1

Parameter identification for deterministic and stochastic differential equations using the "collage method" for fixed point equations

V. Capasso[1], H.E. Kunze[2], D. La Torre[3,2], E.R. Vrscay[4]

[1] *Department of Mathematics, University of Milan, Italy*
[2] *Department of Mathematics and Statistics, University of Guelph, Ontario, Canada*
[3] *Department of Economics, Business and Statistics, University of Milan, Italy*
[4] *Department of Applied Mathematics, University of Waterloo, Ontario, Canada*

Summary. A number of inverse problems may be viewed in terms of the approximation of a target element x in a complete metric space (X, d) by the fixed point $\bar{x}$ of a contraction function $T : X \to X$. In practice, from a family of contraction functions T_λ one wishes to find the parameter $\bar{\lambda}$ for which the approximation error $d(x, \bar{x}_\lambda)$ 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_λ for which the so called *collage distance* $d(x, T_\lambda x)$ is as small as possible. We first show how to solve inverse problems for deterministic and random differential equations and then 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. At the end we show an application of this approach to a system of coupled stochastic differential equations which describes the interaction between particles in a physical system.

Key words: Inverse problems, stochastic differential equations, fixed point equations, Monge-Kantorovich distance, Wasserstein metric, Collage Theorem.

1 Introduction

Many inverse problems may be viewed in terms of the approximation of a target element x in a complete metric space (X,d) by the fixed point $\bar{x}$ of a contraction function $T : X \to X$. Starting from a family of contraction functions T_λ, $\lambda \in \Lambda \subset \mathbb{R}^n$, one wishes to find the parameter $\bar{\lambda}$ for which the distance $d(x, \bar{x}_\lambda)$ is as small as possible. This problem can be solved using a simple consequence of Banach's Fixed Point Theorem; Barnsley has been the first one who showed the importance of this result for solving inverse problems for fractals and other similar objects in [1]. In the fractal literature, this is known as *Collage Theorem* and it has been the basis of most fractal image coding and compression methods.

Theorem 1. *[1] Let (X,d) be a complete metric space and $T : X \times \Lambda \to X$ a family of contraction functions with contraction factor $c_\lambda \in [0,1)$, $\lambda \in \Lambda$. Then for any $x \in X$,*

$$d(x, \bar{x}_\lambda) \leq \frac{1}{1 - c_\lambda} d(x, T_\lambda x), \tag{1}$$

where $\bar{x}_\lambda$ is the fixed point of $T_\lambda = T(\cdot, \lambda)$ and Λ is a set of parameters.

In other words, this result states that one can control the error distance between x and $\bar{x}$ by minimizing the distance $d(\bar{x}, T_\lambda \bar{x})$ (if the contraction factor c_λ is bounded away from one for all λ). This is the so-called *collage error* $d(x, T_\lambda x)$ – in other words, a function that sends the target x as close as possible to itself. This is the essence of the method of *collage coding* which has been the basis of fractal image coding [1]. An easy manipulation of the triangle inequality involving x, Tx and $\bar{x}$ yields the following interesting result.

Theorem 2. *[17] Assume the conditions of Theorem 1 above. Then for any $x \in X$,*

$$d(x, \bar{x}_\lambda) \geq \frac{1}{1 + c_\lambda} d(x, T_\lambda x), \tag{2}$$

The continuity theorem establishes that a small change in a contraction function generally can produce a small change in the associated fixed points.

Theorem 3. *[6] Let (X,d) be a complete metric space and T_1, T_2 be two contractive functions with contraction factors c_1 and c_2 and fixed points $\bar{x}_1$ and $\bar{x}_2$, respectively. Then*

$$d(\bar{x}_1, \bar{x}_2) \leq \frac{1}{1-c} d_{\sup}(T_1, T_2) \tag{3}$$

where

$$d_{\sup}(T_1, T_2) = \sup_{x \in X} d(T_1(x), T_2(x)) \tag{4}$$

and $c = \min\{c_1, c_2\}$.

These results can be extended to the case of *eventually contractive operators*, which we define below.

Definition 1. *Let* (X, d) *be a metric space. A function* $T : X \to X$ *is said to be* eventually contractive *on* X *if it is Lipschitz with Lipschitz constant* c *and if there exists an integer* $m \in \mathbb{N}$ *and a positive constant* $c_m < 1$ *such that:*

$$d(T^m x, T^m y) \leq c_m d(x, y), \tag{5}$$

where $T^m x = T(T^{m-1}x))$ *with* $T^0 x = x$.

It is well known that for eventually contractive operators a generalized version of Banach theorem holds. The inverse problem for the function T can be written by using a result similar to Theorem 1.

Theorem 4. *Let* (X, d) *be a complete metric space,* $x \in X$ *and* $T : X \times \Lambda \to X$ *be an eventually contractive function* $\forall \lambda \in \Lambda$. *Then there exists a constant* C_λ *s.t.*

$$d(x, \bar{x}_\lambda) \leq C_\lambda d(x, T_\lambda x) \tag{6}$$

where $\bar{x}_\lambda$ *is the unique fixed point of* $T_\lambda = T(\cdot, \lambda)$.

Proof. Since T_λ is an eventually contractive map then it is Lipschitz, that is there exists a constant c_λ such that $d(T_\lambda x, T_\lambda y) \leq c_\lambda d(x, y)$ for all $x, y \in X$, and there exists $m_0 \in \mathbb{N}$, $m_0 = m_0(\lambda)$, such that $T_\lambda^{m_0}$ is a contraction with Lipschitz constant $c_{\lambda, m_0} < 1$, for all $\lambda \in \Lambda$. Let $\bar{x}_\lambda$ be its fixed point. Using the Collage Theorem for T^{m_0} we have:

$$\begin{aligned}
(1 - c_{\lambda, m_0}) d(\bar{x}_\lambda, x) &\leq d(T_\lambda^{m_0} x, x) \\
&\leq \sum_{i=0}^{m_0 - 1} d(T_\lambda^{m_0 - i} x, T_\lambda^{m_0 - i - 1} x) \\
&\leq \sum_{i=0}^{m_0 - 1} c_\lambda^{m_0 - i - 1} d(T_\lambda x, x) \qquad (7) \\
& \qquad (8)
\end{aligned}$$

and so

$$d(x, \bar{x}_\lambda) \leq \frac{1 - c_\lambda^{m_0 - 1}}{1 - c_{\lambda, m_0}} d(x, T_\lambda x) \tag{9}$$

In analogous way one can extend theorems 2 and 3.

2 Inverse problems for differential equations using the "collage method"

Many problems in the parameter estimation literature for differential equations can be formulated in such a collage coding framework as showed in [12] and subsequent works [10, 11]. In [12], the authors showed how collage coding could be used to solve inverse problems for systems of differential equations having the form

$$\begin{cases} \dot{x} = f(t, x), \\ x(0) = x_0, \end{cases} \tag{10}$$

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

$$(Tx)(t) = x_0 + \int_0^t f(s, x(s))\, ds. \tag{11}$$

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_∞ metric and assume that f 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 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 δ and $M > 0$.

Theorem 5. *[12] The function T satisfies*

$$\|Tu - Tv\|_2 \leq c\|u - v\|_2 \tag{12}$$

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 (11) 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

$$f_\lambda(s, x) = \sum_{i=1}^{+\infty} \lambda_i \phi_i(s, x). \tag{13}$$

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

Theorem 6. *Let $K, \lambda \in \ell^2(\mathbb{R})$. Then*

$$|f_\lambda(s, x_1) - f_\lambda(s, x_2)| \leq \|K\|_2 \|\lambda\|_2 |x_1 - x_2| \tag{14}$$

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 $\|\lambda\|_2 = \left(\sum_{i=1}^{+\infty} \lambda_i^2\right)^{\frac{1}{2}}$

Proof. Computing, we have

$$
\begin{aligned}
|f_\lambda(s,x_1)-f_\lambda(s,x_2)| &\le \sum_{i=1}^{+\infty}|\lambda_i||\phi_i(s,x_1)-\phi_i(s,x_2)| \\
&\le \left(\sum_{i=1}^{+\infty}|\lambda_i|^2\right)^{\frac{1}{2}}\left(\sum_{i=1}^{+\infty}|\phi_i(s,x_1)-\phi_i(s,x_2)|^2\right)^{\frac{1}{2}} \\
&\le \|\lambda\|_2\|K\|_2|x_1-x_2|, \qquad (15)
\end{aligned}
$$

yielding the desired result.

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

$$
\begin{aligned}
\Delta^2(\lambda) &= \|x - T_\lambda x\|_2^2 \\
&= \int_{-\delta}^{\delta}\left| x(t) - \int_0^t \sum_{i=1}^{+\infty}\lambda_i\phi_i(s,x(s))ds\right|^2 dt \qquad (16)
\end{aligned}
$$

and the inverse problem can be formulated as

$$
\min_{\lambda\in\Lambda}\Delta(\lambda), \qquad (17)
$$

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:

$$
\min_{\lambda\in\tilde{\Lambda}}\tilde{\Delta}^2(\lambda) = \int_{-\delta}^{\delta}\left| x(t) - \int_0^t \sum_{i=1}^{n}\lambda_i\phi_i(s,x(s))ds\right|^2 dt, \qquad (18)
$$

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 $\tilde{\Delta}^n_{\min}$ be the minimum value of $\tilde{\Delta}$ over $\tilde{\Lambda}$. This is a non increasing sequence of numbers (depending on n) and as shown in [8] it is possible to show that $\liminf_{n\to+\infty}\tilde{\Delta}^n_{\min}=0$. This states that the distance between the target element and the unknown solution of the differential equation can be made arbitrary small.

Example 1. Consider the damped harmonic oscillator system

$$
\begin{cases}
\frac{dx_1}{dt} = x_2, \\
\frac{dx_2}{dt} = -kx_2 - bx_1, \\
x_1(0) = x_{10}, \\
x_2(0) = x_{20}.
\end{cases} \qquad (19)
$$

To simulate an experiment, we set $k=0.3$, $b=0.8$, $x_{10}=0.5$, $x_{20}=0.4$, and then solve numerically the system of ODEs. For $t\in[0,20]$, we sample the solutions at 40 uniformly spaced points. Degree-15 polynomials are fitted to the resulting simulated observational data. These two polynomials are our target functions. That is, we seek a Picard operator of the form in (11), with

$$f(x) = \begin{pmatrix} c_1 x_2 \\ c_2 x_2 + c_3 x_1 \end{pmatrix},$$

and with the components of x_0 as parameters, as well. The minimal-collage system to five decimal places is

$$f(x) = \begin{pmatrix} 1.0052 x_2 \\ -0.79935 x_2 - 0.30039 x_1 \end{pmatrix}, \; x_0 = \begin{pmatrix} 0.50133 \\ 0.39992 \end{pmatrix}.$$

3 Inverse problems for random DEs

In [13], Kunze, La Torre and Vrscay considered the case of inverse problems for random differential equations. This kind of problems can be formulated as

$$\begin{cases} \frac{d}{dt} x(\omega, t) = f(t, \omega, x(\omega, t)), \\ x(\omega, 0) = x_0(\omega). \end{cases} \tag{20}$$

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 $T : \Omega \times X \to X$:

$$(T_\omega x)(t) = x_0(\omega) + \int_0^t f(s, \omega, x(s)) \, ds. \tag{21}$$

Solutions to (20) are fixed points of (21), that is solution of the equation $T_\omega x = x$. We recall that a function $T : \Omega \times X \to X$ is called a *random operator* (in a strict sense, see [3], p. 104) if for any $x \in X$ the function $T(\cdot, x)$ is measurable. The random operator T is said to be continuous/Lipschitz/contractive if, for a.e. $\omega \in \Omega$, we have that $T(\omega, \cdot)$ is continuous/Lipschitz/contractive. A measurable mapping $x : \Omega \to X$ is called a *random fixed point* of the random operator T if x is a solution of the equation

$$T(\omega, x(\omega)) = x(\omega), \quad a.e.\ \omega \in \Omega. \tag{22}$$

In order to study the existence of solutions to such equations, let us consider the space Y of all measurable functions $x : \Omega \to X$. If we define the operator $\tilde{T} : Y \to Y$ as $(\tilde{T} x)(\omega) = T(\omega, x(\omega))$ the solutions of this fixed point equation on Y are the solutions of the random fixed point equation $T(\omega, x(\omega)) = x(\omega)$. The space Y is a complete metric space with respect to the following metric (see [13]):

$$d_Y(x_1, x_2) = \int_\Omega d_X(x_1(\omega), x_2(\omega)) dP(\omega). \tag{23}$$

Example 2. Consider the following random differential equation,

$$\begin{cases} \frac{dx}{dt}(t) = A_0 + A_1 x(t) + A_2 x^2(t), \\ x(0) = x_0. \end{cases} \tag{24}$$

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 $x(t, \omega_j)$ 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.

	True Values			
Label	A_0	A_1	A_2	x_0
1	$\mathcal{N}(1.2, 0.09)$	$\mathcal{N}(0.6, 0.04)$	$\mathcal{N}(0.4, 0.01)$	$\mathcal{N}(0.5, 0.09)$
2	$\mathcal{N}(0.5, 0.01)$	$\mathcal{N}(0.3, 0.01)$	$\mathcal{N}(0.3, 0.04)$	$\mathcal{N}(0.2, 0.01)$

Table 1 Distributions used in the inverse problem

		Minimal Collage Values	
Label	N	A_0	A_1
1	10	$(1.1755, 0.1477)$	$(0.5785, 0.0275)$
1	100	$(1.1665, 0.1016)$	$(0.5865, 0.0316)$
1	1000	$(1.2009, 0.0845)$	$(0.5944, 0.0376)$
2	10	$(0.4574, 0.4097)$	$(0.2899, 0.0069)$
2	100	$(0.4362, 0.2798)$	$(0.2956, 0.0082)$
2	1000	$(0.4953, 0.2326)$	$(0.2990, 0.0096)$

Table 2 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_0 and A_1.

		Minimal Collage Values	
Label	N	A_2	x_0
1	10	$(1.4042, 0.0073)$	$(0.4788, 0.0205)$
1	100	$(0.4236, 0.0086)$	$(0.5110, 0.0841)$
1	1000	$(0.3989, 0.0104)$	$(0.4889, 0.0894)$
2	10	$(0.3074, 0.0287)$	$(0.1929, 0.0023)$
2	100	$(0.3441, 0.0349)$	$(0.2037, 0.0094)$
2	1000	$(0.2959, 0.0410)$	$(0.1963, 0.0099)$

Table 3 Results for the inverse problem of Example 2. 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.

Example 3. Let us consider the following system of random equations:

$$\begin{cases} \frac{dx}{dt}(t) = Ax(t) + B_t, \\ x(0) = x_0. \end{cases} \tag{25}$$

where $x : \mathbb{R} \times \Omega \to \mathbb{R}^n$, A is a (deterministic) matrix of coefficients and B_t is a classical vector Brownian motion. As above, an inverse problem for this kind of equations can be formulated as: given an i.d. sample of observations of $x(t, \omega)$, say $(x(t, \omega_1), \ldots, x(t, \omega_n))$, get an estimation of the matrix A. For this purpose, let us take the integral over Ω of both sides of the previous equation and suppose that $x(t, \omega)$ is sufficiently regular; recalling that $B_t \sim \mathcal{N}(0, t)$, we have

$$\int_{\Omega} \frac{dx}{dt} dP(\omega) = \frac{d}{dt}\mathbb{E}(x(t,\cdot)) = A\mathbb{E}(x(t,\cdot)) \tag{26}$$

This is a deterministic differential equation in $\mathbb{E}(x(t,\cdot))$. From the sample of observations of $x(t,\omega)$ we can then get an estimation of $\mathbb{E}(x(t,\cdot))$ and then use of machinery of the previous section to solve the inverse problem for A.

4 Stochastic differential equations as fixed point equations and inverse problems

4.1 Mathematical preliminaries

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 know that the spaces $(C(X), d_H)$ and $(BC(X), d_H)$ are complete with respect to the Hausdorff metric defined as

$$d_H(A,B) = \max\left\{ \sup_{x\in A} \inf_{y\in B} d(x,y), \sup_{x\in B} \inf_{y\in A} d(x,y) \right\}. \tag{27}$$

For a given function f, let $\mathbb{L}ipf$ denote the Lipshitz 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 σ-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 ϕ.

Given $\mu, \nu \in M(X)$, the Prokhorov metric ρ on X is defined by

$$\rho(\mu,\nu) = \inf\{\epsilon > 0, \mu(A) \leq \nu(A^\epsilon) + \epsilon \text{ for all Borel sets } A \subset X\} \tag{28}$$

where $A^\epsilon = \{x \in X : d(x,A) \leq \epsilon\}$. In [2] it is shown that the space $(M(X), \rho)$ is complete and generates the topology of weak convergence.

Given $\mu, \nu \in M(X)$, the Wasserstein metric $D_T(\mu,\nu)$ on X is defined by

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

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 $(M(X), D_T)$ is complete and D_T gives to $M(X)$ the topology of weak convergence.

Let $(M_1(X), d_{MK})$ be the complete metric space (see [2]) consisting of all measures μ 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

$$\begin{aligned} d_{MK}(\mu,\nu) &= \sup_{f\in} \left\{ \int_X f d\mu - \int_X f d\nu : \mathbb{L}ipf \leq 1 \right\} \\ &= \inf_{\gamma} \left\{ \int_X d(x,y) d\gamma(x,y) : \gamma \text{ is a measure on } X\times X, \Pi_1(\gamma) = \mu, \Pi_2(\gamma) = \nu \right\} \end{aligned}$$

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 [7]. Between the convergence in the Monge-Kantorovich metric and the topology of convergence the following relation holds (see [2]):

$$\nu_n \to_{MK} \mu \text{ if and only if } \nu_n \to \mu \text{ and } \int_X d(x,a)d\nu_n(x) \to \int_X d(x,a)d\mu(x)$$

for all $a \in X$.

4.2 Stochastic differential equations as fixed point equations

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:

$$\begin{cases} dX_t = \int_{\mathbb{R}^d} g(X_t, y)d\mu_t(y)dt + dB_t \\ X_{t=0} = X_0 \end{cases} \tag{29}$$

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

$$X_0 + B_t + \int_0^t \int_{C([0,T])} g(X_s, w_s)dm(w_s)ds \tag{30}$$

If X_t is a solution of (31) then its law on $C([0,T])$ is a fixed point of Φ, and vice versa. We have the following theorem which states an existence and uniqueness result for (31).

Theorem 7. *[16] 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:*

$$\begin{cases} dX_t = \int_{\mathbb{R}^d} g(X_t, y)d\mu_t(y)dt + dB_t \\ X_{t=0} = X_0 \end{cases} \tag{31}$$

We have that:

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

$$D_t(\Phi(m_1), \Phi_2(m_2)) \leq c_T \int_0^t D_s(m_1, m_2)ds$$

where c_T is a constant and D_s is the distance between the images of m_1, m_2 on $C([0,s])$;

(ii) Φ is eventually contractive since there is a $k > 0$ so that

$$D_T(\Phi^k(m_1), \Phi^k(m_2)) \leq \frac{c_T^k T^k}{k!} D_T(m_1, m_2) = c_{T,k} D_T(m_1, m_2)$$

with $c_{T,k} < 1$.

(iii) there exists a unique solution, trajectorial and in law, of (31).

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 μ_n the estimated law of the process. We have the following trivial corollary of the Collage Theorem.

Corollary 1. *Let $\mu_n \in M(C[0,T])$ be the estimated law of the process. If μ is the law of the process X_t of (31) then there exists a constant C such that the following estimate holds:*

$$D_T(\mu, \mu_n) \leq C D_T(\Phi(\mu_n), \mu_n) \tag{32}$$

The inverse problem is then reduced to the minimization of $D_T(\Phi(\mu_n), \mu_n)$ which is a function of the unknown coefficients of g.

5 An example: solving an inverse problem for a system of coupled stochastic differential equations using the "collage method"

Consider the following system of coupled stochastic differential equations:

$$\begin{cases} dX_t = \int_{\mathbb{R}^d} g(\lambda(y,t), t) d\mu_t(y) X_t dt + dB_t \\ X_{t=0} = X_0 \end{cases} \tag{33}$$

$$\begin{cases} \frac{d}{dt}\lambda(t,y) = \kappa \delta_{X_t}(y) \lambda(t,y) + \varrho \Delta_y \lambda(t,y), & (t,y) \in [0,T] \times D \\ \lambda(0,y) = \phi_1(y), & \\ \frac{\partial \lambda}{\partial n}(t,y) = \phi_2(t,y), & (t,y) \in [0,T] \times \partial D \end{cases} \tag{34}$$

where μ_t is the law of X_t, δ_{X_t} is the Dirac function at the point X_t and κ and ϱ are parameters. For simplicity we suppose them equal to 1. This kind of models arises, for instance, when studying probabilistic interacting diffusion models between particles in a physical system (see e.g. [16]). If one is interested in the analysis of the system at larger scales, the previous stochastic model can be averaged as in [4] [5]. Averaged quantities allow one to get an easier model which can be studied through the collage method. At larger scales the quantity δ_{X_t} can be approximated using its expectation value, that is $\delta_{X_t}(y) \simeq \mathbb{E}(\delta_{X_t}(y))$. Since X_t is a.c. then it is known that $\mathbb{E}(\delta_{X_t}(y)) = f_{X_t}(y)$ which is the density of the distribution of X_t. So the previous model can be rewritten in an averaged form as

$$\begin{cases} dX_t = \int_{\mathbb{R}^d} g(\tilde{\lambda}(y,t), t) f_{X_t}(y) dy X_t dt + dB_t \\ X_{t=0} = X_0 \end{cases} \tag{35}$$

coupled with the following deterministic PDE

$$\begin{cases} \frac{d}{dt}\tilde{\lambda}(t,y) = f_{X_t}(y)\tilde{\lambda}(t,y) + \Delta_y \tilde{\lambda}(t,y), & (t,y) \in [0,T] \times D \\ \tilde{\lambda}(0,y) = \phi_1(y), & \\ \frac{\partial \tilde{\lambda}}{\partial n}(t,y) = \phi_2(t,y), & (t,y) \in [0,T] \times \partial D \end{cases} \tag{36}$$

Let us consider the following weak formulation of the last parabolic equation

$$\begin{cases} < \frac{d}{dt}\tilde{\lambda}_t, v >= \psi(v) + a(\tilde{\lambda}_t, v) \\ \lambda_0 = f \end{cases} \tag{37}$$

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

$$\psi(v) = \int_D f_{X_t}(y)\tilde{\lambda}(t,y)v(y)dy + \int_{\partial D} \phi_2(y)v(y)ds(y), \tag{38}$$

$a : L^2(D) \times L^2(D) \to \mathbb{R}$ 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 (33) and (34) consists of getting an approximation of g starting from a sample of observations of a target random variable λ^*. To get this, the following Theorem 8 states that the distance between $\tilde{\lambda}^*$ and the solution $\tilde{\lambda}$ of (37) can be reduced by minimizing a functional which depends on f_{X_t}.

Theorem 8. *Let $\tilde{\lambda}^* : [0,T] \to L^2(D)$ be the target solution which satisfies the initial condition in (37) and suppose that $\frac{d}{dt}\tilde{\lambda}^*_t$ exists and belongs to $L^2(D)$. Suppose that there exists $m > 0$ such that $a(v,v) \geq m\|v\|^2$ for all $v \in L^2(D)$. We have the following result:*

$$\int_0^T \|\tilde{\lambda}^*_t - \tilde{\lambda}_t\|^2_{L^2(D)} dt \leq \frac{1}{m^2} \int_0^T \left(\sup_{\|v\|=1} < \tilde{\lambda}^*_t, v > -\psi(v) - a(\tilde{\lambda}^*_t, v) \right)^2 dt \tag{39}$$

where $\tilde{\lambda}_t$ is the solution of (37) s.t. $\tilde{\lambda}_0 = \tilde{\lambda}^*_0$ and $\tilde{\lambda}_T = \tilde{\lambda}^*_T$.

Proof. Computing we have

$$\begin{aligned} m\|\tilde{\lambda}^*_t - \tilde{\lambda}_t\|^2 &\leq a(\tilde{\lambda}^*_t - \tilde{\lambda}_t, \tilde{\lambda}^*_t - \tilde{\lambda}_t) \\ &= a(\tilde{\lambda}^*_t, \tilde{\lambda}^*_t - \tilde{\lambda}_t) - < \frac{d}{dt}(\tilde{\lambda}_t - \tilde{\lambda}^*_t), \tilde{\lambda}^*_t - \tilde{\lambda}_t > \\ &+ \psi(\tilde{\lambda}^*_t - \tilde{\lambda}_t) - < \frac{d}{dt}\tilde{\lambda}^*_t, \tilde{\lambda}^*_t - \tilde{\lambda}_t > \end{aligned}$$

and, by easy calculations, we get

$$m\|\tilde{\lambda}^*_t - \tilde{\lambda}_t\|^2 - \frac{1}{2}\frac{d}{dt}\|\tilde{\lambda}_t - \tilde{\lambda}^*_t\|^2 \leq \tag{40}$$

$$a(\tilde{\lambda}^*_t, \tilde{\lambda}^*_t - \tilde{\lambda}_t) + \psi(\tilde{\lambda}^*_t - \tilde{\lambda}_t) - < \frac{d}{dt}\tilde{\lambda}^*_t, \tilde{\lambda}^*_t - \tilde{\lambda}_t > \tag{41}$$

Integrating both sides with respect to t and recalling that $\tilde{\lambda}_0 = \tilde{\lambda}^*_0$ and $\tilde{\lambda}_T = \tilde{\lambda}^*_T$, we have

$$m \int_0^T \|\tilde{\lambda}^*_t - \tilde{\lambda}_t\|^2_{L^2(D)} \leq$$

$$\int_0^T \|\tilde{\lambda}^*_t - \tilde{\lambda}_t\|_{L^2(D)} \left\{ \sup_{\|v\|=1} a_t(\tilde{\lambda}^*_t, v) + \psi(v) - < \frac{d}{dt}\tilde{\lambda}^*_t, v > \right\} dt \leq$$

$$\left(\int_0^T \|\tilde{\lambda}^*_t - \tilde{\lambda}_t\|^2_{L^2(D)} dt \right)^{\frac{1}{2}} \left(\int_0^T \left(\sup_{\|v\|=1} a(\tilde{\lambda}^*_t, v) + \psi(v) - < \frac{d}{dt}\tilde{\lambda}^*_t, v > \right)^2 dt \right)^{\frac{1}{2}}$$

and now the thesis easily follows.

The previous result states that in order to solve the inverse problem for the parabolic equation (37) one can minimize the following functional

$$\int_0^T \left(\sup_{\|v\|=1} < \tilde{\lambda}^*_t, v > -\psi(v) - a(\tilde{\lambda}^*_t, v) \right)^2 dt \tag{42}$$

over all f_{X_t}. By minimizing the previous functional (42), for instance by projecting the unknown function f_{X_t} on an orthonormal basis, one can get an estimation $\hat{f}_{X_t}$ of it. Of course the estimation $\hat{f}_{X_t}$ depends on $\tilde{\lambda}^*$ but, in practical applications, one can get observations of the solution of the real model, that is λ. The distance between the solution of the real model and the averaged one, that is the distance between λ and $\tilde{\lambda}$, can be interpreted as noise (see [5] for more details about this). So starting from observations of λ^* one can get an estimation of $\tilde{\lambda}^*$ which can be used for solving the inverse problem. Going back to the stochastic differential equation:

$$\begin{cases} dX_t = \int_{\mathbb{R}^d} g(\tilde{\lambda}^*(y,t), t) \hat{f}_{X_t}(y) dy X_t dt + dB_t \\ X_{t=0} = X_0 \end{cases} \tag{43}$$

it is possible to use the collage method for getting an estimation of g by minimizing the distance $D_T(\Phi(\mu_n)), \mu_n)$, where μ_n is the estimated law of the process.

References

1. Barnsley, M.F., *Fractals Everywhere*, Academic Press, New York, 1989.
2. Barnsley, M.F., Hutchinson, J.E., Stenflö: V-variable fractals, *preprint*, (2006).
3. Bharucha-Reid, A. T., *Random integral equations.* Mathematics in Science and Engineering, Vol. 96. Academic Press, New York-London, 1972.
4. Burger, M., Capasso, V., Pizzocchero, L.: Mesoscale averaging of nucleation and growth models. *SIAM Multiscale Model. Simul.*, **5**, 564-592 (2006).
5. Capasso, V., Engl H.W., Kindermann, S.: Parameter identification in a random environment exemplified by a multiscale model for crystal growth. *SIAM Multiscale Model. Simul.*, to appear, (2008).
6. P. Centore and E.R. Vrscay, Continuity of fixed points for attractors and invariant meaures for iterated function systems, *Canadian Math. Bull.*, **37**, 315-329 (1994).

7. Dudley, R.M.: *Real Analysis and Probability*, Cambridge University Press, Cambridge, 2002.
8. B. Forte and E.R. Vrscay, Inverse problem methods for generalized fractal transforms, in *Fractal Image Encoding and Analysis*, NATO ASI Series F, Vol 159, ed. Y.Fisher, Springer Verlag, New York, 1998.
9. Fernandez, B., Meleard, S.: Asymptotic behaviour for interacting diffusion processes with space-time random birth, *Bernoulli*, **6**, 91-111 (2000).
10. Kunze, H., Gomes, S.: Solving an inverse problem for Urison-type integral equations using Banach's fixed point theorem, *Inverse Problems* **19**, 411-418 (2003).
11. Kunze, H., Hicken, J.., Vrscay, E.R.: Inverse problems for ODEs using contraction functions: suboptimality of the “Collage Method”, *Inverse Problems*, **20**, 977-991 (2004).
12. Kunze, H., Vrscay, E.R.: Solving inverse problems for ordinary differential equations using the Picard contraction function, *Inverse Problems*, **15**, 745-770 (1999).
13. Kunze, H., La Torre, D., Vrscay, E.R.: Random fixed point equations and inverse problems using collage method for contraction mappings, *J.Math.Anal.Appl.*, **334**, 1116-1129 (2007).
14. Iacus, S.M., La Torre, D.: A comparative simulation study on the IFS distribution function estimator. *Nonlin. Anal.: Real World Appl.*, **6**, 774-785 (2005)
15. La Torre, D., Mendivil, F.: Iterated function systems on multifunctions and inverse problems. *J.Math.Anal.Appl.*, **340**, 1469-1479 (2008).
16. Sznitman, A.S., Topics in propagation of chaos. In P.L.Hennequin (Ed.), École d'ÉtÉ de Probabilités de Saint-Flour XIX-1989. *Lecture Notes in Mathematics* 1464, Springer Verlag, Berlin, 1991.
17. E.R. Vrscay and D. Saupe, Can one break the ‘collage barrier’ in fractal image coding?, in *Fractals: Theory and Applications in Engineering*, ed. M. Dekking, J. Levy-Vehel, E. Lutton and C. Tricot, Springer Verlag, London, 1999, pp.307-323 .