

# Jornadas de Automática

## Aproximación a la identificación no paramétrica de sistemas muestreados asincrónamente mediante interpolación de Lipschitz

Orihuela, L.<sup>a,\*</sup>, Manzano, J.M.<sup>b</sup>

<sup>a</sup>Dpto. de Ingeniería Electrónica, Sistemas Informáticos y Automática, Universidad de Huelva. Avda. de las Fuerzas Armadas s/n 21007 Huelva, España.

<sup>b</sup>Dpto. de Ingeniería, Universidad Loyola Andalucía. Avda. de las Universidades s/n 41704 Dos Hermanas, Sevilla, España.

**To cite this article:** Orihuela, L., Manzano, J.M. 2024. Approach to non-parametric identification of asynchronously sampled systems with Lipschitz interpolation. *Jornadas de Automática*, 45.  
<https://doi.org/10.17979/ja-cea.2024.45.10952>

### Resumen

En este artículo se presenta una primera aproximación al problema de identificación no paramétrica de sistemas muestreados asincrónamente. La solución propuesta parte, en primer lugar, de la generación de un cuaderno de trayectorias, que se ajustan a partir de los datos asíncronos disponibles. Posteriormente, se presenta un mecanismo de aprendizaje no paramétrico basado en interpolación de Lipschitz. Éste permite realizar predicciones, con error acotado, del valor de los estados del sistema. El mecanismo de aprendizaje, aunque basado en la literatura, es novedoso, al tener que manejar trayectorias  $n$ -dimensionales, en vez de vectores en  $\mathbb{R}^n$ . Se valida el método sobre el sistema dinámico caótico conocido como el atractor de Lorenz.

**Palabras clave:** Identificación de sistemas no lineales, Métodos no paramétricos, Identificación con error acotado, Muestreo asíncrono.

### Approach to non-parametric identification of asynchronously sampled systems with Lipschitz interpolation

#### Abstract

This manuscript presents a first approach to the problem of non-parametric identification of asynchronously sampled systems. The solution is based, firstly, in the generation of a diary of trajectories, which are found by fitting to the available asynchronous data. Then, a non-parametric learning rule based on Lipschitz interpolation is introduced. It allows to make predictions, with bounded error, of the value of the states. The learning rule, in spite of being inspired by the literature, is new, as it needs to handle  $n$ -dimensional trajectories, instead of vectors in  $\mathbb{R}^n$ . The method is validated over the chaotic dynamic system known as Lorenz attractor.

**Keywords:** Nonlinear system identification, Nonparametric methods, Bounded error identification, Asynchronous sampling.

## 1. Introducción

En muchas aplicaciones prácticas en las que se cuenta con redes de sensores, el muestreo de datos no siempre se realiza de manera síncrona. Esto es particularmente relevante en escenarios donde la adquisición de datos depende de eventos aleatorios o no controlables, lo que da lugar a secuencias de datos muestreados de manera asíncrona. Este tipo de muestreo introduce desafíos significativos en la identificación y modelado de sistemas dinámicos, dado que las técnicas tradicionales suelen asumir la disponibilidad de datos a intervalos regulares.

El problema de la identificación de sistemas dinámicos ha sido ampliamente estudiado en la literatura, con una variedad de enfoques. Con el auge de los métodos de aprendizaje automático, últimamente se han obtenido importantes avances en la identificación basada en datos (Quaranta et al., 2020). En ellos, se suele describir el sistema en forma autorregresiva, abriendo luego un abanico de métodos, desde paramétricos (como redes neuronales), hasta no paramétricos, como los procesos gaussianos (Williams and Rasmussen, 1995).

El muestreo asíncrono introduce una capa adicional de complejidad en la identificación de sistemas, ya que los in-

tervalos de tiempo entre las observaciones no son constantes.

Diversos enfoques han sido propuestos para abordar la identificación de sistemas con datos asíncronos, que en la literatura también se conocen como *non-uniformly sampled-data* (NUSD), o datos muestreados de manera no uniforme. Entre ellos, destacan los métodos basados en interpolación y reconstrucción de trayectorias, que buscan generar una representación continua de los datos a partir de las observaciones asíncronas. Estos métodos han sido estudiados desde hace décadas, por grandes expertos en identificación de sistemas, como S. Billings, uno de los creadores de los modelos NARMAX (véase Tsang and Billings (1995)).

Más recientemente, Xie et al. (2017) propusieron un método que usa transformadas específicas para modelar sistemas lineales muestreados asíncronamente. Ouyang and Chen (2023) propusieron usar mínimos cuadrados para ajustar el modelo usando dos sensores para un mismo estado, con período constante, pero no sincronizados; y Xu et al. (2021) propusieron usar métodos frecuenciales. Annamalai et al. (2022) diseñaron un modelo *long short term memory* para visión artificial, con cámaras no sincronizadas. Más ejemplos se proporcionan en Bekiroglu et al. (2017); Johansson (2009); Wang and Xia (2013).

Por otro lado, la interpolación de Lipschitz (Beliakov, 2006) es una técnica no paramétrica que garantiza un error acotado en la predicción de los valores de las trayectorias, lo que la hace particularmente adecuada para sistemas dinámicos, (Manzano et al., 2019). Sin embargo, todos los estudios, hasta lo que conocen los autores, se han centrado en aplicaciones con datos muestreados de manera regular, y no han abordado los desafíos asociados con el muestreo asíncrono.

El presente trabajo aborda el problema de identificación no paramétrica de sistemas muestreados asíncronamente. No se corresponde con un problema de identificación clásico porque no se pretende reconstruir el modelo del sistema, ni estimando sus parámetros a partir de una estructura de modelo conocida, ni con un método generalista capaz de imitar diferentes dinámicas, como una red neuronal. Conceptualmente, el problema encaja en el ámbito de regresión para series temporales. Las dos características diferenciales del problema radican en el carácter asíncrono de los datos, y en el carácter determinista de las predicciones.

La contribución que se realiza en este artículo con respecto al estado del arte es doble: en primer lugar, se propone un método de transformar las observaciones asíncronas para obtener trayectorias del sistema y así generar el denominado cuaderno de trayectorias. Aunque la idea en sí no es nueva, ya que se ha usado en trabajos previos como Tsang and Billings (1995); Brunton et al. (2016), se ha adaptado a los requisitos del método de aprendizaje.

La segunda contribución, y la más importante, es la adaptación del método denominado *kinky inference*, introducido por Calliess (2014), para realizar predicciones del estado del sistema, en base a las trayectorias aprendidas a partir de los datos.

Se valida el método propuesto sobre un caso simulado: la identificación del sistema dinámico caótico conocido como el atractor de Lorenz. El resto del artículo se estructura como sigue: en la Sección 2 se introducen algunos preliminares, como el método de interpolación estándar y su uso para identifica-

ción no paramétrica de sistema dinámicos muestreados síncronamente. En la Sección 3 se plantea el problema a abordar en este trabajo, cuya proposición se presenta en la Sección 4. Por último, se realizan simulaciones ilustrativas de su funcionamiento en la Sección 5.

## 2. Preliminares

En esta sección se presenta el problema y el método de aprendizaje que sirven de punto de partida para el método que se propone en este artículo.

**Problema 1.** *Dada una función desconocida  $z = f(w)$  con espacio de entrada  $w \in \mathcal{W} \subseteq \mathbb{R}^{n_w}$ , y espacio de salida  $z \in \mathcal{Z} \subseteq \mathbb{R}^{n_z}$ , así como un conjunto de  $N$  observaciones  $\mathcal{D} = \{(w_k, z_k), k = 1, \dots, N\}$ , hállese un método que permita obtener predicciones de  $z$ , denotadas  $\hat{z}$ , para entradas  $q \in \mathcal{W}$ , y de la descripción de su incertidumbre, tal que*

$$\hat{z} = \hat{f}(q; \theta, \mathcal{D}), \quad (1)$$

donde  $\theta$  agrupa los parámetros del método de predicción  $\hat{f}$ .

El método denominado *kinky inference* asume que la función  $f$  es continua de Lipschitz.

**Definición 1** (Continuidad de Lipschitz). *Dada una función  $z = f(w)$  con espacio de entrada  $w \in \mathcal{W} \subseteq \mathbb{R}^{n_w}$  y espacio de salida  $z \in \mathcal{Z} \subseteq \mathbb{R}^{n_z}$ , se dice que es continua de Lipschitz para cada componente de la salida si para cualesquiera  $w_1, w_2 \in \mathcal{W}$ :*

$$\|f_j(w_1) - f_j(w_2)\| \leq L_j \|w_1 - w_2\|, \quad \forall j = 1, \dots, n_z, \quad (2)$$

donde  $L > 0 \in \mathbb{R}^{n_z}$  es la denominada constante de Lipschitz de la función.

Si la función a aprender es continua de Lipschitz, el método de aprendizaje se define como sigue:

**Definición 2** (Método de *kinky inference*). *Sea  $z = f(w)$ , con  $w \in \mathcal{W} \subset \mathbb{R}^{n_w}$  y  $z \in \mathcal{Z} \subset \mathbb{R}^{n_z}$ , una función continua de Lipschitz a aprender y  $\mathcal{D} = \{(w_k, z_k), k = 1, \dots, N\}$  un conjunto de observaciones. Entonces, para una entrada  $q \in \mathcal{W}$ , la predicción se obtiene según:*

$$\hat{z} = \hat{f}(q; L, \mathcal{D}) = \frac{1}{2} \left( \min_{k=1, \dots, N} (z_k + L\|q - w_k\|) + \max_{k=1, \dots, N} (z_k - L\|q - w_k\|) \right). \quad (3)$$

Nótese que el único parámetro del modelo es la propia constante de Lipschitz de la función,  $L$ . Si ésta es desconocida, se propone usar una estimada a partir de los datos disponibles,  $L^{\mathcal{D}}$ , que puede obtenerse como se propone en Calliess et al. (2020):

$$L^{\mathcal{D}} = \max_{(w_1, z_1) \in \mathcal{D}} \left( \frac{\|z_k - z_j\|}{\|w_k - w_j\|} \mid k, j = 1, \dots, N \wedge \|w_k - w_j\| > 0 \right). \quad (4)$$

El método de predicción goza de determinadas propiedades que han demostrado ser útiles para desarrollar modelos de sistemas (y controladores basados en ellos) usando esta técnica. Para más detalles, se refiere al lector a Manzano et al. (2019). Entre estas propiedades, se encuentran la simpleza en

el aprendizaje y la predicción; la consistencia de la predicción con las observaciones, proporcional a la cota del ruido (véase (Calliess et al., 2020, Lemma 2)); o la cota en la incertidumbre de la predicción. Calliess et al. (2020) demuestra que dicha cota en el error de predicción,  $\mu \in \mathbb{R}^n$ , viene dada por

$$\|\hat{f}(q; L^{\mathcal{D}}, \mathcal{D}) - f(q)\| \leq \mu = (L + L^{\mathcal{D}})r_{\mathcal{D}}, \quad (5)$$

donde  $r_{\mathcal{D}} \in \mathbb{R}$  denota la máxima distancia entre cualquier punto a predecir ( $q \in \mathcal{W}$ ) y el conjunto de observaciones.

Por último, se introduce a continuación el concepto de norma  $\mathcal{L}_p$  de una función, que generaliza el concepto de norma  $p$  de vectores en un espacio de funciones o dimensión infinita.

**Definición 3.** (Reddy and Rasmussen, 1990) Dada una función  $f : \mathbb{R} \rightarrow \mathbb{R}$ , se define su norma  $\mathcal{L}_p$  como:

$$\|f(x)\|_{\mathcal{L}_p} = \left( \int_{x \in X} |f(x)|^p dx \right)^{1/p}, \quad (6)$$

con  $X \subseteq \mathbb{R}$ .

Dada una función  $g : \mathbb{R} \rightarrow \mathbb{R}^n$ , su norma  $\mathcal{L}_p$  satisface:

$$\|g(x)\|_{\mathcal{L}_p} = \left( \sum_{i=1}^n \|g_i(x)\|_{\mathcal{L}_p}^p \right)^{1/p}, \quad (7)$$

es decir, la potencia  $p$ -ésima de su norma es la suma de las potencias  $p$ -ésimas de las normas de las componentes.

### 3. Descripción del problema

Considérese un sistema dinámico autónomo no lineal descrito por:

$$\frac{dx(t)}{dt} = f(x(t)), \quad (8)$$

donde  $x(t) \in \mathbb{R}^n$  son los estados del sistema, y  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  es su función de transición, desconocida, que se asume invariante en el tiempo.

De la evolución dinámica de dicho sistema se conoce el valor de los estados en determinados instantes de tiempo. Para mayor generalidad, se asume que dichos estados no se muestrean de forma periódica, ni siquiera se muestrean todos los estados a la vez. Por tanto, el conjunto de datos disponible se puede representar mediante:

$$\mathcal{D} = \{(E(t_k)x(t_k); t_k), k = 1, \dots, N_{\mathcal{D}}\}, \quad (9)$$

donde  $E(t_k)$  es una matriz, cuyas filas son vectores euclídeos, que indica qué componentes del estado han sido medidas en el instante  $t_k$ . Si el estado completo fuese medido en  $t_k$ , entonces  $E(t_k) = I$ . La cardinalidad del conjunto  $\mathcal{D}$  está determinada por  $N_{\mathcal{D}}$ . Nótese que, además del valor de los estados medidos, se almacena el tiempo  $t_k$  en el que éstos fueron muestreados.

**Problema 2.** Dado el conjunto de datos (9), diseñar un mecanismo de aprendizaje no paramétrico que permita obtener predicciones  $\hat{x}(t)$  del valor del estado del sistema dinámico desconocido (8) en un instante cualquiera  $t$ , así como proporcionar límites deterministas  $\hat{\underline{x}}_i(t), \hat{\bar{x}}_i(t)$  sobre cada una de las componentes del estado, tal que:

$$\hat{\underline{x}}_i(t) \leq x_i(t) \leq \hat{\bar{x}}_i(t), \quad i = 1, \dots, n. \quad (10)$$

### 4. Descripción de la solución

La solución que se propone se basa en dos ideas: en primer lugar, en el ajuste de funciones que tratan de recoger la trayectoria pasada del sistema y así, capturar el comportamiento dinámico del mismo; en segundo lugar, en la adaptación del método de KI para realizar inferencias a partir de funciones, de forma que pueda predecirse el valor de los estados.

#### 4.1. De los datos al cuaderno de trayectorias

El conjunto de datos  $\mathcal{D}$  no resulta adecuado en su forma original para realizar las predicciones. Por ello, conviene definir un conjunto, que se denominará *cuaderno*, que contiene información útil para el método de aprendizaje. Cada entrada del cuaderno  $C$  es una tupla de la forma:

$$C = \left\{ \left( g^{(k)}(t); \tilde{x}(t_k); t_k \right), t \in T_{\text{AR}}^{(k)}, k \in \{1, \dots, N_{\mathcal{D}}\} \right\}, \quad (11)$$

donde  $g^{(k)}(t) : \mathbb{R} \rightarrow \mathbb{R}^n$  define una trayectoria en un espacio  $n$ -dimensional hallada con los datos disponibles en el intervalo  $T_{\text{AR}}^{(k)} = [t_k - \tau, t_k)$ , siendo  $\tau$  el horizonte autorregresivo. Como las trayectorias están parametrizadas con el tiempo  $t$ ,  $g^{(k)}(t)$  define  $n$  trayectorias unidimensionales  $g_i^{(k)}(t)$  para cada uno de los estados del sistema. El segundo elemento de la tupla,  $\tilde{x}(t_k)$ , representa la estimación del estado en el instante  $t_k$ , que se obtiene también con los datos de  $\mathcal{D}$ . Finalmente, se incluye el tiempo  $t_k$  en el que se realiza la predicción. El número de entradas en el cuaderno se denota  $N_C$  y es, como máximo,  $N_{\mathcal{D}}$ .

Para construir el cuaderno, se establecen las funciones que modelarán las trayectorias, y se ajustan los parámetros de dichas funciones para que la trayectoria se aproxime a los datos. En este artículo se propone el uso de funciones polinómicas de grado  $m$  para cada una de las trayectorias unidimensionales, de forma que  $g_i^{(k)}(t) = \alpha_{0,i} + \alpha_{1,i}t + \dots + \alpha_{m,i}t^m$ ,  $i = 1, \dots, n$ , donde se ha omitido el superíndice ( $k$ ) en los coeficientes  $\alpha_{i,i}$  para simplificar la notación. Las constantes  $\alpha_{i,i}$  se pueden obtener minimizando alguna métrica relacionada con la distancia de la trayectoria  $g_i^{(k)}(t)$  a los datos disponibles en  $T_{\text{AR}}^{(k)}$ .

Con respecto a la estimación  $\tilde{x}(t_k)$ , conviene destacar que, en general,  $\tilde{x}(t_k) \neq x(t_k)$ , pues el estado completo no tiene por qué haberse muestreado en  $t_k$ . La forma de obtener dicha estimación podría ir desde usar los datos muestreados más cercanos en el tiempo a  $t_k$ , hasta una interpolación con datos anteriores y posteriores a  $t_k$ , pasando por un ajuste más complejo con los datos disponibles en un intervalo que incluya a  $t_k$ .

Finalmente, y dado que el sistema (8) se asume invariante en el tiempo, todas las trayectorias pueden retrasarse en el tiempo una cantidad de  $t_k$  segundos, y referenciarlas al intervalo único  $T = [-\tau, 0)$ . Esto permite definir un cuaderno más sencillo, denotado como  $C_o$ , formado por duplas:

$$C_o = \left\{ \left( g^{(k)}(t); \tilde{x}^{(k)}(0) \right), t \in [-\tau, 0), k \in \{1, \dots, N_{\mathcal{D}}\} \right\}, \quad (12)$$

donde la estimación desplazada toma el mismo valor que la estimación en  $t_k$ , es decir,  $\tilde{x}^{(k)}(0) = \tilde{x}(t_k)$ .

**Nota 1.** Para generar una entrada del cuaderno se requieren, al menos,  $m$  observaciones de cada uno de los estados en el intervalo  $T_{\text{AR}}^{(k)}$ . Por ello, en general,  $N_C < N_{\mathcal{D}}$ .

**Nota 2.** Para un sistema en el que  $x(t)$  se muestrea periódicamente cada  $T_s$  segundos, un intervalo dado por  $\tau = mT_s$ , con  $m \in \mathbb{N} > 0$ , incluye datos suficientes para generar trayectorias definidas como polinomios de grado  $m - 1$ . En ese caso, el método coincide con un modelo autorregresivo para series temporales en el que  $x(t_k)$  se hace depender de  $x(t_k - T_s), x(t_k - 2T_s), \dots, x(t_k - mT_s)$ .

**Nota 3.** Tanto el valor del parámetro  $\tau$ , como el orden del polinomio y el mecanismo de ajuste de las trayectorias a los datos, tienen un impacto importante en el método de aprendizaje. El valor de  $\tau$  debe escogerse en relación a las constantes de tiempo del sistema. Sin embargo, no se asume conocimiento alguno del sistema. Por otra parte, con un ajuste mediante mínimos cuadrados ponderados, podría darse más importancia a datos presentes que pasados, o incluso forzar a que la trayectoria pase por alguno de los datos medidos.

#### 4.2. Mecanismo de aprendizaje no paramétrico

Para resolver el Problema 2, se presenta una modificación del método de *kinky inference* descrito en la Definición 2.

En primer lugar, se presenta una reformulación del Problema 1 original que, de ser resuelto, ofrece una solución al Problema 2.

**Problema 3.** Dada una función desconocida  $x = h(g(t))$  con espacio de entrada  $g(t) \in \mathcal{G} \subseteq C^\infty(\mathbb{R}, \mathbb{R}^n)$  y  $t \in T \subset \mathbb{R}$ , y espacio de salida  $x \in \mathcal{X} \subseteq \mathbb{R}^n$ , así como un conjunto de  $N_C$  observaciones  $C_o = \{(g^{(k)}(t); \tilde{x}^{(k)}(0)), t \in [-\tau, 0], k \in \{1, \dots, N_D\}\}$ , hállese un método que permita obtener predicciones de  $x$ , denotadas  $\hat{x}$ , para entradas  $q(t) \in \mathcal{G}$  y  $t \in T$ , y de la descripción de su incertidumbre, tal que

$$\hat{x} = \hat{h}(q(t); \theta, C_o), \quad (13)$$

donde  $\theta$  agrupa los parámetros del método de predicción  $\hat{h}$ .

Siguiendo la presentación del problema original, se presenta a continuación la hipótesis que se le asume a la función  $h(g(t))$ , que es una generalización de la continuidad de Lipschitz presentada en la Definición 1.

**Definición 4.** (Xu et al., 2020) Sean  $(S_1, \|\cdot\|_{S_1})$  y  $(S_2, \|\cdot\|_{S_2})$  dos espacios normados. Una función  $f: S_1 \rightarrow S_2$  se dice que es continua de Lipschitz para conjuntos compactos si, para cualquier conjunto compacto  $\Omega_C \subseteq S_1$ , existe una constante positiva  $L(\Omega_C) > 0$  tal que:

$$\|f(\phi) - f(\psi)\|_{S_2} \leq L(\Omega_C) \|\phi - \psi\|_{S_1}, \quad \forall \phi, \psi \in \Omega_C. \quad (14)$$

Como se detalla en el Problema 3, la función desconocida  $x = h(g(t))$  transforma  $n$  trayectorias definidas en el intervalo genérico  $T = [-\tau, 0)$  a un vector en  $\mathbb{R}^n$ . El método de predicción se va a definir componente a componente, por lo que son las funciones  $h_i(g(t))$ ,  $i = 1, \dots, n$ , las que se asumen continuas de Lipschitz de acuerdo a la definición anterior. Para el espacio de las trayectorias  $C^\infty(\mathbb{R}, \mathbb{R}^n)$  se emplea la norma  $\mathcal{L}_p$  de la Definición 3. Para el espacio  $\mathbb{R}^n$ , se utiliza la norma euclídea estándar.

En este momento, se está en disposición de presentar el mecanismo de aprendizaje propuesto.

**Definición 5** (Método de *kinky inference* para sistemas muestreados asincrónicamente). Sean  $x_i = h_i(g(t))$  para  $i = 1, \dots, n$ , con  $g(t) \in \mathcal{G} \subseteq C^\infty(\mathbb{R}, \mathbb{R}^n)$  para  $t \in T \subset \mathbb{R}$  y  $x_i \in \mathcal{X}_i \subseteq \mathbb{R}$ , funciones continuas de Lipschitz a aprender, y sea  $C_o = \{(g^{(k)}(t); \tilde{x}^{(k)}(0)), t \in [-\tau, 0], k \in \{1, \dots, N_D\}\}$  un conjunto de observaciones. Entonces, para una entrada  $q(t) \in \mathcal{G}$  desplazada al intervalo  $T$ , la predicción se obtiene según:

$$\hat{x}_i(0) = \hat{h}_i(q(t); L_i, C_o) = \frac{1}{2} \left( \min_{k \in \{1, \dots, N_D\}} (\tilde{x}_i^{(k)}(0) + L_i \|q(t) - g^{(k)}(t)\|_{\mathcal{L}_p}) + \max_{k \in \{1, \dots, N_D\}} (\tilde{x}_i^{(k)}(0) - L_i \|q(t) - g^{(k)}(t)\|_{\mathcal{L}_p}) \right), \quad (15)$$

siendo  $L_i$  la constante de Lipschitz de la función  $x_i = h_i(g(t))$ .

Puede parecer que el método (15) solo infiere el valor del estado en  $t = 0$ . Pero esto se debe a que la entrada  $q(t)$  se ha desplazado al intervalo  $T$  en el que están definidas las trayectorias del cuaderno  $C_o$ . Considerando que hay datos suficientes en  $[t_o - \tau, t_o)$  para generar la trayectoria  $q(t)$ , el método podría usarse para predecir en cualquier  $t_o$ .

Al igual que sucedía en el método original, el único parámetro del modelo es la constante de Lipschitz de la función,  $L_i$ . Si se asume desconocida, se podría estimar una constante  $L_i^D$  que satisfaga la condición de continuidad para el conjunto de datos disponibles. De forma similar a como se hacía en Callies et al. (2020), dicha constante se estima como:

$$L_i^{C_o} = \max_{(g^{(1)}(t), \tilde{x}_i^{(1)}(0)) \in C_o} \left( \frac{\|\tilde{x}_i^{(k)}(0) - \tilde{x}_i^{(j)}(0)\|}{\|g^{(k)}(t) - g^{(j)}(t)\|_{\mathcal{L}_p}} \mid k, j \in \{1, \dots, N_D\} \wedge \|g^{(k)}(t) - g^{(j)}(t)\|_{\mathcal{L}_p} > 0 \right), \quad (16)$$

donde  $t \in T$ . Nótese que la obtención de las constantes es una operación que sólo se realiza una vez en la fase de entrenamiento.

**Nota 4.** Tanto el método de aprendizaje de la constante de Lipschitz a partir de los datos, como el mecanismo de inferencia descrito en la Definición 5 requieren calcular la norma  $\mathcal{L}_p$  en muchas ocasiones, sobre todo para un cuaderno muy extenso. La propiedad (7) simplifica la operación, ya que permite obtener la norma de la función a partir de la norma de sus componentes. Sin embargo, para trayectorias  $g_i^{(k)}(t)$  genéricas, la resolución de la integral (6) puede ser una operación costosa. Sin embargo, al escoger polinomios, dichas integrales de Lebesgue se pueden resolver analíticamente.

Por último, y para dar solución completa al Problema 3, se necesita disponer de una descripción de la incertidumbre determinista. Siguiendo a Callies et al. (2020), se puede obtener una cota para el error de predicción:

$$\|\hat{h}_i(q(t); L_i^{C_o}, C_o) - h_i(q(t))\| \leq \mu_i, \quad (17)$$

donde  $\mu_i = (L_i + L_i^{C_o})r_{C_o}$ , siendo  $r_{C_o} \in \mathbb{R}$  la máxima distancia  $\mathcal{L}_p$  entre cualquier trayectoria a predecir ( $q(t) \in \mathcal{G}$ ,  $t \in T$ ) y el conjunto de observaciones.

## 5. Caso de estudio

En esta sección se presenta un caso de estudio donde se aplica el método desarrollado sobre el famoso atractor de Lorenz. El atractor de Lorenz es un sistema dinámico tridimensional, autónomo y no lineal, que fue introducido por Lorenz (1963) como un modelo simplificado de la atmósfera terrestre, que exhibe comportamiento caótico.

Está gobernado por las siguientes ecuaciones:

$$\frac{dx_1}{dt} = \sigma(x_2 - x_1), \quad (18a)$$

$$\frac{dx_2}{dt} = x_1(\rho - x_3) - x_2, \quad (18b)$$

$$\frac{dx_3}{dt} = x_1 x_2 - \beta x_3, \quad (18c)$$

donde  $x = (x_1, x_2, x_3)$  son los estados, y los parámetros son  $\sigma = 10$ ,  $\beta = 8/3$  y  $\rho = 28$ .

### 5.1. Generación del cuaderno de trayectorias

A continuación se describe el mecanismo seguido para la generación del cuaderno  $C_o$  a partir del modelo del sistema. En primer lugar, se ha discretizado usando Euler con un tiempo de integración de 1 ms. Usando unas condiciones iniciales de  $x = (10, 15, 20)$ , y simulando durante 20 s, se obtiene la evolución de los estados que se muestra en la Figura 1.

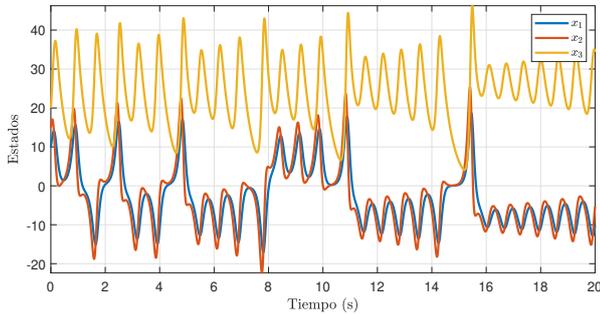


Figura 1: Evolución de los estados para el atractor de Lorenz

Esto corresponde a un primer conjunto de datos, denotado  $\mathcal{D}_1$ , que contiene 20000 estados  $(x_1, x_2, x_3)$  y sus marcas temporales. A continuación, se submuestra cada estado de manera aleatoria, conteniendo entre el 2 y el 5 % de los puntos de las señales originales. A esta base de datos, asincrónicamente muestreada desde  $\mathcal{D}_1$ , se la denota  $\mathcal{D}_2$ .

Posteriormente, se pasa sobre  $\mathcal{D}_2$  una ventana deslizante a lo largo del tiempo, de tamaño  $\tau = 0,1$  s, con un paso de 10 ms. Se chequea si en la ventana hay al menos  $\nu = m + 1 = 3$  muestras en  $\mathcal{D}_2$  de los tres estados.

Una vez encontrada una ventana donde hay  $\nu$  puntos en cada uno de los estados, se ajustan los polinomios  $g_i^{(k)}(t)$  a estos, obteniendo las trayectorias que se muestran en la Figura 2. El resultado se almacena finalmente en el cuaderno, que se denota  $C$ . Posteriormente, la ventana avanza hasta la marca temporal más avanzada, con el objetivo de no repetir ese conjunto de datos en la siguiente ventana.

Por tanto,  $C$  contiene tres polinomios de orden  $m = 2$  (tres coeficientes para cada polinomio) para cada una de las entradas del data set (es decir, cada ventana donde había al menos  $\nu$  puntos en los tres estados). Para el ejemplo de la Figura 2, esto resulta en una dimensión de  $3 \times 3 \times 81$ , siendo  $N_C = 81$  el número total de trayectorias que hay el cuaderno.

La estimación del estado en  $t_k$ , es decir,  $\hat{x}^{(k)}(0)$ , se ha obtenido mediante distintas formas, para evaluar su influencia en el desempeño del algoritmo propuesto en este trabajo. La primera forma, denotada  $\hat{x}_A$ , se obtiene como la interpolación lineal en  $t_k$  entre las observaciones de  $\mathcal{D}_2$  inmediatamente anterior y posterior, de cada estado. Para la segunda forma, denotada  $\hat{x}_B$ , se aplica un mantenedor de orden cero, es decir, se mantiene el valor del estado medido inmediatamente anterior a  $t_k$ ; y para la tercera,  $\hat{x}_C$ , un mantenedor de orden uno (extrapolación lineal con los dos puntos anteriores a  $t_k$ ).<sup>1</sup>

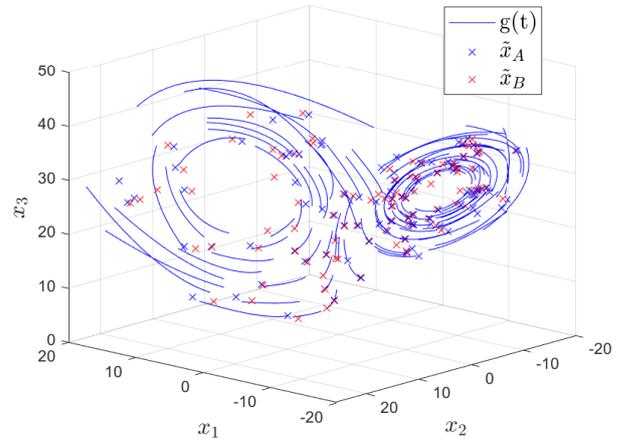


Figura 2: Cuaderno en el espacio de fase

Finalmente, se desplazan las trayectorias a la ventana genérica  $T = [-\tau, 0)$  para obtener el cuaderno  $C_o$ .

### 5.2. Entrenamiento y validación del método de aprendizaje

El cuaderno  $C_o$  se divide en dos conjuntos separados,  $C_{\text{train}}$  y  $C_{\text{test}}$ , usando una partición aleatoria del 80 % de entradas para el primero, y el resto para el segundo. En este caso, el entrenamiento consiste solamente en aprender la constante de Lipschitz  $L^{C_o} = [L_i^{C_o}]$ ,  $i = 1, 2, 3$ , para lo que se usa el dataset  $C_{\text{train}}$  como se indica en la Ecuación (16). Posteriormente, se realizan predicciones sobre  $C_{\text{test}}$ , como se propone en la Ecuación (15). Para evaluar la bondad del método se mide la raíz del error cuadrático medio o RMSE, que se denota como  $\rho$ . Para  $N_o$  observaciones de dos señales  $n$ -dimensionales,  $y_1^{(j)}, y_2^{(j)} \in \mathbb{R}^n$ ,  $j = 1, \dots, N_o$ , éste se calcula como:

$$\rho(y_1, y_2) = \sqrt{\frac{1}{N_o} \sum_{j=1}^{N_o} \|y_1^{(j)} - y_2^{(j)}\|^2}. \quad (19)$$

En este caso, se hace el RMSE de las  $N_{C_{\text{test}}}$  entradas de  $C_{\text{test}}$ . Se calcula este valor comparando los estados predichos mediante KI,  $\hat{x}_A$ ,  $\hat{x}_B$  y  $\hat{x}_C$ , con el valor real del estado  $x$  que, en este caso, está disponible al realizar el ensayo con un

<sup>1</sup>En lo que sigue, se usarán los subíndices  $A$ ,  $B$  y  $C$  para denotar variables obtenidas usando  $\hat{x}_A$ ,  $\hat{x}_C$  y  $\hat{x}_B$ , respectivamente.

sistema conocido. Para validar el método propuesto, se repite el ensayo propuesto en esta sección 500 veces, haciendo cada vez el submuestreo aleatorio y la partición entre  $C_{\text{train}}$  y  $C_{\text{test}}$  aleatoriamente, para dotar de robustez a este ensayo de validación.

La media de las constantes de Lipschitz obtenidas son  $L_A^{C_o} = [4,5, 6,8, 8,6]$ ,  $L_B^{C_o} = [13,3, 19,4, 26,7]$  y  $L_C^{C_o} = [3,9, 6,2, 7,1]$ . Los resultados para el RMSE se muestran en la Figura 3.

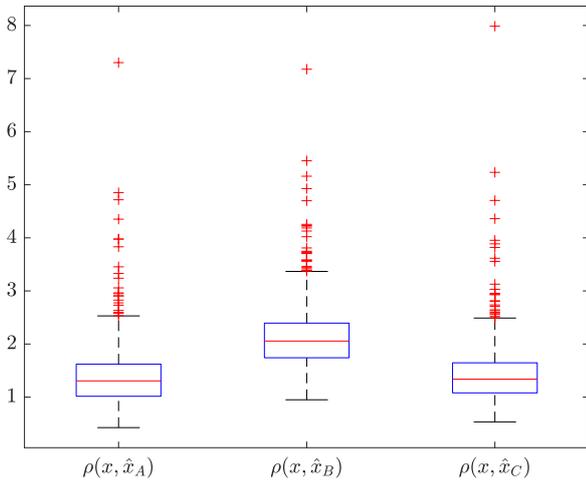


Figura 3: RMSE de 500 ensayos de predicción. Se comparan las diferencias entre los estados reales ( $x$ ) y predichos ( $\hat{x}$ ) usando distintas formas de almacenar el estado estimado en  $t_k$

En vista a los resultados, cabe destacar que, de los tres métodos analizados, la mejor forma de almacenar la estimación del estado en  $t_k$  es usar una interpolación entre los datos existentes. Además, esto provee mejores resultados que almacenar la evaluación de la trayectoria en el instante en el que se desea predecir, esto es  $g(t_k)$ , cuyos resultados arrojan errores tan grandes que no se han incluido en la Figura 3 (mediana del RMSE de 16.8).

## 6. Conclusiones

Se ha presentado un método de identificación no paramétrico para sistemas muestreados asincrónicamente. El método se basa, por una parte, en la definición de un cuaderno de trayectorias a partir de los datos asíncronos para, luego, modificar el mecanismo de aprendizaje denominado *kinky inference* para predecir el estado futuro a partir de dichas trayectorias. Se ha empleado el método para predecir la dinámica del atractor de Lorenz, logrando mejores resultados que predicciones que se podrían hacer en línea con los datos disponibles.

Esta primera aproximación deja numerosos interrogantes a analizar en el futuro, como la influencia del muestreo en la

capacidad de predicción, o el impacto de la generación del cuaderno y, en concreto, de la estimación del estado  $\hat{x}(t_k)$  en el rendimiento. El análisis de las propiedades teóricas que este método logra es importante, ya que el método original satisfacía propiedades muy interesantes para su uso en un bucle de control o estimación.

## Agradecimientos

Este artículo es parte del proyecto de I+D+i PID2020-117800GB-I00, financiado por MCIN/AEI/10.13039/501100011033/.

## Referencias

Annamalai, L., Ramanathan, V., Thakur, C. S., 2022. Event-Istm: An unsupervised and asynchronous learning-based representation for event-based data. *IEEE Robotics and Automation Letters* 7 (2), 4678–4685.

Bekiroglu, K., Lagoa, C., Lanza, S. T., Sznaier, M., 2017. System identification algorithm for non-uniformly sampled data. *IFAC-PapersOnLine* 50 (1), 7296–7301.

Beliakov, G., 2006. Interpolation of lipschitz functions. *Journal of computational and applied mathematics* 196 (1), 20–44.

Brunton, S. L., Proctor, J. L., Kutz, J. N., 2016. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* 113 (15), 3932–3937.

Calliess, J., 2014. Conservative decision-making and inference in uncertain dynamical systems. Ph.D. thesis, Oxford University, UK.

Calliess, J.-P., Roberts, S. J., Rasmussen, C. E., Maciejowski, J., 2020. Lazily adapted constant kinky inference for nonparametric regression and model-reference adaptive control. *Automatica* 122, 109216.

Johansson, R., 2009. Continuous-time model identification and state estimation using non-uniformly sampled data. *IFAC Proceedings Volumes* 42 (10), 1163–1168.

Lorenz, E. N., 1963. Deterministic nonperiodic flow. *Journal of atmospheric sciences* 20 (2), 130–141.

Manzano, J. M., Limon, D., Muñoz de la Peña, D., Calliess, J. P., 2019. Output feedback MPC based on smoothed projected kinky inference. *IET Control Theory & Applications* 13 (6), 795–805.

Ouyang, J., Chen, X., 2023. Least squares solution for system identification with non-uniform data under a coprime collaborative sensing scheme. In: *2023 American Control Conference (ACC)*. IEEE, pp. 1570–1575.

Quaranta, G., Lacarbonara, W., Masri, S. F., 2020. A review on computational intelligence for identification of nonlinear dynamical systems. *Nonlinear Dynamics* 99 (2), 1709–1761.

Reddy, J. N., Rasmussen, M., 1990. *Análisis matemático avanzado: con aplicaciones a ingeniería y ciencias*. Limusa.

Tsang, K., Billings, S., 1995. Identification of systems from non-uniformly sampled data. *International journal of systems science* 26 (10), 1823–1837.

Wang, H. W., Xia, H., 2013. A new identification method for non-uniformly sampled-data systems. In: *Proceedings of the 32nd Chinese control conference*. IEEE, pp. 1792–1796.

Williams, C., Rasmussen, C., 1995. Gaussian processes for regression. *Advances in neural information processing systems* 8.

Xie, L., Yang, H., Ding, F., 2017. Identification of non-uniformly sampled-data systems with asynchronous input and output data. *Journal of the Franklin Institute* 354 (4), 1974–1991.

Xu, K., Wu, X., Liu, X., Wang, D., 2021. Identification of robot joint torsional stiffness based on the amplitude of the frequency response of asynchronous data. *Machines* 9 (9), 204.

Xu, X., Liu, L., Feng, G., 2020. On Lipschitz conditions of infinite dimensional systems. *Automatica* 117, 108947.