

Data assimilation of time under-sampled measurements using observers, the wave-like equation example

Nicolae Cindea, Alexandre Imperiale, Philippe Moireau

▶ **To cite this version:**

Nicolae Cindea, Alexandre Imperiale, Philippe Moireau. Data assimilation of time under-sampled measurements using observers, the wave-like equation example. ESAIM: Control, Optimisation and Calculus of Variations, 2015, pp.35. 10.1051/cocv/2014042 . hal-01054551

HAL Id: hal-01054551

https://inria.hal.science/hal-01054551

Submitted on 7 Aug 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ESAIM: Control, Optimisation and Calculus of Variations Will be set by the publisher
URL: http://www.emath.fr/cocv/

DATA ASSIMILATION OF TIME UNDER-SAMPLED MEASUREMENTS USING OBSERVERS, THE WAVE-LIKE EQUATION EXAMPLE

Nicolae Cîndea[1], Alexandre Imperiale[2] and Philippe Moireau[2]

Abstract. We propose a sequential data assimilation scheme using Luenberger type observers when only some space restricted time under-sampled measurements are available. More precisely, we consider a wave-like equation for which we assume known the restriction of the solution to an open non-empty subset of the spatial domain and for some time samples (typically the sampling step in time is much larger than the time discretization step). To assimilate the available data, two strategies are proposed and analyzed. The first strategy consists in assimilating data only if they are available and the second one in assimilating interpolation of the available data at all the discretization times. In order to tackle the spurious high frequencies which appear when we discretize the wave equation, for both strategies, we introduce a numerical viscous term. In this case, we prove some error estimates between the exact solution and our observers. Numerical simulations illustrate the theoretical results in the case of the one dimensional wave equation.

1991 Mathematics Subject Classification. 35L05; 35L20; 65D05; 65M12; 93B07; 93B70; 93D15.

The dates will be set by the publisher.

Contents

1. Introduction 2
2. Discrete-time observer design 3
2.1. Nudging for wave-like systems 3
2.2. Discrete-time observer 5
3. Convergence estimate for the estimation error 7
3.1. Convergence estimate for the on/off switch 7
3.2. Convergence estimate using interpolated observations 11
4. Numerical illustrations 13
4.1. Spectral analysis 15
4.2. Assessing robustness from numerical examples 22
5. Conclusion 30
Appendix A. Exponential stability result for the on/off switch. 30
References 35

Keywords and phrases: Data assimilation; time under-sampled measurements; Luenberger observers; numerical analysis; interpolation

[1] Laboratoire de Mathématiques, Université Blaise Pascal (Clermont-Ferrand 2), Campus des Cézeaux - BP 80026, 63171 Aubière cedex FRANCE. Email : Nicolae.Cindea@math.univ-bpclermont.fr

[2] Project-Team MΞDISIM, Inria Saclay-Île de France, France

© EDP Sciences, SMAI 1999

1. Introduction

The discipline of data assimilation historically appeared in the context of meteorology and oceanography – see for example the surveys [5,28] – but has also reached new fields of research for instance in life sciences [9]. The two main ingredients in a data assimilation formulation are the model and the data. On the one hand, the model regroups physical information on the system under consideration encapsulated in a mathematical dynamical system – mostly based on (nonlinear) evolution partial differential equations in the data assimilation community – and allows after discretization to simulate various configurations of interest. However, the model contains a various range of potential errors, for example: on the operator driving the dynamical system, on the boundary conditions or on the initial conditions. On the other hand, the data bring complementary and valuable information on the studied system but they are often partial – in space and time – and are likely to be corrupted by the noise inherent to any measurement process. Data assimilation aims at providing a reconstruction of the (observed) real trajectory by coupling the information contained in the model and in the data, hence filtering their respective uncertainties.

To reach this goal several strategies have been considered which could be classified into two main families. The two approaches can be considered in a stochastic or deterministic formalism but we will focus here on the deterministic one. On the one hand, the variational approach popularized by the 4D-Var method [22], consists, in essence, in minimizing a – usually least square based – cost function integrating during a period of time a compromise between (1) some *a priori* on the model and initial conditions and (2) the observation discrepancy between the actual measurements and the synthetic data produced by the model. This cost function is minimized under the constraint of the model dynamics with, in most cases, the help of a descent algorithm involving successive iterations of the model and the dynamics of the so-called adjoint variable. On the other hand, there is the family of filtering methods, where the discrepancy between the simulated system and the data at hand acts in time as a controlled correction of the dynamics of the simulated system to adjust its trajectory to the pursued trajectory. The resulting system is often called an observer of the exact trajectory in the deterministic context – more often referred to as an estimator in the stochastic context. The evolution satisfied by this observer is written in a general form of a dynamical system with a feedback law based on the discrepancy between the model and the data. The most popular observer is the Kalman filter formulated by equivalence with an optimal criterion minimization for finite dimensional systems or infinite dimensional systems [4, 20], hence applicable to any system. However it ultimately leads to operators which are after discretization numerically intractable. As an alternative for this "curse of dimensionality", numerous strategies have been proposed for example the Ensemble Kalman Filter [15] or Reduced Order Kalman-like filters [34]. In the specific context of the conservative wave-like equation, however, several works [7,23,27,31] have rather proposed simplified but effective feedback laws directly based on the physical properties of the system at hand which stabilize at a certain rate – potentially sub-optimal – any errors. This idea follows the path proposed by Luenberger's work [25] for finite dimensional systems and is popularized for PDEs with the *nudging* appellation as initiated in [1,19] – a complete historical perspective can be found in [21].

In general, these observers are defined in an abstract continuous-time framework assuming that the model and the data are available at any time. This clearly represents the asymptotic of any real configuration where in general the data are time-sampled. Eventually the model should be considered with its time-discretization which has no reason to be dependent of the time-sampling of the data. Therefore, we must analyze the impact of the data discretization in the observer definition. In particular, we are concerned by coarse data in time with respect to the model discretization and we will specifically focus in this work on the specific but fundamental case where the data have a constant time-sampling which is much larger than the model time-step. As an illustration, we can cite the case of image sequences assimilation for cardiovascular systems – described *e.g.* in [6] – where the time-sampling of the data is of one or two orders of magnitude larger than the model time-step discretization. Facing this situation, we can think of two alternatives. The first one – considered for example in [27] – consists in interpolating the data in time in order to regenerate a time-continuous sequence which can then be compatible with any time discretization of the model. This approach is very attractive from an abstract standpoint, and, therefore, can be directly analyzed in the light of the literature about the time-discretization of time-continuous

observer – typically in our case following the works [2,12,13,16,30]. However it is at the price of an additional time-interpolation error perturbing the observer dynamics as any other measurement noise. This perturbation has the consistency of the data sampling period. Moreover, the resulting observer is non-causal and, therefore, can not integrate the data in real-time. A second approach – often used in practice without even mentioning it – is to compute the discrepancy only when the data are available. This intermittent correction is potentially error free but may induce correction shocks which limit the stabilization of the error between the observer trajectory and the pursued trajectory. We mention that the stability of some dissipative partial differential evolution equations with intermittent damping was recently studied in [17].

In this paper, we address the issue of analyzing a data assimilation procedure where an intermittent feedback law is defined and compared to a procedure where a time interpolation of the data is considered. The comparison is carried out both theoretically – since convergence estimates are provided in both cases – and numerically – using a simple one-dimensional wave equation model. Indeed, we restrict our analysis to the specific case of a wave-like system and the Luenberger associated observer proposed in [27]. However, this work is intended to illustrate how the data time-sampling influences the definition of any data assimilation sequential strategy.

The outline of the paper is as follows. In Section 2 we introduce the observer methodology in the case of wave-like systems and we propose two types of time discretization presented in a general form. In Section 3 we provide convergence estimates of both observers. Section 4 and Section 4.2 are then devoted to numerical illustrations where (1) we analyze the spectra of the stabilized operators appearing in the dynamical systems satisfied by the estimation error in order to provide an optimal gain for both observers and (2) we provide time simulations which illustrate the robustness of the two strategies in mainly two different contexts: a first one with low data noise but high sampling period, and a second one with reasonable sampling period but noisy data.

2. Discrete-time observer design

2.1. Nudging for wave-like systems

We consider in this work a general class of second order hyperbolic systems in bounded domain characteristic of wave equations or elasticity systems. These models typically correspond to simplified situations of those encountered in the cardiac modeling context, where a heart mechanical model is registered on coarse data obtained from a sequence of few medical images [6]. Formally, we introduce a Hilbert space $\mathcal{H}$ endowed with the inner product $(\cdot,\cdot)$ and we denote by $\|\cdot\|$ the associated norm. Then, we define a self adjoint operator $A_0 : \mathcal{D}(A_0) \to \mathcal{H}$, positive-definite with compact resolvent and we consider the general class of systems

$$\begin{cases} \ddot{w}(t) + A_0 w(t) = 0, \\ \\ w(0) = w_0 + \zeta_0, \quad \dot{w}(0) = w_1 + \zeta_1, \end{cases} \tag{1}$$

where $\dot{x}$ denotes the time derivative of any variable x, (ζ_0, ζ_1) represent some potential errors on the initial conditions and (w_0, w_1) are some known *a priori*. We point out that (1) represents a conservative system and, therefore, any errors on the initial conditions are conserved in time. Denoting by

$$x(t) = \begin{pmatrix} w(t) \\ \dot{w}(t) \end{pmatrix} \in \mathcal{X} = \mathcal{D}(A_0^{\frac{1}{2}}) \times \mathcal{H},$$

we can rewrite (1) as a first-order system

$$\begin{cases} \dot{x}(t) = Ax(t), & t > 0 \\ \\ x(0) = x_0 + \zeta, \end{cases} \tag{2}$$

where $x_0 = \begin{pmatrix} w_0 & w_1 \end{pmatrix}^\intercal$, $\zeta = \begin{pmatrix} \zeta_0 & \zeta_1 \end{pmatrix}^\intercal$ and $A : \mathcal{D}(A) \to \mathcal{X}$ is defined by

$$\mathcal{D}(A) = \mathcal{D}(A_0) \times \mathcal{D}(A_0^{\frac{1}{2}}), \qquad A = \begin{pmatrix} 0 & \mathbb{1} \\ -A_0 & 0 \end{pmatrix}.$$

On this system we consider a particular target system $x_\bullet$ associated with a specific unknown $\zeta_\bullet$. Moreover, we assume that some measurements z – also called observations – are available and defined by

$$z(t) = Hx_\bullet(t), \qquad t > 0, \tag{3}$$

where $H \in \mathcal{L}(\mathcal{X}, \mathcal{Z})$ is the so-called *observation operator* and $\mathcal{Z}$ is the Hilbert space associated with the observations endowed with its norm $\|\cdot\|_{\mathcal{Z}}$. For the sake of simplicity, we restrict ourselves to bounded observation operators but a more general class of admissible observation operators can also be considered [35]. The observation operator can be applied to any solution of (2) and we assume that the pair (A, H) is exactly observable in a time $T_{\text{obs}} > 0$, *i.e.* there exists a constant $C_{\text{obs}} > 0$ such that every solution of (2) satisfies

$$\int_0^{T_{\text{obs}}} \|Hx(t)\|_{\mathcal{Z}}^2 \, \mathrm{d}t \geq C_{\text{obs}} \|x(0)\|_{\mathcal{X}}^2, \qquad x(0) \in \mathcal{X}. \tag{4}$$

In order to benefit from the available data $z(t)$ and considering only the available *a priori* x_0 that we have on the initial condition, we consider the Luenberger observer $\widehat{x}(t)$ [7] – see also similar formulations in [13, 31] – estimating $x_\bullet(t)$ from the dynamics

$$\begin{cases} \dot{\widehat{x}}(t) = A\widehat{x}(t) + \gamma H^* \left(z(t) - H\widehat{x}(t) \right), & t > 0 \\ \widehat{x}(0) = x_0, \end{cases} \tag{5}$$

where $H^* \in \mathcal{L}(\mathcal{Z}, \mathcal{X})$ is the adjoint of the observation operator and $\gamma > 0$ is a gain parameter. We justify the use of this observer by noticing that the estimation error $\widetilde{x}(t) = x_\bullet(t) - \widehat{x}(t)$ satisfies the – damped – dynamics

$$\begin{cases} \dot{\widetilde{x}}(t) = \left(A - \gamma H^* H \right) \widetilde{x}(t), & t > 0 \\ \widetilde{x}(0) = \zeta_\bullet. \end{cases} \tag{6}$$

Provided that H satisfies (4) it is well known – see *e.g.* [24] – that the error is exponentially stable, namely there exist two constants $M > 0$ and $\mu > 0$ such that

$$\|\widetilde{x}(t)\|_{\mathcal{X}} \leq M \exp(-\mu t) \|\widetilde{x}(0)\|_{\mathcal{X}}, \qquad t > 0. \tag{7}$$

Remark 2.1. In the filtering strategy described by (5) we see that the initial dynamics (2) is modified by a feedback law where the so-called *gain* operator is, here, simply given by $G = \gamma H^*$. In comparison, in the context of Kalman filtering [4, 32], the gain is given by $G = P(t)H^*$, where $P(t) \in \mathcal{L}(\mathcal{X}, \mathcal{X})$ satisfies the Riccati equation

$$\dot{P} = AP + PA^* - PH^*HP, \quad P(0) = P_0,$$

in the mild sense.

Remark 2.2. We may distinguish in the wave-like equation context two classes of observation operators. The one corresponding to velocity observations, namely $H = \begin{pmatrix} 0 & H_0 \end{pmatrix}$, and the one corresponding to the direct observation of the field, *i.e.* $H = \begin{pmatrix} H_0 & 0 \end{pmatrix}$. It should be noted that in the context of poor time resolution of the data the two cases are indeed two independent situations since the data cannot be differentiated with respect to

time without dramatically amplifying the measurement noise. We point out that the exponential convergence of these two classes of observers have been demonstrated. The first one is standard see [10,18,26]. The second one is less classical since the adjoint $H^\star$ induces a modification of the identity between the time-derivation of the field and the velocity. This observer can therefore only be considered as a virtual system and has been demonstrated to converge in [7,8]. Ultimately the exponential convergence of both observer systems requires the geometric control conditions to be satisfied – see [3] for the wave equation and [11] for the elasticity system.

2.2. Discrete-time observer

In the first section, we have introduced an abstract observer which assumes that the data at hand are available at any time. However in practice, they are time-sampled and we want to study the influence of their discretization in the definition of the observer. In this respect, we decide to directly introduce the time-discretization avoiding the technicalities induced by potential Dirac distributions associated with the data time-sampling. However, we should consider an observer time-discretization which conserves at the time-discrete level – and uniformly with respect to the time-discretization – the stability properties of the estimation error. Unfortunately, when discretizing in time systems such as (6) spurious high-frequency modes may arise provoking the loss of the uniform decay rate of the error – see *e.g.* [13,36,37]. These modes cannot be captured by the localized observation operator, thus leading to a loss of uniform observability. In other words, a direct discretization of the observer (5) for example using conservative mid-point rule

$$\frac{\widehat{x}^{n+1}-\widehat{x}^{n}}{\Delta \mathrm{t}} = A\Big(\frac{\widehat{x}^{n+1}+\widehat{x}^{n}}{2}\Big) + \gamma H^\star\left(\frac{z^{n+1}+z^{n}}{2} - H\Big(\frac{\widehat{x}^{n+1}+\widehat{x}^{n}}{2}\Big)\right) \tag{8}$$

may not satisfy at the discrete-level a time-discrete counterpart of the observation inequality (4) of the form

$$\sum_{n=0}^{N_{\mathrm{obs}}} \|Hx(n\Delta \mathrm{t})\|_{\mathcal{Z}}^2 \geq C_{\mathrm{obs}}\left(\|w_0\|^2_{\mathcal{D}(A_0^{\frac{1}{2}})} + \|w_1\|^2\right). \tag{9}$$

Note that in (8) we still avoid to consider the time sampling of the data and denote by

$$z^n = Hx_\bullet(n\Delta \mathrm{t}), \tag{10}$$

a discrete (in time) observation potentially available at any model time-step.

To circumvent the loss of observability at a time-discrete level, authors in [13] propose two main options. The first possibility is to impose a (restrictive) CFL condition that discards these undesired high frequency modes. The second option consists in adding an artificial viscous term consistent with the order of the numerical scheme and responsible for the dissipation of the spurious modes – see also [30,33] for a similar proposition. Note that it is also possible to define specific compatible spaces and discretizations – with, for example a mixed formulation for the space discretization – that offer uniform stabilization properties, see, for example [2] or [14] and reference therein. However, these discretizations are less usual and thus more intricate to use in practical applications. Among the various possibilities detailed in [13], we retain for instance the discretization

$$\begin{cases} \dfrac{\widehat{x}_-^{n+1}-\widehat{x}_+^{n}}{\Delta \mathrm{t}} = A\Big(\dfrac{\widehat{x}_-^{n+1}+\widehat{x}_+^{n}}{2}\Big), \qquad n>0 \\[2ex] \dfrac{\widehat{x}_+^{n+1}-\widehat{x}_-^{n+1}}{\Delta \mathrm{t}} = \gamma_{n+1} H^\star\Big(z^{n+1} - H\widehat{x}_+^{n+1}\Big) + \nu_{\Delta \mathrm{t}} A^2 \widehat{x}_+^{n+1}, \qquad n>0 \\[2ex] \widehat{x}_+^{0} = \begin{pmatrix} w_0 \\ w_1 \end{pmatrix}, \end{cases} \tag{11}$$

where some numerical viscosity is introduced with the specific viscous operator (suggested by [13])

$$A^2 = \begin{pmatrix} -A_0 & 0 \\ 0 & -A_0 \end{pmatrix}, \tag{12}$$

and $\nu_{\Delta t}$ controls the amount of this numerical viscosity. In order to respect the order of consistency of this time scheme – and as advised in [13] – we ought to set $\nu_{\Delta t} = O(\Delta t^2)$. This time discretization can be understood as a prediction-correction scheme where the dynamics of the model leading to $\widehat{x}_-^n$ is then corrected by computing a model-data interaction to produce $\widehat{x}_+^n$.

Once the time discretization of the observer is chosen, we can go back to our considerations on the data time-sampling. Here, we should first point out that the observation inequality (9) can be expected to be obtained from its continuous counterpart (4) only with a time discretization step small enough so that there are several time steps included in the characteristic time associated with the smallest frequency of the system. Otherwise, we could imagine a degenerate situation where the measurements are considered at the exact same frequency than a system mode making it unobservable, see also [17] about the observability of intermittent stabilized systems. This condition on the data sampling will be assumed to be satisfied in the rest of the article. We define $(j_r)_{r\in\mathbb{N}} \subset \mathbb{N}$ as a strictly increasing sequence of natural numbers so that the available measurements are

$$z^r = z(j_r \Delta t), \qquad r \in \mathbb{N}. \tag{13}$$

We then consider two strategies to introduce these time-sampled data in (11). The first one is to consider the data only when they are available, hence in essence $\gamma_n = 0$ when the data z^n is not available. The second one consists in interpolating the data to generate an approximated z^n for all n. We summarize these two strategies by

$$\begin{cases} \dfrac{\widehat{x}_-^{n+1} - \widehat{x}_+^n}{\Delta t} = A\Big(\dfrac{\widehat{x}_-^{n+1} + \widehat{x}_+^n}{2}\Big), \qquad n > 0 \\ \dfrac{\widehat{x}_+^{n+1} - \widehat{x}_-^{n+1}}{\Delta t} = \delta^{n+1}\gamma H^* \left(d^{n+1} - H\widehat{x}_+^{n+1}\right) + \nu_{\Delta t} A^2 \widehat{x}_+^{n+1}, \qquad n > 0 \\ \widehat{x}_+^0 = \begin{pmatrix} w_0 \\ w_1 \end{pmatrix} = \widehat{x}_0, \end{cases} \tag{14}$$

where $(\delta^n)_{n\in\mathbb{N}}$ and $(d^n)_{n\in\mathbb{N}}$ will be referred to as the switching coefficients and the interpolated data respectively. For the first idea – named *on/off switch* and where the correction term only appears when measurements are available – we have

$$\delta^n = \begin{cases} 1 \\ 0 \end{cases} \qquad d^n = \begin{cases} z^r & \text{if } \exists r \in \mathbb{N} \ : \ n = j_r \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

The second choice where we interpolate in time the data reads, in the particular case of linear interpolation,

$$\delta^n = 1 \qquad \forall n, \qquad d^n = \frac{n - j_r}{j_{r+1} - j_r} z^{r+1} + \left(1 - \frac{n - j_r}{j_{r+1} - j_r}\right) z^r \qquad j_r \le n \le j_{r+1}. \tag{16}$$

Then in the next sections, we propose to study the convergence of the time-discrete observer given by (14)-(15) – see Theorem 3.2 – and given by (14)-(16) – see Theorem 3.6.

Remark 2.3. There is also an exact counterpart at the time-discrete level of the popular Kalman observer mentioned in Remark 2.1. The resulting time-discrete observer for the system (2)-(3) – called Kalman-Bucy

estimator [20] – is also based on a prediction-correction paradigm reading

$$\text{(Prediction)} \quad \begin{cases} \widehat{x}_-^{n+1} = A_{n+1|n}\widehat{x}_+^n, \\ P_-^{n+1} = A_{n+1|n}P_+^n A_{n+1|n}^*, \end{cases} \tag{17}$$

$$\text{(Correction)} \quad \begin{cases} \widehat{x}_+^{n+1} = \widehat{x}_-^{n+1} + \Delta t P_+^{n+1} H^* \big(z^{n+1} - H\widehat{x}_-^{n+1}\big), \\ P_+^{n+1} = ((P_-^{n+1})^{-1} + H^*H)^{-1}, \end{cases} \tag{18}$$

where we denoted by $A_{n+1|n}$ the state transition operator corresponding to a stable and consistent discretization of the dynamical system (2) for instance in the case of a mid-point discretization

$$A_{n+1|n} = \big(\mathbb{1} - \frac{\Delta t}{2}A\big)^{-1}\big(\mathbb{1} + \frac{\Delta t}{2}A\big).$$

One can prove – see *e.g.* [29] – that the time-discrete observer derives from seeking the minimum value of the functional

$$J(\widetilde{x}_0) = \frac{1}{2}\|\widetilde{x}_0\|^2_{P_0^{-1}} + \frac{1}{2}\sum_{k=1}^{n}\|z^k - Hx^k\|^2_{\mathcal{Z}}\Delta t, \tag{19}$$

with x^k subject to $x^k = A_{k|k-1}x^{k-1}$ and $x^0 = x_0 + \widetilde{x}_0$. Hence, an on/off version of the Kalman-Bucy observer corresponds to the minimum of the adequately adjusted functional

$$J(\widetilde{x}_0) = \frac{1}{2}\|\widetilde{x}_0\|^2_{P_0^{-1}} + \frac{1}{2}\sum_{r=1}^{c_n}\|z^r - Hx^{j_r}\|^2_{\mathcal{Z}}(j_{r+1} - j_r)\Delta t, \tag{20}$$

with

$$c_n = \text{card}\{j,\ 1 \le j \le n,\ \delta_j = 1\}. \tag{21}$$

We expect that the proof that we will present for the nudging observer can be directly adapted to the Kalman approach, hence justifying in the general context of sequential data assimilation methods the choice of interpolating or intermittently filtering the under-sampled data or continuously filtering a reconstructed data by interpolation.

3. Convergence estimate for the estimation error

3.1. Convergence estimate for the on/off switch

Let us define the corrected estimation error by

$$\widetilde{x}_+^n = x_\bullet(n\Delta t) - \widehat{x}_+^n, \tag{22}$$

and the corresponding predicted estimation error by

$$\widetilde{x}_-^n = x_\bullet(n\Delta t) - \widehat{x}_-^n, \tag{23}$$

where $x_\bullet$ is the exact solution of (2) and $\widehat{x}_+^n$ and $\widehat{x}_-^n$ satisfy (14)-(15). We start by giving the dynamical system satisfied by this estimation error in the following proposition.

Proposition 3.1. *Assuming that $x_0 \in \mathcal{D}(A_0^2) \times \mathcal{D}(A_0^{\frac{3}{2}})$, the estimation errors defined by (22) and (23) satisfy the following discrete dynamical system*

$$\begin{cases} \dfrac{\widetilde{x}_-^{n+1} - \widetilde{x}_+^n}{\Delta \mathrm{t}} = A\Big(\dfrac{\widetilde{x}_-^{n+1} + \widetilde{x}_+^n}{2}\Big) + \varepsilon^{n+1}, \\ \dfrac{\widetilde{x}_+^{n+1} - \widetilde{x}_-^{n+1}}{\Delta \mathrm{t}} = -\delta^{n+1}\gamma H^* H \widetilde{x}_+^{n+1} + \nu_{\Delta \mathrm{t}} A^2 \widetilde{x}_+^{n+1} + \varepsilon_\nu^{n+1}, \\ \widetilde{x}_+^0 = x_0 - \widehat{x}_0, \end{cases} \tag{24}$$

where the consistency terms are

$$\begin{cases} \varepsilon^{n+1} = \dfrac{\Delta \mathrm{t}^2}{2} A^3 \Big(\dfrac{1}{3} x_\bullet(\mathrm{t}_n) - \dfrac{1}{2} x_\bullet(\mathrm{r}_n)\Big), \qquad \text{with } \mathrm{t}_n, \mathrm{r}_n \in [n\Delta \mathrm{t}; (n+1)\Delta \mathrm{t}], \\ \varepsilon_\nu^{n+1} = -\nu_{\Delta \mathrm{t}} A^2 x_\bullet((n+1)\Delta \mathrm{t}). \end{cases} \tag{25}$$

Proof. Starting from the definition of the predicted estimation error and from the correction phase of the observer (14) we obtain

$$\begin{aligned} \widetilde{x}_-^{n+1} &= x_\bullet((n+1)\Delta \mathrm{t}) - \widehat{x}_+^{n+1} + \delta^{n+1}\Delta \mathrm{t}\gamma H^*(d^{n+1} - H\widehat{x}_+^{n+1}) + \nu_{\Delta \mathrm{t}}\Delta \mathrm{t} A^2 \widehat{x}_+^{n+1} \\ &= (\mathbb{1} + \delta^{n+1}\Delta \mathrm{t}\gamma H^* H)\widetilde{x}_+^{n+1} + \nu_{\Delta \mathrm{t}}\Delta t A^2 \widetilde{x}_+^{n+1} \\ &= (\mathbb{1} + \delta^{n+1}\Delta \mathrm{t}\gamma H^* H - \nu_{\Delta \mathrm{t}}\Delta \mathrm{t} A^2)\widetilde{x}_+^{n+1} - \Delta \mathrm{t}\varepsilon_\nu^{n+1}, \end{aligned}$$

which is exactly the second equation in (24). Secondly, to obtain the remaining equation it suffices to notice that

$$\frac{\widetilde{x}_-^{n+1} - \widetilde{x}_+^n}{\Delta \mathrm{t}} = \frac{x_\bullet((n+1)\Delta \mathrm{t}) - x_\bullet(n\Delta \mathrm{t})}{\Delta \mathrm{t}} - \frac{\widehat{x}_-^{n+1} - \widehat{x}_+^n}{\Delta \mathrm{t}}.$$

Hence, using a first-order finite difference approximation of $\dot{x}_\bullet((n+1)\Delta \mathrm{t})$, from the above equality – assuming enough regularity on the initial condition – we can assure that there exist a time t_n and a time $\mathrm{r}_n \in [n\Delta \mathrm{t}; (n+1)\Delta \mathrm{t}]$ such that

$$\frac{\widetilde{x}_-^{n+1} - \widetilde{x}_+^n}{\Delta \mathrm{t}} = \frac{\dot{x}_\bullet((n+1)\Delta \mathrm{t}) + \dot{x}_\bullet(n\Delta \mathrm{t})}{2} - \frac{\widehat{x}_-^{n+1} - \widehat{x}_+^n}{\Delta \mathrm{t}} + \frac{\Delta \mathrm{t}^2}{2}\Big(\frac{1}{3}\dddot{x}_\bullet(\mathrm{t}_n) - \frac{1}{2}\dddot{x}(\mathrm{r}_n)\Big).$$

Therefore, from (2) and the first equation of (14) we conclude the proof. □

From this first result we can now give the convergence estimate for the on/off observer in the specific but fundamental case where the data have a constant time-sampling which is much larger than the model time-step.

Theorem 3.2. *Let A be a skew-adjoint operator with compact resolvent and $H \in \mathcal{L}(\mathcal{X}, \mathcal{Z})$ be a bounded linear observation operator such that the observability inequality (4) holds. Assume that there exists a strictly positive integer N such that*

$$\sout{c_n \geq \rho n, \qquad n > 0.} \mapsto j_r = rN\Delta \mathrm{t}, \qquad n \in \mathbb{N}, \tag{26}$$

where $(j_r)_{r\in\mathbb{N}}$ is the sequence appearing in (13). Then, for every $x_0 \in \mathcal{D}(A_0^2) \times \mathcal{D}(A_0^{\frac{3}{2}})$ there exist positive constants M_0, $\mu_0(N)$, C_1 and C_2, independent of $\Delta \mathrm{t} \in (0,1)$ and n, such that $\widetilde{x}_+^n$ solution of (24) satisfies

$$\|\widetilde{x}_+^n\|_{\mathcal{X}} \leq M_0 \exp(-\mu_0 c_n \Delta \mathrm{t})\|\widetilde{x}_0\|_{\mathcal{X}} + \frac{\Delta \mathrm{t}}{1 - \exp(-\mu_0 \rho \Delta \mathrm{t})}\big(\Delta \mathrm{t}^2 C_1 + \nu_{\Delta \mathrm{t}} C_2\big), \tag{27}$$

where c_n is given by (21)*, and, using* (26) *is equal to* $\left\lfloor \frac{n}{N} \right\rfloor$*, and* $\rho = \frac{c_n}{n}$.

Proof. From the system (24) we can explicitly compute $\widetilde{x}_+^{n+1}$ function of $\widetilde{x}_+^n$, ε^{n+1} and ε_ν^{n+1} by the following relation

$$\widetilde{x}_+^{n+1} = \mathcal{P}_{n+1}\mathcal{Q}\,\widetilde{x}_+^n + \Delta t\big(\mathcal{P}_{n+1}\varepsilon^{n+1} + \mathcal{R}_{n+1}\varepsilon_\nu^{n+1}\big), \tag{28}$$

where $\mathcal{P}_n,\ \mathcal{Q},\ \mathcal{R}_n \in \mathcal{L}(\mathcal{X})$ are given by

$$\begin{cases} \mathcal{P}_n = \left(\mathbb{1} + \delta^n\gamma\Delta t H^*H - \Delta t\nu_{\Delta t}A^2\right)^{-1}\left(\mathbb{1} - \dfrac{\Delta t}{2}A\right)^{-1}, \\ \mathcal{Q} = \mathbb{1} + \dfrac{\Delta t}{2}A, \\ \mathcal{R}_n = \left(\mathbb{1} + \delta^n\gamma\Delta t H^*H - \Delta t\nu_{\Delta t}A^2\right)^{-1}. \end{cases}$$

Note that since both semigroups generated by operators $(\nu_{\Delta t}A^2 - \gamma H^*H)$ and A respectively are semigroups of contraction we have that $\mathcal{P}_n$ and $\mathcal{R}_n$ are well defined and $\|\mathcal{P}_n\|_{\mathcal{L}(\mathcal{X})} \le 1$ and $\|\mathcal{R}_n\|_{\mathcal{L}(\mathcal{X})} \le 1$. Hence, using this notation we can write (24) as follows

$$\widetilde{x}_+^n = \left(\prod_{j=1}^{n}\mathcal{P}_j\mathcal{Q}\right)\widetilde{x}_+^0 + \Delta t\sum_{i=0}^{n-1}\left(\prod_{j=1}^{i}\mathcal{P}_j\mathcal{Q}\right)\big(\mathcal{P}_{n-i}\varepsilon^{n-i} + \mathcal{R}_{n-i}\varepsilon_\nu^{n-i}\big). \tag{29}$$

Remark that if $\delta^n = 0$ then $\mathcal{P}_n\mathcal{Q}$ is the operator driving the system

$$\begin{cases} \dfrac{\widetilde{x}_-^{n+1} - \widetilde{x}_+^n}{\Delta t} = A\Big(\dfrac{\widetilde{x}_-^{n+1} + \widetilde{x}_+^n}{2}\Big), \\ \dfrac{\widetilde{x}_+^{n+1} - \widetilde{x}_-^{n+1}}{\Delta t} = \nu_{\Delta t}A^2\widetilde{x}_+^{n+1}, \\ \widetilde{x}_+^0 = x_0 - \widehat{x}_0, \end{cases} \tag{30}$$

The product of non-commutative operators, involving or not the feedback, interferes with a direct use of the exponential results obtained in [13]. However, the proof made in [13], where observations are available at each time step can be directly adapted – using also [12] – with a similar decomposition between low and high frequencies. The low frequencies can still be controlled by our intermittent feedback as soon as the frequency cut-off is chosen with respect to the maximum data sampling time-step. Then, the resulting high frequencies are handled by the numerical viscosity which acts at every time-step Δt. We refer to Appendix A for a complete proof of the existence of the two positive constants M_0 and μ_0, with μ_0 depending on N, such that

$$\left\|\prod_{j=1}^{n}(\mathcal{P}_j\mathcal{Q})\right\|_{\mathcal{L}(\mathcal{X})} \le M_0\exp(-\mu_0 c_n\Delta t), \qquad n \in \mathbb{N}. \tag{31}$$

Combining (29), (31) and the fact that $\|\mathcal{P}_n\|_{\mathcal{L}(\mathcal{X})} \le 1$ and $\|\mathcal{R}_n\|_{\mathcal{L}(\mathcal{X})} \le 1$, we obtain the following estimate

$$\begin{aligned} \|\widetilde{x}_+^n\|_{\mathcal{X}} &\le M_0\exp(-\mu_0 c_n\Delta t)\|\widetilde{x}_+^0\|_{\mathcal{X}} + \Delta t\sum_{i=0}^{n-1}\exp(-\mu_0 c_i\Delta t)\big(\|\varepsilon^{n-i}\|_{\mathcal{X}} + \|\varepsilon_\nu^{n-i}\|_{\mathcal{X}}\big) \\ &\le M_0\exp(-\mu_0 c_n\Delta t)\|\widetilde{x}_+^0\|_{\mathcal{X}} + \frac{\Delta t}{1-\exp(-\mu_0\rho\Delta t)}\max_{1\le i\le n}\big(\|\varepsilon^i\|_{\mathcal{X}} + \|\varepsilon_\nu^i\|_{\mathcal{X}}\big). \end{aligned}$$

Finally, from (25) combined with the conservation of the energy associated with (2), we obtain the final estimate (27) of Theorem 3.2 with

$$C_1 = \frac{5}{12}\|A^3 x_0\|_{\mathcal{X}} \quad \text{and} \quad C_2 = \|A^2 x_0\|_{\mathcal{X}}. \tag{32}$$

□

The estimate (27) gives the convergence (in the energy norm) of the time semi-discrete observer (14) to the solution of the continuous system (2) when $n \to \infty$ and $\Delta t \to 0$. Moreover, this estimate provides an explicit dependence between the error and the ratio ρ associated with the sampling frequency. Considering (26) to be satisfied the data time-sampling is a constant equal to $\Delta T = N\Delta t$, and

$$\rho \sim \frac{\Delta t}{\Delta T}.$$

Therefore if ΔT is large with respect to Δt, then ρ is close to zero , the first term of the estimate (27) right-hand side c_n tends to 0, and the second term is $O(\Delta T(\Delta t^2 + \nu_{\Delta t}))$. Therefore the overall asymptotic estimate reads

$$\|\widetilde{x}_+^n\|_{\mathcal{X}} = O\big(M_0 + \Delta T(\Delta t + \frac{\nu_{\Delta t}}{\Delta t})\big).$$

From this asymptotic estimate, we see in the estimate (27) that the numerical viscosity coefficient appears as a consistency term, which is consistent with the fact that this artificial term is in fact a perturbation of the standard and consistent discretization of (2). For this reason we clearly understand why the viscosity coefficient should be kept $\nu_{\Delta t} = O(\Delta t^2)$. Furthermore, the term M_0 indicates that the initial condition error can not be stabilized. Moreover even if $M_0 = 0$ we see that the data time-sampling ΔT control the error estimate. If ΔT is typically of the same order of magnitude as the global simulation time T, we retrieve the standard numerical error estimate which deteriorates with the simulation time. We point out that when $\rho = 1$ we retrieve the estimate in [8] which first exhibit the gain offered by data assimilation strategies in numerical analysis estimates. More precisely, in this particular case, authors in [8] have shown that the numerical error between the exact solution and the numerical solution is bounded and independent of the total simulation time. The input of data in the observer dynamics balances the accumulation of numerical errors. When $\rho \neq 1$, this remark is still valid but the time-sampling of the data also governed the estimation. We provide in Section 4.2 a numerical illustration of this phenomenon.

The result obtained in Theorem 3.2 can be extended to take into account potential data noise. For instance, let us consider the case of additive noise so that (10) becomes

$$z^n = Hx_\bullet(n\Delta t) + \chi(n\Delta t), \tag{33}$$

where χ is a time dependent function belonging to $\mathcal{Z}$. In the case of the intermittent observer, the second equation of System (24) reads

$$\frac{\widetilde{x}_+^{n+1} - \widetilde{x}_-^{n+1}}{\Delta t} = -\delta^{n+1}\gamma H^* H\widetilde{x}_+^{n+1} + \nu_{\Delta t} A^2 \widetilde{x}_+^{n+1} + \varepsilon_\nu^{n+1} + \delta^{n+1}\gamma H^* \chi((n+1)\Delta t). \tag{34}$$

The data noise directly enters as a source term in the estimation error dynamical system, hence the demonstration proposed in Theorem 3.2 directly applies, entailing the following estimate

$$\|\widetilde{x}_+^n\|_{\mathcal{X}} \leq M_0 \exp(-\mu_0 c_n \Delta t)\|\widetilde{x}_0\|_{\mathcal{X}} + \frac{\Delta t}{1 - \exp(-\mu_0 \rho \Delta t)}\big(\Delta t^2 C_1 + \nu_{\Delta t} C_2 + \gamma C_3 \max_{1\leq i\leq n | \delta^i = 1} \|\chi(i\Delta t)\|_{\mathcal{Z}}\big), \tag{35}$$

where C_1 and C_2 are as in Theorem 3.2 and C_3 is given by

$$C_3 = \|H^*\|_{\mathcal{L}(\mathcal{Z},\mathcal{X})}. \tag{36}$$

Remark 3.3. Theorem 3.2 can be extended to obtain error estimates for a fully discrete observer, combining (14) with Galerkin method. The idea is to adapt the proof of Theorem 3.2 using the method described in [8].

Remark 3.4. We believe that Theorem 3.2 could be extended in order to tackle non-constant time-sampling rates, in particular as soon as the ratio between two measurements and the time discretization step is bounded with respect to time. In any case we should carefully investigate what it is the "good" distribution of the samples such that all the system frequencies are stabilized.

3.2. **Convergence estimate using interpolated observations**

Similarly to what we have proposed for the on/off strategy we analyze the time-discrete observer that assimilates interpolated data – *i.e.* System (14)-(16). For the sake of clarity we will use the same notation for the estimation error, namely

$$\widetilde{x}_+^n = x_\bullet(n\Delta\mathrm{t}) - \widehat{x}_+^n \tag{37}$$

for the corrected estimation error and

$$\widetilde{x}_-^n = x_\bullet(n\Delta\mathrm{t}) - \widehat{x}_-^n \tag{38}$$

for the predicted estimation error.

Proposition 3.5. *Assuming that $x_0 \in \mathcal{D}(A_0^2) \times \mathcal{D}(A_0^{\frac{3}{2}})$ then the estimation error defined by (37) and (38) satisfies the following discrete dynamical system*

$$\begin{cases} \dfrac{\widetilde{x}_-^{n+1} - \widetilde{x}_+^n}{\Delta\mathrm{t}} = A\Big(\dfrac{\widetilde{x}_-^{n+1} + \widetilde{x}_+^n}{2}\Big) + \varepsilon^{n+1}, \\ \dfrac{\widetilde{x}_+^{n+1} - \widetilde{x}_-^{n+1}}{\Delta\mathrm{t}} = -\gamma H^* H \widetilde{x}_+^{n+1} + \nu_{\Delta\mathrm{t}} A^2 \widetilde{x}_+^{n+1} + \varepsilon_\nu^{n+1} + \gamma H^* \varepsilon_d^{n+1}, \\ \widetilde{x}_+^0 = x_0 - \widehat{x}_0, \end{cases} \tag{39}$$

where ε^{n+1} and ε_ν^{n+1} are given by (25) and ε_d is the interpolation error, namely

$$\varepsilon_d^{n+1} = Hx_\bullet((n+1)\Delta\mathrm{t}) - d^{n+1}. \tag{40}$$

Proof. From the definition of the predicted estimation error and from the correction phase of the observer (14) we obtain

$$\begin{aligned} \widetilde{x}_-^{n+1} &= x_\bullet((n+1)\Delta\mathrm{t}) - \widehat{x}_+^{n+1} + \Delta\mathrm{t}\gamma H^*(d^{n+1} - H\widehat{x}_+^{n+1}) + \nu_{\Delta\mathrm{t}}\Delta\mathrm{t}A^2\widehat{x}_+^{n+1} \\ &= (\mathbb{1} + \Delta\mathrm{t}\gamma H^* H)\widetilde{x}_+^{n+1} + \nu_{\Delta\mathrm{t}}\Delta\mathrm{t}A^2\widehat{x}_+^{n+1} + \Delta\mathrm{t}\gamma H^*(d^{n+1} - Hx_\bullet((n+1)\Delta\mathrm{t})) \\ &= (\mathbb{1} + \Delta\mathrm{t}\gamma H^* H - \nu_{\Delta\mathrm{t}}\Delta\mathrm{t}A^2)\widetilde{x}_+^{n+1} - \Delta\mathrm{t}\varepsilon_\nu^{n+1} - \Delta\mathrm{t}\gamma H^*\varepsilon_d^{n+1}. \end{aligned}$$

What remains of the proof follows the demonstration of Proposition 3.1. □

We can now give the convergence estimate in the case of interpolated data and, as a first step, we do not make assumptions on the type of interpolation scheme.

Theorem 3.6. *Making the same assumptions on A, H and x_0 than Theorem 3.2 and denoting by*

$$|\varepsilon_d| = \max_{1\le i\le n} \|\varepsilon_d^i\|_{\mathcal{Z}},$$

we can state that there exist positive constants M_0, μ_0, C_1, C_2 and C_3, independent of $\Delta\mathrm{t} \in (0,1)$ and n, such that,

$$\|\widetilde{x}_+^n\|_{\mathcal{X}} \le M_0 \exp(-\mu_0 n\Delta\mathrm{t})\|\widetilde{x}_0\|_{\mathcal{X}} + \frac{\Delta\mathrm{t}}{1-\exp(-\mu_0\Delta\mathrm{t})}\big(\Delta\mathrm{t}^2 C_1 + \nu_{\Delta\mathrm{t}} C_2 + \gamma C_3|\varepsilon_d|\big). \tag{41}$$

Proof. From the discrete dynamical system (39) satisfied by the estimation error we can extract the explicit relation

$$\widetilde{x}_+^{n+1} = \mathcal{U}\mathcal{V}\mathcal{W}\widetilde{x}_+^n + \Delta\mathrm{t}\big(\mathcal{U}\mathcal{V}\varepsilon^{n+1} + \mathcal{U}\varepsilon_\nu^{n+1} + \gamma H^*\varepsilon_d^{n+1}\big), \tag{42}$$

where $\mathcal{U},\ \mathcal{V},\ \mathcal{W} \in \mathcal{L}(\mathcal{X})$ are given by

$$\begin{cases} \mathcal{U} = \left(\mathbb{1} - \Delta\mathrm{t}\,\nu_{\Delta\mathrm{t}} A^2 + \gamma\Delta\mathrm{t} H^*H\right)^{-1}, \\ \mathcal{V} = \left(\mathbb{1} - \dfrac{\Delta\mathrm{t}}{2}A\right)^{-1}, \\ \mathcal{W} = \mathbb{1} + \dfrac{\Delta\mathrm{t}}{2}A. \end{cases}$$

Following the same arguments as for Theorem 3.2

$$\|\widetilde{x}_+^n\|_{\mathcal{X}} \le M_0 \exp(-\mu_0 n\Delta\mathrm{t})\|\widetilde{x}_+^0\|_{\mathcal{X}} + \frac{\Delta\mathrm{t}}{1-\exp(-\mu_0\Delta\mathrm{t})}\left(\max_{1\le i\le n}\|\varepsilon^i\|_{\mathcal{X}} + \max_{1\le i\le n}\|\varepsilon_\nu^i\|_{\mathcal{X}} + \gamma C_3 \max_{1\le i\le n}\|\varepsilon_d^i\|_{\mathcal{Z}}\right). \tag{43}$$

Hence we obtain the desired results with C_1 and C_2 expressed in (32) and C_3 given by (36). □

Note that estimate (41) directly follows intuition since, as we could imagine, when only few data are available the interpolation error naturally increases hence the right-hand side of (41) grows larger. Moreover, the gain γ appears as a coefficient on this part of the upper-bound which clearly implies that the interpolation error enters in the observer dynamical system as some additive noise on the data. To finally illustrate this phenomenon we propose to give an *a priori* bound of the interpolation error in the case of linearly interpolated data.

Proposition 3.7. *If we associate the sampling time $(t_r)_{r\in\mathbb{N}} \subset \mathbb{N}$ to the sampling time steps $(j_r)_{r\in\mathbb{N}} \subset \mathbb{N}$, by $t_r = j_r\Delta\mathrm{t}$ and if we denote by*

$$\Delta\mathrm{T}_{max} = \max_{r\ge 0}\{t_{r+1} - t_r\},$$

then there exists a constant C_4 depending only on the observation operator H and the initial condition x_0 such that

$$|\varepsilon_d| \ \le\ C_4\,\Delta\mathrm{T}_{max}^2. \tag{44}$$

Proof. To simplify the presentation we first focus our attention on the interval $[t_0, t_1]$. To start with we remark that d^i in (16) can be written as

$$d^i = H\left(\frac{i\Delta\mathrm{t} - t_0}{t_1 - t_0}x_\bullet(t_1) + \frac{t_1 - i\Delta\mathrm{t}}{t_1 - t_0}x_\bullet(t_0)\right). \tag{45}$$

Furthermore, there exists $r_0 \in [t_0, i\Delta\mathrm{t}]$ and $r_1 \in [i\Delta\mathrm{t}, t_1]$ such that

$$\begin{cases} x_\bullet(t_0) = x_\bullet(i\Delta\mathrm{t} + t_0 - i\Delta\mathrm{t}) = x_\bullet(i\Delta\mathrm{t}) + (t_0 - i\Delta\mathrm{t})\dot{x}_\bullet(i\Delta\mathrm{t}) + \dfrac{(t_0 - i\Delta\mathrm{t})^2}{2}\ddot{x}_\bullet(r_0), \\ x_\bullet(t_1) = x_\bullet(i\Delta\mathrm{t} + t_1 - i\Delta\mathrm{t}) = x_\bullet(i\Delta\mathrm{t}) + (t_1 - i\Delta\mathrm{t})\dot{x}_\bullet(i\Delta\mathrm{t}) + \dfrac{(t_1 - i\Delta\mathrm{t})^2}{2}\ddot{x}_\bullet(r_1). \end{cases}$$

Therefore, replacing $x_\bullet(t_0)$ and $x_\bullet(t_1)$ in (45) we obtain

$$\begin{aligned} d^i =& H\Bigg(x_\bullet(i\Delta\mathrm{t}) + \frac{(t_1 - i\Delta\mathrm{t})(t_0 - i\Delta\mathrm{t}) + (i\Delta\mathrm{t} - t_0)(t_1 - i\Delta\mathrm{t})}{t_1 - t_0}\dot{x}_\bullet(i\Delta\mathrm{t}) \\ & \frac{(t_1 - i\Delta\mathrm{t})(t_0 - i\Delta\mathrm{t})^2}{2(t_1 - t_0)}\ddot{x}_\bullet(r_0) + \frac{(i\Delta\mathrm{t} - t_0)(t_1 - i\Delta\mathrm{t})^2}{2(t_1 - t_0)}\ddot{x}_\bullet(r_1)\Bigg) \end{aligned}$$

hence

$$d^i - Hx_\bullet(i\Delta t) = H\left(\frac{(t_1 - i\Delta t)(t_0 - i\Delta t)^2}{2(t_1 - t_0)}\ddot{x}_\bullet(r_0) + \frac{(i\Delta t - t_0)(t_1 - i\Delta t)^2}{2(t_1 - t_0)}\ddot{x}_\bullet(r_1)\right).$$

Using now the conservation of the energy associated with (2) this leads to

$$\|\varepsilon_d^i\|_{\mathcal{Z}} \leq \|H\|_{\mathcal{L}(\mathcal{X},\mathcal{Z})}\frac{(i\Delta t - t_0)^2(t_1 - i\Delta t) + (i\Delta t - t_0)(t_1 - i\Delta t)^2}{2(t_1 - t_0)}\|A^2 x_0\|_{\mathcal{X}}$$

and therefore

$$\|\varepsilon_d^i\|_{\mathcal{Z}} \leq \|H\|_{\mathcal{L}(\mathcal{X},\mathcal{Z})}\frac{(t_1 - t_0)^2}{8}\|A^2 x_0\|_{\mathcal{X}}. \tag{46}$$

We conclude the demonstration by giving the expression of the constant, namely

$$C_4 = \frac{1}{8}\|H\|_{\mathcal{L}(\mathcal{X},\mathcal{Z})}\|A^2 x_0\|_{\mathcal{X}}.$$

□

As previously presented for the intermittent observer, the convergence estimate in Theorem 3.6 can be extended to take into account data noise of the form of (33). In this case, the second equation of the dynamical system (39) verified by the estimation error of the observer using interpolated data becomes

$$\frac{\widetilde{x}_+^{n+1} - \widetilde{x}_-^{n+1}}{\Delta t} = -\gamma H^* H\widetilde{x}_+^{n+1} + \nu_{\Delta t}A^2\widetilde{x}_+^{n+1} + \varepsilon_\nu^{n+1} + \gamma H^*\varepsilon_{d,\chi}^{n+1}, \tag{47}$$

where

$$\varepsilon_{d,\chi}^{n+1} = (\mathbb{1} - \mathcal{I})Hx_\bullet((n+1)\Delta t) - \mathcal{I}\chi((n+1)\Delta t) = \varepsilon_d^{n+1} - \mathcal{I}\chi((n+1)\Delta t). \tag{48}$$

In the previous equation we denoted by $\mathcal{I}$ the interpolation operator which, in the case of linear interpolation, is given by (16) and by ε_d^{n+1} the interpolation error of the observations without noise, as presented in Theorem 3.6. In a general case, assuming that $\mathcal{I} \in \mathcal{L}(\mathcal{Z},\mathcal{Z})$ is a bounded linear operator, the proof of Theorem 3.6 remains valid, so that the estimate (41) reads

$$\|\widetilde{x}_+^n\|_{\mathcal{X}} \leq M_0\exp(-\mu_0 n\Delta t)\|\widetilde{x}_0\|_{\mathcal{X}} + \frac{\Delta t}{1 - \exp(-\mu_0\Delta t)}\Big(\Delta t^2 C_1 + \nu_{\Delta t}C_2 + \gamma C_3|\varepsilon_d| + \gamma C_3\|\mathcal{I}\|_{\mathcal{L}(\mathcal{Z},\mathcal{Z})}\max_{1\leq i\leq n}\|\chi(i\Delta t)\|_{\mathcal{Z}}\Big). \tag{49}$$

In the light of Theorem 3.2 and Theorem 3.6 we understand that choosing a strategy – interpolated or on/off – relies on a compromise between stability and consistency. On the one hand, the observer with intermittent correction phases bear an exact consistency with respect to the data, however corrections shocks may occur thus leading to potential instabilities. On the other hand, the observer fed with interpolated data admits a correction phase at each model time-step – hence its greater stability – but artificially induces noise in the observer dynamical system due to data interpolation error.

4. Numerical illustrations

We proceed in this section with the illustration of our theoretical results on a simple 1D wave model. Our objective is to show how the two strategies can offer two different alternatives depending on the data sampling. Namely, the on/off time discretization is a consistent formulation with respect to the data sampling whereas the interpolated strategy ensures a stability at every time-step. We will proceed with two classes of numerical results. First we will present a spectrum analysis to numerically demonstrate the stability properties of the proposed feedback laws. This will in particular help us to fix the optimal gains of each method with respect

to the other. Secondly, we will present time evolution of the estimation errors for various sources of initial uncertainties in order to show the impact of the data time-sampling in practical data assimilation procedures.

The example of the one-dimensional wave equation reads

$$\begin{cases} \ddot{w}(x,t) - w_{xx}(x,t) = 0, & (x,t) \in (0,1) \times (0,\infty) \\ w_x(0,t) = w_x(1,t) = 0, & t \in (0,\infty) \\ w(x,0) = w_0(x) + \zeta_0(x), & x \in (0,1) \\ \dot{w}(x,0) = w_1(x) + \zeta_1(x), & x \in (0,1). \end{cases} \tag{50}$$

In this setting we have

$$\begin{vmatrix} \mathcal{H} = \mathcal{L}^2(0,1), \\ \mathcal{D}(A_0) = \left\{ w \in \mathcal{H}^1(0,1) \;\middle|\; w_{xx} \in \mathcal{L}^2(0,1) \right\}, \\ A_0 w = -w_{xx}, \qquad w \in \mathcal{D}(A_0). \end{vmatrix}$$

It is easy to see that this particular case fits exactly in the abstract framework described in Section 2. Concerning the observations, we follow the example presented in [8] and we assume that

$$z(t) = w(\cdot,t)|_{\omega_0}, \qquad t \in (0,\infty), \tag{51}$$

with $\omega_0 =]0.3, 0.7[\subset (0,1)$ an open and non-empty interval. Hence, the observation space reads

$$\mathcal{Z} = \mathcal{H}^1(\omega_0). \tag{52}$$

Following [7] we endow $\mathcal{Z}$ with the inner-product

$$\forall \varphi_1,\ \varphi_2 \in \mathcal{H}^1(\omega_0) \qquad (\varphi_1, \varphi_2)_{\mathcal{Z}} = \left(\mathrm{Ext}_{\omega_0}(\varphi_1), \mathrm{Ext}_{\omega_0}(\varphi_2)\right)_{\mathcal{H}^1(0,1)}, \tag{53}$$

where the extension operator $u = \mathrm{Ext}_{\omega_0}(\varphi) \in \mathcal{H}^1(0,1)$ is defined for all $\varphi \in \mathcal{H}^1(\omega_0)$ as the solution of

$$\begin{cases} -u_{xx} = 0, & x \in (0,1) \backslash \omega_0 \\ u = \varphi, & x \in \overline{\omega_0} \\ u_x(0) = u_x(1) = 0. \end{cases}$$

For the inner-products $(\cdot,\cdot)_{\mathcal{Z}}$ and $(\cdot,\cdot)_{\mathcal{H}^1(0,1)}$, one can prove [7] that the adjoint of the observation operator is

$$H_0^\star : \begin{vmatrix} \mathcal{Z} \to \mathcal{H}^1(0,1) \\ \varphi \mapsto \mathrm{Ext}_{\omega_0}(\varphi), \end{vmatrix}$$

and therefore the observer, decomposed into $\widehat{x} = \begin{pmatrix} \widehat{w} & \widehat{v} \end{pmatrix}^\top$, reads

$$\begin{cases} \dot{\widehat{w}} = \widehat{v} + \gamma \mathrm{Ext}_{\omega_0}(z - \widehat{w}|_{\omega_0}), \\ \dot{\widehat{v}} = \widehat{w}_{xx}, \\ \widehat{w}_x|_{x=0} = \widehat{w}_x|_{x=1} = 0, \\ \widehat{w}|_{t=0} = w_0, \\ \widehat{v}|_{t=0} = w_1. \end{cases}$$

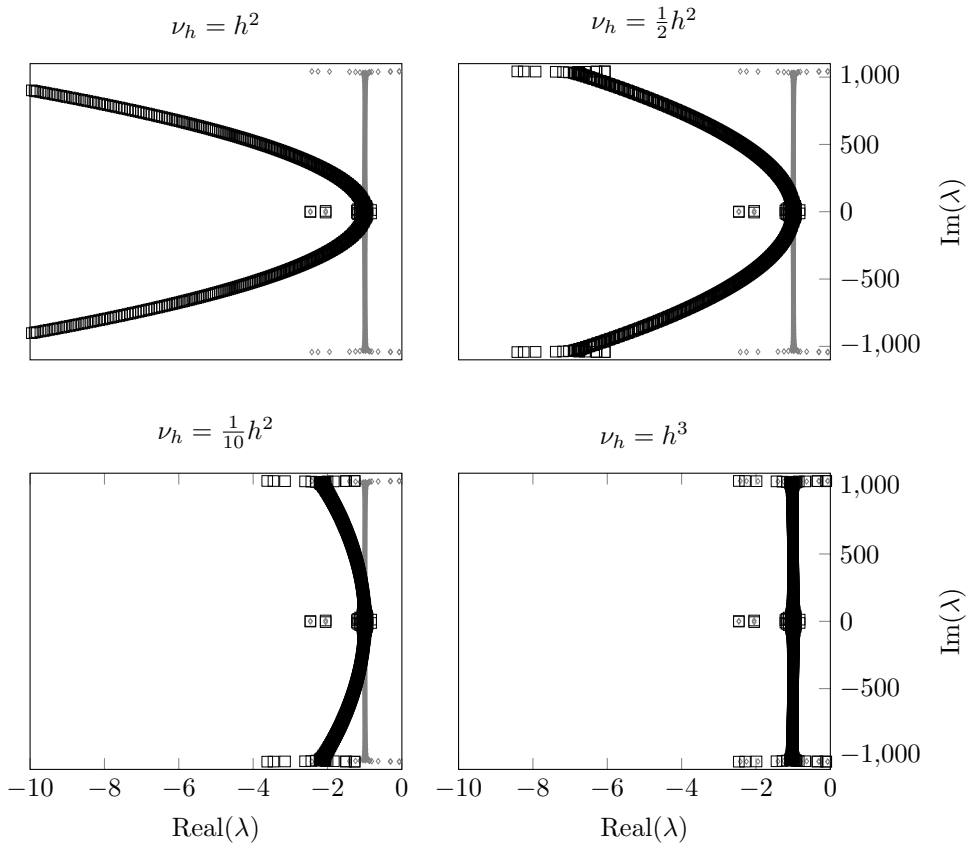


FIGURE 1. Evolution with respect to the viscous coefficient ν_h of the spectrum of the operator driving the continuous-time dynamical system for a fixed value of the gain $\gamma = 5$ and a space step of $h = \frac{1}{200}$. In **(black □)** the operator without numerical viscosity, and in (gray ◊) with numerical viscosity.

4.1. **Spectral analysis**

As the Luenberger filter aims at stabilizing the estimation error system it is natural to assess the quality of the gain filter by evaluating its damping impact on the otherwise conservative system. Following [25], this can be done after full discretization by observing the evolution of the poles of the stabilized operator in both cases, namely using time interpolation or the on/off switch.

To start with, the most natural spectrum to analyze is certainly the spectrum associated with the time-continuous observer formulation, namely with only its spatial discretization. Then, the numerical results can be compared to the existing literature in particular with [8, 27]. This case is the asymptotic limit of the time discretization and should be considered as our ultimate goal when we expect to produce the most accurate simulations. We will then see the impact of the time discretization in a second set of results. Note that we must proceed with caution here since the on/off time discretization is in essence a time-discrete strategy where we have to compose different operators over the sampling period.

4.1.1. *Spectral analysis of the continuous-time observer*

Firstly, we comment the loss of uniform stability after discretization and our choice of an added numerical viscosity to circumvent this limitation. In fact, after spatial discretization, the estimation error dynamical system (6) reads

$$\dot{\widetilde{x}}_h(t) = \big(A_h - \gamma H_h^* H_h\big)\widetilde{x}_h(t), \tag{54}$$

where A_h and H_h are consistent discretization – *w.r.t* a space step h – of the model and the observation operator. In practice, for our numerical experiments we have chosen a reasonable space discretization of $h = \frac{1}{200}$. Even though the observability inequality (4) is satisfied at a continuous level, this discretization procedure already produces – similarly to what we have mentioned in the case of time discretization – some spurious modes inducing the loss of the exponential stability of (54) – see [13]. To circumvent this difficulty we have introduced some artificial viscosity for the time discretization which should also take into account the spatial discretization. When $\Delta t \to 0$ we obtain the time-continuous error dynamics studied in [13], namely

$$\dot{\widetilde{x}}_h(t) = \big(A_h - \gamma H_h^* H_h + \nu_h A_h^2\big)\widetilde{x}_h(t). \tag{55}$$

Therefore, we start our numerical investigations with Figure 1 where we plot the evolution of the spectrum of the operator driving the dynamical system (55) *w.r.t.* various values of the artificial viscosity coefficient. These spectra are compared with the spectrum of the operator without numerical viscosity. By these examples we illustrate – in the particular case of the spatial discretization of 1D wave equation using first order finite element method – the creation of spurious high-frequencies that cannot be stabilized by the feedback operator $H_h^* H_h$. Hence, stabilizing the complete frequency range is then performed by decomposing the tasks. The feedback operator manages the low frequencies whereas the high frequencies are controlled by the numerical viscosity. It should be noted, however, that since this artificial viscosity corresponds to a perturbation of the standard discretization of the problem it should be kept as low as possible. In practice the choice of the coefficient ν_h can be done with an *a priori* knowledge (or estimation) of the frequencies appearing in the specific problem of interest. However, in the rest of our following investigation we will keep a numerical viscosity strong enough to produce unclouded figures.

Secondly, we discuss the well known overdamping effect which constraints the optimal choice of the gain. We plot in Figure 2 the spectra for various values of the gain parameter γ – in **(black □)** the operator without numerical viscosity and in (gray ◊) the operator with numerical viscosity. We fix in these examples $\nu_h = h^2$. As the gain increases, the frequencies, after being shifted towards the left half plane, start to go back to the imaginary axis. This is particularly critical for the low frequencies. This effect is crucial when choosing an adequate value of the gain with an order of magnitude of

$$\gamma \sim \sqrt{2}\omega_0$$

with ω_0 the lower frequency of the system, see [27]. This indicates that, first, even in the unrealistic case of an outstanding data-to-noise ratio, the gain cannot be chosen "as large as possible". In other words, after some particular value of the gain the ability of the observer to stabilize these low frequency modes looses its efficiency. Hence, the choice of the gain is a compromise between the different range of frequencies we aim at stabilizing. In our work we opt for a sort of "mean" strategy by choosing γ such that a maximum of frequencies are shifted significantly to the left half plane without compromising too much the low frequencies which are more present in the potential errors in practice. As an example in Figure 2 we choose $\gamma = 9$.

4.1.2. *Spectral analysis of the discrete-time observers*

Once we have understood the behaviour of the time-continuous observer and chosen a target optimal gain we can now proceed to the numerical study of our different choices of time-discretization. Firstly, we evaluate the relation between the optimal gain found in the time-continuous framework and the sampling period. In this purpose, we need to define the two *transition* operators driving the homogeneous equivalent of system (39)

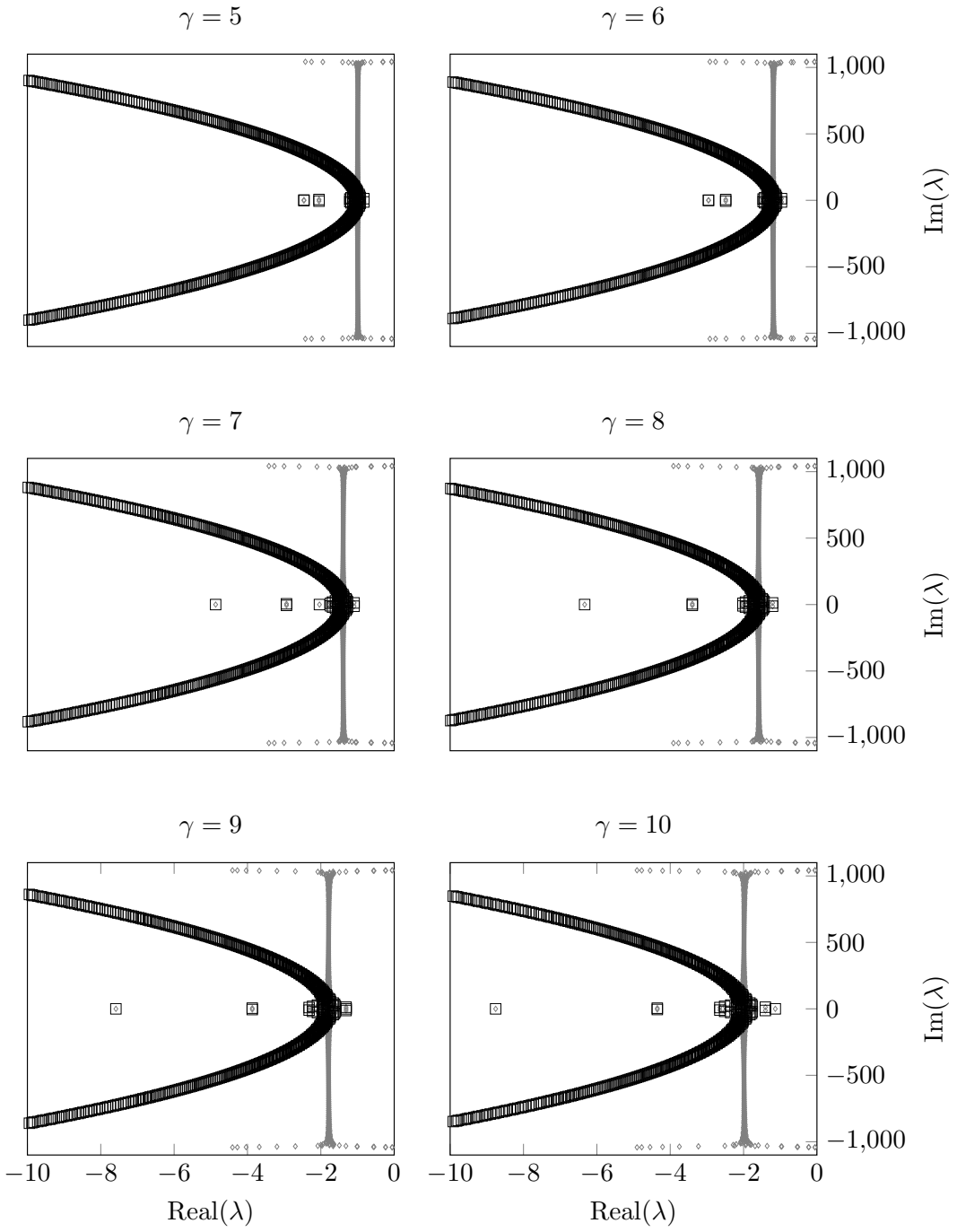


FIGURE 2. Evolution with respect to the gain γ of the spectrum of the operator driving the continuous-time dynamical system with $h = \frac{1}{200}$ and $\nu_h = h^2$. In **(black □)** the operator without numerical viscosity, with numerical viscosity in (gray ◊).

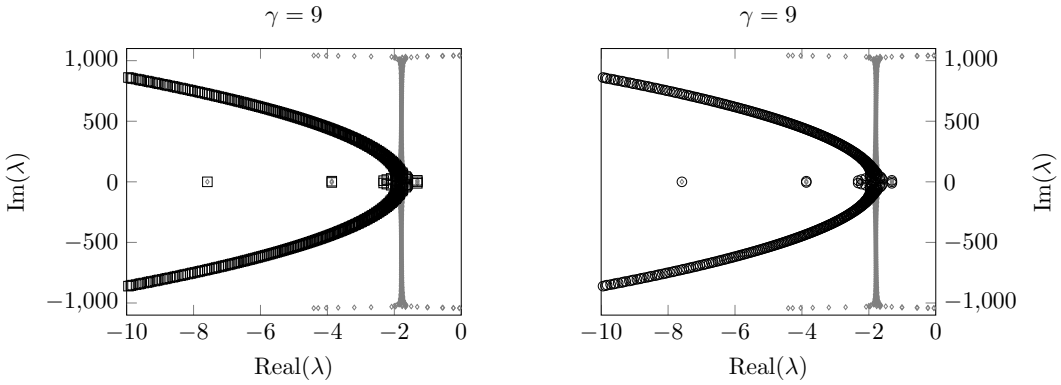


FIGURE 3. Comparison between the spectra of the continuous-time operator (in **(black □)**) and the time-discrete observer using interpolated data (in **(black ○)**) with time overkill. The discretization parameters are $h = \frac{1}{200}$, $\Delta t = h^2$ and $\nu = h^2$.

either with numerical viscosity or not, namely

$$\begin{cases} A^{\gamma}_{n+1|n} = \big(\mathbb{1} + \gamma \Delta t H_h{}^* H_h\big)^{-1}\big(\mathbb{1} - \frac{1}{2}\Delta t A_h\big)^{-1}\big(\mathbb{1} + \frac{1}{2}\Delta t A_h\big), \\ A^{\gamma\nu}_{n+1|n} = \big(\mathbb{1} - \Delta t \nu A_h^2 + \gamma \Delta t H_h{}^* H_h\big)^{-1}\big(\mathbb{1} - \frac{1}{2}\Delta t A_h\big)^{-1}\big(\mathbb{1} + \frac{1}{2}\Delta t A_h\big), \end{cases}$$

where, if $\widetilde{x}^n_+$ is the solution of the system (39) with $\varepsilon = \varepsilon_\nu = \varepsilon_d = 0$, then

$$\widetilde{x}^{n+1}_+ = A^{\gamma\nu}_{n+1|n}\widetilde{x}^n_+,$$

and, if $\nu_{\Delta t} = 0$ in (39), then

$$\widetilde{x}^{n+1}_+ = A^{\gamma}_{n+1|n}\widetilde{x}^n_+.$$

Note that in the definition of $A^{\gamma\nu}_{n+1|n}$, the parameter ν aims at adjusting the amount of numerical viscosity in order to discard spurious modes arising from the time discretization but also the space discretization – see Section 4.1.1. Combining the necessary restrictions arising from both time and space discretization, we understand that we need to choose here the global viscosity coefficient as

$$\nu = O(\max\{\Delta t^2, h^2\}).$$

We now link the eigenvalues α and α_ν of the discrete-time observer to the one corresponding to the continuous-time observer. Therefore, let us denote by λ an eigenvalue of the operator $A_h - \gamma H_h^* H_h$ to be compared with α. If we initialize the discrete-time and continuous-time system with the corresponding eigenvectors v_λ and u_λ then (from the explicit form of the dynamics) we obtain that, at a given time $n\Delta t$

$$\begin{cases} v_\lambda(n\Delta t) = v_\lambda \exp(\lambda n \Delta t) \\ u^n_{\lambda+} = u_\lambda \alpha^n, \end{cases}$$

hence seeking $v_\lambda(n\Delta t) \sim u^n_{\lambda+}$ induces $\lambda = \dfrac{\ln(\alpha)}{\Delta t}$. To numerically validate these arguments we plot in Figure 3 the values

$$\frac{\log(\alpha)}{\Delta t} \text{ (gray } \lozenge\text{)} \quad \text{and} \quad \frac{\log(\alpha_\nu)}{\Delta t} \textbf{ (black } \bigcirc\textbf{)}.$$

In order to be able to compare with the time-continous spectrum we set $\Delta t = h^2$ – this configuration can be understood as an *overkill in time* – and we observe a perfect match between the two plots. In Figure 4 we

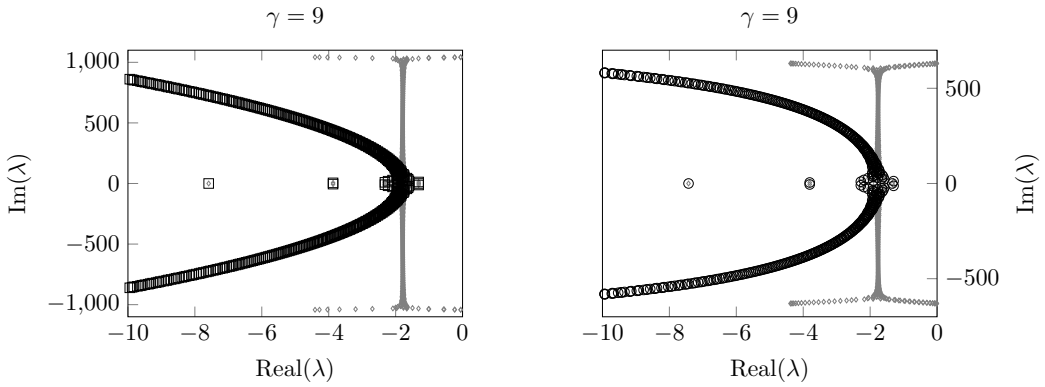


FIGURE 4. Comparison between the spectra of the continuous-time operator (in **(black □)**) and the time-discrete observer using interpolated data (in **(black ◯)**) with $h = \frac{1}{200}$, $\Delta t = h$ and $\nu = \Delta t^2$.

represent a configuration where $\Delta t = h$. Here the spectrum differs from the time-continuous analysis. We see that additional spurious high frequencies are not controlled by the stabilization operator. This justifies even more the use of numerical viscosity which allows to keep the gain chosen in the time-continuous setting. But ultimately, the choice of the gain made with respect to the time-continuous case is very robust to all time-step discretizations.

We can now move to the analysis of the on/off discrete operator in order to determine the associated optimal gain. In this perspective it is convenient to seek for the time-continuous limit, therefore we consider a time overkill situation by setting $\Delta t = h^2$. Then, to fix the ideas, we consider the case where $\Delta T = 5\Delta t$, namely

$$\forall r \in \mathbb{N} \quad (j_{r+1} - j_r) = 5.$$

We introduce the transition operators from time t_n to t_{n+5} as

$$\left|\begin{aligned} B^\gamma_{n+5|n} &= \Big(\big(\mathbb{1} - \frac{1}{2}\Delta t A_h\big)^{-1}\big(\mathbb{1} + \frac{1}{2}\Delta t A_h\big)\Big)^4 \Big(\big(\mathbb{1} + \gamma\Delta t {H_h}^* H_h\big)^{-1}\big(\mathbb{1} - \frac{1}{2}\Delta t A_h\big)^{-1}\big(\mathbb{1} + \frac{1}{2}\Delta t A_h\big)\Big), \\ B^{\nu\gamma}_{n+5|n} &= \Big(\big(\mathbb{1} - \Delta t \nu A_h^2\big)^{-1}\big(\mathbb{1} - \frac{1}{2}\Delta t A_h\big)^{-1}\big(\mathbb{1} + \frac{1}{2}\Delta t A_h\big)\Big)^4 \\ &\qquad \Big(\big(\mathbb{1} - \Delta t \nu A_h^2 + \gamma\Delta t {H_h}^* H_h\big)^{-1}\big(\mathbb{1} - \frac{1}{2}\Delta t A_h\big)^{-1}\big(\mathbb{1} + \frac{1}{2}\Delta t A_h\big)\Big), \end{aligned}\right.$$

and we denote by $\widetilde{\beta}$ and $\widetilde{\beta}_\nu$ the corresponding eigenvalues. These operators should be compared to their equivalent transition operators from time t_n to t_{n+5} in the interpolated case, namely the simple composition

$$A^\gamma_{n+5|n} = \big(A^\gamma_{n+1|n}\big)^5 \quad \text{and} \quad A^{\gamma\nu}_{n+5|n} = \big(A^{\gamma\nu}_{n+1|n}\big)^5,$$

with corresponding eigenvalues $\widetilde{\alpha}$ and $\widetilde{\alpha}_\nu$. Then, in Figure 5 we show

$$\frac{\log(\widetilde{\beta})}{\Delta \mathrm{T}} \text{ (gray } \lozenge\text{)} \quad \text{and} \quad \frac{\log(\widetilde{\beta}_\nu)}{\Delta \mathrm{T}} \textbf{ (black } \times\textbf{)},$$

compared with $\widetilde{\alpha}$ and $\widetilde{\alpha}_\nu$ in the time overkill configuration. We set the gain value for the on/off observer

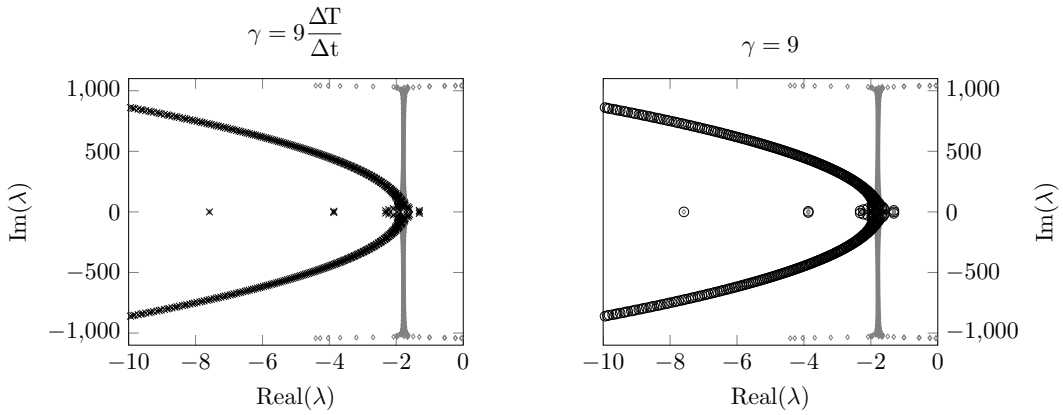


FIGURE 5. Comparison between the time-discrete on/off observer (in **(black** $\times$**)**) and the time-discrete observer using interpolated data (in **(black** $\bigcirc$**)**) to illustrate the relation between the optimal gains in a time overkill context. The complete set of discretization parameters is as follows: $h = \frac{1}{200}$, $\Delta \mathrm{t} = h^2$, $\nu = h^2$ and $\Delta \mathrm{T} = 5\Delta \mathrm{t}$.

five times larger than in the interpolated case and we observe a perfect match. Consequently these numerical arguments confirm the intuition that, in order to obtain similar damping rates between the two strategies, the gain value γ_{switch} for the on/off switch needs to be five times larger than the gain value $\gamma_{\text{int.}}$ for the interpolation strategy. In a general context we can infer the empirical law

$$\frac{\gamma_{\text{switch}}}{\gamma_{\text{int.}}} = \frac{\Delta \mathrm{T}}{\Delta \mathrm{t}} \tag{56}$$

and we point out that we do not face any overdamping phenomena which could have limited the increase of the gain in the on/off switch which is an important contribution of our numerical investigations.

Once the optimal gain of each observer is carefully chosen, we can assess their stabilization properties. We can now analyze a more intricate case where $\Delta \mathrm{t} = h$. We continue to fix the ratio $\Delta \mathrm{T} = 5\Delta \mathrm{t}$ and the spectra of both observers are presented in Figure 6. We observe that additionally to the peculiar form of the spectra more spurious high frequencies appear, validating once again the use of numerical viscosity. Moreover, concerning the on/off observer, we remark some high frequencies that are less stabilized. This can be interpreted as an illustration of the fact that this observer may suffer from a decreased stability on some modes – corresponding to these high frequencies. At last in Figure 7 we set $\Delta \mathrm{t} = h^2$ and $\Delta \mathrm{T} = 5h$. This situation is relevant with practical cases where the time-step of the numerical algorithm is much lower than the sampling time-step of the data but more importantly that it is, in essence, meant to reach small values whereas the sampling rate is fixed. The conclusion stated previously remains valid since we still observe a slight deterioration of the damping of some high frequency modes for the on/off estimator.

To conclude with this section we can say that the time-continuous spectrum analysis remains a very useful tool to fix the optimal gain. It is obvious for the time-interpolated configuration but in fact it is also the case

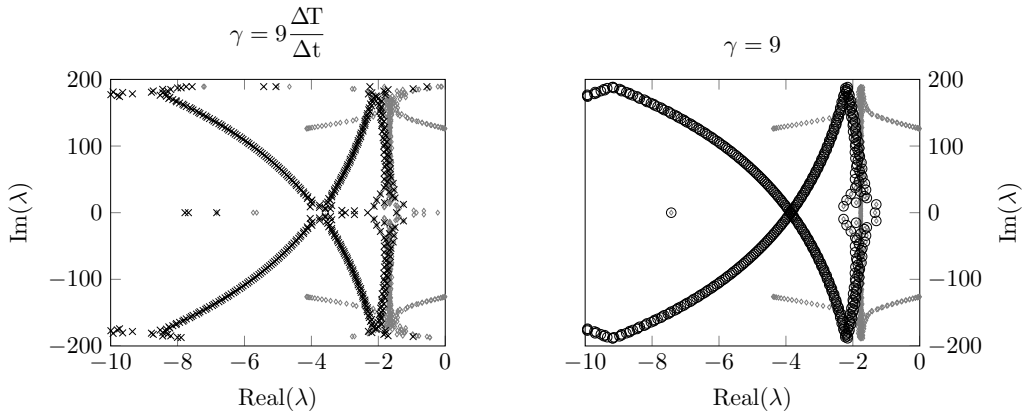


FIGURE 6. Comparison between the time-discrete on/off observer (in **(black ×)**) and the time-discrete observer using interpolated data (in **(black ◯)**) with $h = \frac{1}{200}$, $\Delta t = h$, $\nu = \Delta t^2$ and $\Delta T = 5\Delta t$.

with the on/off switch. The only modification with the on/off switch configuration is that we have to multiply the gain found during the time-continuous spectrum analysis by the ratio of the data time-step over the model time-step. The stability property is preserved globally and we do not face any overdamping phenomena by increasing the value of the gain in this case.

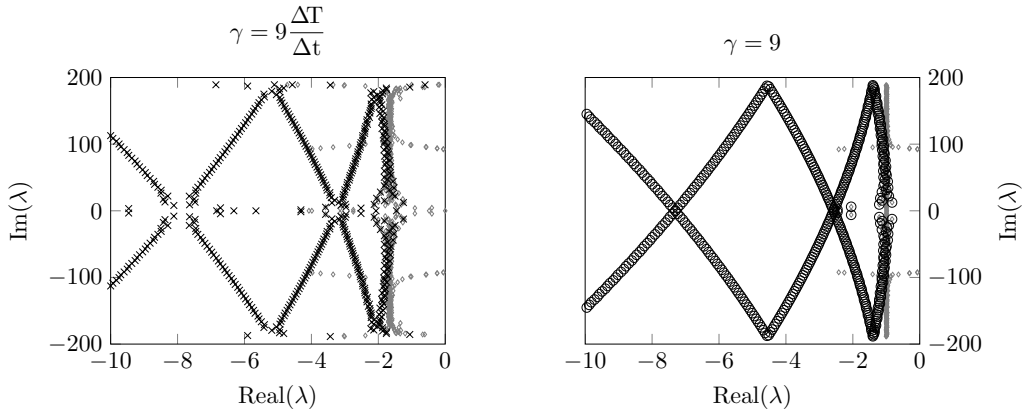


FIGURE 7. Comparison between the time-discrete on/off observer (in **(black ×)**) and the time-discrete observer using interpolated data (in **(black ◯)**) with $h = \frac{1}{200}$, $\Delta t = h^2$, $\nu = h^2$ and $\Delta T = 5h$.

4.2. **Assessing robustness from numerical examples**

We propose in this section to assess the effectiveness of the two observers using *synthetic* data. More precisely, we consider the functions

$$w_0(s) = 16s^2(1-s)^2, \qquad w_1(s) = \begin{cases} 3s - 4s^3 & \text{if } s \in (0, 0.5), \\ 4s^3 - 12s^2 + 9s & \text{if } s \in (0.5, 1), \end{cases} \tag{57}$$

that we use to initiate a *direct model* of the form of (50). The solution of this direct model can be computed analytically

$$w(x,t) = \sum_k \sqrt{2}\big(a_k \cos(\pi k t) + b_k \sin(\pi k t)\big)\sin(\pi k x),$$

with

$$a_k = \frac{2\sqrt{2}(\pi^2 - 12)(\cos(\pi k) - 1)}{\pi^5 k^5} \quad \text{and} \quad b_k = \frac{48\sqrt{2}\sin(\frac{\pi k}{2})}{\pi^4 k^4}$$

from which we extract our synthetic observations. Then we simulate the observer where we artificially introduce some errors by setting

$$\widehat{x}_0 = \begin{pmatrix} w_0 - \alpha\delta w \\ w_1 \end{pmatrix},$$

where δw is a given perturbation of the initial condition (in displacement) and α is a scalar value representing the perturbation amplitude. Hence, the estimation error dynamical system is initialized with

$$\widetilde{x}_0 = \alpha \begin{pmatrix} \delta w \\ 0 \end{pmatrix}.$$

In our numerical simulation, we set $\gamma_{\text{int.}} = 9$, the optimal value discussed in Section 4 and γ_{switch} is obtained from the ratio (56).

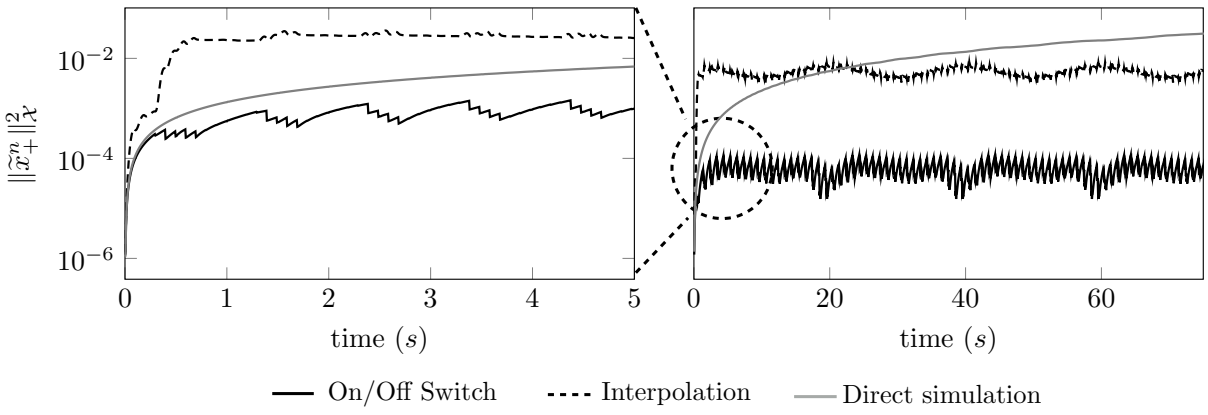


FIGURE 8. Estimation error with $h = \frac{1}{200}$, $\Delta t = h$, $\nu = \Delta t^2$, $\frac{\Delta T}{\Delta t} = 20$ and $\alpha = 0$.

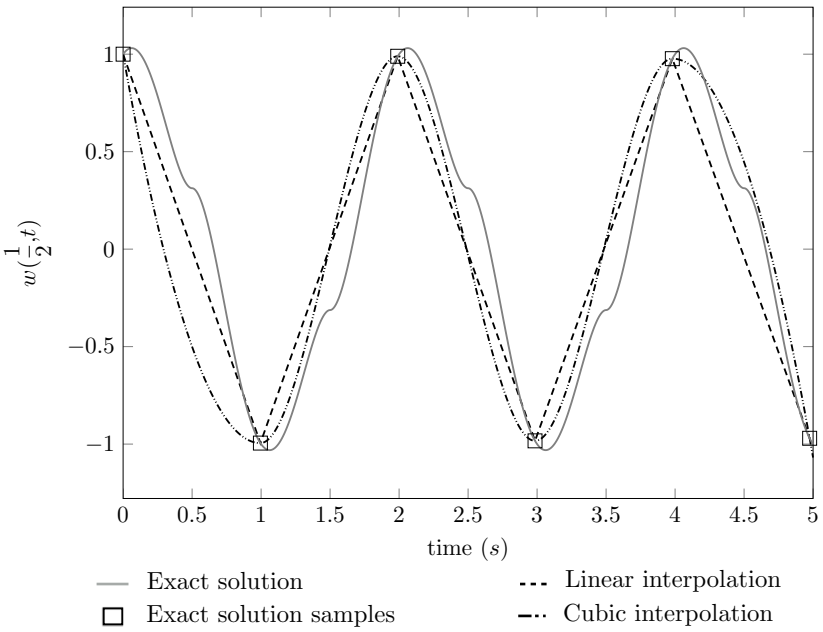


FIGURE 9. Illustration of the impact of time interpolation.

4.2.1. *Improving standard numerical convergence using observers*

First, we propose to consider the case where $\alpha = 0$, namely we initiate the numerical algorithm with the exact – up to some projection errors – initial condition. It is well-known that standard numerical schemes lead to an accumulation of numerical errors thus entailing a deterioration of the numerical solution as the global simulation time grows larger. This phenomenon is illustrated in Figure 8 where we plot the evolution in time of the estimation errors. Concerning the observer we set $\frac{\Delta T}{\Delta t} = 20$ and compare the results of both strategies. We observe that the corresponding estimation errors clearly stabilize to a plateau – during the complete time window and even for large total simulation time – which is a particular behavior already obtained in [8]. In this configuration we observe that, due to the data interpolation error the on/off strategy provides a better numerical solution than the observer using interpolated data.

4.2.2. *Assessing robustness to low data availability*

In our second numerical experiment, we set $\frac{\Delta T}{\Delta t} = 20$, $\alpha = 1$ and $\delta w(x) = \sin(\pi x)$ – the first eigenfunction of the Laplacian operator in the domain. It should be noted that it is not a mode of the stabilized operator driving the dynamics of the estimation error, hence there are multiple excited modes in this dynamical systems. The corresponding results are presented in Figure 10 where we can distinguish several slopes – during early stages of the simulation – in the decay of the estimation error. These slopes correspond to the time constant of stabilization associated with the various modes initially introduced. In the remaining part of the time window we observe that both observers reach a plateau – which is identical to the one observed in Figure 8.

In a third numerical experiment, we consider in Figure 11 the configuration where less data are available on the system by setting $\frac{\Delta T}{\Delta t} = 200$. It is striking to remark that the on/off observer clearly fulfills its task by assimilating data only where they are available thus leading to a great efficiency in the case of highly coarse time distribution of the observations. On the opposite, the other observer oscillates around a plateau due to large interpolation errors. To underline the impact of interpolation error, we have eventually change the type of interpolation by using a cubic interpolation of the data. The resulting sampling is presented in Figure 9.

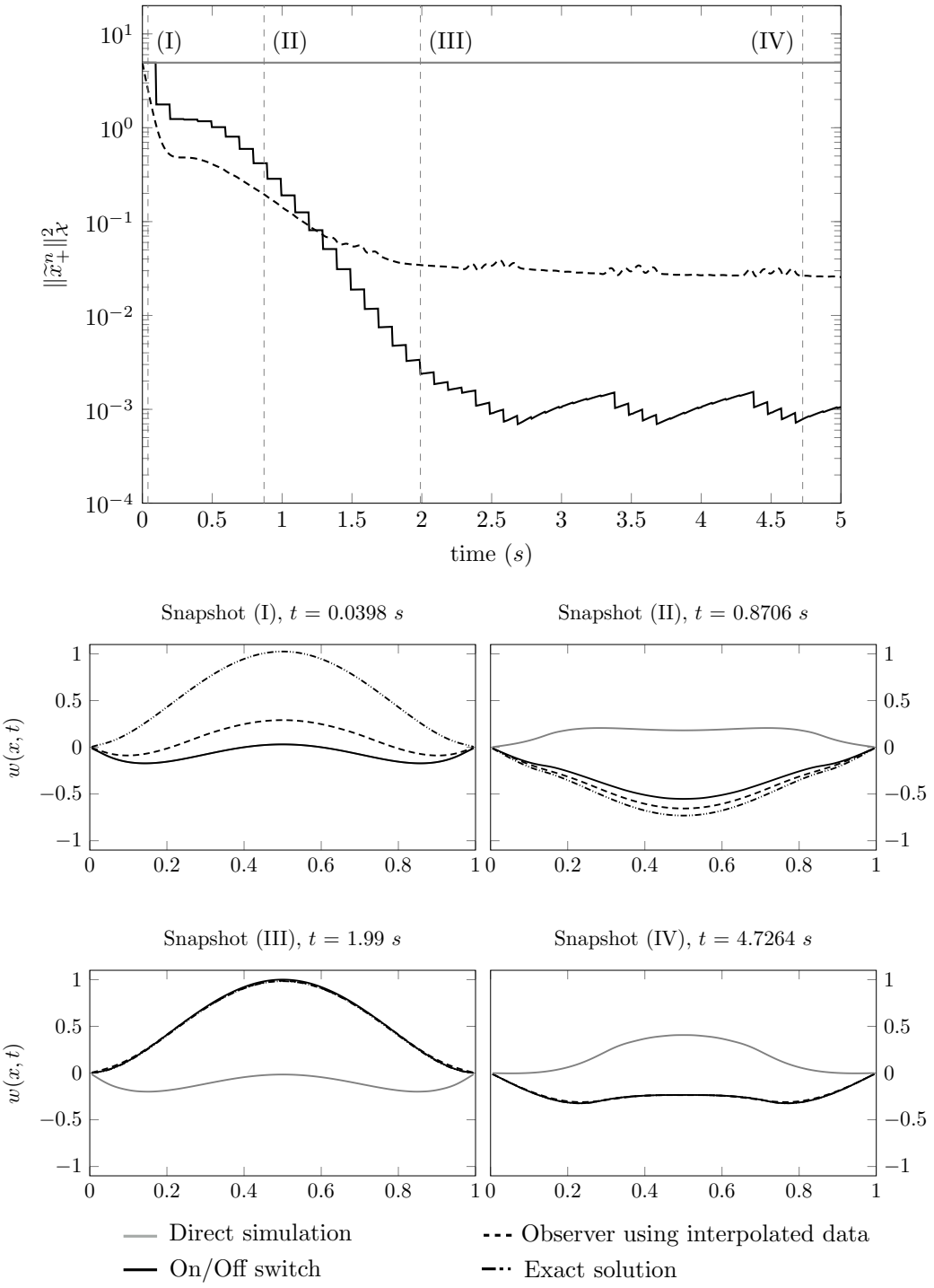


FIGURE 10. Numerical results with $h = \frac{1}{200}$, $\Delta t = h$, $\nu = \Delta t^2, \frac{\Delta T}{\Delta t} = 20$, $\alpha = 1$ and $\delta\varphi(s) = \sin(\pi s)$.

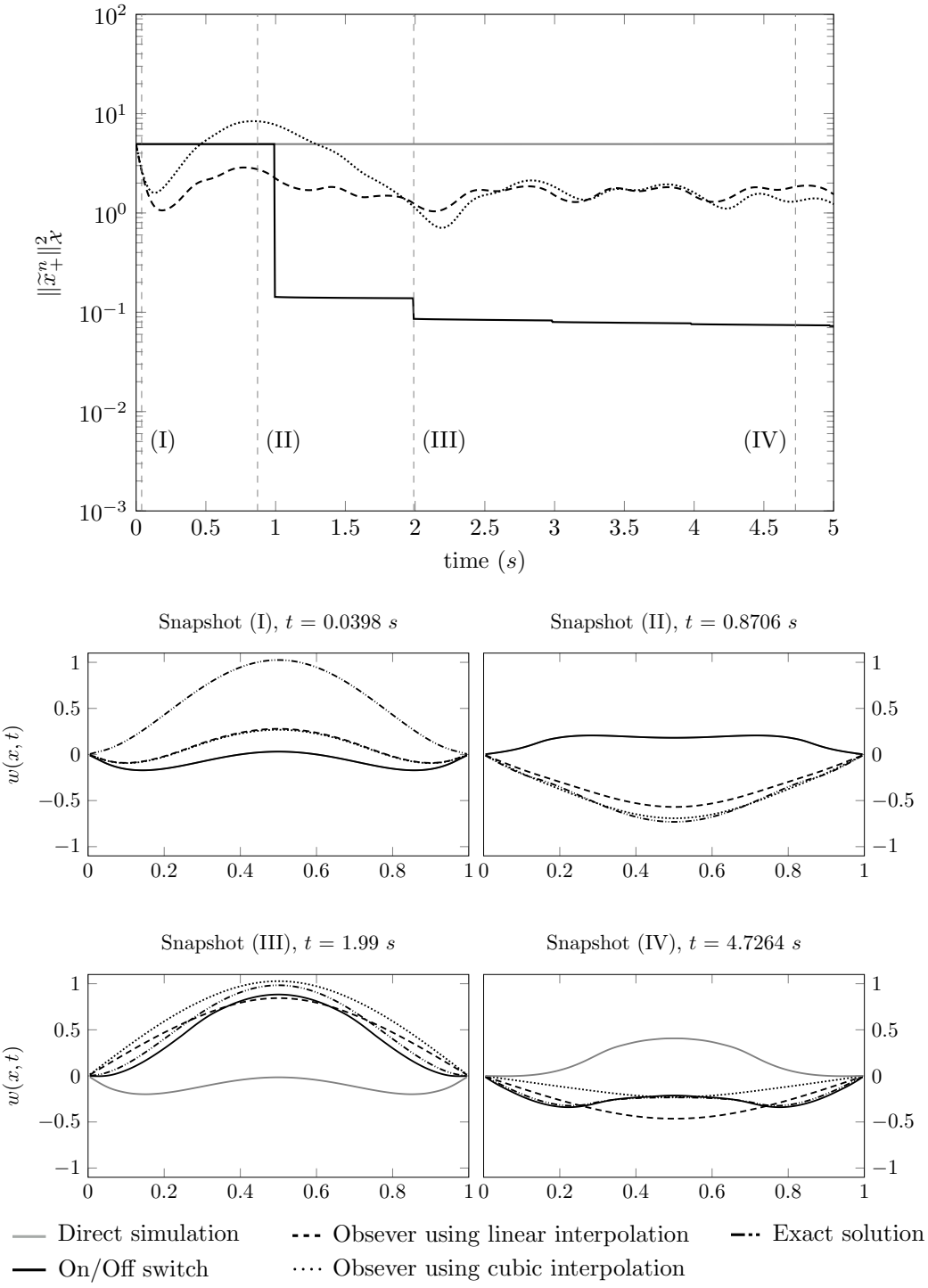


FIGURE 11. Numerical results with $h = \frac{1}{200}$, $\Delta t = h$, $\nu = \Delta t^2$, $\frac{\Delta T}{\Delta t} = 200$, $\alpha = 1$ and $\delta\varphi(s) = \sin(\pi s)$.

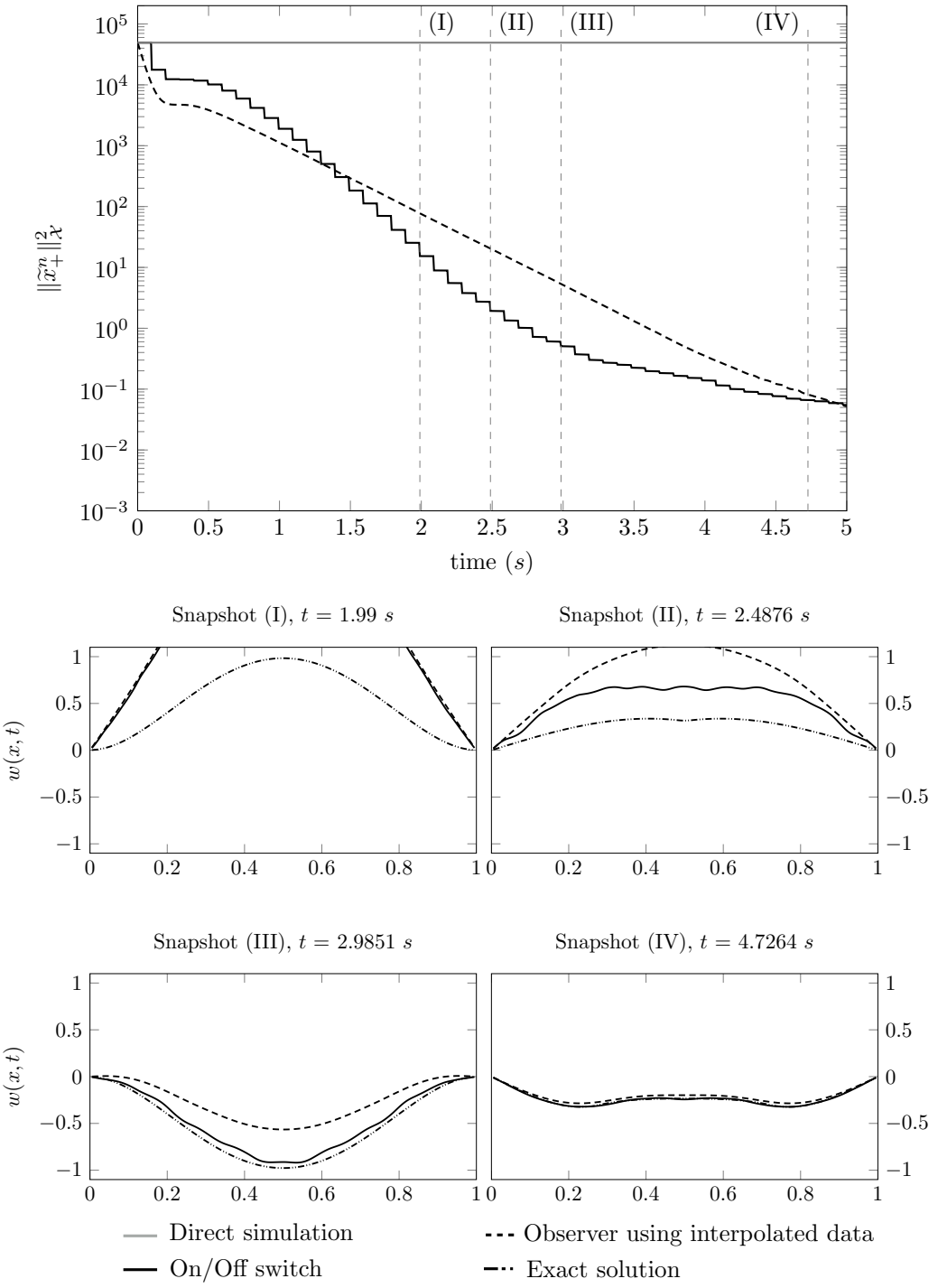


FIGURE 12. Numerical results with $h = \frac{1}{200}$, $\Delta t = h$, $\nu = \Delta t^2, \frac{\Delta T}{\Delta t} = 20$, $\alpha = 10^2$ and $\delta\varphi(s) = \sin(\pi s)$.

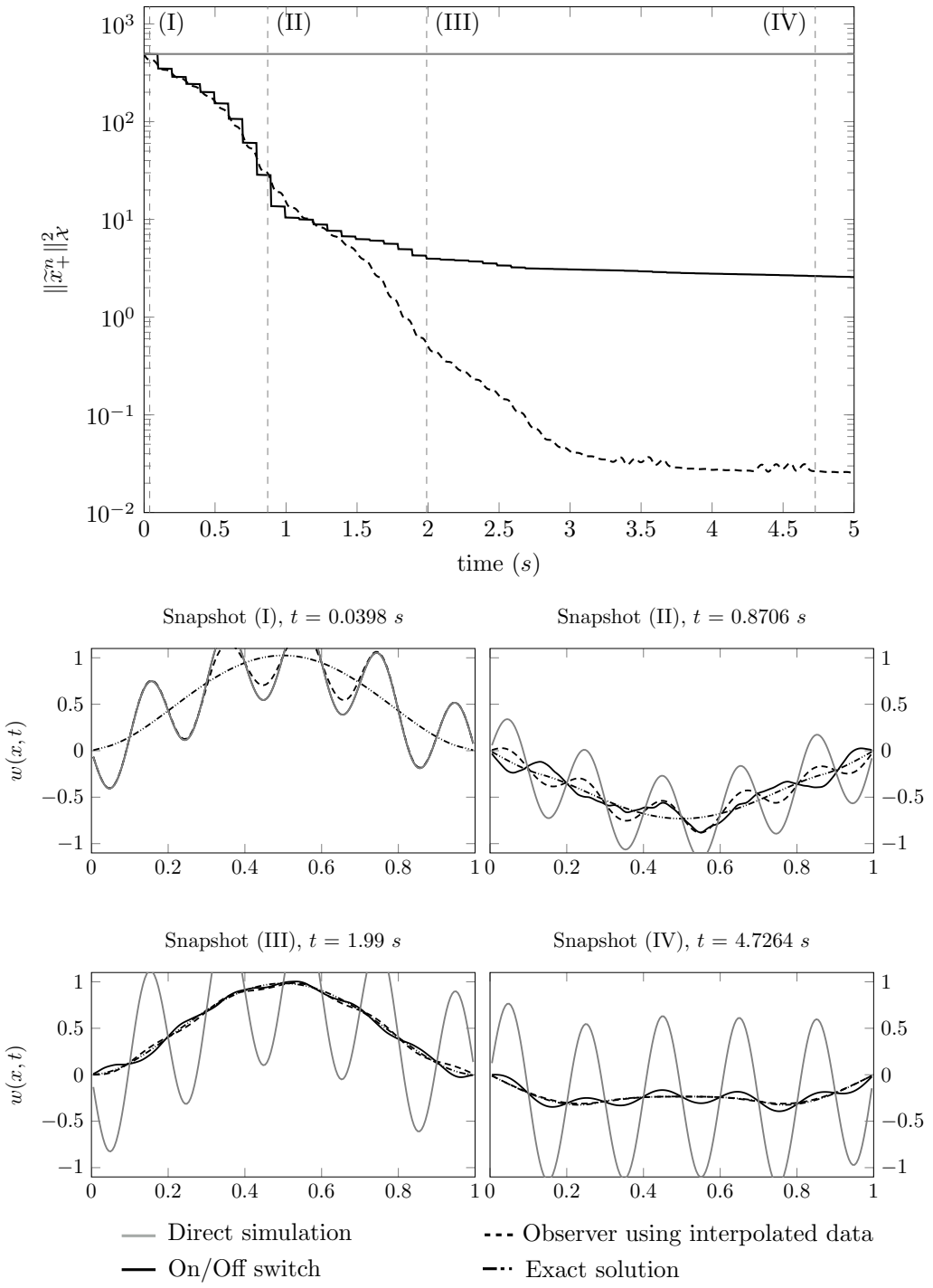


FIGURE 13. Numerical results with $h = \frac{1}{200}$, $\Delta t = h$, $\nu = \Delta t^2, \frac{\Delta T}{\Delta t} = 20$, $\alpha = 1$ and $\delta\varphi(s) = \sin(10\pi s)$.

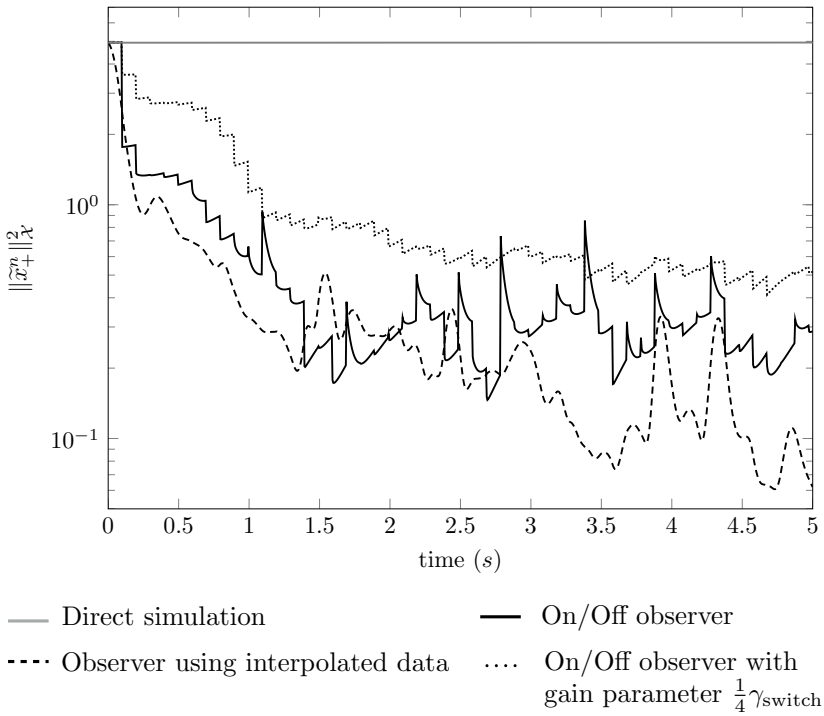


FIGURE 14. Numerical results with $h = \frac{1}{200}$, $\Delta t = h$, $\nu = \Delta t^2$, $\frac{\Delta T}{\Delta t} = 20$, $\alpha = 1$ and $\delta w(s) = \sin(\pi s)$. The data are perturbed with noise as described in (58) with $M = 2$, $\{\sigma_i^F\}_{i=1}^M = \{\frac{1}{20}, \frac{1}{20}\}$ and $\{(\mu_i^K, \sigma_i^K)\}_{i=1}^M = \{(1, \frac{1}{10}), (10, \frac{1}{10})\}$.

We understand that when the cubic reconstruction is better than the linear – here at the end of the time window – then the final estimation error is lower. However at the beginning of the time window, we see that the cubic interpolation error is larger than with a simple linear interpolation which is of dramatic consequence on the estimation error. This also proves that without any idea on the model generating the data, a simple interpolation scheme is sufficient.

Moreover, we are interested in the case where the ratio $\frac{\Delta T}{\Delta t}$ is also set to 200 by decreasing the simulation time-step and using the same sampling period as in Figure 12. This case is representative of the final goal in a numerical procedure, namely for a given configuration we hope to increase the precision by diminishing the model discretization steps. In that case, the results are almost exactly similar to the one presented in Figure 10, namely the observer fed with interpolated data is less efficient due to large interpolation error – which is identical to the one introduced in Figure 10 since it only depends on the sampling period as emphasized in Proposition 3.7.

4.2.3. *Assessing robustness to data noise and large initial condition errors*

In the following experiments we propose to illustrate the robustness of the proposed time-scheme when the magnitude of initial error increases. In Figure 12, we propose to increase the $\mathcal{L}^2$-magnitude of the perturbation by setting $\delta w(s) = \sin(\pi s)$ with $\alpha = 10^2$ and a sampling ratio $\frac{\Delta T}{\Delta t} = 20$. This results in an increase of the potential high frequencies that are initially introduced in the estimation error dynamics. When looking at the snapshots, we see that the on/off observer suffers from high frequency oscillations that do not appear in the observer with the time interpolation scheme. We see here the main advantage of using time interpolation. Namely, since the dissipation brought by the observation operator is present at every time steps, it benefits from a remarkable robustness to initial errors. This feature of both observers can be pointed out more clearly when increasing the $\mathcal{H}^1$-magnitude of the perturbation by setting $\delta w(s) = \sin(10\pi s)$ with $\alpha = 1$. The obtained results are presented

in Figure 13 where we observe that the intermittent observer fails to fully stabilize the initially introduced high frequencies in comparison with the observer using interpolated data.

Finally, our last numerical assessment is dedicated to illustrate the impact of additive noise in the data. Taking into account the space regularity restriction that we have imposed on our measurements, for the analysis we assume the data take the following form

$$z(x,t) = w(x,t)|_{\omega_0} + \sum_{i=1}^{M} F_i \sin(K_i \pi x)|_{\omega_0}, \qquad t \in (0,\infty), \tag{58}$$

where M is a given integer, $\{F_i\}_{i=1}^M$ are random variables following a normal distribution with zero mean and standard variations of $\{\sigma_i^F\}_{i=1}^M$, whereas $\{K_i\}_{i=1}^M$ are random variables following a normal distribution of mean and standard variations of $\{(\mu_i^K, \sigma_i^K)\}_{i=1}^M$. In other words, we randomly add to the data some controlled space high frequencies. The higher the maximum space high frequency M is, the closer we are from a space and time white noise which limit, when the discretization steps goes to 0, remains very specific to study [4]. One typical realization of this random experiment is presented in Figure 14. We can observe that both observers loose in efficiency. The on/off observer has the particularity to present correction shocks, which can be explained by the fact that its larger gain directly amplifies the weight of the noise in the estimation error dynamical system, as presented at the end of Section 3.1. As the gain decreases, we observe that these shocks vanish, however the stabilization properties of the estimation error naturally diminish. This phenomenon is also presented in Figure 14.

4.2.4. *Conclusions from numerical assessment*

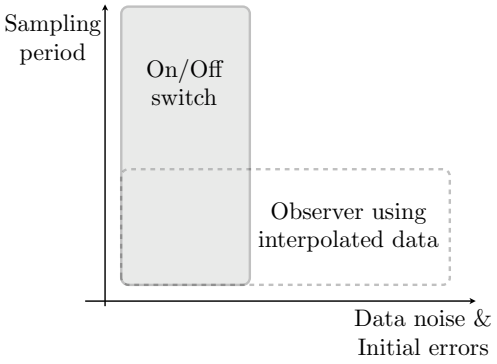


FIGURE 15. Illustration of the different regimes where both observers are more likely to be efficient.

From the various numerical experiments that we have carried out, we can derive general comments on the estimation procedure and also on the particular observers that we have proposed. First of all, in each presented case it is always preferable to use the available data. Secondly, we can not expect a precision improvement better that the time-discretization step of the data, since any procedure designed to fill the gap between two data steps would only create some noise in the estimation procedure. Concerning the differences between the intermittent observer and the observer using interpolated data, these numerical experiments, oriented thanks to the previously presented *a priori* estimates, enabled to characterize the regimes where both strategies should be used. As a matter of fact, the on/off switch has shown an incredible robustness to large data time steps compared to the second observer which efficiency, in that case, is deteriorated by a significant amount of interpolation error. However, in the case of high initial errors or high data noise, it appears to be more efficient to interpolate the data, since the subsequently obtained dynamical system will benefit from a stabilizing operator at each time step. This trade-off is summarized in Figure 15.

5. Conclusion

In t work we have addressed the issue of designing an observer for wave-like systems that is robust when the data have a coarse distribution in time. To circumvent the data-sampling difficulty we have proposed an on/off strategy that filters the observation only at the times when they are available. This strategy was theoretically and numerically analyzed and compared to the case when the data are reconstructed using an interpolation scheme.

The conclusions are twofold. First we have seen that the interpolation remains valid in the case of reasonable repartition (in time) of the data and with potentially high levels of noise. This efficiency directly comes from the presence, at each time step, of the stabilized operator. Secondly, in the case of poor data availability the on/off switch appears to be quite robust since no interpolation error – otherwise entering as a source term in the estimation error dynamical system – is introduced.

The authors would like to deeply thanks Dr. Dominique Chapelle for very useful comments on this work and the reviewers for their careful reading which help us improve the theoretical background of this work.

Appendix A. Exponential stability result for the on/off switch.

We recall that we consider that there exists a constant ratio between the available data and the time discretization step $\Delta T = N\Delta \mathrm{t}$. The general case can be deduced directly as soon as the ratio between two measurements and the time discretization step is bounded with respect to time. For the sake of clarity, we then denote by $\widetilde{x}^{n,k} = \widetilde{x}^{nN+k}$ for $n \geq 0$ and $0 \leq k \leq N-1$. We thus consider the following dynamical system

$$\left\{\begin{array}{l} \dfrac{\widetilde{x}_-^{n,k+1} - \widetilde{x}_+^{n,k}}{\Delta \mathrm{t}} = A\Big(\dfrac{\widetilde{x}_-^{n,k+1} + \widetilde{x}_+^{n,k}}{2}\Big), \qquad n \geq 0, \qquad 0 \leq k \leq N-1 \\ \dfrac{\widetilde{x}_+^{n,k+1} - \widetilde{x}_-^{n,k+1}}{\Delta \mathrm{t}} = \nu_{\Delta \mathrm{t}} A^2 \widetilde{x}_+^{n,k+1} - \delta_{k,N-1}\gamma H^* H \widetilde{x}_+^{n,k+1}, \qquad n \geq 0, \qquad 0 \leq k \leq N-1 \\ \widetilde{x}_+^{n+1,0} = \widetilde{x}_+^{n,N}, \qquad \widetilde{x}_-^{n+1,0} = \widetilde{x}_-^{n,N}, \end{array}\right. \tag{59}$$

where

$$\delta_{k,j} = \left\{\begin{array}{ll} 1, & \text{if } k = j \\ 0, & \text{otherwise.} \end{array}\right.$$

We denote $\widetilde{E}^{n,k} = \dfrac{1}{2}\left\|\widetilde{x}_+^{n,k}\right\|^2$ the energy associated to the state of system (59). First of all, we prove the following energy identity.

Proposition A.1. *Let $n_1 \leq n_2$ be two positive integers and let $0 \leq k_1,\ k_2 \leq N-1$ be two other integers. If $\Delta \mathrm{t}$ is small enough, then the quantity $\widetilde{E}^{n,k}$ satisfies the following energy inequality*

$$\widetilde{E}^{n_2,k_2} + \gamma \Delta t \sum_{j=n_1+\delta_{k_1,0}}^{n_2} \left\|H\widetilde{x}_+^{j,0}\right\|^2 + \Delta \mathrm{t}\nu_{\Delta \mathrm{t}} \sum_{[i,j]=[k_1,n_1]}^{[k_2,n_2]} \left\|A\widetilde{x}_+^{j,i}\right\|^2 + \frac{\Delta \mathrm{t}}{4} \sum_{[i,j]=[k_1,n_1]}^{[k_2,n_2]} \Delta \mathrm{t}\nu_{\Delta \mathrm{t}}^2 \left\|A^2\widetilde{x}_+^{j,i}\right\|^2 \leq \widetilde{E}^{n_1,k_1}. \tag{60}$$

Proof. Taking the inner product in $\mathcal{X}$ of the first line in (59) with $\dfrac{\widetilde{x}_-^{n,k+1}+\widetilde{x}_+^{n,k}}{2}$ and the inner product of the second line in (59) with $\dfrac{\widetilde{x}_+^{n,k+1}+\widetilde{x}_-^{n,k+1}}{2}$, we obtain

$$\begin{cases} \dfrac{\widetilde{E}_-^{n,k+1}-\widetilde{E}_+^{n,k}}{\Delta \mathrm{t}}=0, \qquad n>0, \qquad 0\le k\le N-1 \\ \dfrac{\widetilde{E}_+^{n,k+1}-\widetilde{E}_-^{n,k+1}}{\Delta \mathrm{t}}=\dfrac{\nu_{\Delta \mathrm{t}}}{2}(\widetilde{x}_+^{n,k+1}+\widetilde{x}_-^{n,k+1},A^2\widetilde{x}_+^{n,k+1})-\dfrac{\delta_{k,N-1}\gamma}{2}(\widetilde{x}_+^{n,k+1}+\widetilde{x}_-^{n,k+1},H^*H\widetilde{x}_+^{n,k+1}),\ n>0,\ 0\le k\le N-1 \\ \widetilde{E}_+^{n+1,0}=\widetilde{E}_+^{n,N}, \qquad \widetilde{E}_-^{n+1,0}=\widetilde{E}_-^{n,N}. \end{cases}$$

If $k\neq N-1$, the second relation in the above system reads as

$$\frac{\widetilde{E}_+^{n,k+1}-\widetilde{E}_-^{n,k+1}}{\Delta \mathrm{t}}=-\frac{\nu_{\Delta \mathrm{t}}}{2}\left\|A\widetilde{x}_+^{n,k+1}\right\|^2+\frac{\nu_{\Delta \mathrm{t}}}{2}(\widetilde{x}_-^{n,k+1},A^2\widetilde{x}_+^{n,k+1}).$$

Or, in this case, $\widetilde{x}_-^{n,k+1}=\widetilde{x}_+^{n,k+1}-\Delta \mathrm{t}\nu_{\Delta \mathrm{t}}A^2\widetilde{x}_+^{n,k+1}$, hence, we obtain

$$\frac{\widetilde{E}_+^{n,k+1}-\widetilde{E}_-^{n,k+1}}{\Delta \mathrm{t}}=-\nu_{\Delta \mathrm{t}}\left\|A\widetilde{x}_+^{n,k+1}\right\|^2-\frac{\Delta \mathrm{t}\nu_{\Delta \mathrm{t}}^2}{2}\left\|A^2\widetilde{x}_+^{n,k+1}\right\|^2. \tag{61}$$

If $k=N-1$, then the same relation becomes

$$\frac{\widetilde{E}_+^{n+1,0}-\widetilde{E}_-^{n+1,0}}{\Delta \mathrm{t}}=-\frac{\nu_{\Delta \mathrm{t}}}{2}\left\|A\widetilde{x}_+^{n+1,0}\right\|^2-\frac{\gamma}{2}\|H\widetilde{x}_+^{n+1,0}\|^2+\frac{1}{2}(\widetilde{x}_-^{n+1,0},(\nu_{\Delta \mathrm{t}}A^2-\gamma H^*H)\widetilde{x}_+^{n+1,0}).$$

Or, in this case, $\widetilde{x}_-^{n+1,0}=\widetilde{x}_+^{n+1,0}-\Delta \mathrm{t}(\nu_{\Delta \mathrm{t}}A^2-\gamma H^*H)\widetilde{x}_+^{n+1,0}$, hence, we obtain

$$\begin{aligned} \frac{\widetilde{E}_+^{n+1,0}-\widetilde{E}_-^{n+1,0}}{\Delta \mathrm{t}} &= -\nu_{\Delta \mathrm{t}}\left\|A\widetilde{x}_+^{n+1,0}\right\|^2-\gamma\|H\widetilde{x}_+^{n+1,0}\|^2-\frac{\Delta \mathrm{t}}{2}\left\|(\nu_{\Delta \mathrm{t}}A^2-\gamma H^*H)\widetilde{x}_+^{n+1,0}\right\|^2. \\ &\le -\nu_{\Delta \mathrm{t}}\left\|A\widetilde{x}_+^{n+1,0}\right\|^2-\gamma\|H\widetilde{x}_+^{n+1,0}\|^2-\frac{\Delta \mathrm{t}}{2}\nu_{\Delta \mathrm{t}}^2\|A^2\widetilde{x}_+^{n+1,0}\|-\frac{\Delta \mathrm{t}}{2}\gamma^2\|H^*H\widetilde{x}_+^{n+1,0}\|^2 \\ &\quad+\frac{\Delta \mathrm{t}}{2}2\nu_{\Delta \mathrm{t}}\gamma\|A^2\widetilde{x}_+^{n+10}\|\|H^*H\widetilde{x}_+^{n+1,0}\| \\ &\le -\nu_{\Delta \mathrm{t}}\left\|A\widetilde{x}_+^{n+1,0}\right\|^2-\gamma\|H\widetilde{x}_+^{n+1,0}\|^2-\frac{\Delta \mathrm{t}}{2}\nu_{\Delta \mathrm{t}}^2\|A^2\widetilde{x}_+^{n+1,0}\|-\frac{\Delta \mathrm{t}}{2}\gamma^2\|H^*H\widetilde{x}_+^{n+1,0}\|^2 \\ &\quad+\frac{\Delta \mathrm{t}}{2}\gamma\frac{1}{\varepsilon}\nu_{\Delta \mathrm{t}}^2\|A\widetilde{x}_+^{n+1,0}\|^2+\frac{\Delta \mathrm{t}}{2}\gamma\varepsilon\|H^*H\widetilde{x}_+^{n+1,0}\|^2. \end{aligned}$$

Choosing now $\varepsilon=2\gamma$ and using the boundedness of the operator H the above inequality becomes

$$\frac{\widetilde{E}_+^{n+1,0}-\widetilde{E}_-^{n+1,0}}{\Delta \mathrm{t}}\le -\nu_{\Delta \mathrm{t}}\left\|A\widetilde{x}_+^{n+1,0}\right\|^2-\gamma\|H\widetilde{x}_+^{n+1,0}\|^2-\frac{\Delta \mathrm{t}\nu_{\Delta \mathrm{t}}^2}{4}\left\|A^2\widetilde{x}_+^{n,k+1}\right\|^2-\gamma\left(1-\frac{\Delta \mathrm{t}}{2}\gamma K_{H^*}^2\right)\|H\widetilde{x}_+^{n+1,0}\|^2 \tag{62}$$

Combining (61) and (62) from $[k_1,n_1]$ to $[k_2,n_2]$ and taking $\Delta \mathrm{t}\le\dfrac{2}{\gamma K_{H^*}^2}$, we obtain (60), with $K_{H^*}=\|H^*\|_{\mathcal{L}(\mathcal{Z},\mathcal{X})}$. □

In what follows we prove the following result.

Theorem A.2. *If $\gamma = \nu_{\Delta t} = 0$, then for every $\delta > 0$ there exists $T_\delta > 0$ and $k_{T,\delta} > 0$ such that, for every $T > T_\delta$, the solution $\widetilde{x}_+$ of system (59) satisfies*

$$k_{T,\delta} \left\| \widetilde{x}_+^{0,0} \right\|^2 \leq \Delta t \sum_{n\Delta T \in [0,T]} \left\| H \widetilde{x}_+^{n,0} \right\|^2, \qquad \widetilde{x}_+^{0,0} \in \mathcal{C}_{\delta/\Delta T},$$

where $\mathcal{C}_{\delta/\Delta T} = span\{\Phi_j \text{ such that } |\lambda_j| \leq \frac{\delta}{\Delta T}\}$ and (Φ_j) is an orthonormal basis formed by the eigenfunctions of the operator A corresponding to the eigenvalues (λ_j).

Proof. Here we want to verify that the conditions of Theorem 3.1 in [12] are fulfilled. We introduce

$$A_{n,k+1|k} = \left(\mathbb{1} - \frac{1}{2} \Delta t A \right)^{-1} \left(\mathbb{1} + \frac{1}{2} \Delta t A \right),$$

the transition operator such that

$$\widetilde{x}_+^{n,k+1} = A_{n,k+1|k} \widetilde{x}_+^{n,k}.$$

This transition operator is associated with a conservative system and there exists $h : (-R, R) \mapsto [-\pi, \pi]$ a smooth strictly increasing function, with $R \in (0, \infty]$, i.e.

$$|h(\eta)| < \pi, \quad \text{and for every} \quad \delta < R, \quad \inf\{|h'|(\eta), \quad |\eta| \leq \delta\} > 0.$$

Moreover, $\frac{h(\eta)}{\eta} \to 1$ and if, (μ_j, Φ_j) are the eigenvalues and the corresponding eigenvectors of A, i.e. $(A\Phi_j = i\mu_j \Phi_j)$, we have

$$A_{n,k+1|k} \Phi_j = \exp(i \lambda_{j,\Delta t} \Delta t) \Phi_j \text{ with } \lambda_{j,\Delta t} = \frac{1}{\Delta t} h(\mu_j \Delta t).$$

In fact, we have the following lemma.

Lemma A.3. *The function h is defined by $h(\eta) = 2 \arctan(\frac{\eta}{2})$.*

Proof of Lemma. We seek for $\lambda_{j,\Delta t}$ such that

$$\frac{1 + i \frac{\Delta t}{2} \mu_j}{1 - i \frac{\Delta t}{2} \mu_j} = \exp(i \lambda_{j,\Delta t} \Delta t).$$

Hence by introducing the complex $\alpha = 1 - i \frac{\Delta t}{2} \mu_j$, we look for

$$\frac{\alpha}{\overline{\alpha}} = \exp(-i \lambda_{j,\Delta t} \Delta t) \Rightarrow \lambda_{j,\Delta t} = \frac{2}{\Delta t} \arg(\alpha) = \frac{2}{\Delta t} \operatorname{atan}\left(\frac{\mu_j \Delta t}{2}\right).$$

□

We now introduce the conservative transition operator

$$A_{n+1|n,0} = \left(\left(\mathbb{1} - \frac{1}{2} \Delta t A \right)^{-1} \left(\mathbb{1} + \frac{1}{2} \Delta t A \right) \right)^N,$$

such that

$$\widetilde{x}_+^{n+1,0} = A_{n+1|n,0} \widetilde{x}_+^{n,0},$$

which is the transition operator from one time-step where observations are available to the other. By introducing this time $h_N : \eta \mapsto N h(\frac{\eta}{N})$, we have

$$A_{n+1|n,0} \Phi_j = \exp(i \lambda_{j,\Delta T} \Delta T) \Phi_j \text{ with } \lambda_{j,\Delta T} = \lambda_{j,\Delta t} \frac{1}{\Delta T} h_N(\mu_j \Delta T).$$

Therefore, h_N is also a smooth strictly increasing function satisfying $\frac{h_N(\eta)}{\eta} \to 1$. Hence, the hypotheses of Theorem 3.1 in [12] are satisfied for the transition operator $A_{n+1|n,0}$, which implies that for every $\delta > 0$ there exist $T_\delta > 0$ and $k'_{T,\delta} > 0$ such that, for every $T > T_\delta$, the solution $\widetilde{x}$ of system (59) satisfies

$$k'_{T,\delta} \left\|\widetilde{x}_+^{0,0}\right\|^2 \le \Delta T \sum_{n\Delta T\in[0,T]} \left\| H \frac{\widetilde{x}_+^{n+1,0} + \widetilde{x}_+^{n,0}}{2} \right\|^2, \qquad \widetilde{x}_+^{0,0} \in \mathcal{C}_{\delta/\Delta T}.$$

Recalling that $\Delta T = N\Delta t$, there exists $k_{T,\delta} > 0$ such that

$$k_{T,\delta} \left\|\widetilde{x}_+^{0,0}\right\|^2 \le \Delta t \sum_{n\Delta T\in[0,T]} \left\| H\widetilde{x}_+^{n,0} \right\|^2, \qquad \widetilde{x}_+^{0,0} \in \mathcal{C}_{\delta/\Delta T}.$$

□

The main result of this appendix is the following theorem.

Theorem A.4. *There exist two positive constants μ_0 and ν_0 such that the energy associated to system (59) satisfies*

$$\widetilde{E}^{n,k} \le \mu_0 \widetilde{E}^{n,0} \exp(-\nu_0 k\Delta t), \qquad n \ge 0, \qquad 0 \le k < N.$$

Before we start the proof of Theorem A.4, we state the following lemma.

Lemma A.5. *If $\gamma = 0$ and $\nu_{\Delta t} = \Delta t^2$, then there exist a time $T > 0$ and a positive constant c_T such that the solution $\widetilde{x}$ of (59) satisfies*

$$c_T \|\widetilde{x}_+^{0,0}\|^2 \le 2\Delta t \sum_{n=0}^{n_1} \|H\widetilde{x}_+^{n,0}\|^2 + 2\Delta t \nu_{\Delta t} \sum_{[n,k]=[0,0]}^{[n_1,k_1]} \left\|A\widetilde{x}_+^{n,k}\right\|^2 + \frac{\Delta t}{2} \sum_{[n,k]=[0,0]}^{[n_1,k_1]} \Delta t \nu_{\Delta t}^2 \left\|A^2\widetilde{x}_+^{n,k}\right\|^2, \tag{63}$$

where $n_1 = \lfloor T/\Delta T \rfloor$ and $k_1 = \lfloor T/\Delta t \rfloor - n_1 N$.

Since our problem differs to the one considered in [13] only in the low frequency part and this part is handled by Theorem A.2, the proof of the Lemma A.5 is identical to the one of Lemma 2.4 in [13] and it is omitted here.

Proof of Theorem A.4. We follow the ideas in [13]. Therefore, we write the solution $\widetilde{x}^{n,k}$ of (59) as the sum of the solution $z^{n,k}$ of system (59) with $\gamma = 0$ and $w^{n,k}$, the solution of the following system:

$$\begin{cases} \dfrac{w_-^{n,k+1} - w_+^{n,k}}{\Delta t} = A\Big(\dfrac{w_-^{n,k+1} + w_+^{n,k}}{2}\Big), \qquad n \ge 0, \qquad 0 \le k \le N-1 \\ \dfrac{w_+^{n,k+1} - w_-^{n,k+1}}{\Delta t} = \nu_{\Delta t} A^2 w_+^{n,k+1} - \delta_{k,N-1}\gamma H^* H\widetilde{x}_+^{n,k+1}, \qquad n \ge 0, \qquad 0 \le k \le N-1 \\ w_+^{n+1,0} = w_+^{n,N}, \qquad w_-^{n+1,0} = w_-^{n,N}, \qquad n \ge 0 \\ w^{0,0} = 0. \end{cases} \tag{64}$$

Applying Lemma A.5 to $z^{n,k} = \widetilde{x}_+^{n,k} - w_+^{n,k}$ with $z^{0,0} = \widetilde{x}_+^{0,0}$, we obtain

$$\begin{aligned} c_T \|\widetilde{x}_+^{0,0}\|^2 \le\ & 2\left(2\Delta t \sum_{n=0}^{n_1} \|H\widetilde{x}_+^{n,0}\|^2 + 2\Delta t \nu_{\Delta t} \sum_{[n,k]=[0,0]}^{[n_1,k_1]} \left\|A\widetilde{x}_+^{n,k}\right\|^2 + \frac{\Delta t}{2} \sum_{[n,k]=[0,0]}^{[n_1,k_1]} \Delta t \nu_{\Delta t}^2 \left\|A^2\widetilde{x}_+^{n,k}\right\|^2\right) \\ & 2\left(2\Delta t \sum_{n=0}^{n_1} \|Hw_+^{n,0}\|^2 + 2\Delta t \nu_{\Delta t} \sum_{[n,k]=[0,0]}^{[n_1,k_1]} \left\|Aw_+^{n,k}\right\|^2 + \frac{\Delta t}{2} \sum_{[n,k]=[0,0]}^{[n_1,k_1]} \Delta t \nu_{\Delta t}^2 \left\|A^2 w_+^{n,k}\right\|^2\right). \end{aligned} \tag{65}$$

We want now to bound the terms in w in the above inequality by some terms in $\widetilde{x}$. In order to do this, we multiply the first relation in (64) by $\frac{1}{2}(w_-^{n,k+1}+w_+^{n,k})$ and by $w_+^{n,k+1}+w_-^{n,k+1}$ the second. We obtain

$$\begin{cases} \|w_-^{n,k+1}\| = \|w_+^{n,k}\|, \\ \|w_+^{n,k+1}\|^2 = \|w_-^{n,k+1}\|^2 + \Delta t \nu_{\Delta t}(w_+^{n,k+1}+w_-^{n,k+1}, A^2 w_+^{n,k+1}) - \Delta t \delta_{k,N-1}\gamma(w_+^{n,k+1}+w_-^{n,k+1}, H^*H\widetilde{x}_+^{n,k+1}). \end{cases} \tag{66}$$

If $k \neq N-1$, the second relation in the above system reads as in Proposition 3.7

$$\|w_+^{n,k+1}\|^2 = \|w_-^{n,k+1}\|^2 - \nu_{\Delta t}\Delta t \left\|Aw_+^{n,k+1}\right\|^2 - \frac{\Delta t^2 \nu_{\Delta t}^2}{2}\left\|A^2 w_+^{n,k+1}\right\|^2. \tag{67}$$

If $k = N-1$, we get

$$\begin{aligned}\|w_+^{n+1,0}\|^2 = \|w_-^{n+1,0}\|^2 - \nu_{\Delta t}\Delta t \left\|Aw_+^{n+1,0}\right\|^2 - \gamma\Delta t(Hw_+^{n+1,0}, H\widetilde{x}_+^{n+1,0}) \\ + \Delta t(w_-^{n+1,0}, \nu_{\Delta t}A^2 w_+^{n+1,0} - \gamma H^*H\widetilde{x}_+^{n+1,0})\end{aligned}$$

with here

$$w_-^{n+1,0} + \Delta t(\nu_{\Delta t}A^2 w_+^{n+1,0} + \gamma H^*H\widetilde{x}_+^{n+1,0}) = w_+^{n+1,0}.$$

Therefore, we obtain

$$\begin{aligned}\|w_+^{n+1,0}\|^2 &= \|w_-^{n+1,0}\|^2 - 2\nu_{\Delta t}\Delta t\left\|Aw_+^{n+1,0}\right\|^2 - 2\gamma\Delta t(Hw_+^{n+1,0}, H\widetilde{x}_+^{n+1,0}) - \Delta t^2\|\nu_{\Delta t}A^2 w_+^{n+1,0} - \gamma H^*H\widetilde{x}_+^{n+1,0}\|^2 \\ &\le \|w_-^{n+1,0}\|^2 - 2\nu_{\Delta t}\Delta t\left\|Aw_+^{n+1,0}\right\|^2 - 2\gamma\Delta t(Hw_+^{n+1,0}, H\widetilde{x}_+^{n+1,0}) \\ &\quad - \Delta t^2\nu_{\Delta t}^2\|A^2 w_+^{n+1,0}\|^2 - \Delta t^2\gamma^2\|H^*H\widetilde{x}_+^{n+1,0}\|^2 + 2\Delta t^2\nu_{\Delta t}\gamma\|A^2 w_+^{n+1,k}\|\|H^*H\widetilde{x}_+^{n+1,0}\| \\ &\le \|w_-^{n+1,0}\|^2 - 2\nu_{\Delta t}\Delta t\left\|Aw_+^{n+1,0}\right\|^2 - 2\gamma\Delta t(Hw_+^{n+1,0}, H\widetilde{x}_+^{n+1,0}) \\ &\quad - \Delta t^2\nu_{\Delta t}^2\|A^2 w_+^{n+1,0}\|^2 - \Delta t^2\gamma^2\|H^*H\widetilde{x}_+^{n+1,0}\|^2 + \Delta t^2\gamma\frac{1}{\varepsilon}\nu_{\Delta t}^2\|A^2 w_+^{n+1,0}\|^2 + \Delta t^2\gamma\varepsilon\|H^*H\widetilde{x}_+^{n+1,0}\|^2.\end{aligned}$$

Choosing $\varepsilon = 2\gamma$ in the above Young inequality, we finally obtain

$$\|w_+^{n+1,0}\|^2 + 2\nu_{\Delta t}\Delta t\left\|Aw_+^{n+1,0}\right\|^2 + 2\gamma\Delta t(Hw_+^{n+1,0}, H\widetilde{x}_+^{n+1,0}) + \frac{\Delta t^2\nu_{\Delta t}^2}{2}\left\|A^2 w_+^{n,k+1}\right\|^2 \le \|w_-^{n+1,0}\|^2 + \Delta t^2\gamma^2\|H^*H\widetilde{x}^{n+1,0}\|^2, \tag{68}$$

leading – since H is bounded – to

$$\begin{aligned}\|w_+^{n+1,0}\|^2 + 2\nu_{\Delta t}\Delta t\left\|Aw_+^{n+1,0}\right\|^2 + \frac{\Delta t^2\nu_{\Delta t}^2}{2}\left\|A^2 w_+^{n,k+1}\right\|^2 \le \|w_+^{n,N}\|^2 \\ + \gamma\Delta t\|Hw_+^{n+1,0}\|^2 + \Delta t\gamma(1+\Delta t\gamma K_{H^*}^2)\|H\widetilde{x}_+^{n+1,0}\|^2,\end{aligned}$$

where K_{H^*} is the norm of the operator H^*. Combining the above inequality and (67) for $[n,k]$ from $[0,0]$ to $[n_1,k_1]$, for Δt small enough, there exists a constant c such that

$$\begin{aligned}\sum_{[n,k]=[0,0]}^{[n_1,k_1]}\|w_+^{n,k}\|^2 + 2\gamma\Delta t\sum_{n=0}^{n_1}\|Hw_+^{n,0}\|^2 + 2\Delta t\nu_{\Delta t}\sum_{[n,k]=[0,0]}^{[n_1,k_1]}\left\|Aw_+^{n,k}\right\|^2 + \frac{\Delta t}{2}\sum_{[n,k]=[0,0]}^{[n_1,k_1]}\Delta t\nu_{\Delta t}^2\left\|A^2 w_+^{n,k}\right\|^2 \\ \le c\gamma\Delta t\sum_{n=0}^{n_1}(\|w_+^{n,0}\|^2 + \|H\widetilde{x}_+^{n,0}\|^2).\end{aligned}$$

Following the same steps as in [13] – with Gronwall's inequality – we get that there exists a constant c' such that

$$\sum_{[n,k]=[0,0]}^{[n_1,k_1]} \|w^{n,k}\|^2 + 2\gamma\Delta t \sum_{n=0}^{n_1} \|Hw_+^{n,0}\|^2 + 2\Delta t\nu_{\Delta t} \sum_{[n,k]=[0,0]}^{[n_1,k_1]} \left\|Aw_+^{n,k}\right\|^2 + \frac{\Delta t}{2} \sum_{[n,k]=[0,0]}^{[n_1,k_1]} \Delta t\nu_{\Delta t}^2 \left\|A^2 w_+^{n,k}\right\|^2 \leq c' \sum_{n=0}^{n_1} \|H\widetilde{x}_+^{n,0}\|^2,$$

which, applying (65), gives the existence of a positive constant c'' such that

$$c''\|\widetilde{x}_+^{0,0}\|^2 \leq \gamma\Delta t \sum_{n=0}^{n_1} \|H\widetilde{x}_+^{n,0}\|^2 + \Delta t\nu_{\Delta t} \sum_{[n,k]=[0,0]}^{[n_1,k_1]} \left\|A\widetilde{x}_+^{n,k}\right\|^2 + \frac{\Delta t}{4} \sum_{[n,k]=[0,0]}^{[n_1,k_1]} \Delta t\nu_{\Delta t}^2 \left\|A^2\widetilde{x}_+^{n,k}\right\|^2,$$

where we recall that $\nu_{\Delta t} = \Delta t^2$. The last inequality, combined with the energy inequality (60), gives finally the existence of a constant $c''' \in (0,1)$ such that

$$\|\widetilde{x}_+^{n_1,k_1}\|^2 \leq c'''\|\widetilde{x}_+^{0,0}\|^2$$

which allows us to conclude the proof as in [13].

□

References

[1] R.A. Anthes. Data assimilation and initialization of hurricane prediction model. *Journal of the Atmospheric Sciences*, 31:702–719, 1974.

[2] H T Banks, K. Ito, and C. Wang. Exponentially stable approximations of weakly damped wave equations. In *Estimation and control of distributed parameter systems (Vorau, 1990)*, pages 1–33. Birkhäuser, Basel, 1991.

[3] C. Bardos, G. Lebeau, and J. Rauch. Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary. *SIAM J. Control Optim.*, 30(5):1024–1065, September 1992.

[4] A. Bensoussan. *Filtrage optimal des systèmes linéaires*. Dunod, 1971.

[5] J. Blum, F. X. LeDimet, and I. N. Navon. Data assimilation for geophysical fluids. *Handbook of Numerical Analysis: Computational Methods for the Atmosphere and the Oceans. Elsevier, Amsterdam*, 14:377–434, 2008.

[6] R. Chabiniok, P. Moireau, P.-F. Lesault, A. Rahmouni, J.-F. Deux, and D. Chapelle. Trials on tissue contractility estimation from cardiac cine-MRI using a biomechanical heart model. In *Proceedings of FIMH'11, LNCS 6666*, pages 304–313. Springer, 2011.

[7] D. Chapelle, N. Cîndea, M. De Buhan, and P. Moireau. Exponential convergence of an observer based on partial field measurements for the wave equation. *Math. Probl. Eng.*, 2012:12, October 2012.

[8] D. Chapelle, N. Cîndea, and P. Moireau. Improving convergence in numerical analysis using observers. The wave-like equation case. *Math. Models Methods Appl. Sci.*, 2012.

[9] D. Chapelle, M. Fragu, V. Mallet, and P. Moireau. Fundamental principles of data assimilation underlying the Verdandi library: applications to biophysical model personalization within euHeart. *Med. Biol. Eng. Comput.*, 51(11):1221–1233, 2013.

[10] S. Cox and E. Zuazua. The rate at which energy decays in a damped string. *Comm. Partial Differential Equations*, 19:213–243, 1994.

[11] M. Daoulatli, B. Dehman, and M. Khenissi. Local energy decay for the elastic system with nonlinear damping in an exterior domain. *SIAM J. Control Optim.*, 48(8):5254–5275, 2010.

[12] S. Ervedoza, C. Zheng, and E. Zuazua. On the observability of time-discrete conservative linear systems. *Journal of functional Analysis*, 254(12):3037–3078, 2008.

[13] S. Ervedoza and E. Zuazua. Uniformly exponentially stable approximations for a class of damped systems. *J. Math. Pures Appl.*, 91:20–48, 2009.

[14] Sylvain Ervedoza. Spectral conditions for admissibility and observability of wave systems: applications to finite element schemes. *Numer Math*, 113(3):377–415, January 2009.

[15] G. Evensen. *Data Assimilation – The Ensemble Kalman Filter*. Springer Verlag, 2007.

[16] G. Haine and K. Ramdani. Reconstructing initial data using observers: error analysis of the semi-discrete and fully discrete approximations. *Numer. Math.*, 120(2):307–343, 2012.

[17] F.M. Hante, M. Sigalotti, and M. Tucsnak. On conditions for asymptotic stability of dissipative infinite-dimensional systems with intermittent damping. *J. Differential Equations*, 252(10):5569–5593, 2012.
[18] A. Haraux and E. Zuazua. Decay estimates for some semilinear damped hyperbolic problems. *Arch. Rational Mech. Anal.*, 100(2):191–206, 1988.
[19] J.E. Hoke and R.A. Anthes. The initialization of numerical models by a dynamic-initialization technique(fluid flow models for wind forecasting). *Monthly Weather Review*, 104:1551–1556, 1976.
[20] R.E. Kalman and R.S. Bucy. New results in linear filtering and prediction theory. *J. Basic Eng.*, 83(3):95–108, 1961.
[21] S. Lakshmivarahan and J.M. Lewis. Nudging methods: A critical overview. In S.K. Park and L. Xu, editors, *Data Assimilation for Atmospheric, Oceanic, and Hydrologic Applications*, volume XVIII. Springer, 2008.
[22] F.-X. Le Dimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A*, 38(2):97–110, 2010.
[23] X.-D. Li and C.-Z. Xu. Infinite-dimensional Luenberger-like observers for a rotating body-beam system. *Systems & Control Letters*, 60(2):138–145, 2011.
[24] K. Liu. Locally distributed control and damping for the conservative systems. *SIAM J. Control Optim.*, 35(5):1574–1590, September 1997.
[25] D. G. Luenberger. An introduction to observers. *IEEE T. Automat. Contr.*, 16(6):596–602, 1971.
[26] P. Moireau, D. Chapelle, and P. LeTallec. Joint state and parameter estimation for distributed mechanical systems. *Comput. Methods Appl. Mech. Engrg.*, 197:659–677, 2007.
[27] P. Moireau, D. Chapelle, and P. LeTallec. Filtering for distributed mechanical systems using position measurements: Perspectives in medical imaging. *Inverse Probl.*, 25(3):035010, 2009.
[28] I.M. Navon. Data assimilation for numerical weather prediction : a review. In S.K. Park and L. Xu, editors, *Data Assimilation for Atmospheric, Oceanic, and Hydrologic Applications*, volume XVIII. Springer, 2009.
[29] N.K. Nichols. Mathematical concepts of data assimilation. In William Lahoz, Boris Khattatov, and Richard Menard, editors, *Data Assimilation*, pages 13–39. Springer Berlin Heidelberg, 2010.
[30] K. Ramdani, T. Takahashi, and M. Tucsnak. Uniformly exponentially stable approximations for a class of second order evolution equations—application to LQR problems. *ESAIM Control Optim. Calc. Var.*, 13(3):503–527, 2007.
[31] K. Ramdani, M. Tucsnak, and G. Weiss. Recovering the initial state of an infinite-dimensional system using observers. *Automatica*, pages 1616–1625, 2012.
[32] D. Simon. *Optimal state estimation: Kalman, H_∞ and nonlinear approaches*. Wiley-Interscience, 2006.
[33] Louis T Tebou and Enrique Zuazua. Uniform boundary stabilization of the finite difference space discretization of the 1-d wave equation. *Adv Comput Math*, 26(1-3):337–365, January 2007.
[34] D. Tuan Pham, J. Verron, and L. Gourdeau. Singular evolutive kalman filters for data assimilation in oceanography. *C. R. Acad. Sci. Paris*, pages 255–260, 1997.
[35] M. Tucsnak and G. Weiss. *Observation and control for operator semigroups*. Birkhäuser Basel, 2009.
[36] X. Zhang, C. Zheng, and E. Zuazua. Exact controllability of the time discrete wave equation: a multiplier approach. *Discret. Contin. Dyn. S.*, pages 229–245, 2007.
[37] E. Zuazua. Propagation, observation, and control of waves approximated by finite difference methods. *SIAM Review*, 47(2):197–243, 2005.