AN INVERSE BACKWARD PROBLEM FOR A HEAT EQUATION WITH A MEMORY TERM, SOLVED BY A DEEP LEARNING METHOD

SOUFIANE ABID, KHALID ATIFI AND EL-HASSAN ESSOUFI

Abstract. Our goal in this paper, is the identification of initial condition in a heat equation that contains a memory term from final data. To this aim, we first establish the well-posedness of the direct problem. Then we prove the continuity and the G-derivability of the cost function. Finally we validate the results numerically by using a deep neural network. Our algorithm is meshfree.

1. Introduction

Inverse problems for parabolic PDEs with memory terms can be found in many scientific and engineering applications, and it has become a very active and succeful research area in recent years. This kind of equations occur naturally in geophysics, oil exploration, optical instrumentation, and many other fields where the interior of an object must be imaged by measuring the field in the domains available in real measurement.
In particular in the theory of phase transition for materials with memory, we find several models that have been recently studied from several points of view: for example, as direct problems in Hilbert spaces and as dynamical systems. We list below, without pretension of completeness, some articles and books in which one can find some models and results involving the heat equation : [5], [6], [7], [8], [9], [10], [11].

Extensive research has been conducted on various theoretical and numerical aspects of inverse problems, such as existence, uniqueness, stability and validation of results by numerical simulations, etc...

In our present paper we will adopt a new approach, in the theoretical side we will ensure the existence and uniqueness by showing some criteria verified by the cost function and use deep-learning as numerical aspect to validate our results.

Deep neural networks are machine learning models that have achieved remarkable success in a number of domains from visual recognition and speech, to natural language processing and robotics [13]. Among recent works, [12] proposes to solve PDEs using a meshfree deep learning algorithm. The method is similar in spirit to the Galerkin method, but with several key changes using ideas from machine learning. The Galerkin method is a widely-used computational method which seeks a reduced-form solution to a PDE as a linear combination of basis functions. The deep learning algorithm, or Deep Galerkin Method (DGM), uses a deep neural network instead of a linear combination of basis functions. The deep neural network

2010 *Mathematics Subject Classification.* ...
Key words and phrases. inverse problem, memory term, optimization, deep learning.

is trained to satisfy the differential operator, initial condition, and boundary conditions using stochastic gradient descent at randomly sampled spatial points. By randomly sampling spatial points, the author avoid the need to form a mesh and instead convert the PDE problem into a machine learning problem.

In the present paper, we study the inverse problem of determining the initial state in a singular parabolic equation with a memory term from the theoretical analysis and numerical computation angles. More precisely, we consider the following problem:

$$\begin{cases} \partial_t u + A(u) = f, & \text{in } Q, \\ u(x,t) = 0, & x \in \partial\Omega,\ t \in]0;T[, \\ u(x,0) = u_0(x), & x \in \Omega. \end{cases} \tag{1.1}$$

A is the operator defined as

$$A(u) = -u_{xx}(x,t) - a(x)\int_0^t u(x,s)ds,$$

Where $\Omega := (0,1)$, $Q := \Omega \times (0,T)$, $T > 0$ is an fixed moment of time, $u_0 \in L^2(\Omega)$ is the initial condition, $a \in L^\infty(\Omega)$ is a positive coefficient depends on the space and $f \in L^2(Q)$ is the source term.

In the case where $a(x) = 0$ the problem (1.1) is already treated in [2] and more generally in [3] even in the degenerate case.

Let assume

$$A_{ad} = \{h \in H^1(\Omega) :\ \|h\|_{H^1(\Omega)} \leqslant r\},\ \text{where r is a real strictly positive constant.}$$

Evidently, the set A_{ad} is a bounded, closed, and convex subset of $L^2(\Omega)$.
We have $H^1(\Omega) \underset{compact}{\hookrightarrow} L^2(\Omega)$. Since the set A_{ad} is bounded in $H^1(\Omega)$, then A_{ad} is a compact in $L^2(\Omega)$. Therefore.

Let us define our inverse problem:

Inverse Source Problem (ISP). Let u be the solution to (1.1). Determine the initial state u_0 from the measured data at the final time $u(T,\cdot)$.

Remark 1. It should be mentioned that we do not need the supplement distributed measurements to obtain the numerical solution of the inverse problem.

We treat Problem (ISP) by interpreting its solution as a minimizer of the following problem

$$\text{find } u_0^\star \in A_{ad} \text{ such that } E(u_0^\star) = \min_{u_0 \in A_{ad}} E(u_0), \tag{1.2}$$

where the cost function E is defined as follows

$$E(u_0) = \frac{1}{2T}\left\|u(T) - u^{obs}\right\|^2_{L^2(\Omega)},$$

subject to u is the weak solution of the parabolic problem (1.1) with initial state u_0. $u^{obs} \in L^2(\Omega)$ is the observation data with noise.

The problem (1.2) is ill-posed in the sense of Hadamard, some regularization technique is needed in order to guarantee numerical stability of the computational procedure even with noisy input data. The problem thus consists in minimizing a functional of the form

$$J(u_0) = \frac{1}{2T}\left\|u(T) - u^{obs}\right\|^2_{L^2(\Omega)} + \frac{\varepsilon}{2}\left\|u_0 - u^b\right\|^2_{L^2(\Omega)},$$

here, ε being a small positive regularizing coefficient that provides extra convexity to the functional J. u^b an a priori (background state) knowledge of the state u_0^{exact}. The background error is then defined as: $err = \| u_0^{exact} - u^b \|_2$. u_0^{exact} is called the true state, and is the state to estimate.

2. Well-posedness

Theorem 2.1. . *Assume $u_0 \in L^2(\Omega)$ and $f \in L^2(Q)$, there exists a unique weak solution which solves the problem* (1.1) *such that*

$$u \in C\left([0,T]; L^2(\Omega)\right) \cap L^2\left(0,T; H_0^1(\Omega)\right).$$

and we have the estimate

$$\sup_{t\in[0,T]} \|u(x,t)\|^2_{L^2(\Omega)} + \int_0^T \|u_x(x,t)\|^2_{L^2(\Omega)} dt \leq C_T \left(\|f\|^2_{L^2(Q)} + \|u_0(x)\|^2_{L^2(\Omega)} \right) \tag{2.1}$$

the constant C_1 depending on Ω, and T.$\square$

Proof of Theorem 2.1. the existence and uniqueness of the weak solution of (1.1) is already seen in Proposition 3.1 in [1](with the particular case $\mu = 0$), here we will show only the estimate (2.1).
We multiply the first equation of (1.1) by u and integrate over Ω, wet get

$$\begin{aligned} &\int_0^1 u_t(x,t)\ u(x,t)\ dx - \int_0^1 u_{xx}(x,t)\ u(x,t)\ dx \\ &= \int_0^1 \left(a(x)\ u(x,t) \ \int_0^t u(x,s)\ ds \right)\ dx\ + \int_0^1 f(x,t)\ u(x,t)\ dx. \end{aligned} \tag{2.2}$$

By integration by parts, we obtain

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}\|u(x,t)\|^2_{L^2(\Omega)} + \int_0^1 u_x^2(x,t)\ dx \\ &= \int_0^1 \left(a(x)\ u(x,t) \ \int_0^t u(x,s)\ ds \right)\ dx + \int_0^1 f(x,t)\ u(x,t)\ dx. \end{aligned} \tag{2.3}$$

We have

$$\begin{aligned} &\int_0^1 \left(a(x)\ u(x,t) \int_0^t u(x,s)\ ds \right)\ dx \\ &\leq \left(\int_0^1 (a(x)\ u(x,t))^2\ dx \right)^{\frac{1}{2}} \left(\int_0^1 \left(\int_0^t u(x,s)\ ds \right)^2\ dx \right)^{\frac{1}{2}} \\ &\leq \frac{1}{2}\|a(x)\|^2_{L^\infty(\Omega)}\|u(x,t)\|^2_{L^2(\Omega)} + \frac{T}{2}\int_0^t \|u(x,s)\|^2_{L^2(\Omega)}\ ds, \end{aligned} \tag{2.4}$$

and by the Cauchy-Schwarz inequality we obtain for every $t \in [0,T]$

$$\int_0^1 f(x,t)\ u(x,t)\ dx \leq \frac{1}{2}\|f(x,t)\|^2_{L^2(\Omega)} + \frac{1}{2}\|u(x,t)\|^2_{L^2(\Omega)},$$

by returning to the equation (2.3), we obtain

$$\frac{1}{2}\frac{d}{dt}\|u(x,t)\|^2_{L^2(\Omega)} + \|u_x(x,t)\|^2_{L^2(\Omega)}$$
$$\leq \frac{1}{2}\|f(x,t)\|_{L^2(\Omega)} + \frac{1}{2}\left(1+\|a(x)\|^2_{L^\infty(\Omega)}\right)\|u(x,t)\|^2_{L^2(\Omega)} + \frac{T}{2}\int_0^t\|u(x,s)\|^2_{L^2(\Omega)}\,ds. \tag{2.5}$$

We integrate over $[0,t]$ for all $t \in [0,T]$

$$\|u(x,t)\|^2_{L^2(\Omega)} + 2\int_0^t \|u_x(x,s)\|^2_{L^2(\Omega)}\,ds$$
$$\leq \|f(x,t)\|^2_{L^2(Q)} + \|u_0(x)\|^2_{L^2(\Omega)} + \left(1+\|a(x)\|^2_{L^\infty(\Omega)} + T^2\right)\int_0^t \|u(x,s)\|^2_{L^2(\Omega)}\,ds. \tag{2.6}$$

Since

$$2\int_0^t \|u_x(x,s)\|^2_{L^2(\Omega)}\,ds \geq 0,$$

then

$$\|u(x,t)\|^2_{L^2(\Omega)}$$
$$\leq \|f(x,t)\|^2_{L^2(Q)} + \|u_0(x)\|^2_{L^2(\Omega)} + \left(1+\|a(x)\|^2_{L^\infty(\Omega)} + T^2\right)\int_0^t \|u(x,s)\|^2_{L^2(\Omega)}\,ds.$$

Using Gronwall's inequality, we get

$$\|u(x,t)\|^2_{L^2(\Omega)} \leq \left(1+\|a(x)\|^2_{L^\infty(\Omega)} + T^2\right)e^T\left(\|f(x,t)\|^2_{L^2(Q)} + \|u_0(x)\|^2_{L^2(\Omega)}\right). \tag{2.7}$$

From (2.6) and (2.7), ther exists a constant $M > 0$ such that :

$$\int_0^T \|u_x(x,t)\|^2_{L^2(\Omega)}\,dt \leq M\left(\|f(x,t)\|^2_{L^2(Q)} + \|u_0(x)\|^2_{L^2(\Omega)}\right). \tag{2.8}$$

From (2.7) and (2.8), there exists a constant $C_T > 0$ such that

$$\sup_{t\in[0,T]}\|u(x,t)\|^2_{L^2(\Omega)} + \int_0^T \|u_x(x,t)\|^2_{L^2(\Omega)}dt \leq C_T\left(\|f\|^2_{L^2(Q)} + \|u_0(x)\|^2_{L^2(\Omega)}\right).$$

□

Lemma 2.1. *Let u is the weak solution of* (1.1) *with initial condition u_0. The function*

$$\phi : u_0 \in L^2(\Omega) \longrightarrow u \in C\left([0,T]; L^2(\Omega)\right) \cap L^2\left(0,T; H_0^1(\Omega)\right)$$

is Lipschitz continuous.

An automatic result of the Lemma 2.1 is the following theorem

Theorem 2.2. *Under the same assumptions of the Theorem 2.1 , the functional J is continuous on A_{ad}, and there exist a unique minimizer $u_0^\star \in A_{ad}$, ie*

$$J(u_0^\star) = \min_{u_0\in A_{ad}} J(u_0).$$

Proof of Lemma 2.1. Let $\delta u_0 \in L^2(\Omega)$ be a small perturbation of u_0 such that $u_0 + \delta u_0 \in A_{ad}$.
Consider $\delta u = u^\delta - u$, where u^δ and u are respectively the weak solutions of (1.1)

with initial condition $u_0^\delta = u_0 + \delta u_0$ and u_0. Consequently, δu is the solution of the variational problem :

$$\left\{ \begin{array}{l} \displaystyle\int_0^1 (\delta u)_t \ v \ dx + \int_0^1 (\delta u)_x \ v_x \ dx = \int_0^1 \left(a \ v \int_0^t \delta u \ ds \right) \ dx, \ \forall v \in H_0^1(\Omega), \\ \delta u(0,t) = \delta u(1,t) = 0 \ \ \forall t \in [0,T], \\ \delta u(x,0) = \delta u_0 \ \ \forall x \in \Omega. \end{array} \right. \tag{2.9}$$

Hence, δu is the weak solution of (1.1) with initial condition δu_0 and source term $\delta f = 0$. We apply the estimate in Theorem 2.1, we obtain :
there is a constant $C_T > 0$ such that

$$\sup_{t\in[0,T]} \|\delta u\|^2_{L^2(\Omega)} + \int_0^T \|\delta u(t)\|^2_{L^2(\Omega)} dt \leq C_T \|\delta u_0\|^2_{L^2(\Omega)}.$$

Then

$$\|\delta u\|^2_{C(0,T;L^2(\Omega))} \leq C_T \|\delta u_0\|^2_{L^2(\Omega)},$$

and

$$\|\delta u\|^2_{L^2(0,T;H_0^1)} \leq C_T \|\delta u_0\|^2_{L^2(\Omega)}.$$

This ends the demonstration. □

Proof of Theorem 2.2. The continuity of functional J is deduced from the continuity of the evolution function

$$\phi : u_0 \in L^2(\Omega) \longrightarrow u \in C\left([0,T]; L^2(\Omega)\right) \cap L^2\left(0,T; H_0^1(\Omega)\right),$$

established in the Lemma 2.1.

And since J is continuous, on the compact A_{ad}, then there exist a unique minimizer $u_0^\star \in A_{ad}$ for J. □

Proposition 2.1. *Let u is the weak solution of* (1.1) *with initial condition u_0. The function*

$$\phi : u_0 \in L^2(\Omega) \longrightarrow u \in C\left([0,T]; L^2(\Omega)\right) \cap L^2\left(0,T; H_0^1(\Omega)\right)$$

is G-differentiable, which gives that the functional J is G-derivable on A_{ad}.

Proof of Proposition 2.1. Let δu_0 be a a small amount such that $u_0 + \delta u_0 \in A_{ad}$, we define the function :

$$F'(u_0) := \delta u_0 \in A_{ad} \to \delta u, \tag{2.10}$$

where δu is the solution of the following variational problem :

$$\left\{ \begin{array}{l} \displaystyle\int_0^1 (\delta u)_t \ v \ dx + \int_0^1 (\delta u)_x \ v_x \ dx = \int_0^1 \left(a \ v \int_0^t \delta u \ ds \right) \ dx, \ \forall v \in H_0^1(\Omega), \\ \delta u(0,t) = \delta u(1,t) = 0 \ \ \forall t \in [0,T], \\ \delta u(x,0) = \delta u_0 \ \ \forall x \in \Omega. \end{array} \right. \tag{2.11}$$

We pose

$$\Phi(u_0) = F(u_0 + \delta u_0) - F(u_0) - F'(u_0)\delta u_0. \tag{2.12}$$

We want to show that

$$\Phi(u_0) = O(\delta u_0). \tag{2.13}$$

It is easy to see that Φ verifies the following variational problem :

$$\left\{\begin{array}{l}\displaystyle\int_0^1 \Phi_t\ v\ dx + \int_0^1 \Phi_x\ v_x\ dx = \int_0^1 \left(a\ v \int_0^t \Phi\ ds\right)\ dx,\ \forall v \in H_0^1(\Omega),\\ \Phi(0,t) = \Phi(1,t) = 0\ \ \forall t \in [0,T],\\ \Phi(x,0) = \delta u_0 - (\delta u_0)^2\ \ \forall x \in \Omega.\end{array}\right. \tag{2.14}$$

In the same way used in the proof of continuity, we deduce that

$$\|\Phi\|^2_{C(0,T;L^2(\Omega))} \le C_T \|\delta u_0 - (\delta u_0)^2\|^2_{L^2(\Omega)},$$

and

$$\|\Phi\|^2_{L^2(0,T;H_0^1)} \le C_T \|\delta u_0 - (\delta u_0)^2\|^2_{L^2(\Omega)}.$$

This completes the proof of the proposition. □

3. Stability

. In this section, we will establish the stability of the solution of the inverse problem

Lemma 3.1. *Let $u_0^\star$ be a minimizer of the functional J, then there exists a set of functions $(u^\star, w, u_0)$ such that*

$$\left\{\begin{array}{l} u_t^\star(x,t) - u_{xx}^\star(x,t) = a(x)\displaystyle\int_0^t u^\star(x,s)ds + f(x,t),\ \forall (x,t) \in Q,\\ u^\star(0,t) = u^\star(1,t) = 0\ \forall t \in (0,T),\\ u^\star(x,0) = u_0(x)\ \forall x \in (0,1),\end{array}\right. \tag{3.1}$$

$$\left\{\begin{array}{l} w_t(x,t) - w_{xx}(x,t) = a(x)\displaystyle\int_0^t w(x,s)ds,\ \forall (x,t) \in Q,\\ w(0,t) = w(1,t) = 0\ \forall t \in (0,T),\\ w(x,0) = \kappa(x) - u_0^\star(x)\ \forall x \in (0,1),\end{array}\right. \tag{3.2}$$

and

$$\int_0^T \int_0^1 w(x,T)\ (u^\star(x,T) - \tilde{u}(T))\ dx\ dt + \varepsilon \int_0^1 u_0^\star\ (\kappa - u_0^\star)\ dx \ge 0 \tag{3.3}$$

for any $\kappa \in A_{ad}$

Proof. For any $\kappa \in A_{ad}$ and $0 \le \delta \le 1$, we pose

$$u_0^\delta = (1-\delta)\ u_0^\star + \delta\ \kappa \in A_{ad}.$$

Then there exist a solution u^δ of the equation (1.1) with the initial condition u_0^δ satisfying

$$J_\delta = J(u^\delta) = \frac{1}{2}\int_0^T \|u^\delta(x,T) - \tilde{u}(T)\|^2_{L^2(\Omega)}\ dt + \frac{\varepsilon}{2}\|u_0^\delta\|^2_{L^2(\Omega)}.$$

Now taking Frchet derivative of J_δ with optimal solution $u_0^\star$, we have

$$\frac{dJ_\delta}{d\delta}|_{\delta=0} = \int_0^T \int_0^1 (u^\star(x,T) - \tilde{u})\hat{u}_\delta\ dx\ dt + \varepsilon \int_0^1 u_0^\star(\kappa - u_0^\star)\ dx \ge 0, \tag{3.4}$$

where $\hat{u}_\delta = \dfrac{du}{d\delta}|_{\delta=0}$ the Frchet derivative of u, which verifies the following equation :

$$\left\{\begin{array}{ll} (\hat{u}_\delta)_t(x,t) - (\hat{u}_\delta)_{xx}(x,t) - a(x)\displaystyle\int_0^t \hat{u}_\delta(x,s)\ ds = 0 & (x,t) \in \Omega \times [0,T], \\ \hat{u}_\delta(0,t) = \hat{u}_\delta(1,t) = 0, \quad \forall t \in [0;T], & \\ \hat{u_\delta}(x,0) = \kappa(x) - u_0^\star(x), \quad \forall x \in [0;1]. & \end{array}\right. \tag{3.5}$$

Set $w = \hat{u}_\delta$, then w satisfies :

$$\left\{\begin{array}{l} w_t(x,t) - w_{xx}(x,t) = a(x)\displaystyle\int_0^t w(x,s)ds, \ \forall (x,t) \in Q, \\ w(0,t) = w(1,t) = 0 \ \forall t \in (0,T), \\ w(x,0) = \kappa(x) - u_0^\star(x) \ \forall x \in (0,1), \end{array}\right. \tag{3.6}$$

Combining (3.4) and (3.6), one can easily obtain that

$$\int_0^T \int_0^1 w(x,T)\ (u^\star(x,T) - \tilde{u}(T))\ dx\ dt + \varepsilon \int_0^1 u_0^\star\ (\kappa - u_0^\star)\ dx \geq 0.$$

□

Theorem 3.1. *Suppose that $\tilde{u}_1(T)$ and $\tilde{u}_2(T)$ are two given functions in $L^2(\Omega)$. Let v_1 and v_2 the minimzer of J corresponding to $\tilde{u}_1(T)$ and $\tilde{u}_2(T)$ respectively, If there exists a point $x_0 \in \Omega$ such that $v_1(x_0) = v_2(x_0)$, then we have the following estimate :*

$$\|v_1(x) - v_2(x)\|^2_{L^2(\Omega)} \leq \Lambda \int_0^T \|\tilde{u}_1(T) - \tilde{u}_2(T)\|_{L^2(\Omega)} dt,$$

where the constant Λ only depends on Ω and ε.

Proof. In the estimate (3.3) of Lemma 3.1 we take $\kappa = v_2$ and $u_0^\star = v_1$ and we take $\kappa = u_0^1$ and $u_0^2 = u_0^\star$, we get

$$\int_0^T \int_0^1 w_1(x,T)\ (u_1^\star(x,T) - \tilde{u}_1(T))\ dx\ dt + \varepsilon \int_0^1 v_1\ (v_2 - v_1)\ dx \geq 0, \tag{3.7}$$

and

$$\int_0^T \int_0^1 w_2(x,T)\ (u_2^\star(x,T) - \tilde{u}_2(T))\ dx\ dt + \varepsilon \int_0^1 v_2\ (v_1 - v_2)\ dx \geq 0. \tag{3.8}$$

Where $\{u_i^\star; w_i\}$ $(i = 1,2)$ are solutions of systems 3.1 and 3.2 with initial condition $u_0^\star = v_i$ $(i = 1,2)$, respectively. Setting

$$U = u_1^\star - u_2^\star, \qquad W = w_1 + w_2,$$

Then U and W satisfy by taking $\kappa = v_2$ and $\kappa = v_1$

$$\left\{\begin{array}{l} U_t(x,t) - U_{xx}(x,t) - a(x)\displaystyle\int_0^t U(x,s)ds = 0, \ \forall (x,t) \in Q, \\ U(0,t) = U(1,t) = 0 \ \forall t \in (0,T), \\ U(x,0) = v_1 - v_2 \ \forall x \in (0,1), \end{array}\right. \tag{3.9}$$

$$\left\{\begin{array}{ll} W_t(x,t) - W_{xx}(x,t) - a(x)\displaystyle\int_0^t W(x,s)\ ds = 0, & \forall (x,t) \in Q, \\ W(0,t) = W(1,t) = 0 \ \forall t \in (0,T), & \\ W(x,0) = 0 \ \forall x \in (0,1), & \end{array}\right. \tag{3.10}$$

By the extremum principle we know that (3.10) only has zero solution and thus

$$w_1(x,t) = -w_2(x,t). \tag{3.11}$$

Moreover, w_1 satisfies the following equation

$$\begin{cases} (w_1)_t(x,t) - (w_1)_{xx}(x,t) = a(x) \displaystyle\int_0^t w_1(x,s)ds, \ \forall (x,t) \in Q, \\ w_1(0,t) = w_1(1,t) = 0 \ \forall t \in (0,T), \\ w_1(x,0) = v_2 - v_1 \ \forall x \in (0,1), \end{cases} \tag{3.12}$$

By noticing (3.9) and (3.12) we have

$$U(x,t) = -w_1(x,t). \tag{3.13}$$

From (3.7), (3.8), (3.11) and (3.13) we have

$$\begin{aligned} &\varepsilon \int_0^1 |v_1(x) - v_2(x)|^2 dx \\ &\le \int_0^T \int_0^1 w_1(x,T)(u_1^\star(x,T) - \tilde{u}_1(T)) \ dx \ dt + \int_0^T \int_0^1 w_2(x,T)(u_2^\star(x,T) - \tilde{u}_2(T)) \ dx \ dt \\ &\le \int_0^T \int_0^1 U(x,T) \ w_1(x,t) \ dx \ dt + \int_0^T \int_0^1 (\tilde{u}_2(T) - \tilde{u}_1(T)) \ w_1(x,T) \ dx \ dt \\ &\le - \int_0^T \int_0^1 |w_1(x,t)|^2 \ dx \ dt + \frac{1}{2} \int_0^T \int_0^1 |w_1(x,t)|^2 \ dx \ dt \\ &\quad + \frac{1}{2} \textstyle\int_0^T \int_0^1 |\tilde{u}_1(T) - \tilde{u}_2(T)|^2 \ dx \ dt \\ &\le -\frac{1}{2} \int_0^T \int_0^1 |w_1(x,t)|^2 \ dx \ dt + \frac{1}{2} \int_0^T \int_0^1 |\tilde{u}_1(T) - \tilde{u}_2(T)|^2 \ dx \ dt. \end{aligned} \tag{3.14}$$

Then

$$\|v_1(x) - v_2(x)\|^2_{L^2(\Omega)} \le \Lambda \int_0^T \|\tilde{u}_1(T) - \tilde{u}_2(T)\|^2_{L^2(\Omega)} dt, \tag{3.15}$$

with $\Lambda = \dfrac{1}{2 \ \varepsilon}$ □

Remark 2. From the Teorem 3.1, we can easily deduce that if the final measurements of the systems (1.1) and (3.1) are equal, then the data u_0 can be determined uniquely almost everywhere.

4. Numerical experiments

Consider the objective function:

$$\begin{aligned} J(y) = &\frac{1}{2}\|A(y) - f(x,t)\|^2_{L^2(\Omega)} + \frac{1}{2}\|y(x=0)\|^2_{L^2(0,T)} + \frac{1}{2}\|y(x=1)\|^2_{L^2(0,T)} \\ &+ \frac{1}{2}\|y(t=T) - \tilde{u}\|^2_{L^2(\Omega)} + \frac{\varepsilon}{2}\|y(t=0) - u^b\|^2_{L^2(\Omega)}. \end{aligned} \tag{4.1}$$

The *DGM* algorithm approximates $u(t,x)$ with a deep neural network $y(t,x;\theta)$ where $\theta \in \mathbb{R}^k$ are the neural network's parameters. The goal is to find a set of parameters θ such that the function $y(t,x;\theta)$ minimizes the error $J(y)$. If the error $J(y)$ is small, then $y(t,x;\theta)$ will closely satisfy the PDE differential operator, boundary conditions, and initial condition. Therefore, a θ which minimizes $J(y(\cdot;\theta))$ produces a reduced-form model $y(t,x;\theta)$ which approximates the PDE solution $u(t,x)$.

In this sense, we recall the following theorem

Theorem 4.1. [12] *Let the L^2 error $J(f)$ measure how well the neural network f satisfies the differential operator, boundary, initial and observability condition.*

Define $\mathbb{C}^n$ as the class of neural networks with n hidden units and let f^n be a neural network with n hidden units which minimizes $J(f)$.

there exists $f^n \in \mathbb{C}^n$ such that $J(f^n) \longrightarrow 0$, as $n \longrightarrow \infty$, and $f^n \longrightarrow u$ as $n \longrightarrow \infty$.

Network architecture:

The first layer and the last of this neural network are fully connected. the rest is made up of GRU cells [4], which is a simplified version of the LSTM cell (Figure 1):

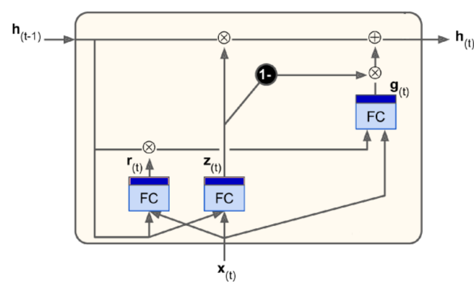

FIGURE 1. GRU cell

We found the following network architecture:

$$
\begin{aligned}
z_{(t)} &= tanh(w_{xz}^T x_{(t)} + w_{hz}^T h_{(t-1)} + b_z) \\
r_{(t)} &= tanh(w_{xr}^T x_{(t)} + w_{hr}^T h_{(t-1)} + b_r) \\
g_{(t)} &= tanh(w_{xg}^T x_{(t)} + w_{hg}^T (r_{(t)} \otimes h_{(t-1)}) + b_g) \\
h_{(t)} &= z_{(t)} \otimes h_{(t-1)} + (1 - z_{(t)}) \otimes g_{(t)}
\end{aligned}
$$

The main steps for descent method at each iteration are the following:
- Generate random points (t_n, x_n) from Q and θ_n.
- Take a descent step at the random point (t_n, x_n):
$\theta_{n+1} = \theta_n - \alpha_n \nabla_\theta J(t_n, x_n; \theta_n)$
- Repeat until convergence criterion is satisfied.

Parameters are updated using the well-known ADAM algorithm with a decaying learning rate schedule.

4.1. **The noise resistance of the proposed method.** The data u^b and u^{obs} are assumed to be corrupted by measurement errors, which we will refer to as noise. In particular, we suppose that $u^b = u^{exact}(t=0) + e$ and $u^{obs} = u^{exact} + e^{obs}$. Let $err = \frac{\|e\|_2}{\|u^{exact}(t=0)\|_2}$ and $err^{obs} = \frac{\|e^{obs}\|_2}{\|u^{exact}\|_2}$. We did two tests:
In the first, we suppose $err^{obs} = 0$, and we study the impact of err on construction of initial state. In the second test, we suppose $err = 0$, and we study the impact of err^{obs} on construction of initial state.

4.1.1. *Impact of err on construction of initial state.*

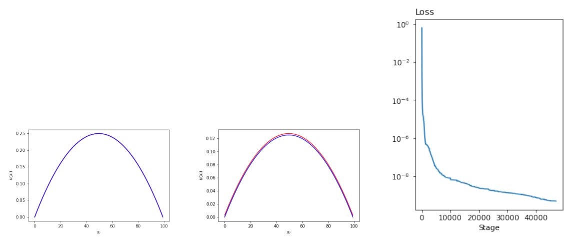

Figure 2. Test with $err = 0\%$. This figure shows that we can rebuild u_0 (left), and J converges to 0 (right).

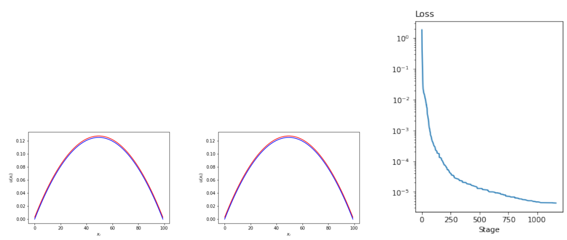

Figure 3. Test with $err = 5\%$. This figure shows that we can rebuild u_0 (left), and J converges to 0 (right).

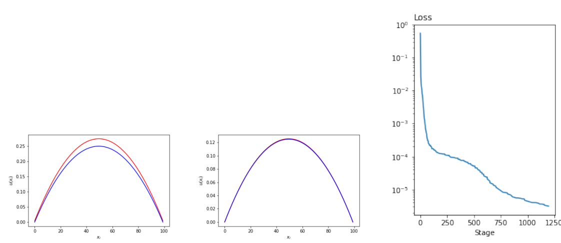

Figure 4. Test with $err = 10\%$. u_0 begins to move away from $u^{exact}(t = 0)$ (left).

These tests (Fig. 2 to 4) show that the proposed algorithm is uniformly stable to noise.

4.1.2. *Impact of err^{obs} on the construction of the initial state.*

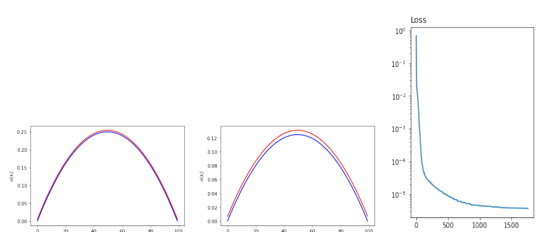

Figure 5. Test with $err^{obs} = 3\%$. This figure shows that we can rebuild u_0 at time (t=0) (left) and at (t=T) (middle), and J converges to 0 (right).

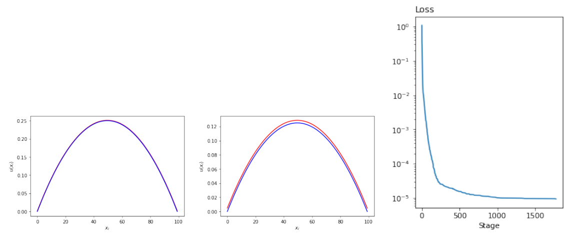

Figure 6.Test with $err^{obs} = 5\%$. This figure shows that we can rebuild u_0 at time (t=0) (left) and at (t=T) (middle), and J converges to 0 (right).

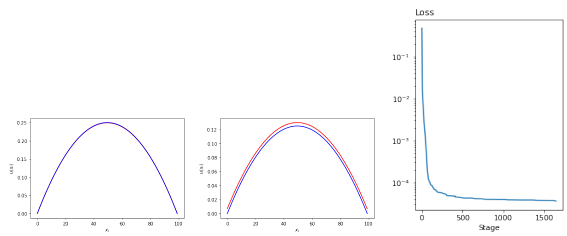

Figure 7. Test with $err^{obs} = 10\%$. u_0 begins to move away from $u^{exact}(t = 0)$ at time $(t = T)$ (middle), and J converges to 0 (right).

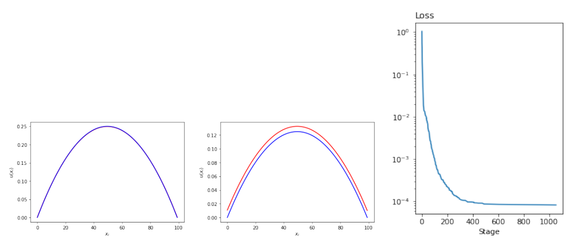

Figure 8. Test with $err^{obs} = 15\%$. u_0 begins to be far from $u^{exact}(t = 0)$ at time $(t = T)$ (middle) , and J converges to 0 (right).

These tests (5 to 8) show that the proposed algorithm is uniformly stable to noise. And we can rebuild u_0 with $err^{obs} \leq 10\%$ and $err \leq 5\%$. Also we notice that J always converges to 0.

5. Conclusion

In this paper a new approach has been adopted, in the theoretical side we have shown the existence, uniqueness and stability of the inverse problem concerning the determination of the initial condition of parabolic problem with memory term from final observations, verifying some criteria verified by the cost function. In the numerical part, we used deep neural networks to validate our results, which proves that the proposed algorithm is uniformly stable to noise. This method proves to be effective in reducing the execution time.

References

[1] Brahim Allal, Genni Fragnelli, and Jawad Salhi. Null controllability for the singular heat equation with a memory term. *arXiv preprint arXiv:2005.04442*, 2020.

[2] Khalid Atifi, Idriss Boutaayamou, Hamed Ould Sidi, and Jawad Salhi. An inverse source problem for singular parabolic equations with interior degeneracy. In *Abstract and Applied Analysis*, volume 2018. Hindawi, 2018.

[3] Khalid Atifi, El-Hassan Essoufi, and Bouchra Khouiti. Null controllability of degenerate/singular parabolic equations with degeneracy and singularity occurring in the interior of the spatial domain. *Eurasian Journal of Mathematical and Computer Applications*, 8(1):4–29, 2020.

[4] Gron Aurlien. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow - Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd Edition.*

[5] Deborah Brandon and William J Hrusa. Construction of a class of integral models for heat flow in materials with memory. *The Journal of Integral Equations and Applications*, 1(2):175–201, 1988.

[6] Martin Brokate and Jürgen Sprekels. *Hysteresis and phase transitions*, volume 121. Springer Science & Business Media, 1996.

[7] G Caginalp. Stefan and hele-shaw type models as asymptotic limits of the phase-field equations. *Physical Review A*, 39(11):5887, 1989.

[8] Gunduz Caginalp. An analysis of a phase field model of a free boundary. *Archive for Rational Mechanics and Analysis*, 92(3):205–245, 1986.

[9] Bernard D Coleman and Morton E Gurtin. Equipresence and constitutive equations for rigid heat conductors. *Zeitschrift für angewandte Mathematik und Physik ZAMP*, 18(2):199–208, 1967.

[10] Pierluigi Colli, Gianni Gilardi, Maurizio Grasselli, et al. Global smooth solution to the standard phase-field model with memory. *Advances in Differential Equations*, 2(3):453–486, 1997.

[11] Fabrizio Colombo and Vincenzo Vespri. A semilinear integrodifferential inverse problem. In *Evolution equations*, pages 91–104. CRC Press, 2019.

[12] Sirignano Justin and Spiliopoulos Konstantinos. Dgm: A deep learning algorithm for solving partial differential equations. *Journal of computational physics, Elsevier*, 2018.

[13] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature 521(7553), 436444*, 2015.

Khalid Atifi

Université Cadi Ayyad, Faculté des sciences et techniques, Département de Mathématiques, Laboratoire de Mathmatiques Appliques et Informatique, Marrakech, Maroc.

E-mail address: k.atifi.uhp@gmail.com

SOUFIANE ABID

Laboratoire de Mathématiques, Informatique et Sciences de l'ingénieur (MISI), Université Hassan 1, Settat 26000, Morocco

E-mail address: s.abid@uhp.ac.ma

El-Hassan Essoufi

Laboratoire de Mathématiques, Informatique et Sciences de l'ingénieur (MISI), Université Hassan 1, Settat 26000, Morocco

E-mail address: e.h.essoufi@gmail.com