Solving inverse problems for differential equations by a “generalized
collage” method and application to a mean field stochastic model

V. Capasso a, H.E. Kunze b, D. La Torre c,∗, E.R. Vrscay d

a Department of Mathematics, University of Milan, Italy
b Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada
c Department of Economics, University of Milan, Italy
d Department of Applied Mathematics, University of Waterloo, Ontario, Canada

A R T I C L E I N F O

Article history:
Received 13 May 2011
Accepted 13 May 2011

Keywords:
Collage theorem
Generalized collage theorem
Inverse problem
Boundary value problem
Mean field model

A B S T R A C T

In the first part of this paper, after recalling how to solve inverse problems for deterministic
and random differential equations using the collage method, we switch to the analysis of
stochastic differential equations. Here inverse problems can be solved by minimizing the
collage distance in an appropriate metric space. In the second part, we develop a general
collage coding framework for inverse problems for boundary value problems. Although a
general inverse problem can be very complicated, via the Generalized Collage Theorem
presented in this paper, many such problems can be reduced to an optimization problem
which can be solved at least approximately. We recall some previous results by some of
the authors on the same topic, but we provide more numerical examples to analyze the
stability of the generalized collage method under perturbation of data. We then extend
these results to the case of diffusion equations. Finally, we show an application of this
methodology to a system of coupled stochastic differential equations which describes the
interaction between particles in a physical system.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Many inverse problems may be viewed in terms of the approximation of a target element \( u \) in a complete metric space
\((X, d)\) by the fixed point \( \bar{u} \) of a contraction mapping \( T : X \rightarrow X \). In practical applications, from a family of contraction
mappings \( T_\lambda, \lambda \in \Lambda \subset \mathbb{R}^n \), one wishes to find the parameter \( \lambda \) for which the approximation error
\( d(u, \bar{u}_\lambda) \) is as small as possible. Thanks to a simple consequence of Banach’s fixed point theorem known as the “Collage Theorem”, most practical
methods of solving the inverse problem for fixed point equations seek to find an operator \( T \) for which the collage distance
\( d(u, Tu) \) is as small as possible.

Theorem 1.1 (“Collage Theorem” [1]). Let \((X, d)\) be a complete metric space and \( T : X \rightarrow X \) a contraction mapping with
contraction factor \( c \in [0, 1) \). Then for any \( u \in X \),

\[
d(u, \bar{u}) \leq \frac{1}{1-c} d(u, Tu),
\]

where \( \bar{u} \) is the fixed point of \( T \).
One now seeks a contraction mapping $T$ that minimizes the so-called collage error $d(u, Tu)$—in other words, a mapping that sends the target $u$ as close as possible to itself. This is the essence of the method of collage coding which has been the basis of most, if not all, fractal image coding and compression methods.

In [2] (and subsequent works [3–9]), the authors showed how collage coding could be used to solve inverse problems for systems of differential equations having the form

$$
\begin{align*}
\dot{u} &= f(t, u), \\
u(0) &= u_0,
\end{align*}
$$

(2)

when $f$ is a polynomial and by reducing the problem to the corresponding Picard integral operator associated with it,

$$
(Tu)(t) = u_0 + \int_0^t f(s, u(s)) \, ds.
$$

(3)

Here we show how one can attack this problem in the general case when $f$ belongs to $L^2$. Let us consider the complete metric space $C([0, T])$ endowed with the usual $d_\infty$ metric and assume that $f(t, x)$ is Lipschitz in the variable $x$, that is there exists a $K \geq 0$ such that $|f(s, x_1) - f(s, x_2)| \leq K|x_1 - x_2|$, for all $x_1, x_2 \in \mathbb{R}$. For simplicity we suppose that $x \in \mathbb{R}$ but the same consideration can be developed for the case of several variables. Under these hypotheses $T$ is Lipschitz on the space $C([-\delta, \delta] \times [-M, M])$ for some $\delta$ and $M > 0$.

**Theorem 1.2** ([2]). The function $T$ satisfies

$$
\|Tu - Tv\|_2 \leq c\|u - v\|_2
$$

(4)

for all $u, v \in C([-\delta, \delta] \times [-M, M])$ where $c = \delta K$.

Now let $\delta' > 0$ be such that $\delta'K < 1$. In order to solve the inverse problem for (3) we take the $L^2$ expansion of the function $f$. Let $\{\phi_i\}$ be a basis of functions in $L^2([-\delta, \delta] \times [-M, M])$ and consider

$$
f_i(s, x) = \sum_{i=1}^{+\infty} \lambda_i \phi_i(s, x).
$$

(5)

Each sequence of coefficients $\lambda = \{\lambda_i\}_{i=1}^{+\infty}$ then defines a Picard operator $T_\lambda$. Suppose further that each function $\phi_i(s, x)$ is Lipschitz in $x$ with constants $K_i$.

**Theorem 1.3** ([2]). Let $K, \lambda \in \ell^2(\mathbb{R})$. Then

$$
|f_i(s, x_1) - f_i(s, x_2)| \leq \|K\|_2 \|\lambda\|_2 |x_1 - x_2|
$$

(6)

for all $s \in [-\delta', \delta']$ and $x_1, x_2 \in [-M, M]$ where $\|K\|_2 = \left(\sum_{i=1}^{+\infty} K_i^2\right)^{\frac{1}{2}}$ and $\|\lambda\|_2 = \left(\sum_{i=1}^{+\infty} \lambda_i^2\right)^{\frac{1}{2}}$.

Given a target solution $x$, we now seek to minimize the collage distance $\|u - T_\lambda u\|_2$. The square of the collage distance becomes

$$
\Delta^2(\lambda) = \|u - T_\lambda u\|_2^2
$$

(7)

and the inverse problem can be formulated as

$$
\min_{\lambda \in \Lambda} \Delta(\lambda),
$$

(8)

where $\Lambda = \{\lambda \in \ell^2(\mathbb{R}) : \|\lambda\|_2 \|K\|_2 < 1\}$. To solve numerically this problem, let us consider the first $n$ terms of the $L^2$ basis; in this case the previous problem can be reduced to:

$$
\min_{\lambda \in \tilde{\Lambda}} \tilde{\Delta}^2(\lambda) = \int_{-\delta}^{\delta} \left| x(t) - \int_0^t \sum_{i=1}^{n} \lambda_i \phi_i(s, x(s)) \, ds \right|^2 \, dt,
$$

(9)

where $\tilde{\Lambda} = \{\lambda \in \mathbb{R}^n : \|\lambda\|_2 \|K\|_2 < 1\}$. This is a classical quadratic optimization problem which can be solved by means of classical numerical methods. Let $\Delta_{\text{min}}^n$ be the minimum value of $\Delta$ over $\tilde{\Lambda}$. This is a non-increasing sequence of numbers (depending on $n$) and as shown in [10] it is possible to show that $\lim_{n \to +\infty} \Delta_{\text{min}}^n = 0$. This states that the distance between the target element and the unknown solution of the differential equation can be made arbitrary small.
Random integral operator $T: \Omega \to Y$.

Solutions to (10) are fixed points of (11), that is, solutions of the equation

$$
\begin{align*}
T^* u(t) &= u_0(\omega) + \int_0^t f(s, \omega, u(s)) \, ds.
\end{align*}
$$

where both the vector field $f$ and the initial condition $x_0$ are random variables defined on an appropriate probability space $(\Omega, \mathcal{F}, P)$. Analogous to the deterministic case, for $X = C([0, T])$ this problem can be reformulated by using the following random integral operator $T: \Omega \times X \to X$:

$$(T_u u)(t) = u_0(\omega) + \int_0^t f(s, \omega, u(s)) \, ds.$$  

Solutions to (10) are fixed points of (11), that is, solutions of the equation $T_u u = u$. We recall that a function $T: \Omega \times X \to X$ is called a random operator (in a strict sense, see [11, p. 104]) if for any $u \in X$ the function $T(\cdot, u)$ is measurable. The random operator $T$ is said to be measurable/Lipschitz/contractive if, for a.e. $\omega \in \Omega$, we have that $T(\omega, \cdot)$ is measurable/Lipschitz/contractive. A measurable mapping $u: \Omega \to X$ is called a random fixed point of the random operator $T$ if $u$ is a solution of the equation

$$
T(\omega, u(\omega)) = u(\omega), \quad \text{a.e. } \omega \in \Omega.
$$

In order to study the existence of solutions to such equations, let us consider the space $Y$ of all measurable functions $u: \Omega \to X$. If we define the operator $T: Y \to Y$ as $(T u)(\omega) = T(\omega, u(\omega))$ the solutions of this fixed point equation on $Y$ are the solutions of the random fixed point equation $T(\omega, u(\omega)) = u(\omega)$. The space $Y$ is a complete metric space with respect to the following metric (see [6]):

$$
d_Y(u_1, u_2) = \int_{\Omega} d_X(u_1(\omega), u_2(\omega)) \, dP(\omega).
$$

Example 1.1. Consider the following random differential equation,

$$
\begin{align*}
\frac{dx}{dt}(t) &= A_0 + A_1 x(t) + A_2 x^2(t), \\
x(0) &= x_0
\end{align*}
$$

where $A_0, A_1, A_2, x_0$ are real valued random variables on the same probability space $(\Omega, \mathcal{F}, P)$. The realizations are calculated by solving numerically the related differential equation, sampling the solution at 10 uniformly distributed points, and fitting the polynomial $x(t, \omega)$ to the data. In Table 1, we list the distributions used for the parameters. The collage coding results are presented in Tables 2 and 3.

Example 1.2. Suppose now that the stochastic process $X_t$ is driven by a logistic stochastic process; then it satisfies the stochastic differential equation

$$
\begin{align*}
\frac{dX_t}{dt} = X_t(a - bX_t) \, dt + X_t \, dW_t,
\end{align*}
$$

The distributions used in the inverse problem are given in Table 1.

Table 1

<table>
<thead>
<tr>
<th>Label</th>
<th>True Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_0$</td>
</tr>
<tr>
<td></td>
<td>$A_1$</td>
</tr>
<tr>
<td></td>
<td>$A_2$</td>
</tr>
<tr>
<td></td>
<td>$x_0$</td>
</tr>
<tr>
<td>1</td>
<td>$\mathcal{N}(1.2, 0.09)$</td>
</tr>
<tr>
<td>2</td>
<td>$\mathcal{N}(0.5, 0.01)$</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>Label</th>
<th>$N$</th>
<th>Minimal collage values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$A_0$</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>(1.1755, 0.1477)</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>(1.1665, 0.4136)</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>(1.2009, 0.0845)</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>(0.4574, 0.4997)</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>(0.4362, 0.2798)</td>
</tr>
<tr>
<td>2</td>
<td>1000</td>
<td>(0.4953, 0.2326)</td>
</tr>
</tbody>
</table>

In [6], Kunze et al. considered the case of inverse problems for random differential equations. This kind of problems can be formulated as

$$
\begin{align*}
\frac{du}{dt}(t) = f(t, \omega, u(\omega, t)), \\
u(\omega, 0) &= u_0(\omega)
\end{align*}
$$

Please cite this article in press as: V. Capasso et al., Solving inverse problems for differential equations by a “generalized collage” method and application to a mean field stochastic model, Nonlinear Analysis: Real World Applications (2011), doi:10.1016/j.nonrwa.2011.05.017
Table 3
Results for the inverse problem. The first column indicates the distribution from Table 1 from which realizations are generated. N is the number of realizations, and the final two columns give the (mean, variance) obtained via collage coding for $A_2$ and $x_0$.

<table>
<thead>
<tr>
<th>Label</th>
<th>N</th>
<th>Minimal collage values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$A_2$</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>(1.4042, 0.0073)</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>(0.4236, 0.0086)</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>(0.3989, 0.0104)</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>(0.3074, 0.0287)</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>(0.3441, 0.0349)</td>
</tr>
<tr>
<td>2</td>
<td>1000</td>
<td>(0.2959, 0.0410)</td>
</tr>
</tbody>
</table>

Table 4
Minimal collage distance parameters for different $N$ and $M$, to five decimal places.

<table>
<thead>
<tr>
<th>N</th>
<th>M</th>
<th>$a$</th>
<th>$b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>300</td>
<td>20.03020</td>
<td>10.01753</td>
</tr>
<tr>
<td>100</td>
<td>600</td>
<td>19.99748</td>
<td>9.99778</td>
</tr>
<tr>
<td>100</td>
<td>900</td>
<td>20.00240</td>
<td>10.00031</td>
</tr>
<tr>
<td>300</td>
<td>300</td>
<td>20.02060</td>
<td>10.00162</td>
</tr>
<tr>
<td>300</td>
<td>600</td>
<td>19.99667</td>
<td>9.99889</td>
</tr>
<tr>
<td>300</td>
<td>900</td>
<td>19.98866</td>
<td>9.99262</td>
</tr>
</tbody>
</table>

Fig. 1. Different paths of the stochastic logistic process.

where $a$ and $b$ are two parameters and $W_t$ is a Wiener process. We consider the inverse problem: given realizations/paths $X_i^t$, $1 \leq i \leq N$, estimate $a$ and $b$. As an example and following the same approach as above, we set $a = 20$ and $b = 10$ and $X_0 = 0.1$, and then generate paths on $[0, 1]$. Fig. 1 shows ten paths for this process $X_t$.

Table 4 presents the numerical results of the example.

2. Stochastic differential equations as fixed point equations and related inverse problems

Let $(X, d)$ be a separable complete metric space, $C(X)$ be the collection of nonempty compact subsets of $X$ and $BC(X)$ be the family of nonempty bounded closed subsets of $X$. It is well known that the spaces $(C(X), d_H)$ and $(BC(X), d_H)$ are complete with respect to the Hausdorff metric defined as

$$d_H(A, B) = \max \left\{ \sup_{x \in A} \inf_{y \in B} d(x, y), \sup_{x \in B} \inf_{y \in A} d(x, y) \right\}.$$  \hspace{1cm} (16)

For a given function $f$, let $\text{Lip} f$ denote the Lipschitz constant for $f$, that is the least $L$ such that $d(f(x), f(y)) \leq Ld(x, y)$. Let $M(X)$ be the collection of probability measures on $(X, \mathcal{B}(X))$, where $\mathcal{B}(X)$ is the Borel $\sigma$-algebra of $X$. We recall that $\nu_n \to \mu$ in the weak convergence means that $\int_X \phi d\nu_n \to \int_X \phi d\mu$ for all bounded continuous $\phi$. Given $\mu, \nu \in M(X)$, the Wasserstein metric $d_W(\mu, \nu)$ on $X$ is defined by

$$d_W(\mu, \nu) = \inf_{\gamma} \left\{ \int_{X \times X} \min\{d(x, y), 1\} d\gamma : \gamma \text{ is a measure on } X \times X, \Pi_1(\gamma) = \mu, \Pi_2(\gamma) = \nu \right\}$$

where $\Pi_1, \Pi_2 : X \times X \to X$ are the projections onto the first and the second coordinates and $\Pi_j(\gamma)(A) = \gamma(\Pi_j^{-1}(A)), A \subset X, j = 1, 2$. It is known that the space $(M(X), d_W)$ is complete and $d_W$ gives to $M(X)$ the topology of weak convergence.

Please cite this article in press as: V. Capasso, et al., Solving inverse problems for differential equations by a “generalized collage” method and application to a mean field stochastic model, Nonlinear Analysis: Real World Applications (2011), doi:10.1016/j.nonrwa.2011.05.017
Let \((M_1(X), d_{MK})\) be the complete metric space (see [12]) consisting of all measures \(\mu\) with finite first moment (this means \(\int_X d(a, x) d\mu(x) < \infty\) for any \(a \in X\)) with the Monge–Kantorovich metric \(d_{MK}\) defined as

\[
d_{MK}(\mu, \nu) = \sup_{f \in F} \left\{ \int_X f d\mu - \int_X f d\nu : \text{Lip} f \leq 1 \right\} = \inf_{\gamma} \left\{ \int_{X \times X} d(x, y) d\gamma(x, y) : \gamma \text{ is a measure on } X \times X, \Pi_1(\gamma) = \mu, \Pi_2(\gamma) = \nu \right\}.
\]

It is well known that the moment condition is automatically satisfied if \((X, d)\) is bounded. The equivalence between the two previous definitions is shown in [13]. Between the convergence in the Monge–Kantorovich metric and the topology of convergence the following relation holds (see [12]):

\[
v_n \to_{MK} \mu \quad \text{if and only if} \quad v_n \to \mu \quad \text{and} \quad \int_X d(x, a) d\nu_n(x) \to \int_X d(x, a) d\mu(x)
\]

for all \(a \in X\).

Let \((\Omega, \mathcal{F}, P)\) be a probability space, \(\{\mathcal{F}_t\}_{t \geq 0}\) be a filtration, \(\{B_t\}_{t \geq 0}\) be a classical \(\mathbb{R}^d\) Brownian motion, \(X_0\) be a \(\mathcal{F}_0 - \mathbb{R}^d\)-measurable random vector, \(g : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d\), we look at the solution of the SDE:

\[
\begin{align*}
\text{d}X_t &= \int_{\mathbb{R}^d} g(X_t, y) d\mu_t(y) \text{d}t + dB_t \\
X_0 &= X_0
\end{align*}
\]

where \(\mu_t = P_{X_t}\) is the law of \(X_t\). Given \(T > 0\), it is well known that associated with each process \(X_t\) one can define a random variable from \(\Omega\) to \(C([0, T])\). Consider the complete metric space \((C([0, T]), d_{\infty})\) and the space \(M(C([0, T]))\) of probability measures on \(C([0, T])\). \(X_t\) induces a probability measure on \(M(C([0, T]))\). Let \(\Phi : M(C([0, T])) \to M(C([0, T]))\) be the function which associates with each element \(m \in M(C([0, T]))\) the law of the process \(\Pi_t = \Phi_m\).

If \(X_t\) is a solution of (17) then its law on \(C([0, T])\) is a fixed point of \(\Phi\), and vice versa. We have the following theorem which states an existence and uniqueness result for (17).

**Theorem 2.1 ([14]).** Let \((\Omega, \mathcal{F}, P)\) be a probability space, \(\{\mathcal{F}_t\}_{t \geq 0}\) be a filtration, \(\{B_t\}_{t \geq 0}\) be a classical \(\mathbb{R}^d\) Brownian motion, \(X_0\) be a \(\mathcal{F}_0 - \mathbb{R}^d\)-measurable random vector, \(g : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d\) be a bounded Lipschitz function. Consider the following stochastic differential equation:

\[
\begin{align*}
\text{d}X_t &= \int_{\mathbb{R}^d} g(X_t, y) d\mu_t(y) \text{d}t + dB_t \\
X_{t=0} &= X_0
\end{align*}
\]

We have that:

(i) for \(t \leq T\), \(m_1, m_2 \in M(C([0, T]))\),

\[
d_{MK}(\Phi(m_1), \Phi(m_2)) \leq c_T \int_0^t d_{MK}(m_1, m_2) \text{d}s
\]

where \(c_T\) is a constant and \(d_{MK}\) is the distance between the images of \(m_1, m_2\) on \(C([0, s])\);

(ii) \(\Phi\) is eventually contractive since there is a \(k > 0\) so that

\[
d_{MK}(\Phi^k(m_1), \Phi^k(m_2)) \leq \frac{k^2}{k} d_{MK}(m_1, m_2) = c_T k^k \text{d}T \quad d_{MK}(m_1, m_2) = c_T k^k \text{d}T
\]

for \(c_T < 1\).

(iii) there exists a unique solution, pathwise and in law, of (17).

The aim of the inverse problem consists of finding an estimation of \(g\) starting from a sample of observations of \(X_t\). Let \(\{X_{t}^1, X_{t}^2, \ldots, X_{t}^n\}, t \in [0, T]\), be an independent sample (i.d.) and \(\mu_n\) the estimated law of the process. We have the following trivial corollary of the Collage Theorem.

**Corollary 2.1.** Let \(\mu_n \in M(C[0, T])\) be the estimated law of the process. If \(\mu\) is the law of the process \(X_t\) of (17) then there exists a constant \(C\) such that the following estimate holds:

\[
d_{MK}(\mu, \mu_n) \leq C d_{MK}(\Phi(\mu), \mu_n).
\]

The inverse problem is then reduced to the minimization of \(d_{MK}(\Phi(\mu_n), \mu_n)\) which is a function of the unknown coefficients of \(g\).
3. A generalized collage method

A linear functional on a real Hilbert space $H$ is a linear map from $H$ to $\mathbb{R}$. A linear functional $\phi$ is bounded and hence continuous, if there exists a constant $M$ such that

$$|\phi(x)| \leq M\|x\|$$

for all $x \in H$. By the linearity of $\phi$ it is trivial to prove that we may choose

$$M = \max_{x \in H, \|x\| = 1} \phi(x).$$

(21)

Let $H$ be a Hilbert space and $\phi$ be a bounded linear nonzero functional. Suppose that $a(u, v)$ is a bilinear form on $H \times H$ which satisfies the following:

- There exists a constant $M > 0$ such that $|a(u, v)| \leq M\|u\|\|v\|$ for all $u, v \in H$.
- There exists a constant $m > 0$ such that $|a(u, u)| \geq m\|u\|^2$ for all $u \in H$.

Then it is well known that the Lax–Milgram Theorem states there is a unique vector $u^* \in H$ such that $\phi(v) = a(u^*, v)$ for all $v \in H$.

Suppose that we have a given Hilbert space, a “target” element $u \in H$ and a family of bilinear functionals $a_\lambda$. Then by the Lax–Milgram Theorem, there exists a unique vector $u_\lambda$ such that $\phi(v) = a_\lambda(u_\lambda, v)$ for all $v \in H$. We would like to determine if there exists a value of the parameter $\lambda$ such that $u_\lambda = u$ or, more realistically, such that $\|u_\lambda - u\|$ is small enough. The following theorem will be instrumental for the solution of this problem.

**Theorem 3.1 (Generalized Collage Theorem [8,9])**. Suppose that $a_\lambda(u, v) : \mathcal{F} \times H \times H \to \mathbb{R}$ is a family of bilinear forms for all $\lambda \in \mathcal{F}$ and $\phi_\lambda : \mathcal{F} \times H \to \mathbb{R}$ is a given family of linear functionals. Let $u_\lambda$ denote the solution of the equation $a_\lambda(u, v) = \phi_\lambda(v)$ for all $v \in H$ as guaranteed by the Lax–Milgram theorem. Given a target element $u \in H$ then

$$\|u - u_\lambda\| \leq \frac{1}{m_\lambda} F(\lambda),$$

(23)

where

$$F(\lambda) = \sup_{v \in H, \|v\| = 1} |a_\lambda(u, v) - \phi_\lambda(v)|.$$  

(24)

In order to ensure that the approximation $u_\lambda$ is close to a target element $u \in H$, we can, by the Generalized Collage Theorem, try to make the term $F(\lambda)/m_\lambda$ in Eq. (23) as close to zero as possible. The appearance of the $m_\lambda$ factor complicates the procedure as does the factor $1/(1 - c)$ in standard collage coding, i.e., Eq. (1). As such, we shall follow the usual practice and ignore the $m_\lambda$ factor, assuming, of course, that all allowable values are bounded away from zero. So if $\inf_{\lambda \in \mathcal{F}} m_\lambda \geq m > 0$ then the inverse problem can be reduced to the minimization of the function $F(\lambda)$ on the space $\mathcal{F}$, that is,

$$\inf_{\lambda \in \mathcal{F}} F(\lambda).$$

(25)

Let us now extend the previous generalized collage theorem to the case of diffusion equation of the form

$$\begin{cases}
\frac{d}{dt} u, v = \psi(v) + a(u, v) \\
u(0) = f
\end{cases}$$

(26)

where $\psi : H \to \mathbb{R}$ is a linear functional, $a : H \times H \to \mathbb{R}$ is a bilinear form, and $f \in H$ is an initial condition. The aim of the inverse problem for the above system consists of getting an approximation of the coefficients and parameters starting from a sample of observations of a target. To do this, let us consider a family of bilinear functionals $a_\lambda$ and $\psi_\lambda$, let $u_\lambda$ be the solution to

$$\begin{cases}
\frac{d}{dt} u_\lambda, v = \psi_\lambda(v) + a_\lambda(u_\lambda, v) \\
u_\lambda(0) = f
\end{cases}$$

(27)

We would like to determine if there exists a value of the parameter $\lambda$ such that $u_\lambda = u$ or, more realistically, such that $\|u_\lambda - u\|$ is small enough. To this end, **Theorem 3.2** states that the distance between the target solution $u$ and the solution $u_\lambda$ of (27) can be reduced by minimizing a functional which depends on parameters.
Theorem 3.2. Let \( u : [0, T] \to H \) be the target solution which satisfies the initial condition in (26) and suppose that \( \frac{d}{dt} u \) exists and belongs to \( H \). Suppose that \( a_\lambda(u, v) : \mathcal{F} \times H \times H \to \mathbb{R} \) is a family of bilinear forms satisfying the hypotheses of the Lax–Milgram Theorem and \( \psi_\lambda : \mathcal{F} \times H \to \mathbb{R} \) is family of a bounded linear functionals for all \( \lambda \in \mathcal{F} \). We have the following result:

\[
\int_0^T \| u(t) - u_\lambda(t) \|_H^2 \, dt \leq \frac{1}{m^2_\lambda} \int_0^T \left( \sup_{\|v\| = 1} \left( \frac{d}{dt} u(t), v \right) - \psi_\lambda(v) - a_\lambda(u(t), v) \right)^2 \, dt
\]

(28)

where \( u_\lambda \) is the solution of (27) s.t. \( u_\lambda(0) = u(0) \) and \( u_\lambda(T) = u(T) \).

Proof. Computing we have

\[
m_\lambda \| u(t) - u_\lambda(t) \|_H^2 \leq a(u(t) - u_\lambda(t), u(t) - u_\lambda(t))
\]

\[
= a(u(t), u(t) - u_\lambda(t)) + \left( \frac{d}{dt} (u(t) - u_\lambda(t)), u(t) - u_\lambda(t) \right) + \psi_\lambda(u(t) - u_\lambda(t))
\]

\[
- \left( \frac{d}{dt} u(t), u(t) - u_\lambda(t) \right)
\]

and, by simple calculations, we get

\[
m_\lambda \| u(t) - u_\lambda(t) \|_H^2 - \frac{1}{2} \frac{d}{dt} \| u(t) - u_\lambda(t) \|_H^2 \leq a(u(t), u(t) - u_\lambda(t)) + \psi_\lambda(u(t) - u_\lambda(t)) - \left( \frac{d}{dt} u(t), u(t) - u_\lambda(t) \right).
\]

Integrating both sides with respect to \( t \) and recalling that \( u(0) = u_\lambda(0) \) and \( u(T) = u_\lambda(T) \), we have

\[
m_\lambda \int_0^T \| u(t) - u_\lambda(t) \|_H^2 \, dt \leq \int_0^T \| u(t) - u_\lambda(t) \|_H \left( \sup_{\|v\| = 1} a_\lambda(u(t), v) + \psi_\lambda(v) - \left( \frac{d}{dt} u(t), v \right) \right) \, dt
\]

\[
\leq \left( \int_0^T \| u(t) - u_\lambda(t) \|_H^2 \, dt \right)^{1/2} \left( \int_0^T \left( \sup_{\|v\| = 1} a(u(t), v) + \psi_\lambda(v) - \left( \frac{d}{dt} u(t), v \right) \right)^2 \, dt \right)^{1/2}
\]

and now the thesis easily follows. \( \square \)

Whenever \( \inf_{\lambda \in \mathcal{F}} m_\lambda \geq m > 0 \) then the previous result states that in order to solve the inverse problem for the parabolic equation (26) one can minimize the following functional

\[
\int_0^T \left( \sup_{\|v\| = 1} \left( \frac{d}{dt} u(t), v \right) - \psi_\lambda(v) - a_\lambda(u(t), v) \right)^2 \, dt
\]

(29)

over all \( \lambda \in \mathcal{F} \).

4. Numerical examples

In this section we illustrate two numerical examples of the previous inverse problem techniques based on the collage and the generalized collage results.

4.1. An inverse problem for the steady-state equation \( (\kappa(x)u_\lambda)_x = f(x) \)

Let us consider the following one-dimensional steady-state diffusion equation

\[
- \frac{d}{dx} \left( \kappa(x) \frac{du}{dx} \right) = f(x), \quad 0 < x < 1,
\]

(30)

\[
u(0) = u_{\text{left}},
\]

(31)

\[
u(1) = u_{\text{right}},
\]

(32)

where the diffusivity \( \kappa(x) \) varies in \( x \). The inverse problem of interest is: given \( u(x) \), possibly in the form of an interpolation of data points, and \( f(x) \) on \([0, 1]\), determine an approximation of \( \kappa(x) \). As discussed in the introduction, in [15] this problem is studied and solved via a regularized least squares minimization problem. It is important to stress that the approach in [15] seeks to directly minimize the error between the given \( u(x) \) and the solutions \( v(x) \) to Eq. (30). The collage coding approach allows us to perform a different minimization to solve the inverse problem. A natural goal is to recover \( \kappa(x) \) from observations of the response \( u(x) \) to a point source \( f(x) = \delta(x - x_i) \), a Dirac delta function at \( x_i \in (0, 1) \). In what follows,
we consider \(u_{\text{left}} = u_{\text{right}} = 0\), although the approach can be modified to treat nonzero values. We multiply Eq. (30) by a test function \(\xi_i(x) \in H^1_0([0, 1])\) and integrate by parts to obtain \(a(u, \xi_i) = \phi(\xi_i)\), where

\[
a(u, \xi_i) = \left\{ \int_0^1 \kappa(x)u'(x)\xi'_i(x) \, dx - \xi_i(x)\kappa(x)u'(x) \right\}_0^1
\]

\[
= \int_0^1 \kappa(x)u'(x)\xi'_i(x) \, dx, \quad \text{and}
\]

\[
\phi(\xi_i) = \int_0^1 f(x)\xi_i(x) \, dx.
\]

For a fixed choice of \(n\), partition \([0, 1]\) at \(x_i = \frac{i}{n+1}\), \(i = 0, \ldots, n + 1\), with \(n\) interior points, and define for \(j = 0, 1, 2, \ldots\)

\[
V_n^r = \{ v \in C[0, 1] : v(0) = v(1) = 0 \text{ and } v \text{ is a polynomial of degree } r \text{ on } [x_{i-1}, x_i], \ i = 1, \ldots, n + 1 \}.
\]

Denote a basis for \(V_n^r\) by \(\{\xi_1, \ldots, \xi_n\}\). When \(r = 1\), our basis consists of the hat functions

\[
\xi_i(x) = \begin{cases} 
(n + 1) (x - x_{i-1}), & x_{i-1} \leq x \leq x_i \\
 - (n + 1) (x - x_{i+1}), & x_i \leq x \leq x_{i+1}, \\
0, & \text{otherwise}
\end{cases}
\]

and when \(r = 2\), our hats are replaced by parabolas, and so on. Suppose that \(\kappa(x) > 0\) for all \(x \in [0, 1]\). Then the \(m_n\) in our formulation, which we denote by \(m_n\), can be chosen equal to \(\inf_{x \in [0, 1]} \kappa(x)\). In fact, we have

\[
a(u, u) = \int_0^1 \kappa(x)u'(x)u'(x) \, dx \geq \inf_{x \in [0, 1]} \kappa(x) \int_0^1 (u'(x))^2 \, dx = m_n \|u\|^2_{H^1_0},
\]

where the norm on \(H^1_0\) is defined by the final equality. As a result, because we divide by \(m_n\), we expect our results will be good when \(\kappa(x)\) is bounded away from 0 on \([0, 1]\).

We shall consider two different scenarios: (i) a continuous framework and (ii) a discretized framework. In a final discussion (iii), we consider the case of \(f(x)\) being a point source in each of the two frameworks. Finally, we discuss the incorporation of multiple data sets by our method in (iv).

(i) Continuous framework

Assume that we are given data points \(u_i\) measured at various \(x\)-values having no relation to our partition points \(x_i\). These data points are interpolated to produce a continuous target function \(u(x)\), a polynomial, say. Let us now assume a polynomial representation of the diffusivity, i.e.,

\[
\kappa(x) = \sum_{j=0}^N \lambda_j x^j.
\]

In essence, this introduces a regularization into our method of solving the inverse problem. Working on \(V_n^r\), we have

\[
a_s(u, \xi_i) = \sum_{j=0}^N \lambda_j A_{ij}, \quad \text{with } A_{ij} = \int_{x_{i-1}}^{x_i+1} x^j u'(x)\xi'_i(x) \, dx.
\]

Letting

\[
b_i = \int_0^1 f(x)\xi_i(x) \, dx = \int_{x_{i-1}}^{x_i+1} f(x)\xi_i(x) \, dx, \quad i = 1, \ldots, n,
\]

we now minimize

\[
(F_n(\lambda))^2 = \sum_{i=1}^n \left[ \sum_{j=0}^N \lambda_j A_{ij} - b_i \right]^2.
\]

Various minimization techniques can be used; in this work we used the quadratic program solving package in Maple’s Maple.

As a specific experiment, consider \(f(x) = 8x\) and \(\kappa_{\text{true}}(x) = 2x + 1\), in which case the solution to the steady-state diffusion equation is \(u_{\text{true}}(x) = x - x^2\). We shall solve this solution at 10 data points, add Gaussian noise of small amplitude \(\varepsilon\) to these values and then fit the data points to a polynomial of degree 2, to be denoted as \(u_{\text{target}}(x)\). Given \(u_{\text{target}}(x)\) and \(f(x)\), we seek a degree 10 polynomial \(\kappa(x)\) with coefficients \(\lambda_i\) so that the steady-state diffusion equation admits \(u_{\text{target}}(x)\) as an approximate solution. We now construct \(F_30(\lambda)\) and minimize it with respect to the \(\lambda_i\). Table 5 presents the results. In all cases, the recovered coefficients for all terms of degree two and higher are zero to five decimal places, so we do not report

Please cite this article in press as: V. Capasso, et al., Solving inverse problems for differential equations by a “generalized collage” method and application to a mean field stochastic model, Nonlinear Analysis: Real World Applications (2011), doi:10.1016/j.nonwa.2011.05.017
them in the table. \(d_2\) denotes the standard \(L^2\) metric on \([0, 1]\). In Table 6 are presented results obtained when \(N = 40\) basis functions and a degree 4 polynomial for \(\kappa(x)\) are employed.

(ii) Discretized framework

In a practical example, we generally obtain discrete data values for \(u\). If we are given values \(u_i\) at the partition points \(x_i, \ i = 1, \ldots, n\), and set \(u_0 = u_{\text{left}} = 0\), and \(u_{n+1} = u_{\text{right}} = 0\), then, working on \(V_n\), we write

\[
u(x) = \sum_{i=1}^{n} u_i \xi_i(x).
\]

Then Eq. (37) becomes

\[
A_{ij} = \sum_{l=1}^{n} u_l \int_{x_{l-1}}^{x_{l+1}} x^i \xi_{ij}(x) \xi_{lj}(x) \, dx.
\]

Notice that we face a problem regardless of our approach. In the earlier approach, we interpolate the points to obtain a target \(u(x)\) to use in the above formulas; it is quite possible that small errors in that interpolation can lead to large errors in the derivative \(u'(x)\) that we need to calculate the \(A_{ij}\). Here, as well, small errors in our data values \(u_i\) can be amplified. If, in addition, we are given values of \(f(x_i) = f_i, \ i = 0, \ldots, n + 1\), then we extend our basis of \(V_n\) by adding the two “half-hat” functions at the end points. We represent \(f(x)\) in this extended basis, writing

\[
f(x) = \sum_{l=0}^{n+1} f_l \xi_l(x)
\]

to approximate \(b_i\) and, thereafter, \(c_i\).

In Table 8, we repeat the same experiment as in framework (i), this time without interpolating the data points and instead approximating the \(A_{ij}\) as discussed above.

(iii) Point sources

Finally, we consider the case when \(f(x)\) is a point source at one of our partition points,

\[
f(x) = \delta(x - x_s), \quad \text{where } s \in \{1, 2, \ldots, n\}.
\]

Working on \(V_n\), Eq. (38) becomes

\[
b_i = \phi(\xi_i) = \int_0^1 f(x) \xi_i(x) \, dx = \begin{cases} 1, & \text{if } i = s \\ 0, & \text{otherwise.} \end{cases}
\]

In framework (i), we can use Eq. (39) to solve our inverse problem, where the right hand side of the equation now simplifies to \(A_{ik}\). A similar change occurs in framework (ii).

But suppose that we seek an expansion of \(\kappa(x)\) in the extended \(\xi_i\) basis:

\[
\kappa(x) = \sum_{j=0}^{n+1} \lambda_j \xi_j(x).
\]

---

Table 5

Collage coding results when \(f(x) = 8x, \ \kappa_{\text{true}}(x) = 1 + 2x\), data points = 10, number of basis functions = 30, and degree of \(\kappa_{\text{collage}} = 10\). In the first four rows, we work on \(V_0\); in the last four rows, we work on \(V_4\). In each case, \(\kappa_{\text{true}}.\kappa_{\text{collage}}\) is equal to 0–10 decimal places.

<table>
<thead>
<tr>
<th>Noise (\varepsilon)</th>
<th>(d_2(u_{\text{true}}, u_{\text{target}}))</th>
<th>(\kappa_{\text{collage}})</th>
<th>(d_2(\kappa_{\text{collage}}, \kappa_{\text{true}}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.00000</td>
<td>1.00000 + 2.00000x</td>
<td>0.00000</td>
</tr>
<tr>
<td>0.01</td>
<td>0.00353</td>
<td>1.03050 + 2.05978x</td>
<td>0.06281</td>
</tr>
<tr>
<td>0.05</td>
<td>0.01770</td>
<td>1.17365 + 2.33952x</td>
<td>0.35712</td>
</tr>
<tr>
<td>0.10</td>
<td>0.03539</td>
<td>1.42023 + 2.81788x</td>
<td>0.86213</td>
</tr>
<tr>
<td>0.00</td>
<td>0.00000</td>
<td>1.00000 + 2.00000x</td>
<td>0.00000</td>
</tr>
<tr>
<td>0.01</td>
<td>0.00353</td>
<td>1.00832 + 2.03967x</td>
<td>0.03040</td>
</tr>
<tr>
<td>0.05</td>
<td>0.01770</td>
<td>1.03981 + 2.21545x</td>
<td>0.16011</td>
</tr>
<tr>
<td>0.10</td>
<td>0.03539</td>
<td>1.07090 + 2.46292x</td>
<td>0.34301</td>
</tr>
</tbody>
</table>

Table 6

Collage coding results when \(f(x) = 8x, \ \kappa_{\text{true}}(x) = 1 + 2x\), data points = 100, number of basis functions = 40, and degree of \(\kappa_{\text{collage}} = 4\), working on \(V_4\).

<table>
<thead>
<tr>
<th>Noise (\varepsilon)</th>
<th>(\kappa_{\text{collage}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00000</td>
<td>1.00000 + 2.00000x + (0.00000x^2 + 0.00000x^3 + 0.00000x^4)</td>
</tr>
<tr>
<td>0.00001</td>
<td>1.00001 + 2.00058x - 0.05191x^2 + 0.10443x^3 - 0.5928x^4</td>
</tr>
<tr>
<td>0.00010</td>
<td>1.00004 + 2.05768x - 0.52113x^2 + 1.04778x^3 - 0.59470x^4</td>
</tr>
</tbody>
</table>
Plugging in Eq. (40) and the basis expansion for \( u(x) \), we get

\[
A_{ij} = \sum_{k=1}^{n} u_k \int_{0}^{1} \xi_j(x) \xi_k'(x) \xi_i'(x) \, dx.
\]

Clearly, \( A_{ij} = 0 \) if \(|i - j| > 1\); the corresponding matrix \( A \) is tridiagonal. The problem is once again to minimize \((F_n(\lambda))^2\), as given in (39). And again, a number of techniques can be used. (Note, however, that if we impose the stationarity condition \( \frac{d(F_n(\lambda))^2}{d\lambda} = 0 \), we obtain the linear system presented by Vogel in [15].)

As an example, we set \( f(x) = \delta \left( x - \frac{3}{10} \right) \), and pick \( n \) so that \( \frac{3}{10} \) is a partition point. We choose \( \kappa_{\text{true}}(x) \), solve the boundary value problem, sample the data at the \( x_i \), and possibly add noise to produce our observed data. The goal is now to recover \( \kappa_{\text{true}}(x) \).

**Table 7**

Collage coding results when \( f(x) = \delta \left( x - \frac{3}{10} \right) \), \( \kappa_{\text{true}}(x) = 1 + 2x \), data points = 10, number of basis functions = 10. The form of \( \kappa_{\text{collage}} \) indicates which method of discussion (iii) has been used.

<table>
<thead>
<tr>
<th>Noise ( \varepsilon )</th>
<th>Space ( \kappa_{\text{collage}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00 ( V_1 )</td>
<td>0.996829 + 2.00363x</td>
</tr>
<tr>
<td>0.00 ( V_2 )</td>
<td>0.99669 + 2.00572x − 0.00617x^2 + 0.00266x^3</td>
</tr>
<tr>
<td>0.001 ( V_3 )</td>
<td>1.20541 + 1.61207x</td>
</tr>
<tr>
<td>0.001 ( V_4 )</td>
<td>0.959467 + 5.70939x − 12.9070x^2 + 8.08981x^3</td>
</tr>
</tbody>
</table>

**Table 8**

Collage coding results for the parabolic equation.

<table>
<thead>
<tr>
<th>Noise ( \varepsilon )</th>
<th>( N )</th>
<th>( k_0 )</th>
<th>( k_1 )</th>
<th>( k_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
<td>0.87168</td>
<td>2.90700</td>
<td>0.21353</td>
</tr>
<tr>
<td>0</td>
<td>20</td>
<td>0.93457</td>
<td>2.97239</td>
<td>1.49201</td>
</tr>
<tr>
<td>0</td>
<td>30</td>
<td>0.94479</td>
<td>2.98304</td>
<td>1.76421</td>
</tr>
<tr>
<td>0</td>
<td>40</td>
<td>0.94347</td>
<td>2.97346</td>
<td>1.85572</td>
</tr>
<tr>
<td>0.01</td>
<td>10</td>
<td>0.87573</td>
<td>2.82810</td>
<td>0.33923</td>
</tr>
<tr>
<td>0.01</td>
<td>20</td>
<td>0.92931</td>
<td>2.91536</td>
<td>1.32864</td>
</tr>
<tr>
<td>0.01</td>
<td>30</td>
<td>0.92895</td>
<td>2.84553</td>
<td>0.59199</td>
</tr>
<tr>
<td>0.10</td>
<td>10</td>
<td>0.90537</td>
<td>1.97162</td>
<td>0.59043</td>
</tr>
<tr>
<td>0.10</td>
<td>20</td>
<td>0.77752</td>
<td>−0.12677</td>
<td>−0.14565</td>
</tr>
<tr>
<td>0.10</td>
<td>30</td>
<td>0.60504</td>
<td>−0.12677</td>
<td>−0.14565</td>
</tr>
</tbody>
</table>

Fig. 2. The solution to our test problem and the 9 data points used to generate the results of Table 7 (right to left) no noise added, and Gaussian noise with amplitude 0.001 added.
As a final experiment, we consider the case when

\[ \kappa_{\text{true}}(x) = \begin{cases} 
1, & x \leq \frac{1}{3} \\
2, & x > \frac{1}{3}.
\end{cases} \]

simulating a scenario when two rods with different constant diffusivities are melded at \( x = \frac{1}{3} \). We solve several instances of our boundary value problem, each with a single point source a distinct location, and then produce observational data by sampling the solution and possibly adding low amplitude Gaussian noise. As one can see in Fig. 4, collage coding with 40 uniformly placed unnoised data values taken from the solution with a point source at \( x_s = \frac{3}{10} \) produces a \( \kappa_{\text{collage}}(x) \) in fine agreement with our \( \kappa_{\text{true}}(x) \). However, when low amplitude noise is added to data points, the result worsens quite dramatically, as we see in rightmost plot of the figure.

(iv) Multiple data sets

It is most likely that several experiments would be performed in order to determine the conductivity \( \kappa(x) \)—for example, measuring the steady-state responses to point sources located at several positions \( x_k, i = 1, 2, \ldots, M \), on the rod. These results could be combined into one determination by considering the minimization of a linear combination of squared collage errors of the form (39), i.e.,

\[
(G_n(\lambda))^2 = \sum_{k=1}^{M} \mu_k(F_n^{(k)}(\lambda))^2
\]

\[
= \sum_{k=1}^{M} \mu_k \left( \sum_{i=1}^{n} \left[ \sum_{j=0}^{N} \lambda_j A_{ij}^{(k)} - b_i^{(k)} \right]^2 \right),
\]

(41)

where the \( k \)th set of elements \( A_{ij}^{(k)} \) and \( b_i^{(k)} \), \( k = 1, 2, \ldots, M \), is obtained from the response to the source at \( x_k \). The nonnegative \( \mu_k \) are weighting factors.

In the almost certain case of noisy data sets, it is possible that such a combination of data may improve the estimates of \( \kappa(x) \) obtained by minimizing (41). This is indeed demonstrated by the plots in Fig. 5, where a number of equally noisy sets of observed data have been combined. (Here \( \mu_k = 1 \).)
As an illustrative example, let us consider the following system:

**Example 5.1.** Describes the birth/growth rate of crystals coupled with an underlying temperature field. This is a multiple-scale model, in which the temperature evolves at a larger scale than many particles systems. The authors propose a stochastic model for the processes of nucleation and crystallization strongly coupled with a continuous underlying field. In [18], a multiple-scale structure model is introduced to analyze polymer crystallization processes via an underlying field. This is a multiple-scale model, in which the temperature evolves at a larger scale according to a classical diffusion equation (43), strongly coupled with a lower scale stochastic process (42) which describes the birth/growth rate of crystals.

**4.2. An inverse problem for the diffusion equation** \( u_t = (k(x)u_x)_x + f(x,t) \)

In this paragraph we consider an example of an inverse problem for the following diffusion equation \( u_t = (k(x)u_x)_x + f(x,t), \quad 0 < x < 1 \). If a steady-state solution \( \bar{u} \) exists, one could proceed by using it and the methodology developed for the steady-state case. In this example, let us suppose \( f(x,t) = tx(1-x) \), subject to \( u(x,0) = 10 \sin(\pi x) \) and \( u(0,t) = u(1,t) = 0 \). We set \( k(x) = 1 + 3x + 2x^2 \), solve for \( u(x,t) \), and sample the solution at \( N^2 \) uniformly positioned grid points for \( (x,t) \in [0,1]^2 \) to generate a collection of targets. Given this data and \( f(x,t) \), we then seek an estimation of \( k(x) \) in the form \( k(x) = k_0 + k_1 x + k_2 x^2 \). The results we obtain through the generalized collage method are summarized in Table 8. As for the steady-state case, the table shows that the method subject to noisy perturbations is stable.

**5. An application: solving an inverse problem for a mean field stochastic model**

Consider the following system of coupled stochastic differential equations:

\[
\begin{aligned}
\dot{X}_t &= \int_0^1 g(\lambda(y,t), X_t, y) f_{X_t}(y) dy \, dt + dB_t \\
X_{t=0} &= X_0 \\
\frac{d}{dt} \lambda(t,y) &= \kappa f_X(y) \lambda(t,y) + \varrho \Delta \lambda(t,y), \quad (t,y) \in [0,T] \times D \\
\lambda(0,y) &= \phi_1(y), \\
\frac{\partial \lambda}{\partial n}(t,y) &= \phi_2(t,y), \quad (t,y) \in [0,T] \times \partial D
\end{aligned}
\]

(42)

where \( f_X \) is the density of the distribution of \( X_t \) and \( \kappa \) and \( \varrho \) are parameters. The main difference between (17) and (42) is that the function \( g \) depends not only on the law of the process \( X_t \) but also on the function \( \lambda \) whose behavior is described through the diffusionequation (43). The previous model (42) and (43) describes a mean field stochastic model (see [16, 17]) which arises, for instance, when studying a large system of interacting particles (see [14]) strongly coupled with a continuous underlying field. In [18], a multiple-scale structure model is introduced to analyze polymer crystallization processes via many particle systems. The authors propose a stochastic model for the processes of nucleation and crystallization strongly coupled with an underlying temperature field. This is a multiple-scale model, in which the temperature evolves at a larger scales according to a classical diffusion equation (43), strongly coupled with a lower scale stochastic process (42) which describes the birth/growth rate of crystals.

**Example 5.1.** As an illustrative example, let us consider the following system:

\[
\begin{aligned}
\dot{X}_t &= \mu dt + dB_t \\
X_{t=0} &= X_0 \\
\frac{\partial}{\partial t} \lambda(t,y) &= f_X(y) \lambda(t,y) + \varrho \frac{\partial^2}{\partial y^2} \lambda(t,y), \quad (t,y) \in [0,T] \times [0,1] \\
\lambda(0,y) &= 10 \sin(\pi y), \\
\frac{\partial \lambda}{\partial n}(t,0) &= \frac{\partial \lambda}{\partial n}(t,1) = 0.
\end{aligned}
\]

(44)

(45)

Please cite this article in press as: V. Capasso, et al., Solving inverse problems for differential equations by a “generalized collage” method and application to a mean field stochastic model, Nonlinear Analysis: Real World Applications (2011), doi:10.1016/j.nonwva.2011.05.017
Fig. 6. The long-run behavior of $\lambda$.

In this case it is well known that the density of the process is:

$$f_X(t, y) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y - X_0 - \mu t)^2}{2t}\right).$$  \hspace{1cm} \text{(46)}$$

Fig. 6 shows the long-run behavior of $\lambda$ ($X_0 = 0$, $\mu = 1$, $\rho = 2$).

Let us consider the following weak formulation of the parabolic equation

$$\frac{d}{dt} \lambda(t, v) = \psi(v) + a(\lambda(t, v))$$

$$\lambda(0) = \phi_1$$  \hspace{1cm} \text{(47)}$$

where $\psi : L^2(D) \to \mathbb{R}$ is the linear functional given by

$$\psi(v) = \int_D f_X(y) \lambda(t, y)v(y)dy + \int_{\partial D} \phi_2(y)v(y)ds(y),$$

$$a : L^2(D) \times L^2(D) \to \mathbb{R}$$

be a bilinear form and $f \in L^2(D)$ be the initial condition. The aim of the inverse problem for the above system of equations \text{(42)} and \text{(43)} consists of getting an approximation of a target $\lambda^*$. To get this, Theorem 3.2 states that the distance between $\lambda^*$ and the solution $\lambda$ of \text{(47)} can be reduced by minimizing a functional which depends on $f_X$. In order to solve the inverse problem for the parabolic equation \text{(47)} one can minimize the following functional

$$\int_0^T \left( \sup_{\|v\|=1} \left( \frac{d}{dt} \lambda^*(t, v) - \psi(v) - a(\lambda^*(t, v)) \right)^2 dt \right)$$

over all $f_X$. By minimizing the previous functional \text{(49)}, for instance by projecting the unknown function $f_X$ on an orthonormal basis, one can get an estimation $\hat{f}_X$ of it which depends on $\lambda^*$. Going back to the stochastic differential equation:

$$dX_t = \left[ \int_{\mathbb{R}^d} g(\lambda^*(y, t), X_t, y)f_X(y)dy \right] dt + dB_t$$

$$X_{t=0} = X_0$$  \hspace{1cm} \text{(50)}$$

it is possible to use the methods developed in the previous sections and based on the collage theorem for getting an estimation of $g$. To clarify this step, let us consider the case in which $g$ is linear in $\lambda^*$ and $X_t$, namely $g(\lambda^*(y, t), X_t, y) = \alpha_0 \lambda^*(y, t) + \alpha_1 X_t$. By using the estimation of the density of the process $\hat{f}_X$, one can easily find an estimation of the expected value $\mathbb{E}(X_t)$ of $X_t$. Taking the expectation of both sides in \text{(50)} we get

$$\frac{d}{dt} \mathbb{E}(X_t) = \alpha_0 \left[ \int_{\mathbb{R}^d} \lambda^*(y, t)f_X(y)dy \right] + \alpha_1 \mathbb{E}(X_t)$$

$$X_{t=0} = X_0$$  \hspace{1cm} \text{(51)}$$

The unknown coefficients of this expansion can be found by minimizing the collage error as shown in Section 1.

Please cite this article in press as: V. Capasso et al., Solving inverse problems for differential equations by a “generalized collage” method and application to a mean field stochastic model, Nonlinear Analysis: Real World Applications (2011), doi:10.1016/j.nonwa.2011.05.017
Acknowledgments

This work was supported in part by research grants (HEK and ERV) from the Natural Sciences and Engineering Research Council of Canada (NSERC), which are hereby gratefully acknowledged.

References