s)\big) ds$$

be approximated by interpolation as $G_n = (-\mu S_n, 0, 0)$ where

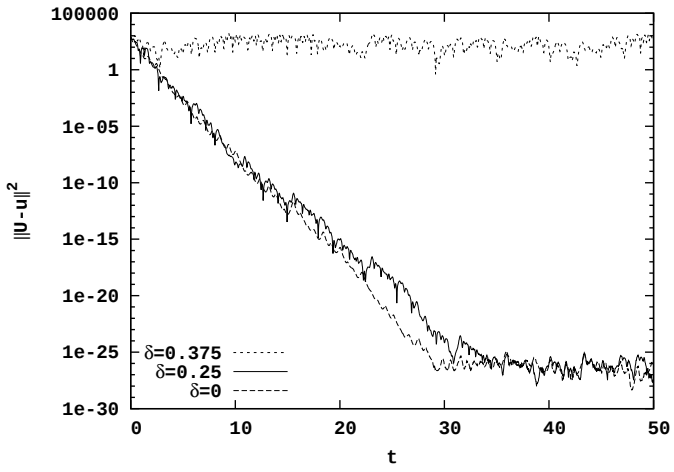
$$S_n = \frac{\Delta t}{\delta_*} \Big\{ \Big( \frac{\delta_*}{\Delta t} + 1 - m_* \Big) x_{n+1-m_*} + \sum_{k=0}^{m_*-2} x_{n-k} \Big\}. \tag{4.3}$$

Now, upon setting $n_0 = \max(m, m_*) - 1$, we compute $u_n \approx u(t_n)$ for different values of $\varepsilon$, $\delta$, $\delta_*$ and $\omega$ by

$$u_{n+1} = \begin{cases} u_n + F(t_n, u_n)\Delta t & \text{for } n < n_0 \\ u_n + f(t_n, u_n)\Delta t + G_n \Delta t & \text{for } n \geq n_0 \end{cases} \tag{4.4}$$

with initial condition $u_0 = (0, 0, 0)$.

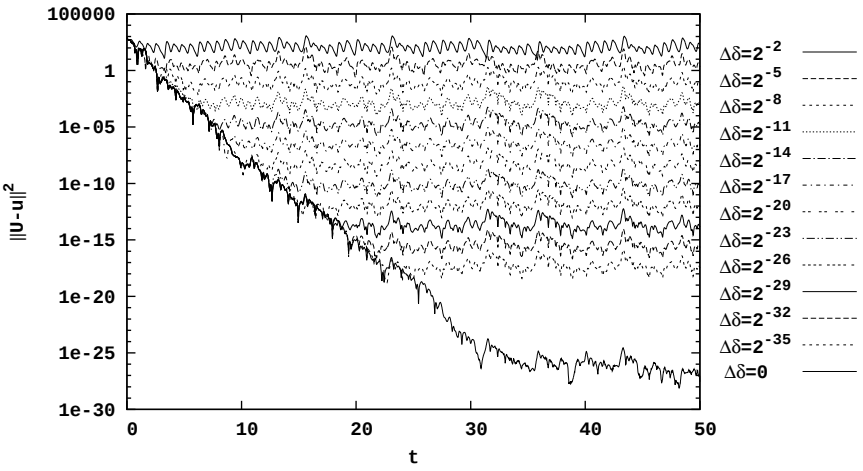
Figure 1: Results when coupling noiseless time-averaged measurements with window size $\delta$ using a feedback term with averaging window size $\delta_* = \delta$. Note that $\mu = 30$.




We remark the definition of $n_0$ in (4.4) effectively replaces the conditions $t < \delta_*$ and $t > \delta_*$ in (1.4) by $t < t_0$ and $t > t_0$ for some $t_0 \geq \max(\delta_*, \delta)$. Mathematically this is equivalent to changing the initial condition $u_0$ in the original nudging equations from $(0, 0, 0)$ to something slightly different—the

forward evolution of $(0,0,0)$ by the dynamics of the Lorenz system over a time interval of length $t_0 - \delta_*$. As the hypotheses of Theorem 1.2 do not depend on $u_0$, which was assumed to be arbitrary to begin with, we proceed with this numerically convenient modification in the computations presented here. Note also that the case in which there is no blur in time is numerically equivalent to taking $\delta = \Delta t$ in (4.1) and (4.3). In this case we shall make a slight abuse of notation for clarity and instead write $\delta = 0$ to more intuitively indicate a time averaging window of length zero.

Figure 2: Results when coupling noiseless time-averaged measurements with window size $\delta = 0.25$ using a feedback term with averaging window size $\delta_* \leq \delta$. Note that $\mu = 30$ and $\Delta\delta = \delta - \delta_*$.




Our first numerical experiments are performed for the case of noiseless blurred-in-time observations in which the size of the averaging window is known. Thus, we take $\delta_* = \delta$ and $\varepsilon = 0$. Figure 1 depicts the time evolution of $\|U - u\|^2$ when $\delta \in \{0, 0.25, 0.375\}$ and $\mu = 30$. While this choice for $\mu$ doesn't satisfy the condition $\mu \geq K$ required by Theorem 1.2, it is comparable to the value $\mu = K/40 \approx 38.5$ that was numerically effective in [16]. For $\delta = 0$ and $\delta = 0.25$ the error is asymptotically bounded by $10^{-25}$, which to within the limits of the floating point arithmetic used for the computation, is numerically equivalent to zero. Note also the rate of convergence is nearly the same when $\delta = 0.25$ as when $\delta = 0$. On the other hand, for $\delta = 0.375$ the difference between $U$ and $u$ remains large. Moreover, further computations not depicted here indicate there are no values of $\mu$ for which $\|U - u\|$ numerically converges to zero when $\delta = 0.375$. Although the limits on the

size of $\delta$ which appear in Theorem 1.2 are more restrictive than what works in practice, our numerics confirm limits on the maximum size of $\delta$ exist.

Before adding noise, we first determine what happens when the length of the averaging window responsible for blurring the observations in time is unknown and $\delta_* \neq \delta$. Figure 2 depicts the evolution of $\|U - u\|^2$ when $\delta = 0.25$ and $\delta_*$ is smaller but near $\delta$. Similar calculations performed for $\delta_*$ greater than $\delta$ result in a nearly identical graph. Bounds on the asymptotic error may be approximated numerically as a function of $\Delta\delta = \delta - \delta_*$ by

$$E(\Delta\delta) = \max\big\{\, \|U - u\|^2 : t \in [50, 100] \,\big\}.$$

A summary of our computations is presented in Table 1. As expected, the error levels decrease as $\delta_*$ gets closer to $\delta$. We now consider the functional dependency how $E(\Delta\delta)$ depends on $\Delta\delta$.

Table 1: Numeric values of $E(\Delta\delta) = \max\big\{\, \|U - u\|^2 : t \in [50, 100] \,\big\}$ as a function of $\Delta\delta = \delta - \delta_*$. Here $\delta$ represents the unknown size of the averaging present in the observational measurements and $\delta_*$ an approximation of $\delta$ used in the feedback control.

| $\Delta\delta$ | $E(\Delta\delta)$ |
|---|---|
| $2^{-2}$ | 1267.31 |
| $2^{-5}$ | 194.081 |
| $2^{-8}$ | 5.61624 |
| $2^{-11}$ | 0.0906275 |
| $2^{-14}$ | 0.00140762 |
| $2^{-17}$ | $2.19772 \times 10^{-5}$ |
| $2^{-20}$ | $3.43361 \times 10^{-7}$ |
| $2^{-23}$ | $5.36495 \times 10^{-9}$ |
| $2^{-26}$ | $8.38272 \times 10^{-11}$ |
| $2^{-29}$ | $1.30979 \times 10^{-12}$ |
| $2^{-32}$ | $2.04656 \times 10^{-14}$ |
| $2^{-35}$ | $3.19773 \times 10^{-16}$ |
| 0 | $2.67042 \times 10^{-25}$ |

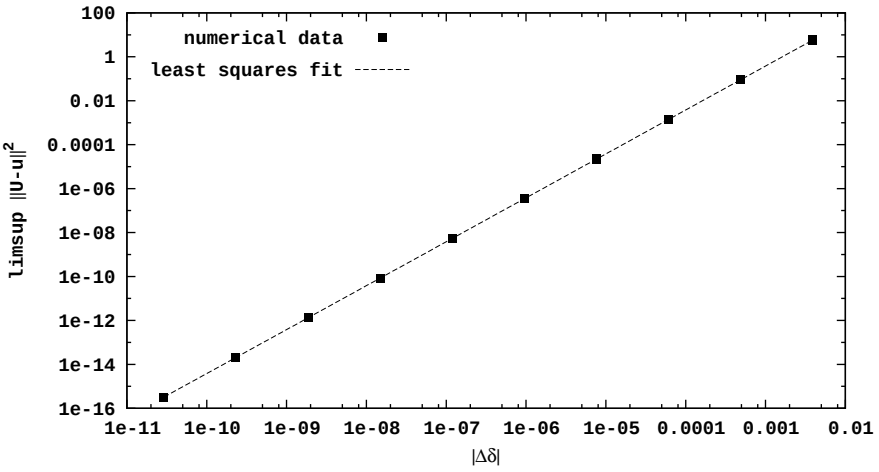
| $\Delta\delta$ | $E(\Delta\delta)$ |
|---|---|
| $-2^{-2}$ | 2663.41 |
| $-2^{-5}$ | 249.552 |
| $-2^{-8}$ | 5.86352 |
| $-2^{-11}$ | 0.0911086 |
| $-2^{-14}$ | 0.00143265 |
| $-2^{-17}$ | $2.24025 \times 10^{-5}$ |
| $-2^{-20}$ | $3.50073 \times 10^{-7}$ |
| $-2^{-23}$ | $5.46996 \times 10^{-9}$ |
| $-2^{-26}$ | $8.54684 \times 10^{-11}$ |
| $-2^{-29}$ | $1.33543 \times 10^{-12}$ |
| $-2^{-32}$ | $2.08664 \times 10^{-14}$ |
| $-2^{-35}$ | $3.26052 \times 10^{-16}$ |
| 0 | $2.67042 \times 10^{-25}$ |

The theoretical bounds given by Corollary 3.2 suggests that $E(\Delta\delta)$ is proportional to $(\Delta\delta)^2$. In order to test this hypothesis we performed a least squares fit to find $m$ and $s$ such that

$$\log E(\Delta\delta) \approx \log m + s \log(\Delta\delta).$$

Figure 3 shows that the least squares fit agrees with $E(\Delta\delta)$ to remarkable precision over a range of ten decimal orders of magnitude. Since the fitted value $s = 1.99964$ is so close to 2, we presume that the $\Delta\delta$ dependency given by our analytic bound is physically relevant even though the hypotheses on $\mu$ and $\delta$ are much more restrictive than numerically needed.

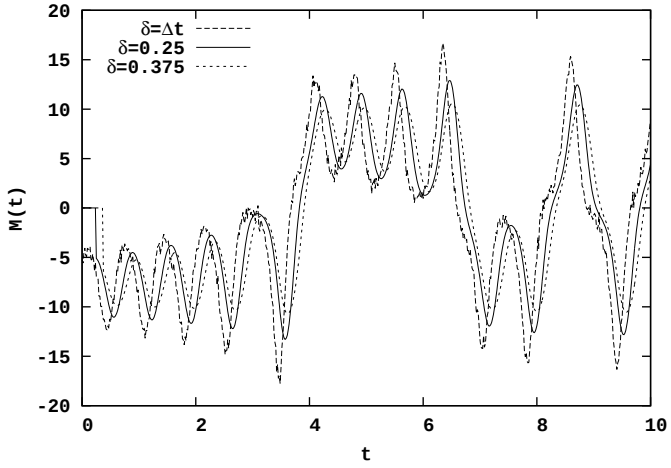
Figure 3: The numerical data from Table 1 when plotted with the least squares fit of $m(\Delta\delta)^s$ given by $m = 378647$ and $s = 1.99964$.




We now turn to the numerical simulation of noisy measurements that have been blurred in time. Computations in [16] without time blur demonstrate numerical bounds on $\mathbf{E}\big[\|U - u\|^2\big]$ which are proportional to $\varepsilon^2$ and therefore consistent with the theoretical bounds derived in that work. Similar results may be obtained in the presence of time blur that are consistent with the bounds in Corollary 3.1 proved here. Instead of detailing those results, we instead fix the noise level at $\varepsilon = 0.01$ and focus on the interaction between the time averaging and the noisy measurements.

Although the standard Euler-Maruyama time-stepping scheme for the stochastic case $\delta = 0$ may be obtained by taking $\delta = \Delta t$ in our numerical methods, there is a significant theoretical difference between these two values of $\delta$ with respect to the signal-to-noise ratio. When $\delta = 0$ the amplitude of the noise is infinite and the signal-to-noise ratio zero since Brownian motions are not differentiable. When $\delta = \Delta t$ the amplitude of the noise is finite and the signal-to-noise ratio is strictly positive; however, both are non-physical as they depend on the size of the numerical time-steps. In either of these cases it makes sense to characterize the noise level present

Figure 4: Time-averaged noisy measurements for different values of $\delta$ corresponding to a representative Brownian path with $\varepsilon = 0.01$. Note that the time step $\Delta t \approx 0.0005$.




in the measurements by comparing $\varepsilon$ to the root-mean-squared value of $X$. For the particular free running solution used in our experiments

$$\Big(\frac{1}{100}\int_0^{100} X(s)^2 ds\Big)^{1/2} \approx 7.927.$$

Thus, $\varepsilon = 0.01$ is 0.13 percent the root-mean-squared value of $X$.

Figure 4 depicts the time evolution of a typical realization of the noisy observational measurements subject to different amounts of time blur. Note that when $\delta \gg 0$ the noise has been smoothed out at the expense of a phase shift that delays the measurements in time. This trade off is a direct result of matching the averaging window in the feedback term with the averaging in the measurements. As discussed in (1.3) this phase shift could be eliminated; however, we instead interpret the time-averaged measurements as an approximation of the state of the system at the end of the averaging window in order to ensure that the driven equations (1.4) exactly reproduce the free-running solution given identical initial conditions when $\delta_* = \delta$ in the time-averaged but noise-free case.

When $\delta > 0$ we can quantify the tradeoff between the noise and the phase shift by computing the blurred signal-to-noise ratio

$$\mathrm{SNR}(\delta) = \int_\delta^{100} X(s)^2 ds \Big/ \mathbf{E}\Big[\int_\delta^{100} \big(\mathcal{M}(s) - X(s)\big)^2 ds\Big].$$

Table 2 lists the blurred signal-to-noise ratios corresponding to the values of $\delta$ depicted in Figure 4. Note that $\mathrm{SNR}(\delta) \to 0$ as $\delta \to 0$ and decreases to near unity for large $\delta$. It would be interesting to consider values of $\delta$ which maximize the blurred signal-to-noise ratio as a possible noise filter. For example, $\mathrm{SNR}(0.004) \approx 2017.4$. In this paper, however, we focus on values of $\delta$ which are larger and therefore present a closer analogy to the time averages present in real-world scientific instrumentation.

Table 2: Numerical values of the blurred signal-to-noise ratio when $\varepsilon = 0.01$ as a function of $\delta$. Note that $\Delta t \approx 0.0005$.

| $\delta$ | 0 | $\Delta t$ | 0.25 | 0.375 |
|---|---|---|---|---|
| SNR | 0 | 306.9 | 3.118 | 1.948 |

We shall use 500 independent realizations of the noisy measurements to compute the expected value

$$\mathbf{E}\big[\|U-u\|^2\big] \approx \frac{1}{500}\sum_{i=1}^{500} \|U(t)-u(\omega_i;t)\|^2$$

where $\omega_i \in \Omega$ and $u_n \approx u(\omega_i, t_n)$ is computed according to (4.4) using 500 different seeds for the pseudo-random number generator. The resulting evolution of $\mathbf{E}\big[\|U-u\|^2\big]$ is depicted in Figure 5. Note that the expected value varies over time in accordance with the deterministic dynamics of the free running solution $U$. Moreover, since computing with additional trajectories does not significantly change our numerical results, we suppose that that 500 realizations of Brownian motion are sufficient.

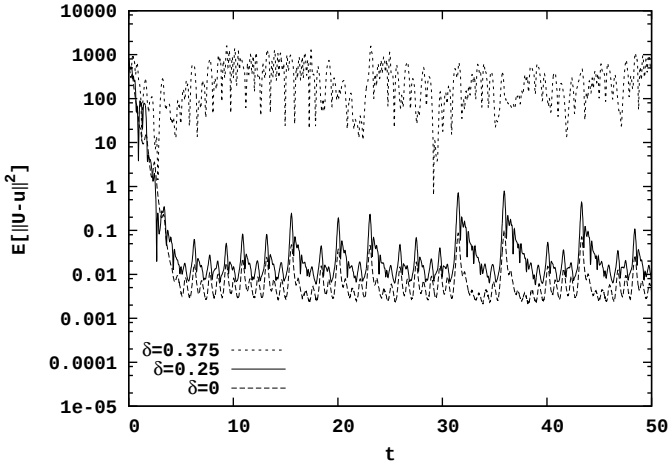
As in the noiseless case, the difference between $U$ and $u$ remains large when $\delta = 0.375$. Similarly, for $\delta = 0.25$ and $\delta = 0$ the error decreases over time; however, unlike the noiseless case where the error in both solutions asymptotically decreased to numerical zero, the error level for $\delta = 0.25$ in the presence of noise is noticeably larger than for $\delta = 0$. In particular,

$$\max\big\{\, \mathbf{E}\big[\|U-u\|^2\big] : t \in [50,100] \,\big\} \approx \begin{cases} 0.0089 & \text{for } \delta = 0 \\ 0.49 & \text{for } \delta = 0.25. \end{cases}$$

We conjecture that the time shift observed in Figure 4 when $\delta = 0.25$ results in a decreased ability of the feedback control to overcome noise in the measurements. Without noise, a time blur of $\delta = 0.25$ in the observations can be compensated for by taking $\delta_* = 0.25$ in the feedback term so that the difference between the free running and driven solution tends to zero over

time. In the presence of noise, however, the nudging must also overcome the noise. The time shift which results from the averaging leads leads less efficient nudging that in turn leads to a higher error level in the difference between $U$ and $u$ over time. When $\delta_* = \delta$ and $\delta$ is small, the size of the averaging window makes only a slight difference to how accurately the free running solution can be recovered in the presence of noise.

Figure 5: Results when coupling time-averaged noisy measurements with window size $\delta$ and noise level $\varepsilon = 0.01$ using a feedback term with averaging window size $\delta_* = \delta$. The expected values were computed using 500 independent realizations of Brownian motion.




Before considering the final case when the length of the averaging window responsible for the blurred noisy measurements is unknown and $\delta_* \neq \delta$, we explore further the statistics of $\|U - u\|^2$ when $\delta_* = \delta$ and $\delta = 0.25$. Let $L(t)$ be the smallest value such that

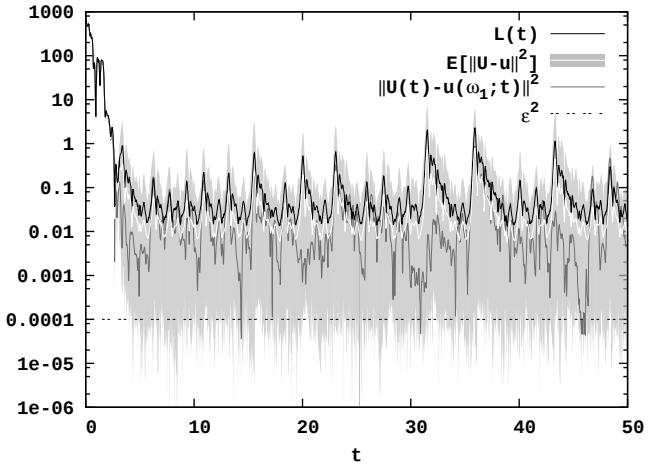
$$\text{card}\big\{\, i \in \{1, 2, \ldots, 500\} : \|U(t) - u(\omega_i; t)\|^2 \leq L(t) \,\big\} = 450.$$

Thus, $L(t)$ bounds 90 percent of the paths at each point in time. Assuming 500 samples is statistically large enough, we have

$$\mathbf{P}\big\{\, \omega \in \Omega : \|U(t) - u(\omega, t)\|^2 \leq L(t) \,\big\} \approx 0.9.$$

Figure 6 compares $L(t)$ with the expected value of $\|U - u\|^2$. Note that $\mathbf{E}\big[\|U-u\|^2\big]$ oscillates around 0.01 and that $L(t)$ is about five times greater. Thus, $L(t)$ is about two times smaller than the factor-of-ten bound guaranteed in (1.6) by Chebyshev's inequality. As indicated by the grey region,

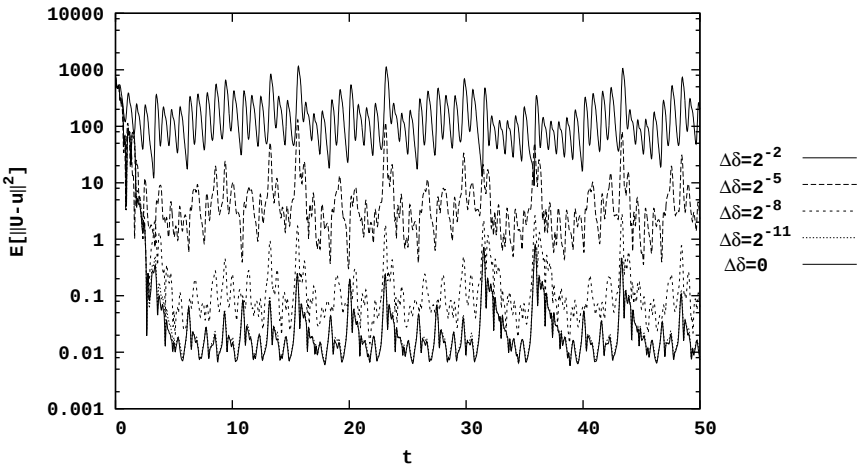
Figure 6: Results when coupling time-averaged noisy measurements with window size $\delta = 0.25$ and noise level $\varepsilon = 0.01$ using a feedback term with averaging window size $\delta_* = \delta$. The expected values were computed using 500 independent realizations of Brownian motion and $L(t)$ is the numerical bound on 90 percent of the paths at each point in time. The gray region represents the 500 pathwise solutions, one of which has been plotted in a darker shade for illustration.




there are some pathwise trajectories at each point in time which are significantly more accurate than expected and a few which satisfy $\|U - u\|^2 \leq \varepsilon^2$. This reflects the fact that over any finite period of time there is a positive probability that the random noise in the measurements is very small. That few trajectories appear much larger than the expected value is an illusion due to the logarithmic scale used for the vertical axis in the graph. When plotted on a linear scale the shaded region above the expected value appears much larger. In fact, as indicated by (1.5) in the introduction, over any finite period of time there is also a positive probability that the random noise will be arbitrarily large. Thus, as more trajectories resulting from additional realizations of the measurement error are computed, the height of the shaded region will continue to grow while the values of $L(t)$ and $\mathbf{E}\big[\|U-u\|^2\big]$ converge to become more accurate.

The last set of computations performs simulations when there is noise and when the length of the averaging window responsible for blurring the observations in time is unknown. This is the setting for Theorem 1.2, our main analytical result. Figure 7 depicts the evolution of the expected value

Figure 7: Results when coupling time-averaged noisy measurements with $\varepsilon = 0.01$ and window size $\delta = 0.25$ using a feedback term with averaging window size $\delta_* \leq \delta$. Note that $\mu = 30$ and $\Delta\delta = \delta - \delta_*$.




of $\|U - u\|^2$ when $\delta = 0.25$ and $\delta_*$ is smaller but near $\delta$. As in the noiseless case, similar calculations performed for $\delta_*$ greater than $\delta$ result in a nearly identical graph. The error level decreases as $\delta_*$ gets closer to $\delta$, but unlike the noiseless case, there is very little difference between $\Delta\delta = 0$ and $\Delta\delta = 2^{-11}$. Bounds on the expected error may be approximated numerically as

$$\widetilde{E}(\Delta\delta) = \max\big\{\, \mathbf{E}\big[\|U - u\|^2\big] : t \in [50, 100] \,\big\}.$$

A summary of our computations are presented in Table 3. These numeric results show, as do our theoretical bounds, that $\delta_*$ can be tuned to match $\delta$ only up to the point where the noise term dominates. Specifically, when $\delta = 0.25$ and $\epsilon = 0.01$ it is possible to tune $\delta_*$ so that $|\delta - \delta_*| < 2^{-8}$.

# 5 Conclusions

In the context of the Lorenz equations we have shown the data-assimilation method introduced in [1] is well-posed when the observational data is contaminated by a moving average. In the noiseless case, analysis and numeric simulation shows that synchronization of the driven solution to the free-running solution occurs when the time averaging window is known and small enough. Moreover, when the time averaging window is unknown, the bound

Table 3: Values of $\widetilde{E}(\Delta\delta) = \max\left\{ \mathbf{E}\left[\|U-u\|^2\right] : t \in [50,100] \right\}$ as a function of $\Delta\delta = \delta - \delta_*$ when $\varepsilon = 0.01$. Here $\delta$ represents the unknown size of the averaging present in the observational measurements and $\delta_*$ an approximation of $\delta$ used in the feedback control.

| $\Delta\delta$ | $\widetilde{E}(\Delta\delta)$ |
|---|---|
| $2^{-2}$ | $1.60729 \times 10^6$ |
| $2^{-5}$ | 38419.5 |
| $2^{-8}$ | 44.133 |
| $2^{-11}$ | 0.693591 |
| 0 | 0.491966 |

| $\Delta\delta$ | $\widetilde{E}(\Delta\delta)$ |
|---|---|
| $-2^{-2}$ | $6.09399 \times 10^6$ |
| $-2^{-5}$ | 62104.1 |
| $-2^{-8}$ | 38.5955 |
| $-2^{-11}$ | 0.56483 |
| 0 | 0.491966 |

on the error is proportional to $(\Delta\delta)^2$ where $\Delta\delta$ is the difference between the size of the actual averaging window and the guess for the size of the window used in the feedback control.

These results suggest that it is possible to determine the size $\delta$ of the unknown averaging window by varying $\delta_*$ in a way that minimizes the resulting error. In the context of data assimilation, the only information actually available concerning $U$ are the time averaged measurements of $X$. Therefore, it is not possible to directly compute $\|U-u\|$ over time to determine the optimal choice of $\delta_*$. However, it would be physically possible to compute

$$E_*(\delta_*) = \Big(\mathcal{M}(t) - \frac{1}{\delta_*}\int_{t-\delta_*}^{t} x(s)ds\Big)^2$$

and tune $\delta_*$ so that $E_*$ is on average as small as possible. To what extent this procedure can be used to find a good value for $\delta_*$ is an interesting question.

In the presence of noise, time averaging smooths out the noise in the observational measurements at the expense of a phase shift that delays the measurements in time. Although the averaging considered here gives no improvement in the analytic and numeric bounds on $\|U-u\|$, there is also no significant deterioration in those bounds provided the averaging window is not too big. It remains an open question for future work to check whether smaller values of $\delta$ which maximize the blurred signal-to-noise ratio could function as a noise filter and lead to improved bounds. In the present case, the time averaging appears to slightly weaken the coupling between the driven and the free-running solution which, in turn, yields a slight increase in the expected error levels.

In this paper we have focused on coupling two identical copies of the Lorenz equations. This coupling can be interpreted as a type of data assimilation when the exact dynamics governing the observed system are known. In applications, the dynamics of the free-running solution may not be known exactly and the driven solution evolved using some sort of approximate dynamics. In the case of the Lorenz equations, we might imagine the exact values of $\sigma$, $r$ and $b$ are only known approximately. A more general technique, amenable to further statistical analysis, is to represent uncertainties in the exact dynamics by additional noise terms. For example, white-in-time noise with a known covariance was imposed on the dynamics of $X$, $Y$ and $Z$ in [16] to represent model error. Data assimilation using a similar framework for model error was analyzed by Hoang, Law and Stuart [14] for the two-dimensional Navier–Stokes equations. This framework for model error is further supported by the observation in [20] that deterministic approximations of chaotic non-linear systems—such as using the LANS-alpha turbulence model to approximate the Navier–Stokes equations—can lead to model errors which are correlated over small time scales but appear uncorrelated over large enough time scales. Although measurement error is the only way that noise enters in the present study, similar noise terms could be used to represent model error with similar results. Moreover, given the short-time correlations observed in computational simulations, such model errors may actually be better represented as short-time averages.

It is our hope that the techniques used in our analysis can be adapted to apply to other model problems such as the two-dimensional Navier–Stokes equations and the surface quasi-geostrophic equations. The main difficulty in those cases is that the resulting pathwise integro-differential inequality is non-linear. This makes estimating the expected value of $\|U - u\|^2$ difficult because the same techniques used in the proof of Theorem 1.2 would immediately lead to a non-closed system. Work on a modified coupling algorithm that filters statistical outliers from the observational measurements to treat this case is currently in progress.

## Acknowledgements

The work of Eric Olson and Jordan Blocher were supported in part by NSF grant DMS-1418928. We also thank the anonymous referees for their contributions to this paper.

## References


[1] A. Azouani, E. Olson, and E. Titi. Continuous data assimilation using general interpolant observables. *J. Nonlinear Sci.*, Vol. 24, 277–304, 2014.

[2] H. Bessaih, E. Olson and E. Titi. Continuous Data Assimilation with Stochastically Noisy Data, *Nonlinearity*, Vol. 28, 729–753, 2015.

[3] J. Blocher, *Synchronization of the Lorenz System using Time-averaged Measurements*, Master's Thesis, Department of Mathematics and Statistics, University of Nevada, Reno, 2016.

[4] Emine Celik, Eric Olson and Edriss Titi, Spectral Filtering of Interpolant Observables for Discrete-in-time Data Assimilation, preprint.

[5] C. Doering and J. Gibbon. *Applied Analysis of the Navier–Stokes Equations*. Cambridge University Press, 1995.

[6] C. Doering and J. Gibbon. On the shape and dimension of the Lorenz attactor. *Dynam. Stability Systems*, Vol. 10, 255–268, 1995.

[7] A. Farhat, E. Lunasin, E.S. Titi, Abridged continuous data assimilation for the 2D Navier–Stokes equations utilizing measurements of only one component of the velocity field, *J. Math. Fluid Mech.*, Vol. 18, No. 1, 2016, pp. 1-23.

[8] A. Farhat, E. Lunasin, E.S. Titi, Continuous data assimilation for a 2D Bénard convection system through horizontal velocity measurements alone, *J. Nonlinear Sci.*, No. 3, 2017, pp. 1065–1087.

[9] C. Foias, C. Mondaini, and E.S. Titi, A discrete data assimilation scheme for the solutions of the 2D Navier-Stokes equations and their statistics, *SIAM J. Appl. Dyn. Syst.*, Vol. 15, 2016, pp. 2109-2142.

[10] M. Gesho, *A Numerical Study of a Continuous Data Assimilation for the Two-dimensional Navier–Stokes Equations Using Nodal Points*, Masters Thesis, University of Nevada, Department of Mathematics and Statistics, 2013.

[11] M. Gesho, E. Olson, E.S. Titi, A Computational Study of a Data Assimilation Algorithm for the Two-dimensional Navier–Stokes Equations, *Comm. in Comp. Physics*, Vol. 19, No. 04, pp. 1094–1110, 2016.

[12] K. Hayden, *Synchronization in the Lorenz System*, Masters Thesis, University of Nevada, Department of Mathematics and Statistics, 2007.

[13] K. Hayden, E. Olson, and E.S. Titi. Discrete data assimilation in the Lorenz and 2d Navier-Stokes equations. *Physica D*, pages 1416–1425, 2011.

[14] V.H. Hoang, K.J.H. Law, A.M. Stuart, Determining white noise forcing from Eulerian observations in the Navier–Stokes equations, *Stoch. Partial Differ. Equ. Anal. Comput.*, Vol. 2, No. 2, pp. 233–261, 2014.

[15] M.S. Jolly, V.R. Martinez, E.J. Olson and E.S. Titi, Continuous Data Assimilation with Blurred-in-time Measurements of the Surface Quasi-geostrophic Equation, preprint.

[16] K. Law, A. Shukla, and A. Stuart. Analysis of the 3DVAR filter for the partially observed Lorenz '63 model. *Disc. Contin. Dyn. Sys.*, 34:1061–1078, 2014.

[17] J. Levin and J. Nohel. On a nonlinear delay equation. *J. Math. Anal. Appl.*, 8:31–44, 1964.

[18] E. Lorenz. Deterministic nonperiodic flow. *J. Atmospheric Sci.*, 20:130–141, 1963.

[19] G. Marsaglia, Xorshift RNGs, *J. of Stat. Software*, Vol. 8, 2003.

[20] E. Olson, Stochastic model error in the LANS-alpha and NS-alpha deconvolution models of turbulence, to appear in *Int. J. Numer. Anal. Model.*

[21] L. Pecora and T. Carroll. Synchronization in chaotic systems. *Phys. Rev. Lett.*, 64(8):821–824, 1990.

[22] R. Schilling and L. Partzsch, *Brownian Motion: An Introduction to Stochastic Processes* , De Gruyter, 2012.

[23] W. Tucker. The Lorenz attractor exists. *C. R. Acad. Sci. Paris*, pages 1197–1202, 1999.