Evolution Equations and Control Theory
Vol. 13, No. 2, April 2024, pp. 329-348
doi:10.3934/eect.2023048

NONLINEAR CONTINUOUS DATA ASSIMILATION

Adam Larios[✉*1] and Yuan Pei[✉1,2]

[1]University of Nebraska-Lincoln, USA

[2]Western Washington University, USA

(Communicated by Justin T. Webster)

Abstract. We introduce three new nonlinear continuous data assimilation algorithms. These models are compared with the linear continuous data assimilation algorithm introduced by Azouani, Olson, and Titi (AOT). As a proof-of-concept for these models, we computationally investigate these algorithms in the context of the 1D Kuramoto-Sivashinsky equations. We observe that the nonlinear models experience super-exponential convergence in time, and converge to machine precision significantly faster than the linear AOT algorithm in our tests. For both simplicity and completeness, we provide the key analysis of the exponential-in-time convergence in the linear case.

1. **Introduction.** Recently, a promising new approach to data assimilation was pioneered by Azouani, Olson, and Titi [5, 6] (see also [17, 62, 94] for early ideas in this direction). This approach, which we call AOT data assimilation or the linear AOT algorithm, is based on feedback control at the partial differential equation (PDE) level, described below. In the present work, we propose several nonlinear data assimilation algorithms based on the AOT algorithm, that exhibit significantly faster convergence in our simulations; indeed, the convergence rate appears to be super-exponential.

Let us describe the general idea of the AOT algorithm. Consider a dynamical system in the form,

$$\begin{aligned} \frac{du}{dt} &= F(u), \\ u(0) &= u_0. \end{aligned} \tag{1}$$

For example, this could represent a system of partial differential equations modeling fluid flow in the atmosphere or the ocean. A central difficulty is that, even if one were able to solve the system exactly, the initial data u_0 is largely unknown. For example, in a weather or climate simulation, the initial data may be measured at certain locations by weather stations, but the data at locations in between these stations may be unknown. Therefore, one might not have access to the complete initial data u_0, but only to the observational measurements, which we denote by $I_h(u_0)$. (Here, I_h is assumed to be a linear operator that can be taken, for example, to be an interpolation operator between grid points of maximal spacing h, or as an

2020 *Mathematics Subject Classification.* Primary: 34D06, 35K41, 35K61, 35Q93, 93C20.

Key words and phrases. Nonlinear data assimilation, Azouani-Olson-Titi, feedback control, Kuramoto-Sivashinsky equation, nudging.

*Corresponding author: Adam Larios.

orthogonal projection onto Fourier modes no larger than $k \sim 1/h$.) Moreover, the data from measurements may be streaming in moment by moment, so in fact, one often has the information $I_h(u) = I_h(u(t))$, for a range of times t. Data assimilation is an approach that eliminates the need for complete initial data and also incorporates incoming data into simulations. Dated back to the 1970s, the technique called nudging or Newtonian relaxation were discussed and analyzed in simpler cases in [3, 63] and the references therein. Essentially the idea was to incorporate the coarse mesh (low-resolution) measurements into certain dynamical systems in the form of a linear feedback control term. Other classical approaches to data assimilation are typically based on the Kalman filter. See, e.g., [4, 31, 74, 84] and the references therein for more information about the Kalman filter. In 2014, an entirely new approach to data assimilation—the AOT algorithm—was introduced in [5, 6]. This new approach overcomes some of the drawbacks of the Kalman filter approach (see, e.g., [13] for further discussion). Moreover, it is implemented directly at the PDE level and is valid for a wide class of dissipative dynamical systems. The approach has been the subject of much recent study in various contexts. For example, it was adapted to the setting of geophysical convection in [7, 35, 36, 37, 39, 40, 41, 42, 96, 97], weather models in [2, 32, 33, 60, 61], magnetohydrodynamics (MHD) in [13, 64], porous media in [8, 89], the surface quasi-geostrophic (SQG) equations in [71, 72], and sabra shell models in [24]. It has also been studied in the context of turbulence models or modified versions of Navier–Stokes [1, 18, 43, 54, 79, 82, 105], stochastic models or noisy observations [9, 12, 20], moving observers, [11, 10, 51] recovery and sensitivity of parameters, [19, 21, 38, 90].

Discrete-in-time observations and other discretizations were studied in [20, 47, 67, 83]. The algorithm in has been studied in the 3D case in [15, 82]. The algorithm itself has been studied, modified, generalized, or improved in various works, including [14, 23, 53, 93]. Computational studies of the AOT algorithm appeared in [26, 27, 52, 55, 80, 83, 99].

The following system was proposed and studied in [5, 6]:

$$\begin{aligned} \frac{dv}{dt} &= F(v) + \mu(I_h(u) - I_h(v)), \\ v(0) &= v_0, \end{aligned} \tag{2}$$

where μ is called the nudging parameter whose value has significant influence on the performance of the nudging technique/AOT algorithm as well as in our nonlinear data assimilation. The central message of the current paper is the emphasis on the nonlinearity in the nudging/data assimilation term and details discussions and analysis on the nudging parameters for more complicated dynamical systems will be the contents of a forthcoming paper. The above system, used in conjunction with (1), is the AOT algorithm for data assimilation of system (1). In the case where the dynamical system (1) is the 2D Navier-Stokes equations, it was proven in [5, 6] that, for any divergence-free initial data $v_0 \in L^2$, $\|u(t) - v(t)\|_{L^2} \to 0$, exponentially in time, under the assumption on the interpolation operator $I_h(\cdot)$ either as

$$\|\phi - I_h(\phi)\|_{L^2} \leq ch\|\phi\|_{H^1}, \tag{3}$$

for all $\phi \in H^1$, or, as

$$\|\phi - I_h(\phi)\|_{L^2} \leq ch\|\phi\|_{H^1} + ch^2\|\phi\|_{H^2}, \tag{4}$$

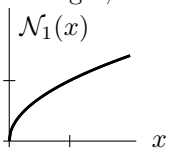
for all $\phi \in H^2$. In particular, even without knowing the initial data u_0, the solution u can be approximately reconstructed for large times. We emphasize that, as noted in

[5], the initial data for (2) can be any L^2 function, even $v_0 = 0$. Thus, no information about the initial data is required to reconstruct the solution asymptotically in time.

The current work is based on the authors' preliminary manuscript [78], in which the following new algorithm was invented for the first time, and which has inspired subsequent work such as [34, 64]. Our principal goal in the current article, is to provide the full scope, after overcoming some computational difficulty encountered in [78], of a new class of nonlinear algorithms for data assimilation. The main idea is to use a nonlinear modification of the AOT algorithm for data assimilation to try to drive the algorithm toward the true solution at a faster rate. In particular, for a given, possibly nonlinear function $\mathcal{N}$, we consider a modification of (2) in the form:

$$\begin{aligned} \frac{dv}{dt} &= F(v) + \mu \mathcal{N}(I_h(u) - I_h(v)), \\ v(0) &= v_0. \end{aligned} \tag{5}$$

To begin, we first focus on the following form of the nonlinearity:



$$\mathcal{N}(x) = \mathcal{N}_1(x) := x|x|^{-\gamma}, \quad x \neq 0, \quad 0 < \gamma < 1, \tag{6}$$

with $\mathcal{N}_1(0) = 0$.

Remark 1.1. Note that by formally setting $\gamma = 0$, one recovers the linear AOT algorithm (2). The main idea behind using such a nonlinearity is that, when $I_h(u)$ is close to $I_h(v)$, the solution v is driven toward the true solution u more strongly than in the linear AOT algorithm. In particular, for any $c > 0$, if $x > 0$ is small enough, then $\mathcal{N}_1(x) > cx$, so no matter how large μ is chosen in the linear AOT algorithm, the nonlinear method with $\mathcal{N} = \mathcal{N}_1$ will always penalize small errors more strongly.

As a preliminary test of the effectiveness of this approach, in this work we demonstrate the nonlinear data assimilation algorithm (5) on a one-dimensional PDE; namely, the Kuramoto-Sivashinsky equation (KSE), given in dimensionless units by:

$$\begin{aligned} u_t + uu_x + \lambda u_{xx} + u_{xxxx} &= 0, \\ u(x,0) &= u_0(x), \end{aligned} \tag{7}$$

in a periodic domain $\Omega = [-L/2, L/2] = \mathbb{R}/L\mathbb{Z}$ of length L. Here, $\lambda > 0$ is a dimensionless parameter. For simplicity, we assume that the initial data is sufficiently smooth (made more precise below) and mean-free, i.e., $\int_{-L/2}^{L/2} u_0(x)\,dx = 0$, which implies $\int_{-L/2}^{L/2} u(x,t)\,dx = 0$ for all $t \geq 0$. This equation has many similarities with the 2D Navier-Stokes equations. It is globally well-posed; it has chaotic large-time behavior; and it has a finite-dimensional global attractor, making it an excellent candidate for studying large-time behavior. It governs various physical phenomena, such as the evolution of flame-fronts, the flow of viscous fluids down inclined planes, and certain types of crystal growth (see, e.g., [77, 101, 102]). Much of the theory of the 1D Kuramoto-Sivashinsky equation was developed in the periodic case in [29, 30, 59, 103, 104, 68] (see also [5, 16, 25, 46, 48, 49, 57, 58, 66, 69, 70, 73, 77, 85, 86, 92, 95, 101, 102, 65]). For a discussion of other boundary conditions for (7), see, e.g., [77, 101, 100, 98, 81]. Discussions about the numerical simulations of the

KSE, can be found in, e.g., [44, 50, 45, 87, 75]. Data assimilation in several different contexts for the 1D Kuramoto-Sivashinsky equation was investigated in [69, 88, 91], who also recognized its potential as an excellent test-bed for data assimilation.

Using the nonlinear data assimilation algorithm (5) in the setting of the Kuramoto-Sivashinsky equation, with the choice (6) for the nonlinearity for some $0 < \gamma < 1$, we arrive at:

$$\begin{aligned} v_t + vv_x + \lambda v_{xx} + v_{xxxx} &= \mu \operatorname{sign}(I_h(u) - I_h(v))|I_h(u) - I_h(v)|^{1-\gamma}, \\ v(x,0) &= v_0(x). \end{aligned} \tag{8}$$

We take I_h to be either orthogonal projection onto the low Fourier modes k with $k \leq c/h$, for some constant $c > 0$, or linear interpolation of nodal observations separated by a distance $h > 0$. Other physically relevant choices of I_h, such as local volume averages, have been considered in the case of the linear AOT algorithm (see, e.g., [5, 55]).

Remark 1.2. We emphasize a crucial point of this work. Indeed, the initial impression of some of our computational work might show that the linear AOT algorithm outperforms our newly-introduced nonlinear DA algorithm. However, we point out that the faster convergence (at exponential rate) of the linear AOT is due to the optimal μ-value, which in practice is not easily obtained. Thus, in view of

1. the convergence at the same precision level, as well as the super-exponential rate of our nonlinear algorithm, without any optimization or fine-tuning of the parameter μ;
2. the significantly fast convergence of our nonlinear algorithm in comparison with that of the linear AOT, both with the same μ-value;
3. the fact that this new nonlinear scheme opens the door to a dynamic μ that may depend on the measured error;

we may infer the nontrivial superiority of the new nonlinear algorithm over the linear AOT algorithm.

Remark 1.3. We note that we performed calculations with nodal interpolation (i.e., I_h is piecewise-linear interpolation on a small number of nodes in physical space), but the results were very similar; hence they are not reproduced here for the sake of brevity.

Remark 1.4. We note that in our convergence plots, we plot with respect to simulation time, as opposed to CPU time. This is due to the fact that our codes make no significant effort to optimize calculations. However, we did track CPU time on every simulation, and the differences in the plots using CPU time instead of simulation time were minor and showed no major qualitative differences. Hence, they are not reproduced here for the sake of brevity.

Remark 1.5. Nonlinear AOT-style data assimilation was first introduced by the authors in a preprint [78] of this paper, originally released on the arXiv [https://arxiv.org/abs/1703.03546] in March 2017, where the super-exponential convergence of the scheme was observed for the first time. For various reasons, the paper was not submitted until much later. We note that several other works [64, 34, 22] have been released since then (which cited our original preprint), continuing investigations into nonlinear data assimilation. Moreover, [34] only investigates the Lorenz equations (not any partial differential equations), and [64, 22] use a different form of the nonlinearity involving an (inherently nonlocal) L^2 norm rather than an

absolute value. Therefore, to the best of our knowledge, the present work is the only one to investigate nonlinear AOT-style data assimilation using a local nonlinearity.

2. **Preliminaries.** In this work, we compute norms of the error given by the difference between the data assimilation solution and the reference solution. We focus on the L^2 and H^1 norms, defined by

$$\|u\|_{L^2}^2 = \frac{1}{L}\int_{-L/2}^{L/2} |u(x)|^2\,dx, \qquad \|u\|_{H^1} = \|\nabla u\|_{L^2}.$$

(Note that, by Poincaré's inequality, $\|u\|_{L^2} \le C\|u\|_{H^1}$, which holds on any bounded domain, $\|u\|_{H^1}$ is indeed a norm.)

We briefly mention the scaling arguments used to justify the form (7). For $a > 0$, $b > 0$, $L > 0$, consider an equation in the form

$$u_t + uu_x + au_{xx} + bu_{xxxx} = 0, \qquad x \in [-L/2, L/2]. \tag{9}$$

Choose time scale $T = L^4/b$, characteristic velocity $U = L/T = b/L^3$, and dimensionless number $\lambda = aL^2/b$. Write $u' = u/U$, $x' = (x + L/2)/L$, $t' = t/T$, where the prime denotes a dimensionless variable. Then

$$\frac{U}{T}u'_{t'} + \frac{U^2}{L}u'u'_{x'} + \frac{aU}{L^2}u'_{x'x'} + \frac{bU}{L^4}u'_{x'x'x'x'} = 0, \qquad x' \in [0, 1] \tag{10}$$

Multiply by $L^4/(bU)$. The equation in dimensionless form then becomes

$$u'_{t'} + u'u'_{x'} + \lambda u'_{x'x'} + u'_{x'x'x'x'} = 0, \qquad x' \in [0, 1]. \tag{11}$$

Thus, λ acts as a parameter which influences the dynamics, in the same way that the Reynolds number influences dynamics in turbulent flows.

Another approach is to set $\ell = (b/a)^{1/2}$, $T = \ell^4/b = b/a^2$, and $U = \ell/T = a^{3/2}/b^{1/2}$. Then, we define dimensionless quantities (denoted again by primes) $u' = u/U$, $x' = (x + L/2)/\ell$, $t' = t/T$. The equation now takes the form

$$\frac{U}{T}u'_{t'} + \frac{U^2}{\ell}u'u'_{x'} + \frac{aU}{\ell^2}u'_{x'x'} + \frac{bU}{\ell^4}u'_{x'x'x'x'} = 0, \qquad x' \in [0, \tfrac{L}{\ell}]. \tag{12}$$

Multiplying by $\ell^4/(bU)$ yields

$$u'_{t'} + u'u'_{x'} + u'_{x'x'} + u'_{x'x'x'x'} = 0, \qquad x' \in [0, \tfrac{L}{\ell}]. \tag{13}$$

Thus, equation (12) is similar to equation (10), with $\lambda = 1$, except that the dynamics are influenced by the dimensionless parameter L/ℓ. In particular, the dynamics can be thought of as influenced by parameter λ with L fixed, or equivalently influenced by the length of the domain L with λ fixed, where $\lambda \sim (L/\ell)^2$. In this work, for the sake of matching the initial data used in [76], we choose the domain to be $[-16\pi, 16\pi]$, so $L = 32\pi$ is fixed, and we let λ be the parameter affecting the dynamics.

Next, we summarize and specify the well-posedness and regularity results of system 7 from the references listed in the previous section. We also state the main theorems regarding the exponential-in-time convergence in L^2- and H^1-norms of the *linear* AOT data assimilation solution to that of the reference solution, of which the proof will be given in Section 4, which is similar to the proof of null controlability given in [88]. Note that the global existence, uniqueness, and convergence of the nonlinear schemes (5), (8), etc., presented here for the KSE are still open. However,

we also note that very recently in [22], super-exponential convergence for a related[1] nonlinear data assimilation scheme was rigorously established in the context of the Navier-Stokes equations.

We first state the frequently-used Agmon's inequality in one dimension. For $\psi \in H^1$, there exists some constant K such that

$$\|\psi\|_{L^\infty}^2 \leq K\|\psi\|_{L^2}\|\psi\|_{H^1}. \tag{14}$$

Next, we define the weak and strong solutions of system 7 as follows.

Definition 2.1. For mean-free initial data $u_0 = u(0) \in L^2$, we say $u(t)$ is a weak solution to system 7 on the time interval $[0,T), T > 0$, if

$$u \in L^\infty(0,T;L^2) \cap L^2(0,T;H^2),$$

and if it also satisfies the following weak formulation for the test function $\phi \in H^2$:

$$\frac{d}{dt}(u,\phi) - \frac{1}{2}(u^2,\phi_x) + \lambda(u,\phi_{xx}) + (u_{xx},\phi_{xx}) = 0.$$

If we further assume $u_0 \in H^1$, then we say $u(t)$ is a strong solution to system 7, if

$$u \in L^\infty(0,T;H^1) \cap L^2(0,T;H^3).$$

We state the following well-posedness and regularity results regarding system 7. Global well-posedness was proved in [103] (see also [16, 56] and the references therein). Also, by Theorem 3.1 of [28], the Gevrey norm is uniformly bounded in time, which automatically implies a uniform bound on the L^∞ norm and all Sobolev norms (see also [29, 59] and the references therein).

Theorem 2.2. *For all $T > 0$ and mean-free initial condition $u(\cdot,0) = u_0 \in H^1$, there exists a unique strong solution $u(t)$ defined on the time interval $[0,T)$ as in Definition 2.1, to systems 7. If we further assume $u_0 \in H^k$ for any $k \in \mathbb{N}^+$, then the solution*

$$u \in L^\infty(0,T;H^k) \cap L^2(0,T;H^{k+2}).$$

Moreover, the H^k-norms (k=1,2,...) are all bounded uniformly, i.e., independent of time T.

By similar arguments from the reference mention above Theorem 2.2, which we omit for the sake of simplicity, we have the well-posedness results of the linear AOT data assimilation equation of the 1D Kuramoto-Sivashinsky equation as follows.

Theorem 2.3. *For all $T > 0$ and for (smooth) initial condition $v(\cdot,0) = v_0 = 0$, there exists a unique strong solution (defined analogously to Definition 2.1) $v(t) \in L^\infty([0,T],H^1) \cap L^2([0,T],H^3)$ to systems 8 with $\gamma = 0$.*

Our main convergence results for linear AOT are the following theorems.

Theorem 2.4. *Let $\gamma = 0$, i.e., consider the linear AOT case. Suppose $\mu > C(\lambda^2 + \|u\|_{L^\infty(0,\infty;H^1)}^2)$ and $h > 0$ is such that $h^4\mu < c$, where c, C are absolute constants (determined below). Suppose I_h satisfies condition 3. Then*

$$\lim_{t\to\infty} \|u(t) - v(t)\|_{L^2} = 0$$

[1] In [22], the authors use a nonlinearity of the form $\mathcal{N}(u) = \beta u + \mu u\|u\|_{L^2}^{-\gamma}$, whereas this nonlinearity in the present work does not involve an L^2 norm, only an absolutle value. Note also that [64] used an L^2 norm rather than an absolute value.

exponentially fast in time, where $u(t), v(t)$ are solutions to (7) and (8), respectively, for any mean-free initial data u_0, $v_0 \in L^2$.

Theorem 2.5. *Under the same assumption of Theorem (2.4), and if we further assume*

$$\mu > C(\lambda^2 + \|u\|^2_{L^\infty(0,\infty;H^1)}),$$

and h is such that

$$h^2\mu < \min\left\{\frac{1}{2}, \frac{\lambda_1}{8c}\right\},$$

where c, C are absolute constants, and with I_h satisfying condition 4, then, we have

$$\lim_{t\to\infty} \|u(t) - v(t)\|_{H^1} = 0$$

exponentially fast in time, where $u(t), v(t)$ are solutions to (7) and (8), respectively, respectively, for any mean-free initial data u_0, $v_0 \in H^1$.

3. **Computational results.** In this section, we demonstrate some computational results for the nonlinear data assimilation algorithm given by (8).

3.1. **Numerical methods.** It was observed in [55] that no higher-order multi-stage Runge-Kutta-type method exists for solving (8) due to the need to evaluate at fractional time steps, for which the data $I_h(u)$ is not available. Therefore, we use a semi-implicit spectral method with Euler time stepping. The linear terms are treated via a first-order exponential time differencing (ETD) method (see, e.g., [76] for a detailed description of this method). The nonlinear term is computed explicitly, and in the standard way, i.e., by computing the derivatives in spectral space, and products in physical space, respecting the usual 2/3's dealiasing rule. We use $N = 2^{13} = 8192$ spatial grid points on the interval $[-16\pi, 16\pi)$, so $\Delta x = 32\pi/N \approx 0.0123$. We use a fixed time-step respecting the advective CFL; in fact, we choose $\Delta t = 1.2207 \times 10^{-4}$. For simplicity, we choose $\mu = 1$, however, the results reported here are qualitatively similar for a wide range of μ values. For example, when $\mu = 10$, convergence times are shorter for all methods, but the error plots are qualitatively similar. In [76], the case $\lambda = 1$ is examined. However, to examine a slightly more chaotic setting, we take $\lambda = 2$, which is still well-resolved with $N = 8192$. Our results are qualitatively similar for smaller values of λ.

Let I_h be the projection onto the lowest $M = \lfloor 32\pi/h \rfloor$ Fourier modes. In this work, we set $M = 32$ (i.e, $h = \pi$); so only the lowest 32 modes of u are passed to the assimilation equation via $I_h(u)$.

One can consider a more general interpolation operator as well, such as nodal interpolation, but we focus on projection onto low Fourier modes.

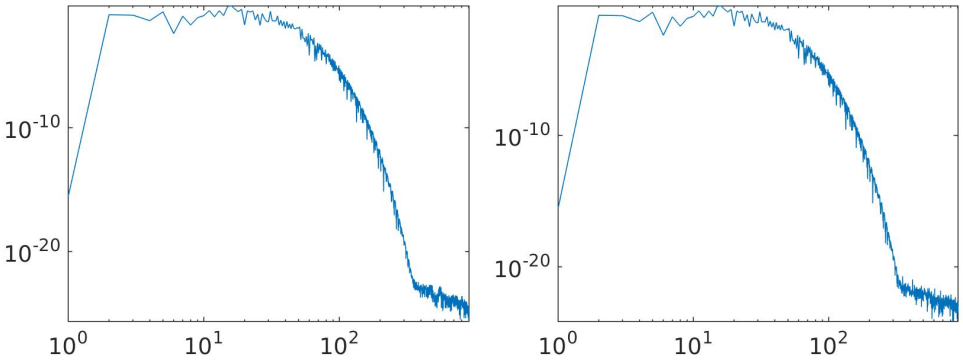
To fix ideas, in this paper we mainly use the initial data used in [76] to simulate (7); namely

$$u_0(x) = \cos(x/16)(1 + \sin(x/16)); \tag{15}$$

on the interval $[-16\pi, 16\pi]$. However, we also investigated several other choices of initial data. In all cases, the results were qualitatively similar to the ones reported here.

Note that explicit treatment of the term $\mu(I_h(u) - I_h(v))$ imposes a constraint on the time step, namely $\Delta t < 2/\mu$ (which follows from a standard stability analysis for Euler's method). This is not a serious restriction in this work, since we choose $\mu = 1$.

All the simulations in the present work are well-resolved. In Figure (1) we show plots of time-averaged spectra of all the PDEs simulated in the present work. One can see that all relevant wave-modes are captured to within machine precision.



(a) Time-averaged spectrum of the reference solution u to the 1D-Kuramoto-Sivashinsky equation.

(b) Time-averaged spectrum of the data assimilation solution v with nonlinear pure-power ($\mathcal{N}_1$) algorithm.

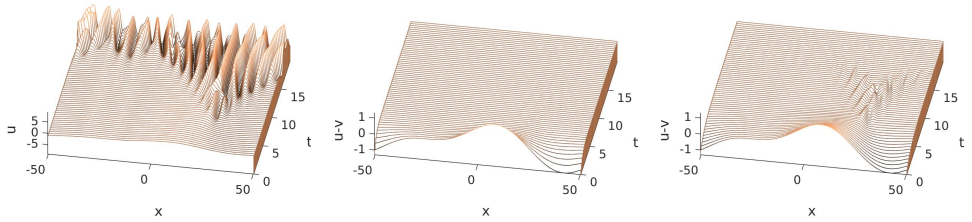
(c) Time-averaged spectrum of the data assimilation solution v with nonlinear hybrid ($\mathcal{N}_2$) algorithm.

(d) Time-averaged spectrum of the data assimilation solution v with nonlinear concave-convex ($\mathcal{N}_3$) algorithm.

FIGURE 1. Log-log plots of the energy spectra for the above scenarios. the mode amplitude $|\widehat{u}_k|$ is on the y-axis, and the wave number k is on the x-axis (shifted to $k+1$ to accommodate the $k=0$ mode on a log scale). Plots are averaged over all time steps between times $t=20$ and $t=60$. ($\lambda=2$)

3.2. **Simple power nonlinearity.** We compare the error in the nonlinear data assimilation algorithm (5) with the error in the linear AOT algorithm. We first focus on nonlinearity given by a power according to (6); i.e., we consider equation (8) together with equation (7). In Figure 2(a), the solution to (7) (which we call the "reference" solution) evolves from the smooth, low-mode initial condition (15) to a chaotic state after about time $t=20$. In Figure 2(b), the difference between this solution and the AOT data assimilation solution is plotted. It rapidly decays to zero in a short time.

We observe in Figure 3 that errors in the linear AOT algorithm (2) and the nonlinear algorithm (5) solutions both decay. The error in the nonlinear algorithm



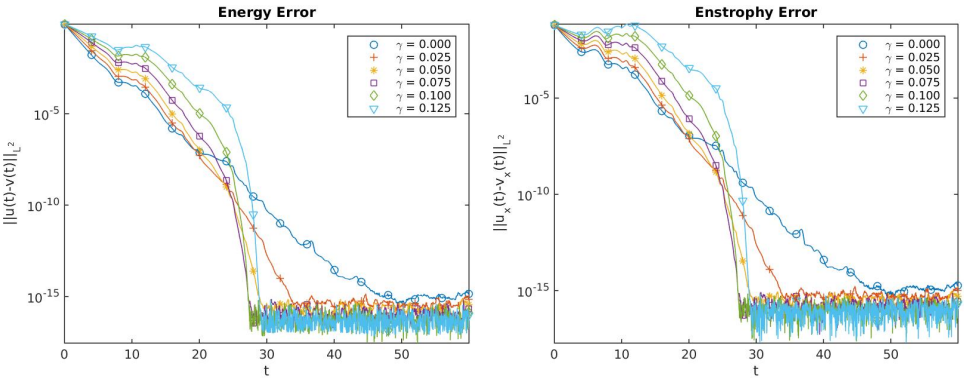
(a) A chaotic solution to the Kuramoto-Sivashinsky equation evolving in time.

(b) Error in data assimilation solution using linear AOT algorithm ($\gamma = 0$).

(c) Error in data assimilation solution using nonlinear algorithm (5) ($\gamma = 0.125$).

FIGURE 2. Data Assimilation for the Kuramoto-Sivashinsky equation ($\lambda = 2$) using linear and nonlinear algorithms. The difference rapidly decays to zero in time, and visually the errors look similar. The assimilation equations were initialized with $v_0(x) = 0$. I_h is the orthogonal projection onto the lowest 32 Fourier modes. Similar results appear in tests of a wide variety of initial data, and for $0 < \gamma < 0.125$.

has oscillations for roughly $5 \lesssim t \lesssim 15$ which are not present in the error for the AOT algorithm. However, by tracking norms of the difference of the solutions, one can see in Figure 3 that the nonlinear algorithm reaches machine precision significantly faster than the linear AOT algorithm, for a range of γ values. When



(a) Errors in L^2-norm vs. time.

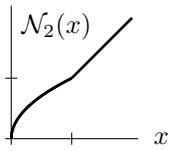
(b) Errors in H^1-norm vs. time.

FIGURE 3. Error for the linear AOT ($\gamma = 0$) solution and the nonlinear (5) ($\gamma > 0$) solution for various values of γ. Resolution 8192. (Log-linear scale.)

$\gamma > 0.2$, our simulations appear to no longer converge (not shown here). The error in the linear AOT algorithm (i.e., $\gamma = 0$) reaches machine precision at roughly time $t \approx 49.8$. For $0 < \gamma < 0.2$, there seems to be an optimal choice in our simulations around $0.075 \lesssim \gamma \lesssim 0.1$, reaching machine precision around $t \approx 27.3$, a speedup factor of roughly $49.8/27.3 \approx 1.8$. Moreover, the shape of the curves with $\gamma > 0$ indicate super-exponential convergence, indicated by the concave curve on the log-linear plot in Figure 3, while for the linear AOT algorithm, the convergence is only exponential, indicated by the linear shape on the log-linear plot. Currently, the

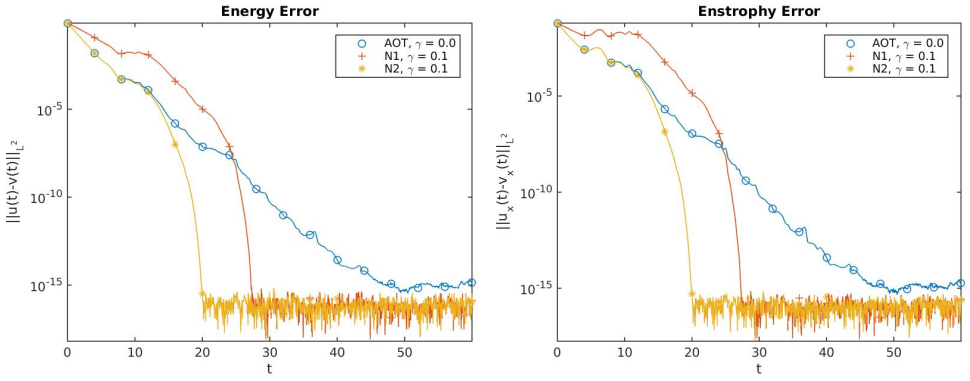
super exponential convergence is only an observation in simulations. An analytical investigation of the convergence rate will be the subject of a forthcoming work.

3.3. **Hybrid linear/nonlinear methods.** In this subsection, we investigate a family of hybrid linear/nonlinear data assimilation algorithms. One can see from Figure 3) in the previous subsection that, although the nonlinear methods converge to machine precision at earlier times than the linear method, the nonlinear method suffers from larger errors than the linear method for short times. This motivates the possibility of using a hybrid linear/nonlinear method. For example, one could look for an optimal time to switch between the models, say, perhaps around time $t \approx 18 \pm 2$, according to Figure 3), but this seems highly situationally dependent and difficult to implement in general. Instead, the approach we consider here is to let $\mathcal{N}(x)$ be given by (6) for $|x| \leq 1$ but let it be linear for $|x| > 1$. The idea is that, when the error is small, deviations are strongly penalized, as in Remark (1.1). However, where the error is large, the linear AOT algorithm should give the greater penalization (i.e., $\mathcal{N}_1(x) < x$ when $x > 1$). Therefore, we consider algorithm (5) with the following choice of nonlinearity, for some choice of γ, $0 < \gamma < 1$ (we take $\gamma = 0.1$ in all following simulations).



$$\mathcal{N}(x) = \mathcal{N}_2(x) := \begin{cases} x, & |x| \geq 1, \\ x|x|^{-\gamma}, & 0 < |x| < 1, \\ 0, & x = 0. \end{cases} \tag{16}$$

We see the improvement this yields in the error in the plots below.



(a) Error in L^2-norm vs. time.

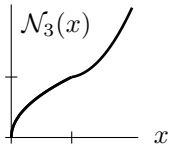
(b) Error in H^1-norm vs. time.

FIGURE 4. Error in linear ($\gamma = 0$), nonlinear ($\gamma = 0.1$), and hybrid ($\gamma = 0.1$) algorithms ($\lambda = 2$). Resolution 8192. (Log-linear scale.)

In Figure 4, we compare the linear algorithm (2) with the nonlinear algorithm (5) with pure-power nonlinearity $\mathcal{N}_1$, given by (6), and also with hybrid nonlinearity $\mathcal{N}_2$, given by (16). The convergence to machine precision happens approximately at $t \approx 49.8$ (for AOT), $t \approx 27.3$ (for $\mathcal{N}_1$), and $t \approx 20.0$ (for $\mathcal{N}_2$), respectively. In addition, one can see that the hybrid algorithm remains close to the linear AOT algorithm for short times. Moreover, after a short time, the hybrid algorithm undergoes super-exponential convergence, converging faster than every algorithm analyzed so far. The benefits of this splitting of the nonlinearity between $|x| > 1$ and $|x| < 1$ seem

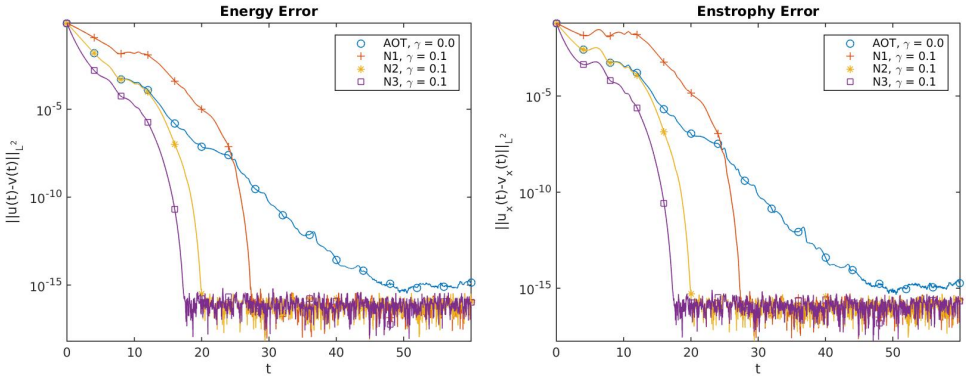
clear. Moreover, this approach can be exploited further, which is the topic of the next subsection.

3.4. **Concave-convex nonlinearity.** Inspired by the success of the hybrid method, in this subsection, we further exploit the effect of the feedback control term $\mu\mathcal{N}(I_h(u) - I_h(v))$ by accentuating the nonlinearity for $|x| > 1$. We consider the following nonlinearity in conjunction with (5) for the Kuramoto-Sivashinsky equation.



$$\mathcal{N}(x) = \mathcal{N}_3(x) := \begin{cases} x|x|^{\gamma}, & |x| \geq 1, \\ x|x|^{-\gamma}, & 0 < |x| < 1, \\ 0, & x = 0. \end{cases} \tag{17}$$

Note that this choice of $\mathcal{N}_3$ is concave for $|x| < 1$, and convex for $|x| \geq 1$. The convexity for $|x| \geq 1$ serves to more strongly penalize large deviations from the reference solution. In Figure 5, we see that *at every positive time* this method has significantly smaller error than the linear AOT method, and the methods involving $\mathcal{N}_1$ and $\mathcal{N}_2$. Convergence to machine precision happens at roughly $t \approx 17.4$, a speedup factor of roughly $49.8/17.4 \approx 2.8$ compared to the linear AOT algorithm.

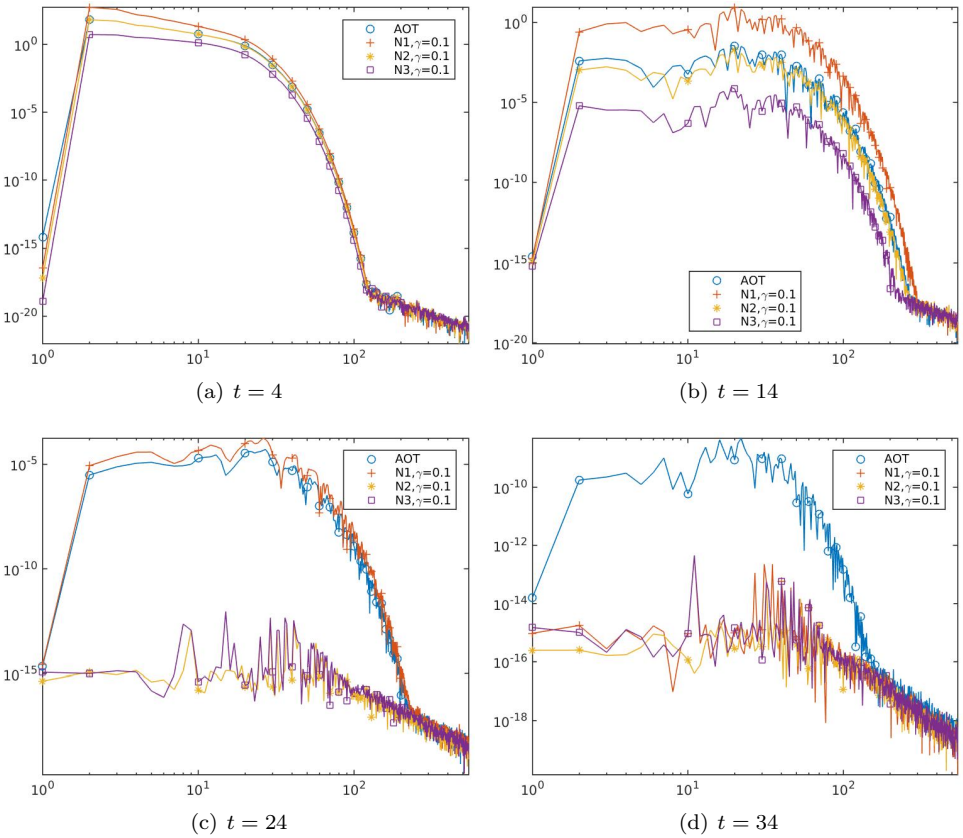


(a) Error in L^2-norm vs. time. (b) Error in H^1-norm vs. time.

FIGURE 5. Error in linear AOT ($\gamma = 0$) algorithm, and the non-linear algorithm with nonlinearities $\mathcal{N}_1$, $\mathcal{N}_2$, and $\mathcal{N}_3$ (each with $\gamma = 0.1$) (Modal case). ($\lambda = 2$) Resolution 8192. (Log-linear scale.)

3.5. **Comparison of all methods (Modal observations).** Let us also consider the error at every Fourier mode. In Figure 6, one can see these errors at various times. We examine a time before the transition to fully-developed chaos ($t \approx 4$), at a time during the transition ($t \approx 14$), a time after the solution has settled down to an approximately statistically steady state ($t \approx 24$), and a later time ($t \approx 34$). At each mode, and at each positive time, the error in the solution with nonlinearity (17) is the smallest.

Next, we point out that our results hold qualitatively with different choices of initial data for the reference equation (7). Therefore, we wait until the solution to (7) with initial data (15) has reached an approximately statistically steady state (this happens roughly at $t \approx 20$). Then, we use this data to re-initialize the solution



(a) $t = 4$ (b) $t = 14$

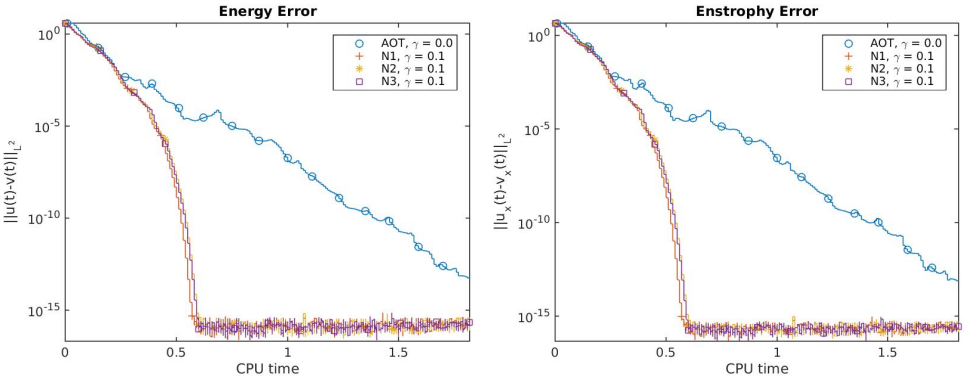
(c) $t = 24$ (d) $t = 34$

FIGURE 6. Error in spectrum (mode amplitude vs. wave number) at different times for all methods.

to (7) (in fact, we use the solution at $t = 30$ to be well within the time interval of fully developed chaos). We still initialize (8) with $v_0 \equiv 0$. Norms of the errors are shown in Figure 7. We observe that, although convergence time is increased for all methods, the qualitative observations discussed above still hold.

3.6. **Comparison of all methods (Nodal observations).** For the sake of comparison, we also consider nodal observations in Figure 8. That is, we take observations of u on a uniform grid in space with separation $h > 0$. The operator I_h is then taken to a piece-wise linear interpolation of these observations. In our simulations on the domain $[-16\pi, 16\pi)$, we choose $h = \pi$ for a total of $M = 32$ grid points, the same number of observations as in the modal case. Note that, because the domain is periodic, the linear interpolation on the last point ($x = 15\pi$) must wrap around to the first point ($x = -16\pi$). This was done using Matlab's `interp1` function by duplicating the data of the first point $x = -16\pi$ and placing it at location $x = 16\pi$ for a total of $M+1 = 33$ observation points, performing the interpolation, and then deleting the redundant point.

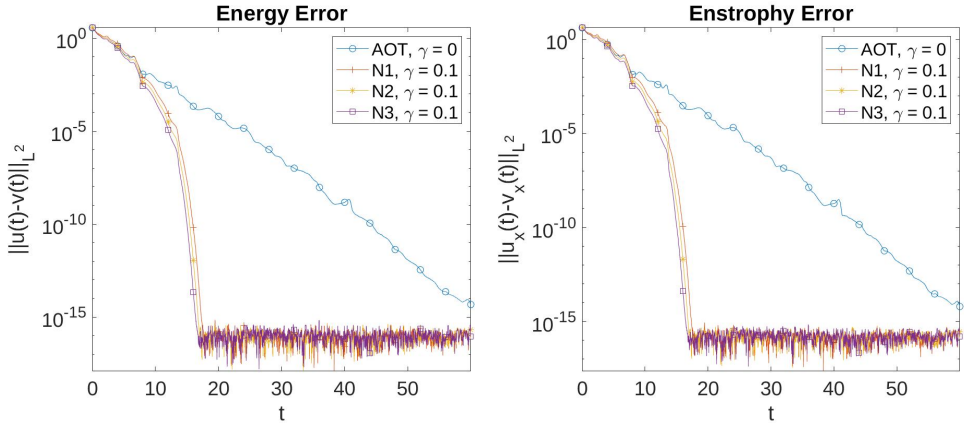
The results for the nodal case are qualitatively similar to the modal case. Hence to avoid redundancy but still give some validation, we only include one figure (Figure



(a) Error in L^2-norm vs. time. (b) Error in H^1-norm vs. time.

FIGURE 7. Error in all algorithms with chaotic initialization for reference solution, plotted against CPU time. Modal observations. ($\lambda = 2$) Resolution 8192. (Log-linear scale.)

8) showing the results in the nodal case, which we note is qualitatively quite similar to the modal case in Figure 5.



(a) Error in L^2-norm vs. time. (b) Error in H^1-norm vs. time.

FIGURE 8. Error in all algorithms with chaotic initialization for reference solution. Nodal observations. ($\lambda = 2$) Resolution 8192. (Log-linear scale.)

4. **Analytic results: Proof of Theorem 2.4 and Theorem 2.5.** In this section, for the sake of completeness, we provide the key *a priori* estimates for the proof of both of our convergence theorems, which follow almost line-by-line the estimates in [88], though the context here is slightly different in that we consider a data assimilation algorithm rather than the problem of null controlability. We point out that constants c and C below may vary from line to line, but are absolute, and are compatible with the condition of the theorem.

As above, u is the reference solution and v is the solution to the linear data assimilation system. Let $w = u - v$, and subtract equation (8) from equation (7), we have

$$w_t + uw_x + wu_x - ww_x + \lambda w_{xx} + w_{xxxx} = -\mu\, I_h(w), \tag{18}$$

with $w_0 = w(\cdot, 0) = u_0 - v_0$.

Taking the inner-product of equation (18) with w in L^2, integrating by parts, using the fact that $(ww_x, w) = 0$ due to the periodic boundary conditions, and rearranging terms, we obtain

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|w\|_{L^2}^2 + \|w_{xx}\|_{L^2}^2 = (uw_x, w) - \lambda(w_{xx}, w) - \mu(I_h(w), w) \\
&\leq \|u\|_{L^\infty}\|w_x\|_{L^2}\|w\|_{L^2} + C\lambda\|w_{xx}\|_{L^2}\|w\|_{L^2} \\
&\quad - \mu\|w\|_{L^2}^2 + \mu|w - I_h(w)\|_{L^2}\|w\|_{L^2} \\
&\leq C\|u\|_{H^1}^2\|w\|_{L^2}^2 + \frac{1}{8}\|w_{xx}\|_{L^2}^2 + C\lambda^2\|w\|_{L^2}^2 + \frac{1}{8}\|w_{xx}\|_{L^2}^2 \\
&\quad - \mu\|w\|_{L^2}^2 + \mu\|w - I_h(w)\|_{L^2}\|w\|_{L^2},
\end{aligned}$$

where we used Cauchy-Schwarz and Young's inequalities, as well as Poincaré's inequality and Morrey's inequality in the last step. Then, by condition (3) and Sobolev-type inequality, we bound the last term in the above inequality as

$$\begin{aligned}
&\mu\|w - I_h(w)\|_{L^2}\|w\|_{L^2} \leq ch\mu\|w_x\|_{L^2}\|w\|_{L^2} \\
&\leq ch\mu\|w\|_{L^2}^{3/2}\|w_{xx}\|_{L^2}^{1/2} \leq \frac{1}{4}\|w_{xx}\|_{L^2}^2 + c(h\mu)^{4/3}\|w\|_{L^2}^2.
\end{aligned}$$

Thus, combining the above estimates, we obtain

$$\frac{d}{dt}\|w\|_{L^2}^2 + \|w_{xx}\|_{L^2}^2 \leq C\big(\|u\|_{H^1}^2 + \lambda^2 + c(h\mu)^{4/3} - \mu\big)\|w\|_{L^2}^2.$$

Therefore, by our assumption on h and μ, and in view of Grönwall's inequality, we have

$$\|u(t) - v(t)\|_{L^2}^2 \leq e^{(c(h\mu)^{4/3} - \frac{\mu}{2})t}\|u_0 - v_0\|_{L^2}^2,$$

which gives the desired exponential rate of the convergence in L^2-norm.

Next, we show H^1 convergence in Theorem 2.5. Taking the inner product of (18) with $-w_{xx}$ in L^2, and integrating by parts, we obtain

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|w_x\|_{L^2}^2 + \|w_{xxx}\|_{L^2}^2 \\
&= -(uw, w_{xxx}) + \frac{1}{2}(w^2, w_{xxx}) - \lambda(w_{xxx}, w_x) + \mu(I_h(w), w_{xx}) \\
&\leq C\|u\|_{H^1}^2\|w\|_{L^2}^2 + C\|w\|_{L^2}^{3/2}\|w_x\|_{L^2}^{1/2}\|w_{xxx}\|_{L^2} + C\lambda^2\|w_x\|_{L^2}^2 + \frac{1}{8}\|w_{xxx}\|_{L^2}^2 \\
&\quad - \mu\|w_x\|_{L^2}^2 + \mu\|w - I_h(w)\|_{L^2}\|w_{xx}\|_{L^2},
\end{aligned}$$

where we used Cauchy-Schwarz, Young's inequalities, as well as Agmon's inequality 14. For the second term on the right side of the above inequality, we further bound it by

$$C\|w\|_{L^2}^{3/2}\|w_x\|_{L^2}^{1/2}\|w_{xxx}\|_{L^2} \leq \frac{C}{\sqrt{\lambda_1}}\|w\|_{L^2}^2\|w_x\|_{L^2}^2 + \frac{1}{8}\|w_{xxx}\|_{L^2}^2$$

while for the last term, we apply condition (4) on the interpolator and bound it by

$$\mu\|w - I_h(w)\|_{L^2}\|w_{xx}\|_{L^2} \le ch\mu\|w_x\|_{L^2}\|w_{xx}\|_{L^2} + ch^2\mu\|w_{xx}\|_{L^2}^2$$
$$\le \frac{Ch^2\mu^2}{\lambda_1}\|w_x\|_{L^2}^2 + \frac{1}{8}\|w_{xxx}\|_{L^2}^2 + \frac{ch^2\mu}{\lambda_1}\|w_{xxx}\|_{L^2}^2,$$

where we used Poincáre's inequality and λ_1 is the first eigenvalue of the negative Laplacian operator.

Therefore, by combining all the above estimates, we have

$$\frac{d}{dt}\|w_x\|_{L^2}^2 + \|w_{xxx}\|_{L^2}^2 \le C_{\lambda_1}\big(\|u\|_{H^1}^2 + \lambda^2 + \|w\|_{L^2}^2 + h^2\mu^2 - \mu\big)\|w_x\|_{L^2}^2,$$

where we used the second condition on our parameters

$$\frac{ch^2\mu}{\lambda_1} \le \frac{1}{8}.$$

Note that our first condition on the parameters implies $h^2\mu^2 < \frac{\mu}{2}$ and

$$\frac{\mu}{2} > \|u\|_{L^\infty(0,\infty;H^1)}^2 + \lambda^2,$$

and after some time $T > 0$, $\|w\|_{L^2}^2 < \epsilon$ becomes sufficiently small. Hence, Grönwall's inequality implies the desired H^1-convergence result after inflating the absolute constant C. (c.f., details in [88], Section 4.2). □

5. **Conclusions.** Our results indicate that significant advantages might be gained by using nonlinear data assimilation. We used the Kuramoto-Sivashinsky equation as a proof-of-concept for this method; however, in future work, we will extend the method to more challenging equations, including the Navier-Stokes equations of fluid flow. Mathematical analysis of these methods will also be subject of future work (and see [22] for analysis in the 2D Navier-Stokes case of a norm-based nonlinear data assimilation algorithm).

Our findings seem to indicate that, of the nonlinear functions we considered, namely $\mathcal{N}_1$, $\mathcal{N}_2$, and $\mathcal{N}_3$, there seems to be a clear progression of improvement in convergence times from $\mathcal{N}_1$ through $\mathcal{N}_3$. However, for other more complex systems, this might not be the case. Moreover, depending on discretization, implementation, etc., some algorithms may have better computational performance than others, which is why we presented all three algorithms here. It may also be the case that certain versions of $\mathcal{N}$ are easier to handle analytically. Indeed, in [22], only for $\mathcal{N}_2$ (modified to divide by a norm rather than an absolute value) could super-exponential convergence be proved, while proofs for the $\mathcal{N}_1$ and $\mathcal{N}_3$ cases were intractable. In addition, this progression of improvements indicates that even "better" choices of nonlinearity may be possible. Indeed, one may imagine a functional given by

$$\mathfrak{F}(\mathcal{N}) = t_*$$

where t_* is the time of convergence to within a certain error tolerance, such as to within machine precision. (One would need to show that admissible functions $\mathcal{N}$ are in some sense independent of the parameters and initial data, say, after some normalization.) One could also consider a functional whose value at $\mathcal{N}$ is given by a particular norm of the error. By minimizing such functionals, one might discover even better data-assimilation methods.

Acknowledgments. The authors would like to thank Prof. Edriss S. Titi for interesting discussions. The research of A.L. was supported in part by NSF grants DMS-2206762 and CMMI-1953346, and USGS grant G23AS00157 number GRANT13798170.

REFERENCES

[1] D. A. Albanez, H. J. Nussenzveig Lopes and E. S. Titi, Continuous data assimilation for the three-dimensional Navier–Stokes-α model, *Asymptotic Anal.*, **97** (2016), 139-164.

[2] M. U. Altaf, E. S. Titi, O. M. Knio, L. Zhao, M. F. McCabe and I. Hoteit, Downscaling the 2D Benard convection equations using continuous data assimilation, *Comput. Geosci*, **21** (2016), 393-410.

[3] R. A. Anthes, Data assimilation and initialization of hurricane prediction models, *J. Atmos. Sci.*, **31** (1974), 702-719.

[4] M. Asch, M. Bocquet and M. Nodet, *Data Assimilation: Methods, Algorithms, and Applications*. Fundam. Algorithms, 11. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2016.

[5] A. Azouani, E. Olson and E. S. Titi, Continuous data assimilation using general interpolant observables, *J. Nonlinear Sci.*, **24** (2014), 277-304.

[6] A. Azouani and E. S. Titi, Feedback control of nonlinear dissipative systems by finite determining parameters—a reaction-diffusion paradigm, *Evol. Equ. Control Theory*, **3** (2014), 579-594.

[7] A. Balakrishna and A. Biswas, Determining map, data assimilation and an observable regularity criterion for the three-dimensional Boussinesq system, *Appl. Math. Optim.*, **86** (2022), Paper No. 28, 53 pp.

[8] H. Bessaih, V. Ginting and B. McCaskill, Continuous data assimilation for displacement in a porous medium, *Numer. Math.*, **151** (2022), 927-962.

[9] H. Bessaih, E. Olson and E. S. Titi, Continuous data assimilation with stochastically noisy data, *Nonlinearity*, **28** (2015), 729-753.

[10] A. Biswas, Z. Bradshaw and M. Jolly, Convergence of a mobile data assimilation scheme for the 2D Navier–Stokes equations, *Discrete and Continuous Dynamical Systems*, (2022), `arXiv:2210.11282`.

[11] A. Biswas, Z. Bradshaw and M. S. Jolly, Data assimilation for the Navier–Stokes equations using local observables, *SIAM J. Appl. Dyn. Syst.*, **20** (2021), 2174-2203.

[12] A. Biswas, C. Foias, C. F. Mondaini and E. S. Titi, Downscaling data assimilation algorithm with applications to statistical solutions of the Navier–Stokes equations, *Annales de L'Institut Henri Poincaré C, Analyse non Linéaire*, **36** (2019), 295-326.

[13] A. Biswas, J. Hudson, A. Larios and Y. Pei, Continuous data assimilation for the 2D magnetohydrodynamic equations using one component of the velocity and magnetic fields, *Asymptot. Anal.*, **108** (2018), 1-43.

[14] A. Biswas and V. R. Martinez, Higher-order synchronization for a data assimilation algorithm for the 2D Navier–Stokes equations, *Nonlinear Anal. Real World Appl.*, **35** (2017), 132-157.

[15] A. Biswas and R. Price, Continuous data assimilation for the three-dimensional Navier–Stokes equations, *SIAM Journal on Mathematical Analysis*, **53** (2021), 6697-6723.

[16] J. C. Bronski and T. N. Gambill, Uncertainty estimates and L_2 bounds for the Kuramoto–Sivashinsky equation, *Nonlinearity*, **19** (2006), 2023-2039.

[17] C. Cao, I. G. Kevrekidis and E. S. Titi, Numerical criterion for the stabilization of steady states of the Navier–Stokes equations, *Indiana Univ. Math. J.*, **50** (2001), 37-96.

[18] Y. Cao, A. Giorgini, M. Jolly and A. Pakzad, Continuous data assimilation for the 3D Ladyzhenskaya model: Analysis and computations, *Nonlinear Anal. Real World Appl.*, **68** (2022), Paper No. 103659, 29 pp.

[19] E. Carlson, J. Hudson and A. Larios, Parameter recovery for the 2 dimensional Navier–Stokes equations via continuous data assimilation, *SIAM J. Sci. Comput.*, **42** (2020), A250-A270.

[20] E. Carlson, J. Hudson, A. Larios, V. R. Martinez, E. Ng and J. P. Whitehead, Dynamically learning the parameters of a chaotic system using partial observations, *Discrete Contin. Dyn. Syst.*, **42** (2022), 3809–3839.

[21] E E. Carlson and A. Larios, Sensitivity analysis for the 2D Navier–Stokes equations with applications to continuous data assimilation, *J. Nonlinear Sci.*, **31** (2021), Paper No. 84, 30 pp.

[22] E. Carlson, A. Larios and E. S. Titi, Super-exponential convergence rate of a nonlinear continuous data assimilation algorithm: The 2D Navier–Stokes equations paradigm, submitted, (2023), arXiv:2304.01128.
[23] E. Celik, E. Olson and E. S. Titi, Spectral filtering of interpolant observables for a discrete-in-time downscaling data assimilation algorithm, *SIAM J. Appl. Dyn. Syst.*, **18** (2019), 1118-1142.
[24] N. Chen, Y. Li and E. Lunasin, An efficient continuous data assimilation algorithm for the sabra shell model of turbulence, *Chaos*, **31** (2021), 103123.
[25] A. Cheskidov and C. Foias, On the non-homogeneous stationary Kuramoto–Sivashinsky equation, *Phys. D*, **154** (2001), 1-14.
[26] P. Clark Di Leoni, A. Mazzino and L. Biferale, Inferring flow parameters and turbulent configuration with physics-informed data assimilation and spectral nudging, *Phys. Rev. Fluids*, **3** (2018), 104604.
[27] P. Clark Di Leoni, A. Mazzino and L. Biferale, Synchronization to big data: Nudging the Navier–Stokes equations for data assimilation of turbulent flows, *Physical Review X*, **10** (2020), 011023.
[28] P. Collet, J.-P. Eckmann, H. Epstein and J. Stubbe, Analyticity for the Kuramoto–Sivashinsky equation, *Phys. D*, **67** (1993), 321-326.
[29] P. Collet, J.-P. Eckmann, H. Epstein and J. Stubbe, A global attracting set for the Kuramoto–Sivashinsky equation, *Comm. Math. Phys.*, **152** (1993), 203-214.
[30] P. Constantin, C. Foias, B. Nicolaenko and R. Temam, *Integral Manifolds and Inertial Manifolds for Dissipative Partial Differential Equations*, Appl. Math. Sci., 70. Springer-Verlag, New York, 1989.
[31] R. Daley, *Atmospheric Data Analysis*, Cambridge Atmospheric and Space Science Series, Cambridge University Press, 1993.
[32] S. Desamsetti, H. Dasari, S. Langodan, O. Knio, I. Hoteit and E. S. Titi, Efficient dynamical downscaling of general circulation models using continuous data assimilation, *Quarterly Journal of the Royal Meteorological Society*, **145** (2019), 3175-3194.
[33] S. Desamsetti, H. Prasad Dasari, S. Langodan, Y. Viswanadhapalli, R. Attada, T. M. Luong, O. Knio, E. S. Titi and I. Hoteit, Enhanced simulation of the Indian summer monsoon rainfall using regional climate modeling and continuous data assimilation, *Frontiers in Climate*, **4** (2022), 817076.
[34] Y. J. Du and M.-C. Shiue, Analysis and computation of continuous data assimilation algorithms for Lorenz 63 system based on nonlinear nudging techniques, *J. Computat. and Appl. Math.*, **386** (2021), 113246, 17 pp.
[35] A. Farhat, N. E. Glatt-Holtz, V. R. Martinez, S. A. McQuarrie and J. P. Whitehead, Data assimilation in large Prandtl Rayleigh-Bénard convection from thermal measurements, *SIAM J. Appl. Dyn. Syst.*, **19** (2020), 510-540.
[36] A. Farhat, H. Johnston, M. Jolly and E. S. Titi, Assimilation of nearly turbulent Rayleigh–Bénard flow through vorticity or local circulation measurements: A computational study, *Journal of Scientific Computing*, **77** (2018), 1519-1533.
[37] A. Farhat, M. S. Jolly and E. S. Titi, Continuous data assimilation for the 2D Bénard convection through velocity measurements alone, *Phys. D*, **303** (2015), 59-66.
[38] A. Farhat, A. Larios, V. R. Martinez and J. P. Whitehead, Identifying the body force from partial observations of a 2D incompressible velocity field, submitted, (2023), arXiv:2302.04701.
[39] A. Farhat, E. Lunasin and 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.*, **18** (2016), 1-23.
[40] A. Farhat, E. Lunasin and E. S. Titi, Data assimilation algorithm for 3D Bénard convection in porous media employing only temperature measurements, *J. Math. Anal. Appl.*, **438** (2016), 492-506.
[41] A. Farhat, E. Lunasin and E. S. Titi, On the Charney conjecture of data assimilation employing temperature measurements alone: The paradigm of 3D planetary geostrophic model, *Mathematics of Climate and Weather Forecasting*, **2** (2016).
[42] A. Farhat, E. Lunasin and E. S. Titi, Continuous data assimilation for a 2D Bénard convection system through horizontal velocity measurements alone, *J. Nonlinear Sci.*, **27** (2017), 1065–1087.

[43] A. Farhat, E. Lunasin and E. S. Titi, A data assimilation algorithm: The paradigm of the 3D Leray-α model of turbulence, *Partial Differential Equations Arising from Physics and Geometry*, **450** (2019), 253-273.
[44] C. Foias, M. S. Jolly, I. G. Kevrekidis and E. S. Titi, Dissipativity of numerical schemes, *Nonlinearity*, **4** (1991), 591-613.
[45] C. Foias, M. S. Jolly, I. G. Kevrekidis and E. S. Titi, On some dissipative fully discrete nonlinear Galerkin schemes for the Kuramoto–Sivashinsky equation, *Phys. Lett. A*, **186** (1994), 87-96.
[46] C. Foias and I. Kukavica, Determining nodes for the Kuramoto–Sivashinsky equation, *J. Dynam. Differential Equations*, **7** (1995), 365-373.
[47] C. Foias, C. F. Mondaini and E. S. Titi, A discrete data assimilation scheme for the solutions of the two-dimensional Navier–Stokes equations and their statistics, *SIAM J. Appl. Dyn. Syst.*, **15** (2016), 2109-2142.
[48] C. Foias, B. Nicolaenko, G. R. Sell and R. Temam, Variétés inertielles pour l'équation de Kuramoto-Sivashinski, *C. R. Acad. Sci. Paris Sér. I Math.*, **301** (1985), 285-288.
[49] C. Foias, B. Nicolaenko, G. R. Sell and R. Temam, Inertial manifolds for the Kuramoto–Sivashinsky equation and an estimate of their lowest dimension, *J. Math. Pures Appl.*, **67** (1988), 197-226.
[50] C. Foias and E. S. Titi, Determining nodes, finite difference schemes and inertial manifolds, *Nonlinearity*, **4** (1991), 135-153.
[51] T. Franz, A. Larios and C. Victor, The bleeps, the sweeps, and the creeps: Convergence rates for dynamic observer patterns via data assimilation for the 2D Navier-Stokes equations, *Comput. Methods Appl. Mech. Engrg.*, **392** (2022), 114673, 19 pp.
[52] B. García-Archilla and J. Novo, Error analysis of fully discrete mixed finite element data assimilation schemes for the Navier–Stokes equations, *Adv. Comput. Math.*, **46** (2020), Paper No. 61, 33 pp.
[53] B. García-Archilla, J. Novo and E. S. Titi, Uniform in time error estimates for a finite element method applied to a downscaling data assimilation algorithm for the Navier–Stokes equations, *SIAM J. Numer. Anal.*, **58** (2020), 410-429.
[54] M. Gardner, A. Larios, L. G. Rebholz, D. Vargun and C. Zerfas, Continuous data assimilation applied to a velocity-vorticity formulation of the 2D Navier–Stokes equations, *Electron. Res. Arch.*, **29** (2021), 2223-2247.
[55] M. Gesho, E. Olson and E. S. Titi, A computational study of a data assimilation algorithm for the two-dimensional Navier–Stokes equations, *Commun. Comput. Phys.*, **19** (2016), 1094-1110.
[56] L. Giacomelli and F. Otto, New bounds for the Kuramoto–Sivashinsky equation, *Comm. Pure Appl. Math.*, **58** (2005), 297-318.
[57] M. Goldman, M. Josien and F. Otto, New bounds for the inhomogenous Burgers and the Kuramoto–Sivashinsky equations, *Comm. Partial Differential Equations*, **40** (2015), 2237-2265.
[58] A. A. Golovin, S. H. Davis, A. A. Nepomnyashchy and M. A. Zaks, Convective Cahn-Hilliard models for kinetically controlled crystal growth, *International Conference on Differential Equations, World Sci. Publ., River Edge, NJ*, **1,2** (2000), 1281-1283.
[59] J. Goodman, Stability of the Kuramoto–Sivashinsky and related systems, *Comm. Pure Appl. Math.*, **47** (1994), 293-306.
[60] M. A. E. R. Hammoud, O. Le Maitre, E. S. Titi, I. Hoteit and O. Knio, Continuous and discrete data assimilation with noisy observations for the Rayleigh–Bénard convection: A computational study, *Comput. Geosci.*, **27** (2023), 63-79.
[61] M. A. E. R. Hammoud, E. S. Titi, I. Hoteit and O. Knio, Cdanet: A physics-informed deep neural network for downscaling fluid flows, *Journal of Advances in Modeling Earth Systems*, **14** (2022), e2022MS003051.
[62] K. Hayden, E. Olson and E. S. Titi, Discrete data assimilation in the Lorenz and 2D Navier–Stokes equations, *Phys. D*, **240** (2022), 1416-1425.
[63] J. E. Hoke and R. A. Anthes, The initialization of numerical models by a dynamic-initialization technique, *Monthly Weather Review*, **104** (1976), 1551-1556.
[64] J. Hudson and M. Jolly, Numerical efficacy study of data assimilation for the 2D magneto-hydrodynamic equations, *J. Comput. Dyn.*, **6** (2019), 131-145.
[65] J. M. Hyman and B. Nicolaenko, The Kuramoto–Sivashinsky equation: A bridge between PDEs and dynamical systems, *Phys. D*, **18** (1986), 113-126.

[66] J. M. Hyman, B. Nicolaenko and S. Zaleski, Order and complexity in the Kuramoto–Sivashinsky model of weakly turbulent interfaces, *Phys. D*, **23** (1986), 265-292.
[67] H. A. Ibdah, C. F. Mondaini and E. S. Titi, Fully discrete numerical schemes of a data assimilation algorithm: Uniform-in-time error estimates, *IMA Journal of Numerical Analysis*, **40** (2020), 2584-2625.
[68] J. S. Il'yashenko, Global analysis of the phase portrait for the Kuramoto–Sivashinsky equation, *J. Dynam. Differential Equations*, **4** (1992), 585-615.
[69] M. Jardak, I. M. Navon and M. Zupanski, Comparison of sequential data assimilation methods for the Kuramoto–Sivashinsky equation, *Internat. J. Numer. Methods Fluids*, **62** (2010), 374-402.
[70] M. S. Jolly, I. G. Kevrekidis and E. S. Titi, Approximate inertial manifolds for the Kuramoto–Sivashinsky equation: Analysis and computations, *Phys. D*, **44** (1990), 38-60.
[71] 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, *Chin. Ann. Math. Ser. B*, **40** (2019), 721-764.
[72] M. S. Jolly, V. R. Martinez and E. S. Titi, A data assimilation algorithm for the subcritical surface quasi-geostrophic equation, *Adv. Nonlinear Stud.*, **17** (2017), 167-192.
[73] M. S. Jolly, R. Rosa and R. Temam, Evaluating the dimension of an inertial manifold for the Kuramoto–Sivashinsky equation, *Adv. Differential Equations*, **5** (2000), 31-66.
[74] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, 2003.
[75] A. Kalogirou, E. E. Keaveny and D. T. Papageorgiou, An in-depth numerical study of the two-dimensional Kuramoto–Sivashinsky equation, *Proc. A.*, **471** (2015), 20140932, 20 pp.
[76] A.-K. Kassam and L. N. Trefethen, Fourth-order time-stepping for stiff PDEs, *SIAM J. Sci. Comput.*, **26** (2005), 1214-1233.
[77] Y. Kuramoto and T. Tsuzuki, Persistent propagation of concentration waves in dissipative media far from equilibrium, *Prog. Theor. Phys.*, **55** (1976), 365-369.
[78] A. Larios and Y. Pei, Nonlinear continuous data assimilation, submitted, (2017), arXiv:1703.03546.
[79] A. Larios and Y. Pei, Approximate continuous data assimilation of the 2D Navier–Stokes equations via the Voigt-regularization with observable data, *Evol. Equ. Control Theory*, **9** (2020), 733-751.
[80] A. Larios, L. G. Rebholz and C. Zerfas, Global in time stability and accuracy of IMEX-FEM data assimilation schemes for Navier–Stokes equations, *Comput. Methods Appl. Mech. Engrg.*, **345** (2019), 1077-1093.
[81] A. Larios and E. S. Titi, Global regularity versus finite-time singularities: Some paradigms on the effect of boundary conditions and certain perturbations, *Recent Progress in the Theory of the Euler and Navier-Stokes Equations, London Math. Soc. Lecture Note Ser., Cambridge University Press, Cambridge*, **430** (2016), 96-125.
[82] A. Larios and C. Victor, Continuous data assimilation for the 3D and higher-dimensional Navier–Stokes equations with higher-order fractional diffusion, submitted, (2023), arXiv:2307.00096.
[83] A. Larios and C. Victor, The second-best way to do sparse-in-time continuous data assimilation: Improving convergence rates for the 2D and 3D Navier–Stokes equations, submitted, (2023), arXiv:2303.03495.
[84] K. Law, A. Stuart and K. Zygalakis, *A Mathematical Introduction to Data Assimilation*, Texts Appl. Math., 62. Springer, Cham, 2015.
[85] P.-L. Lions, B. Perthame and E. Tadmor, A kinetic formulation of multidimensional scalar conservation laws and related equations, *J. Amer. Math. Soc.*, **7** (1994), 169-191.
[86] X. Liu, Gevrey class regularity and approximate inertial manifolds for the Kuramoto–Sivashinsky equation, *Phys. D*, **50** (1991), 135-151.
[87] M. Á. López Marcos, Numerical analysis of pseudospectral methods for theKuramoto–Sivashinsky equation, *IMA J. Numer. Anal.*, **14** (1994), 233-242.
[88] E. Lunasin and E. S. Titi, Finite determining parameters feedback control for distributed nonlinear dissipative systems–a computational study, *Evol. Equ. Control Theory*, **6** (2017), 535-557.
[89] P. A. Markowich, E. S. Titi and S. Trabelsi, Continuous data assimilation for the three-dimensional Brinkman-Forchheimer-extended Darcy model, *Nonlinearity*, **29** (2016), 1292-1328.

[90] V. R. Martinez, Convergence analysis of a viscosity parameter recovery algorithm for the 2D Navier–Stokes equations, *Nonlinearity*, **35** (2022), 2241-2287.
[91] R. Mojgani, A. Chattopadhyay and P. Hassanzadeh, Discovery of interpretable structural model errors by combining Bayesian sparse regression and data assimilation: A chaotic Kuramoto–Sivashinsky test case, *Chaos*, **32** (2022), 061105, 9 pp.
[92] L. Molinet, Local dissipativity in L^2 for the Kuramoto–Sivashinsky equation in spatial dimension 2, *J. Dynam. Differential Equations*, **12** (2000), 533-556.
[93] C. F. Mondaini and E. S. Titi, Uniform-in-time error estimates for the postprocessing Galerkin method applied to a data assimilation algorithm, *SIAM J. Numer. Anal.*, **56** (2018), 78-110.
[94] E. Olson and E. S. Titi, Determining modes for continuous data assimilation in 2D turbulence, *J. Statist. Phys.*, **113** (2003), 799-840.
[95] F. Otto, Optimal bounds on the Kuramoto–Sivashinsky equation, *J. Funct. Anal.*, **257** (2009), 2188-2245.
[96] B. Pachev, J. P. Whitehead and S. A. McQuarrie, Concurrent multiparameter learning demonstrated on the Kuramoto–Sivashinsky equation, *SIAM J. Sci. Comput.*, **44** (2022), A2974-A2990.
[97] Y. Pei, Continuous data assimilation for the 3D primitive equations of the ocean, *Commun. Pure Appl. Anal.*, **18** (2019), 643-661.
[98] S. I. Pokhozhaev, On the blow-up of solutions of the Kuramoto–Sivashinsky equation, *Mat. Sb.*, **199** (2008), 97-106.
[99] L. G. Rebholz and C. Zerfas, Simple and efficient continuous data assimilation of evolution equations via algebraic nudging, *Numer. Methods Partial Differ. Equations*, **37** (2021), 2588-2612.
[100] J. C. Robinson, *Infinite-Dimensional Dynamical Systems*. Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2001.
[101] G. I. Sivashinsky, Nonlinear analysis of hydrodynamic instability in laminar flames. I. Derivation of basic equations, *Acta Astronaut.*, **4** (1977), 1177-1206.
[102] G. I. Sivashinsky, On flame propagation under conditions of stoichiometry, *SIAM J. Appl. Math.*, **39** (1980), 67-82.
[103] E. Tadmor, The well-posedness of the Kuramoto–Sivashinsky equation, *SIAM J. Math. Anal.*, **17** (1986), 884-893.
[104] R. Temam, *Infinite-Dimensional Dynamical Systems In Mechanics and Physics*, Appl. Math. Sci., 68. Springer-Verlag, New York, 1997.
[105] C. Zerfas, L. G. Rebholz, M. Schneier and T. Iliescu, Continuous data assimilation reduced order models of fluid flow, *Comput. Methods Appl. Mech. Engrg.*, **357** (2019), 112596, 18 pp.

Received April 2023; 1st revision July 2023; 2nd revision August 2023; early access September 2023.