

Time series analysis of the Bacillus subtilis sporulation network reveals low dimensional chaotic dynamics

*Original*

Time series analysis of the Bacillus subtilis sporulation network reveals low dimensional chaotic dynamics / Lecca, Paola; Mura, Ivan; Re, Angela; Barker, Gary C.; Ihekwebaba, Adaoha E. C.. - In: FRONTIERS IN MICROBIOLOGY. - ISSN 1664-302X. - ELETTRONICO. - 7:NOV(2016), pp. 1760-1777. [10.3389/fmicb.2016.01760]

*Availability:*

This version is available at: 11583/2970528 since: 2022-08-08T09:31:29Z

*Publisher:*

Frontiers media

*Published*

DOI:10.3389/fmicb.2016.01760

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)



# Time Series Analysis of the *Bacillus subtilis* Sporulation Network Reveals Low Dimensional Chaotic Dynamics

Paola Lecca<sup>1\*</sup>, Ivan Mura<sup>2</sup>, Angela Re<sup>3</sup>, Gary C. Barker<sup>4</sup> and Adaoha E. C. Ihekweba<sup>4\*</sup>

<sup>1</sup> Department of Mathematics, University of Trento, Trento, Italy, <sup>2</sup> Department of Industrial Engineering, Universidad de los Andes, Bogotá, Colombia, <sup>3</sup> Laboratory of Translational Genomics, Centre for Integrative Biology, University of Trento, Trento, Italy, <sup>4</sup> Gut Health and Food Safety, Institute of Food Research, Norwich, UK

## OPEN ACCESS

### Edited by:

Imrich Barak,  
Slovak Academy of Sciences, Slovakia

### Reviewed by:

Anthony Joseph Wilkinson,  
University of York, UK  
Andreas Kremling,  
Technische Universität München,  
Germany

### \*Correspondence:

Paola Lecca  
paola.lecca@unitn.it  
Adaoha E. C. Ihekweba  
Adaoha.Ihekweba@ifr.ac.uk

### Specialty section:

This article was submitted to  
Microbial Physiology and Metabolism,  
a section of the journal  
Frontiers in Microbiology

**Received:** 21 July 2016

**Accepted:** 19 October 2016

**Published:** 07 November 2016

### Citation:

Lecca P, Mura I, Re A, Barker GC and  
Ihekweba AEC (2016) Time Series  
Analysis of the *Bacillus subtilis*  
Sporulation Network Reveals Low  
Dimensional Chaotic Dynamics.  
*Front. Microbiol.* 7:1760.  
doi: 10.3389/fmicb.2016.01760

Chaotic behavior refers to a behavior which, albeit irregular, is generated by an underlying deterministic process. Therefore, a chaotic behavior is potentially controllable. This possibility becomes practically amenable especially when chaos is shown to be low-dimensional, i.e., to be attributable to a small fraction of the total systems components. In this case, indeed, including the major drivers of chaos in a system into the modeling approach allows us to improve predictability of the systems dynamics. Here, we analyzed the numerical simulations of an accurate ordinary differential equation model of the gene network regulating sporulation initiation in *Bacillus subtilis* to explore whether the non-linearity underlying time series data is due to low-dimensional chaos. Low-dimensional chaos is expectedly common in systems with few degrees of freedom, but rare in systems with many degrees of freedom such as the *B. subtilis* sporulation network. The estimation of a number of indices, which reflect the chaotic nature of a system, indicates that the dynamics of this network is affected by deterministic chaos. The neat separation between the indices obtained from the time series simulated from the model and those obtained from time series generated by Gaussian white and colored noise confirmed that the *B. subtilis* sporulation network dynamics is affected by low dimensional chaos rather than by noise. Furthermore, our analysis identifies the principal driver of the networks chaotic dynamics to be sporulation initiation phosphotransferase B (Spo0B). We then analyzed the parameters and the phase space of the system to characterize the instability points of the network dynamics, and, in turn, to identify the ranges of values of Spo0B and of the other drivers of the chaotic dynamics, for which the whole system is highly sensitive to minimal perturbation. In summary, we described an unappreciated source of complexity in the *B. subtilis* sporulation network by gathering evidence for the chaotic behavior of the system, and by suggesting candidate molecules driving chaos in the system. The results of our chaos analysis can increase our understanding of the intricacies of the regulatory network under analysis, and suggest experimental work to refine our behavior of the mechanisms underlying *B. subtilis* sporulation initiation control.

**Keywords:** systems biology, computational modeling, sensitivity analysis, low dimensional chaos, signal transduction, sporulation, *Bacillus subtilis*

## 1. INTRODUCTION

Bacterial spores are important contaminants in food, and the spore forming bacteria are often implicated in food safety and food quality considerations (Carlin, 2011). Most microbial spore forming bacteria respond to stress (e.g., nutrient deprivation) by inducing the expression of an appropriate suit of adaptive (stress-response) genes to help them cope with adverse environmental circumstances; an extreme example is endospore formation (Ihekwaba et al., 2014).

Since sporulation is an energy consuming process that requires a significant reorganization of cellular activity, the decision to commit to spore formation is subject to the result of integration of multiple signals by a complex gene regulation network.

The initiation of sporulation is one of the decisive moments in spore formation, as exemplified by the bacterium *Bacillus subtilis*. The changes in gene expression and morphology induced by sporulation are regulated in *B. subtilis* by a complex network involving more than 120 genes (Stragier and Losick, 1996; Fawcett et al., 2000).

The DNA-binding protein Spo0A is the master regulator for entry into sporulation in *B. subtilis*. The concentration level and the phosphorylation state determine the ability of Spo0A to alter transcription. Upon phosphorylation, Spo0A undergoes an allosteric change that re-orientates a phenylalanine residue and allows the molecule to bind DNA (Muchová et al., 2004) and activate key genes that drive the positive regulation of sporulation, particularly the *spoIIA*, *spoIIIE*, and *spoIIG* genes involved in establishing compartment-specific transcription under the control of  $\sigma F$  (*spoIIA* operon and the *spoIIIE* gene) and  $\sigma E$  (*spoIIG* operon) (Satola et al., 1991, 1992; York et al., 1992). Phosphorylated Spo0A also acts as a repressor, blocking the expression of the *abrB* gene. This repression has the consequence of setting up a self-reinforcing cycle that contributes to the further accumulation of Spo0A at the start of sporulation (Fujita and Losick, 2005; Tojo et al., 2013). Indeed, the inhibition that phosphorylated Spo0A exerts on *abrB* gene expression leads to the depletion of the AbrB protein from the cell and to the accumulation of  $\sigma H$ , with the net result of enhancing the expression of KinA, Spo0F and of the Spo0A gene itself (Strauch et al., 1993; Tojo et al., 2013).

Spo0A activation is under the control of a complex network capable of integrating diverse physiological and environmental signals, and relaying signals through a three-level phosphorelay down to the response regulator Spo0A. Various mathematical models of *B. subtilis* sporulation mechanisms can be found in the literature, among the most recent ones we mention (Kuchina et al., 2011; Sen et al., 2011; Narula et al., 2012; Kothamachu et al., 2013; Vishnoi et al., 2013; Ihekwaba et al., 2014).

This paper is based on the model proposed in Ihekwaba et al. (2014), which integrates most of previous mathematical modeling works on *B. subtilis* sporulation initiation. The model we consider encodes the relationships among the time-dependent concentrations of sporulation signals, histidine kinases, phosphorelay proteins and sporulation initiation proteins in the form of a deterministic differential model

having 27 variables. Simulation of the differential equations via numerical integration provides predictions about the evolution of the *B. subtilis* sporulation initiation regulation network, given the initial state of variables. Ihekwaba et al. (2014) also performed a sensitivity analysis of the model to explore the set of possible behaviors with varying the values of its parameters (i.e., the kinetic rate constants).

In this paper, we continue the analysis of model behavior, with the aim of investigating whether the time series of the variables, as predicted by the differential equations model, are affected by deterministic chaos, or simply *chaos*. A chaotic system is a system that is predictable up until a given time, after which it becomes unpredictable (i.e., long term unpredictability) due to its sensitivity to initial conditions (Kellert, 1993). Even if the initial state is known at a very accurate level of detail, any imprecision in its quantification, no matter how small, grows quickly (exponentially) with time, rendering long-term prediction impossible.

Identifying chaos and its drivers in a biological system provides useful information (i) to understand the origins of the observed dynamics (Weiss et al., 1994; Lecca et al., 2016), and (ii) to shed light into the control mechanisms that a biological system may have implemented to maintain a stable activity even when subject to perturbations of its initial conditions (Sinha, 1997). Both chaotic dynamics and stochastic dynamics exhibit a complex phase space structure and are not predictable, but chaos is not stochastic noise (Lecca et al., 2016). Indeed, a chaotic dynamics is governed by deterministic laws in which no randomness is involved, whereas a stochastic dynamics is governed by rules involving random variables. As a consequence, if the laws and the drivers of the dynamics of a chaotic system are known, its unpredictability can potentially be controlled (Sinha, 1997; Lai, 2014).

*Low-dimensional* chaos occurs when a reduced number of contributing species are responsible for the complex dynamics. Such a low-dimensional chaos is of particular interest in biology (Skinner, 1994; Kaneko, 2006; Vasseur, 2015). Since in a biological system affected by low-dimensional chaos the variables governing the spatial and temporal dynamics are few in number, a low-energy control of unpredictability of the system dynamics can be implemented and a simpler model of a complex dynamics can be provided. Low-dimensional chaos is expected to be common in systems with few degrees of freedom (Skinner, 1994), but is expected to be rare in systems with many degree of freedom such as the sporulation network of *B. subtilis*. The results of our analysis show that only few molecular species are contributing to the appearance of deterministic chaos in the dynamics of the modeled network.

## 2. THE MODEL

In Ihekwaba et al. (2014), a mathematical model of the network regulating *Bacillus subtilis* sporulation initiation was proposed. The model represents at the molecular level the sequence of events that lead to the activation of the early genes under control of the master regulator molecule Spo0A, distilling and extending the results obtained in various modeling studies focused on

systems where sporulation is induced by an artificial inducer, the Isopropyl-D-1-thiogalactopyranoside (IPTG) (see for instance Narula et al., 2012), and modeling as well the induction of sporulation that occurs in wild-type cells.

A high-level diagrammatic description of the molecular network governing sporulation initiation in *B. subtilis* is provided in **Figure 1**, where pointed arrows represent activation and blunt arrows indicate repression.

The *B. subtilis* sporulation network model considered in this study is the published model by Lecca et al. (2016) and Ihekweba et al. (2014). It follows the topology of the network shown in **Figure 1**, thereby encompassing three distinct sub-models:

1. input signal, representing the sporulation initiation processes induced by the signals on the histidine kinases;
2. phosphorelay, encoding the signal transduction along the phosphorelay components, from histidine kinases downwards to the master regulator Spo0A;
3. gene expression, modeling the target gene expression activation operated by the activated effector Spo0A.

In the following section, we explain the structure of each sub-model, and use a graphical notation to represent activation/repression (arrows with non-solid ends) which is introduced in **Figure 1**. In our modeling, we consider both transcription and translation of proteins. For each species involved in a synthesis process (i.e., transcripts and proteins), the model includes a degradation reaction, not shown in the model diagrams for clarity.

## 2.1. Input Signal Sub-model

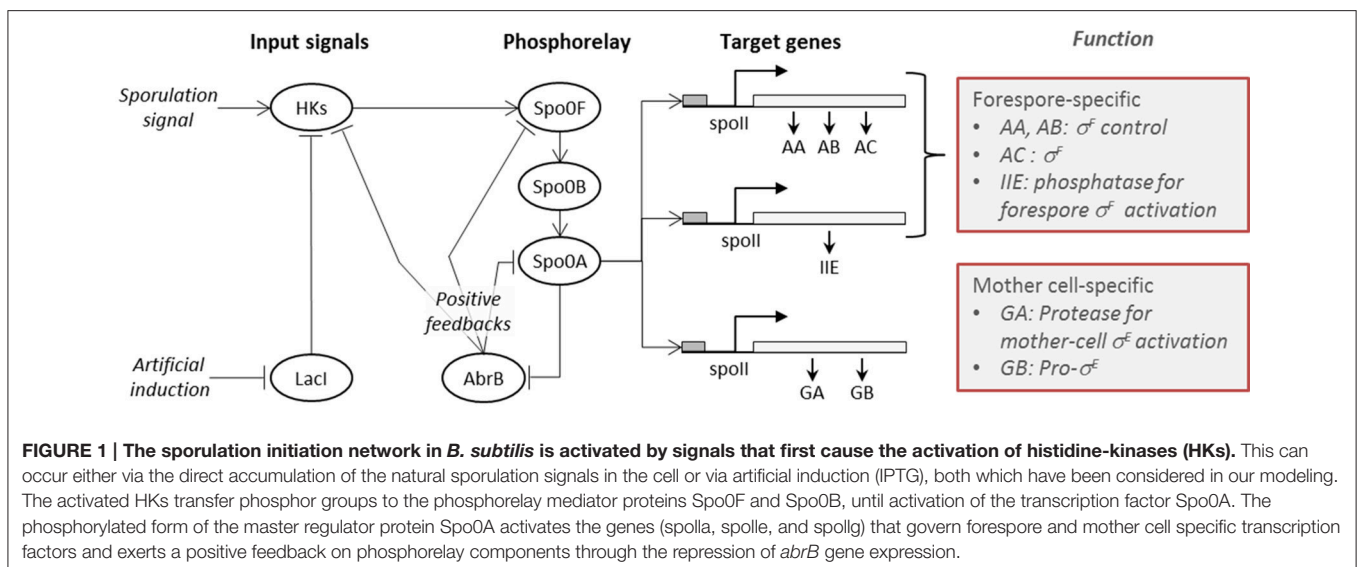
The input signal sub-model, shown in **Figure 2**, represents the regulation effects that artificial inducers (in this case IPTG) and cell produced sporulation signals (modeled by species SS) have on the HKs. In the model, reactions are consecutively numbered. In our model, of the five known kinases that have been identified

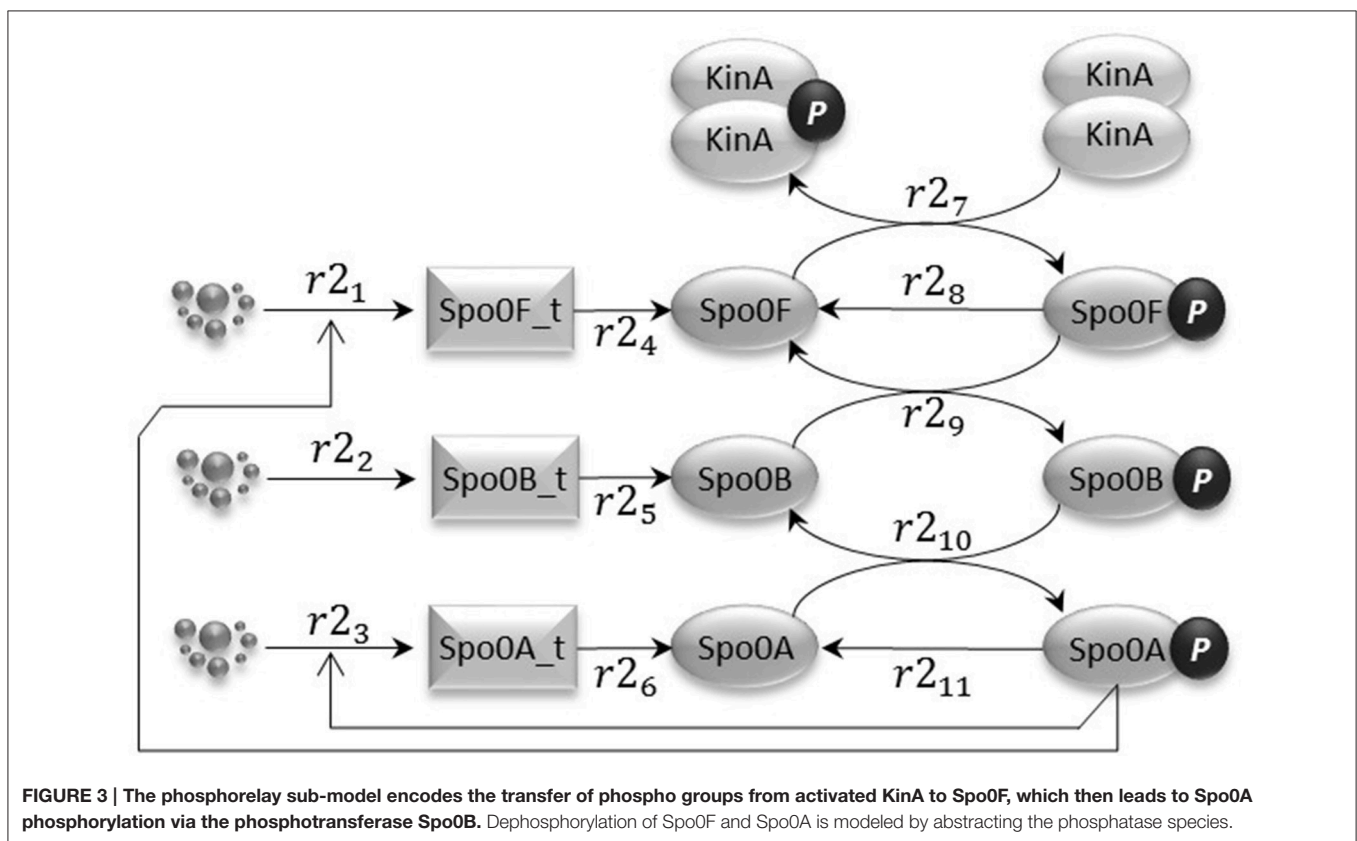
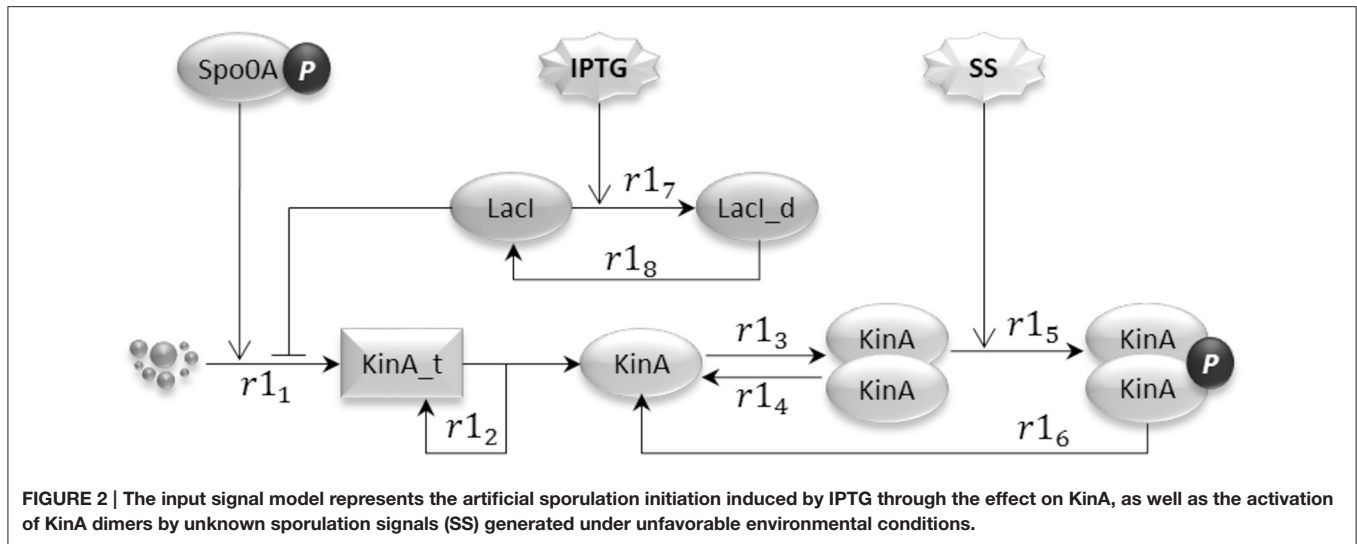
as being capable of initiating sporulation in *B. subtilis* (Jiang et al., 2000), we only considered the histidine kinase KinA. This is the major kinase responsible for initiation of sporulation and its overexpression during exponential growth is sufficient to induce entry into sporulation (Fujita and Losick, 2005).

The IPTG regulation of sporulation is rendered by the indirect release of inhibition for the transcription initiation of KinA (Eswaramoorthy et al., 2009; Narula et al., 2012), exerted by the lactose repressor (LacI) on the binding site incorporated into the promoter. The addition of IPTG causes a conformational change in the LacI protein, bringing it (reaction  $r_{17}$ ) to an inactive form (LacI\_d) that has very low affinity for the KinA promoter. The consequence of the inhibition release is an increased level of KinA transcription (reaction  $r_{11}$ ). In the diagrammatic representation, we use a “droplet” notation for species that are not explicitly represented, such as genes in transcription reactions. The LacI conformational change is however reversible (see reaction  $r_{18}$ ). Molecules of KinA transcript are translated into protein molecules (reaction  $r_{12}$ ), which can reversibly bind (reactions  $r_{13}$  and  $r_{14}$ ) to form dimers. KinA dimers have the ability to autophosphorylate (reaction  $r_{15}$ ), producing the active species that initiates the phosphorelay signaling (Wang et al., 2001; Eswaramoorthy et al., 2009). We model the dephosphorylation and the unbinding of the KinA dimer as a single reaction ( $r_{16}$ ). The model also considers the activation of the histidine kinase caused by the naturally occurring sporulation signals (SS), which accelerate the KinA autophosphorylation reaction and can lead *B. subtilis* into sporulation alone. Last, the model includes the positive effect that active Spo0A has on the transcription of KinA, via the double repression feedback loop that links phosphorylated Spo0A with AbrB, and AbrB with KinA.

## 2.2. Phosphorelay Sub-model

The phosphorelay sub-model depicted in **Figure 3** is based on phosphorylation, dephosphorylation and phosphotransfer





reactions. Our model includes the main phosphorelay species Spo0F, Spo0B, and Spo0A, which together form a cascading phosphotransfer (de Jong et al., 2010; Sen et al., 2011). For each of these proteins, the model includes a gene transcription reaction ( $r_{21}$ ,  $r_{22}$ , and  $r_{23}$ ), and a translation reaction ( $r_{24}$ ,  $r_{25}$ , and  $r_{26}$ ). The phosphorylated KinA dimer transfers the phosphate group to Spo0F (reaction  $r_{27}$ ), phosphorylated Spo0F transfers the phosphate group to Spo0B (reaction  $r_{29}$ ),

and finally phosphorylated Spo0B transfers the phosphate group to Spo0A (reaction  $r_{210}$ ). In the model, phosphorylated Spo0F and phosphorylated Spo0A spontaneously lose the phosphate group (reactions  $r_{28}$  and  $r_{211}$ ). Finally, we include in the model the phosphorelay self-activation loop induced by phosphorylated Spo0A, which as already described positively affects the transcription of both Spo0F and Spo0A species.



### 2.3. Gene Expression Sub-model

Phosphorylated Spo0A up-regulates transcription from spoIIA, spoIIIE, and spoIIIG promoters. The gene expression sub-model shown in **Figure 4** encodes the activation of transcription exerted by Spo0A, and includes transcription reactions ( $r_{31}$ ,  $r_{32}$ , and  $r_{33}$ ) and translation reactions for AA, AB, and AC proteins ( $r_{34}$ ,  $r_{35}$ , and  $r_{36}$ ), IIE protein ( $r_{37}$ ) and GA and GB protein molecules ( $r_{38}$  and  $r_{39}$ ). Notice that AA, AB, and AC, and also GA and GB, are transcribed polycistronically from the spoIIA and spoIIIG operons, respectively (Narula et al., 2012).

In the rest of the paper and in the Supplementary Material, we adopt the following notation to indicate the variables corresponding with the molecular species of the model: the name of the protein is written in lowercase (e.g., Spo0A, *spo0a*), the transcript of the gene is denoted by the suffix “\_t” (e.g., the transcript of gene Spo0A is denoted by *spo0a\_t*), and the phosphorylated form of the protein is indicated by the suffix “p” (e.g., the phosphorylated form of protein Spo0A is *spo0ap*). The mathematical specification of the model and its parameters are given in Tables S1–S4. The time is measured in seconds (s), and the molecular species concentration in nM.

## 3. DETECTING CHAOS IN *B. SUBTILIS* SPORULATION NETWORK DYNAMICS

A system is affected by deterministic chaos if its dynamics is governed by deterministic rules and any change in the initial state, no matter how small, grows quickly with time, rendering long-term prediction of the system behavior impossible. A system is affected by *low-dimensional* chaos if only a small number of variables exhibits a chaotic dynamics, i.e., an aperiodic irregular time-behavior (Tél and Gruiz, 2005; Layek, 2015).

The presence of low-dimensional chaos in biological systems is of particular interest, because it indicates that the variables governing the spatial and temporal behavior of the system may be few in number. This means that the dynamics of the system might

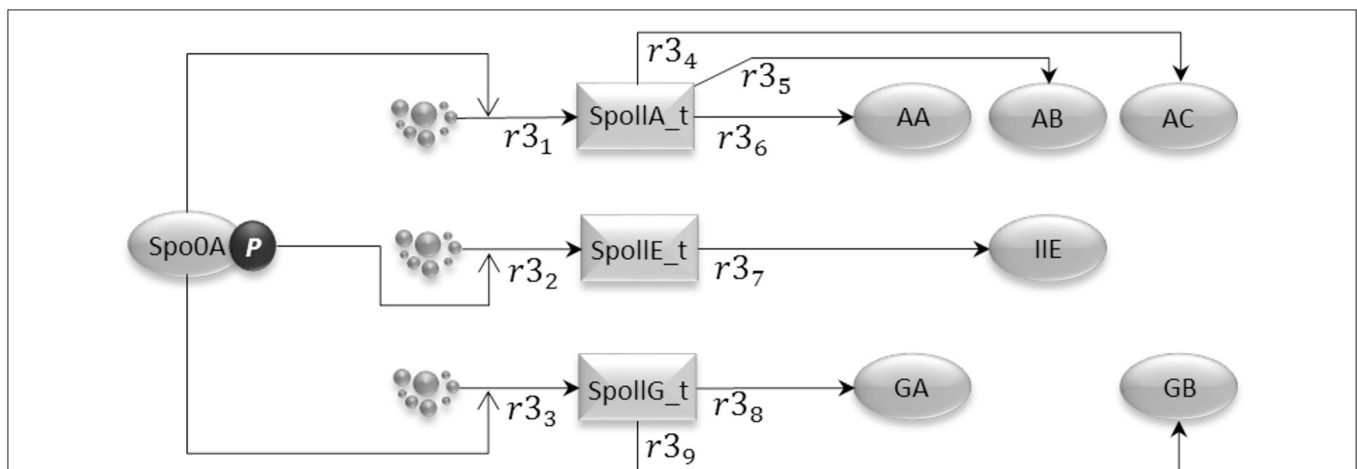
be controlled by only a few crucial variables. The complexity of control inherent in chaotic systems may be important in the dynamics of gene expression regulation. Therefore, it is of particular interest to assess the presence of low-dimensional chaos in a complex system such as the *B. subtilis* sporulation network, as this analysis allows the identification of the variables (few in numbers) that control the predictability of the dynamics of the whole system.

There are two-established methods to explore chaotic behavior of a dynamical biological system. The first one is a direct analysis of the experimental time series, combined with the development of algorithms for computing relevant indices quantifying the features of the system dynamics. The second is the implementation of a model developed directly from the experimental observations that aims to account for the essential mechanisms at work in the real system and explain the dominant behavior. Then, a subsequent analysis focuses on the simulated time series obtained by model solution and its phase space in order (i) to evaluate the control parameters, (ii) to detect the system components (e.g., genes, proteins, chemical species, etc.) that exhibit a chaotic dynamics, and (iii) to investigate the robustness of the dynamics against perturbations. We implemented the second approach, because the inclusion in the analysis of a model of the systems dynamics affords not only the identification of the drivers of chaotic dynamics, but also the conceptualization of their role and of their effects within the mechanisms of interaction with other molecular species. In the next section, we provide a detailed explanation of this analysis.

### 3.1. Sensitivity Analysis

We undertook local sensitivity analysis to assess the sensitivity of the solutions of the system’s equations to the variation of individual parameter in the system. We discussed the feasibility of global sensitivity analysis of the system to variation in parameters in the Supplementary Information.

We randomly sampled  $N_p$  values from a uniform distribution for each parameter (i.e., kinetic rate constant). The uniform



**FIGURE 4 |** The gene expression model represents the transcriptional activity of phosphorylated Spo0A, which binds to the SpoII promoters and promotes transcription initiation for the important sporulation initiation proteins, AA, AB, AC, IIE, GA, and GB.

distributions were positively defined on the maximal range of parameter variability in which the system of ordinary differential equations has a unique solution (i.e., it is not underdetermined). We determined this range  $I_{q^*}(p_h)$  by attempting to solve the system of equations for different sets of parameters  $P(q) = \{p_h\}$ ,  $h = 1, 2, \dots, N_p$ , obtained by varying the value of  $q \in \mathbb{N}^+$  in the interval  $[2, 30]$  in the following expression

$$I_q(p_h) = \left[ \frac{p_h}{q}, q \times p_h \right], \quad h = 1, 2, \dots, N_p \quad (1)$$

The maximal interval of parameter variation is defined by the maximal value of  $q$  for which the system of differential equations has a unique solution.

The parameters were changed one at a time while keeping the values of the others fixed. Since for each parameter  $p_h$  we sampled  $N_p$  values, we performed  $N_p$  model simulations, i.e., one simulation for each sampled value in the range of parameter variability  $I_q(p_h)$ . The index of sensitivity of the time series  $x_s(t)$ , ( $s = 1, 2, \dots, d$ , where  $d$  is the number of molecular species in the system), with respect to the change of the  $h$ -th parameter from the value  $p_h$  to the value  $p'_h$  is calculated as the mean of the standard deviations of the distributions of the simulated values of the variable over the range of parameter variability and over time, i.e.:

$$SI_{s,h} = \frac{1}{N} \sum_{k=1}^N \left( \frac{1}{N_p - 1} \sum_{r=1}^{N_p} (x_s^{(r)}(t_k | p_h \leftarrow p'_h) - \bar{x}_s^{(r)}(p_h \leftarrow p'_h))^2 \right)^{\frac{1}{2}} \quad (2)$$

where  $N$  is the length of the time series, and

$$\bar{x}_s(p_h \leftarrow p'_h) = \frac{1}{N_p} \sum_{k=1}^{N_p} x_s(t_k | p_h \leftarrow p'_h). \quad (3)$$

With the expression  $p_h \leftarrow p'_h$ , we denote the replacement of value  $p_h$  with the value  $p'_h$ .

### 3.2. Complexity Indices

In order to detect the presence of chaos in *B. subtilis* network dynamics, for the time series of each molecular species we calculated a set of indices that capture different aspects of the complexity. This set of indices includes:

1. **Lyapunov exponents** ( $\lambda$ ): they measure the rate of separation of infinitesimally close trajectories in the phase space generated by slightly different values in the initial state of the system. The largest Lyapunov exponent is usually considered important in the determination of chaotic behavior. A positive value for the largest Lyapunov exponent indicate orbital instability and chaos (Kaneko and Tsuda, 2001; Sprott, 2003; Kalitin, 2004).
2. **fractal dimension** ( $D_F$ ): a statistical index for pairwise distances of the points of a time series; it indicates how a set of points fills its space and thus quantifies the complexity of the behavior of a trajectory;

3. **sample entropy** ( $S_E$ ): a measure of data regularity; a smaller value of sample entropy indicates more self-similarity in the data of the time series and a less noise (less disorder);
4. **time lag** ( $T_L$ ): the time after which the auto-correlation of the time series is negligible;
5. **embedding dimension** ( $D_E$ ): similarly to the fractal dimension, it measures topological complexity of a time series. A set of points has embedding dimension  $D_E$  if  $D_E$  is the smallest integer for which it can be embedded into  $\mathbf{R}^{D_E}$  without intersecting itself. So,  $D_E$  is the minimum dimension of a space in which a trajectory in the phase space reconstructed from the observed time series does not cross itself (in this case the dynamics is deterministic) (Abarbanel, 1996; Tamma and Khubchandani, 2016).

In chaotic systems, small differences in the initial condition result in strongly different solutions. Therefore, a chaotic system is unpredictable in the sense that the variability of the prediction induced by small changes in the initial conditions is unacceptably high in comparison to the difference of the initial states.

In deterministic systems, complete knowledge of the rules of the dynamics and of the initial state (i.e., values for the abundance of the system's components at initial time  $t_0$ —sometimes called *initial conditions*)  $\mathbf{x}(t_0)$ , is sufficient to determine  $\mathbf{x}(t)$  at each  $t > t_0$ . In chaotic deterministic systems, if the initial state is changed by a small value  $\epsilon$ , two trajectories that were initially close, will exponentially separate. Formally, if  $\mathbf{x}(t)$  and  $\mathbf{x}'(t)$  are the two trajectories generated by the initial states  $\mathbf{x}(t_0)$  and  $\mathbf{x}'(t_0)$ , and if  $|\mathbf{x}(t_0) - \mathbf{x}'(t_0)| < \epsilon$ , we have that

$$|\mathbf{x}(t) - \mathbf{x}'(t)| \sim \epsilon e^{\lambda t} \quad (4)$$

where  $\lambda$  is the angular coefficient of the straight line defining  $\ln |\mathbf{x}(t) - \mathbf{x}'(t)|$  as a function of time  $t$ :

$$\ln |\mathbf{x}(t) - \mathbf{x}'(t)| = \lambda t + \ln \epsilon.$$

Using Equation (4) it is possible to predict the time  $t^*$  after which the predicted trajectory is too imprecise. Indeed, if  $\delta$  is the tolerance on the precision of the prediction, then from Equation (4)  $\epsilon e^{\lambda t^*} \sim \delta$ , and therefore

$$t^* \sim \frac{1}{\lambda} \ln \frac{\delta}{\epsilon}. \quad (5)$$

The expression in Equation (5) suggests that  $t^*$  can be arbitrarily increased by decreasing  $\epsilon$ . However, de facto, it is not possible to obtain a value of  $t^*$  much greater than  $\frac{1}{\lambda}$ . For instance, if we want to increase  $t^*$  by one order of magnitude, we have to decrease  $\epsilon$  by a factor  $e^{10} \sim 10^4$ . This example points out that the dependence of  $t^*$  on the ratio  $\frac{\delta}{\epsilon}$  is so weak that in Equation (5), the only term that strongly influences  $t^*$  is  $\lambda$  (Vulpiani, 2004; Cencini et al., 2009; Lecca et al., 2016).

The system of differential equations describing the dynamics of the *B. subtilis* sporulation network is a  $d$ -dimensional system, where  $d$  is the number of molecular species involved in the system. At each instant of time  $t$  the system is contained in a  $d$ -dimensional sphere in the phase-space. In particular, this  $d$ -dimensional sphere is centered at  $\mathbf{x}(0)$  ( $\mathbf{x}(0)$  belonging to the

attractor) and has radius  $\epsilon$ . The time evolution of the system dictated by the equations deforms the sphere into an ellipse. If  $l_i(t)$  denotes the length of the  $i$ -th semi-axis of the ellipse at time  $t$ , the characteristic *Lyapunov exponents* ( $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$ ) are defined as follows:

$$\lambda_i = \frac{1}{t} \ln \frac{l_i(t)}{\epsilon}, \quad i = 1, 2, \dots, d. \quad (6)$$

If  $\lambda_i > 0$ , the  $i$ -th semi-axis grows with time; in contrast, if  $\lambda_i < 0$  the  $i$ -th semi-axis shrinks with time. In a system extremely sensitive to the initial conditions, at least one of the Lyapunov exponents is greater than zero.

For each molecular species  $i$  in the *B. subtilis* network we have calculated the maximal Lyapunov exponent from the corresponding simulated time series, i.e., the Lyapunov exponent at the maximum observed time, formally defined as follows

$$\lambda_i^{\max} = \lim_{t \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \lambda_i \quad (7)$$

The greater a positive maximal Lyapunov exponent, the faster the rate of divergence of the two trajectories  $\mathbf{x}(t)$  and  $\mathbf{x}'(t)$ . Thus, the Lyapunov coefficients were used to measure the contribution of each molecular species to the system's dynamics. In this study, we used the Rosenstein et al. (1993) algorithm to estimate the maximal Lyapunov exponent.

The Lyapunov exponents capture the unpredictability in a system's evolution which can be generated by slightly different initial states. However, unpredictability could depend also on an irregular aperiodic behavior of the abundance of some molecular species in the system.

To capture this aspect of a chaotic dynamical system, and, most importantly to distinguish it from noise, we have estimated the *fractal* and the *embedding* dimensions of the time series of each gene and protein in the system. Both fractal and embedding dimensions are generalizations of the topological dimension and measure the dimensionality of the space occupied by the set of points of the time series. The more complex and irregular the distribution of the points in space is, the higher the fractal and embedding dimensions of the system.

We estimate the fractal dimension as a *correlation dimension* (Theiler, 1990; Ding et al., 1993), defined in terms of the correlation integral  $C(\epsilon)$ :

$$C(\epsilon) = \lim_{N \rightarrow \infty} \frac{g\epsilon}{N^2} \quad (8)$$

where  $N$  is the number of points in the time series, and  $g$  is the total number of pairs of points that dist from each other is less than  $\epsilon$  (a graphical representation of such close pairs is the *recurrence plot* (Marwan et al., 2016)). The correlation integral estimates the probability that a pair of points of the time series is separated by a distance less than  $\epsilon$ . For  $\epsilon \ll 1$  it can be shown (Theiler, 1990) that

$$C(\epsilon) \sim \epsilon^{D_F} \quad (9)$$

where  $D_F$  is the correlation dimension. For a sufficiently large, and evenly distributed, number of points in a time series, a log-log graph of the correlation integral vs.  $\epsilon$  can be used to estimate

$D_F$  (Kantz, 2004). The more complex and irregular a time series is, the higher its correlation dimension, as the number of ways for points to be close to each other is greater (Higuchi, 1988). Indeed the fractal dimension corresponds to the number of the degrees of freedom of the time series (Mera and Morán, 2002).

Unlike topological dimension, the fractal dimension can take non-integer values, indicating that a set of points of a trajectory can fill its space qualitatively, and quantitatively, in a different way from an ordinary geometrical set. For instance, a curve with fractal dimension very near to 1, behaves quite like an ordinary line, but a curve with fractal dimension greater than 2 winds convolutedly through space very nearly like a surface or a volume. As a consequence, if a time series of a gene or protein has a fractal dimension significantly greater than 1, the dynamics of that gene or protein is more likely affected by chaos than by noise.

The sample entropy  $S_E$  adds further information to that provided by the Lyapunov exponents and the fractal dimension, as it is a direct measure of the unpredictability of a time series (Mao, 2011). Indeed,  $S_E$  estimates how much a given data point depends on the values of a number  $m$  of preceding data points, averaged over the whole time series.  $S_E$  is computed as the negative logarithm of the conditional probability that two similar samples from the time series remain similar at the next point (Richman and Moorman, 2000; Azar and Vaidyanathan, 2016). To calculate the sample entropy, points matching within a tolerance  $\epsilon$  are computed until there is no match according to this condition. Formally, if

$$X(t) = \{x(t_1), x(t_2), \dots, x(t_N)\} \equiv \{x_1, x_2, \dots, x_N\}$$

is a time series of length  $N$ , the sample entropy is defined as in the following by Azar and Vaidyanathan (2016), Sokunbi (2014), and Richman and Moorman (2000).

$$S_E(m, r, N, \tau) = -\ln \frac{U^{(m+1)}(\epsilon)}{U^{(m)}(r)} \quad (10)$$

where  $B_i$  is the number of  $j$  where  $|X(i) - X(j)| \leq r$ , and

$$U^{(m)}(\epsilon) = \frac{1}{N - m\tau} \sum_{i=1}^{N-m\tau} \frac{B_i}{N - (m+1)\tau}$$

$$X_m(i) = \{x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau}\}$$

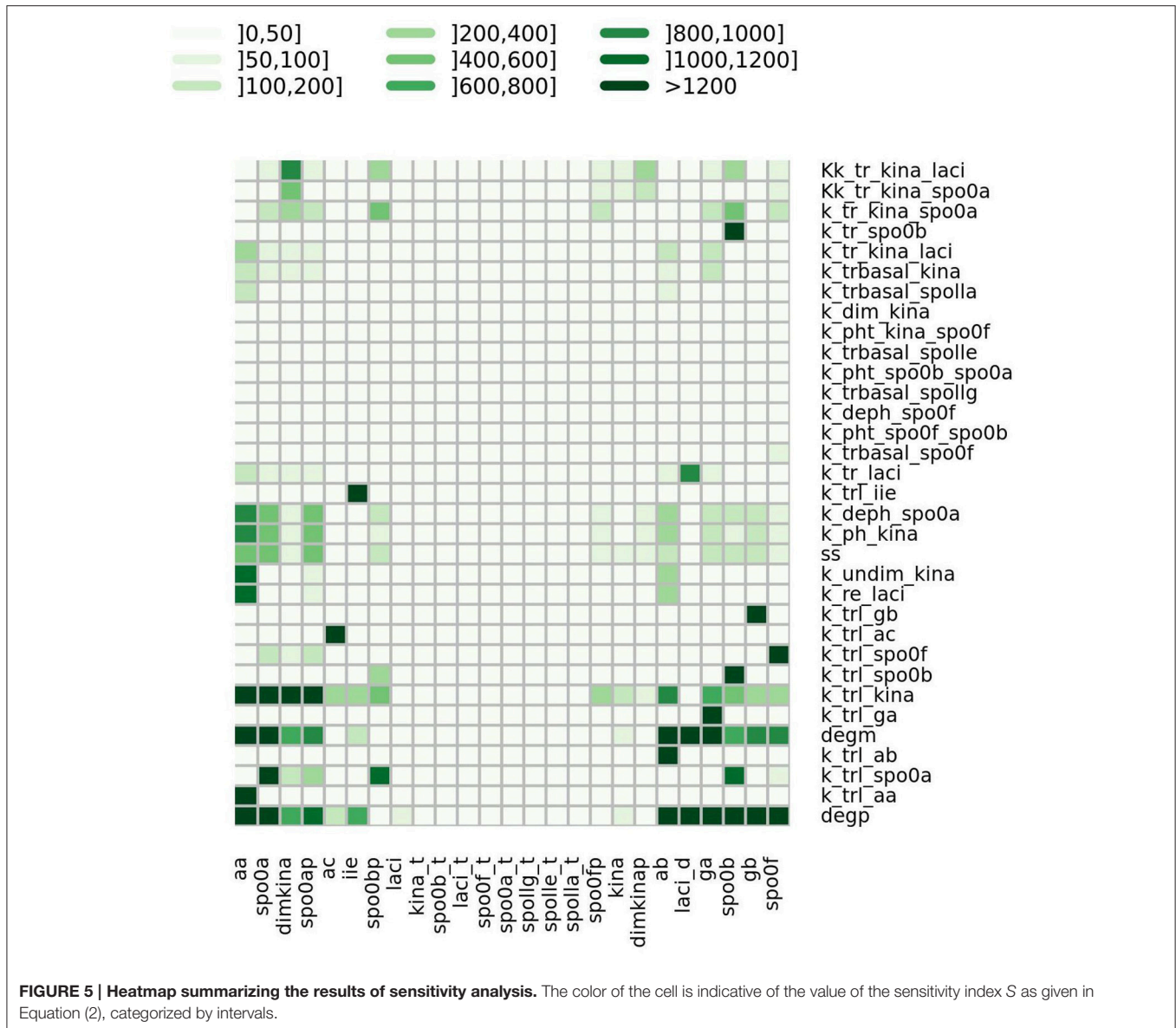
$$X_m(j) = \{x_j, x_{j+\tau}, \dots, x_{j+(m-1)\tau}\}$$

$$1 \leq j \leq N - m\tau, \quad j \neq i.$$

$X_m(i)$  is called *template vector* of length  $m$  of the time series  $X(t)$ , and an instance where a vector  $X_m(j)$  is within  $\epsilon$  of  $X_m(i)$  is called a *template match*. The quantity  $\frac{B_i}{N - (m+1)\tau}$  is the probability that any vector  $X_m(j)$  is within  $r$  of  $X_m(i)$ . Finally,  $\tau$  is called *time delay*, and in our analysis it has been set equal to the time lag  $T_L^*$ , that is an estimate of the time at which the time series behavior becomes unpredictable. It can be computed using the auto-correlation function method (Zeraoulia, 2011) and is taken as the lag time  $T_L^*$  at which the auto-correlation function

$$r_{T_L} = \frac{\sum_{i=1}^{N-T_L} (x_i - \bar{x})(x_{i+T_L} - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2}. \quad (11)$$





first crosses zero. This choice of  $\tau$  in the estimation of sample entropy is motivated by the need to capture also non-linear autocorrelation properties of the time series (Kaffashia et al., 2008). For instance, it has been proved that with a unity time delay (Kaffashia et al., 2008), the sample entropy measures only the linear autocorrelation properties of the time series. A lower value of  $S_E$  (and a higher value of  $T_L^*$ ) indicates higher predictability of the time series, while a higher value of  $S_E$  (and a lower value of  $T_L^*$ ) indicates lower predictability.

Finally, we also considered the embedding dimension as a measure of time series complexity. The embedding dimension of a time series is the smallest dimension required to embed it, and it can be estimated by the Cao's algorithm (Cao, 1997). In our analysis, the parameter  $m$  in the definition of sample entropy has been set equal to the embedding dimension.

### 3.2.1. Distinguishing Noise from Chaos

Since both the presence of chaos and the presence of noise are manifested as topological and statistical complexity of a time series, our analysis aims to distinguish chaos from noise. In the past decade many methods in a different application domains have been proposed to make this distinction, the most recent are reported in Skiadas and Skiadas (2016), Ravetti et al. (2014), Rohde (2008), Gao et al. (2006), and Rosso et al. (2007).

We adopted a simple well established method based on the comparison of the complexity indices identified above and obtained from the time series simulated by the model with those obtained from Gaussian white noise, colored noise and power-law noise (Skiadas and Skiadas, 2016). The expectation is that sample entropy, time lag, and embedding dimension for the non-noisy candidate chaotic times series are significantly different from those estimated for the white and colored noise time

**TABLE 1 | Molecular species and the parameters mostly controlling their dynamics.**

Molecular species	Parameters
aa	degp,k_trl_aa
spo0a	degp,k_trl_spo0a
dimkina	k_trl_kina,Kk_tr_kina_laci
spo0ap	k_trl_kina,degp
ac	k_trl_ac,k_trl_kina
iee	k_trl_iee,degp
spo0bp	k_trl_spo0a,k_trl_kina
ab	degp,k_trl_ab
laci_d	degp,degm
ga	degp,k_trl_ga
spo0b	degp,k_trl_spo0b
gb	degp,k_trl_gb
spo0f	degp, k_trl_spo0f

series. Moreover, the time behavior of the Lyapunov exponents is expected to be non-linear for the noise and at least linear for chaotic non-noisy time series (Gao and Zheng, 1994).

### 3.3. Analysis of the Jacobian Matrix: The Time Evolution of the Phase Space

In order to explore the phase space of the systems and calculate its equilibria and its time evolution we analyzed the Jacobian matrix **J** of the system of ordinary differential equations describing the dynamics.

$$\mathbf{J} = \begin{bmatrix} \frac{\partial f_1}{\partial s_1} & \frac{\partial f_1}{\partial s_2} & \dots & \frac{\partial f_1}{\partial s_d} \\ \frac{\partial f_2}{\partial s_1} & \frac{\partial f_2}{\partial s_2} & \dots & \frac{\partial f_2}{\partial s_d} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_d}{\partial s_1} & \frac{\partial f_d}{\partial s_2} & \dots & \frac{\partial f_d}{\partial s_d} \end{bmatrix} \quad (12)$$

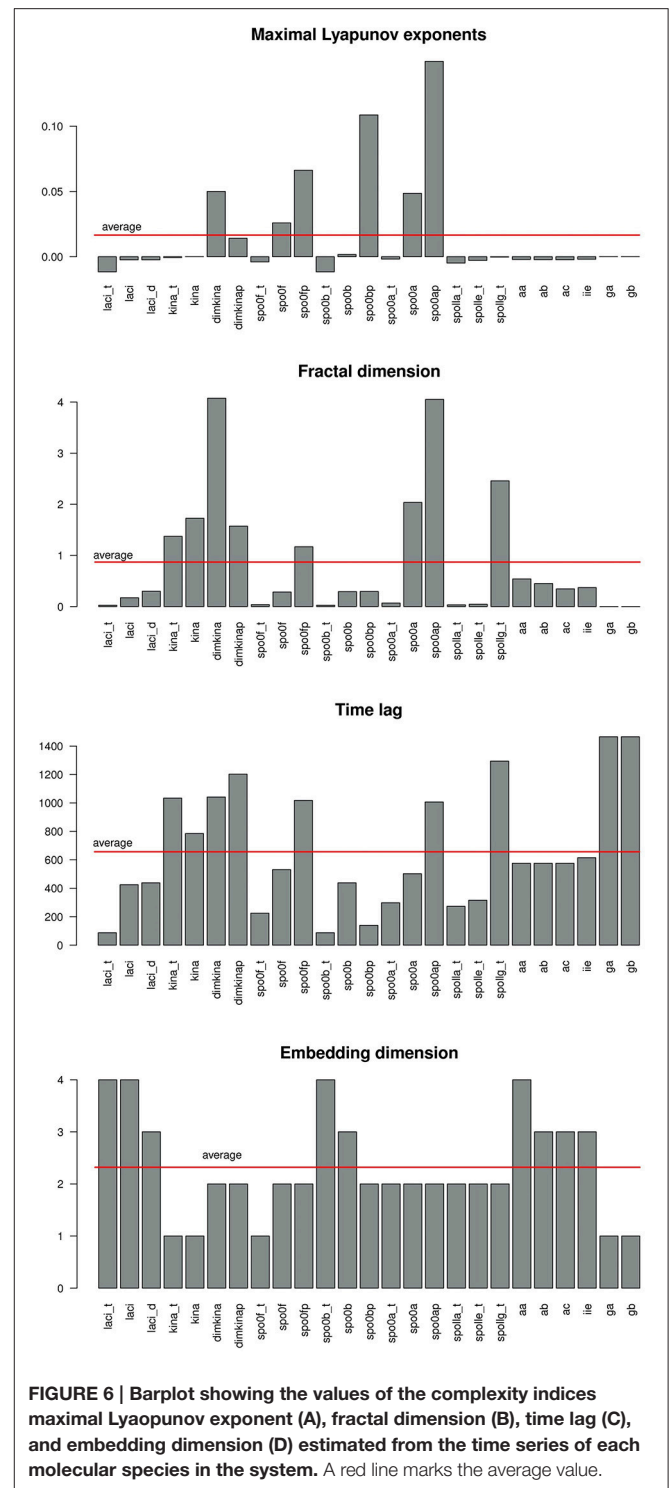
where  $f_i = \frac{ds_i}{dt}$ , and  $s_i$  is the abundance of the  $i$ -th molecular species in the system ( $i = 1, 2, \dots, d$ ).

A steady state point  $s_{eq} = \{s_i^{eq}\}$ ,  $i = 1, 2, \dots, d$ , of the systems is defined by a solution of the system of algebraic equations as in the follows:

$$\frac{ds_i}{dt} = 0, \quad i = 1, 2, \dots, d.$$

The stability of a steady state point is determined by the sign of the real part of the eigenvalues of the Jacobian matrix. In particular, if the real parts of all eigenvalues are negative, the steady state point is stable. It's termed *sink*, because, there is a basin around it, and any initial condition in that basin will result in a trajectory falling in toward the steady state point.

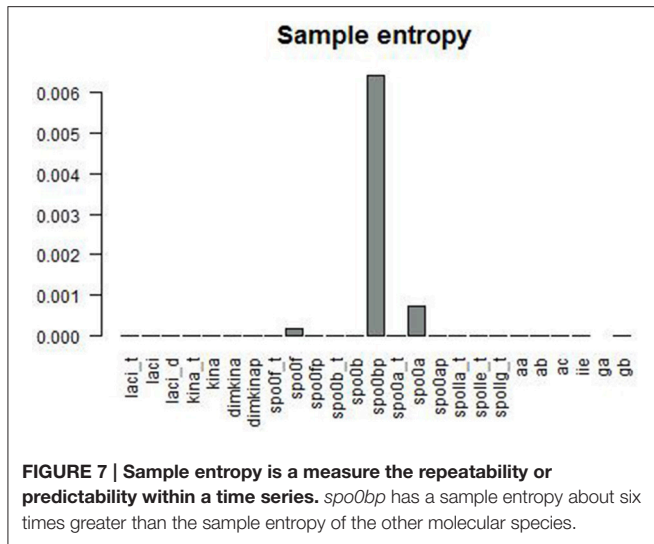
If the real parts of all the eigenvalues are positive, the steady state point is *unstable*. It is termed *source*, because, starting from an initial point close to it, the trajectory will move away from it. If the real parts of the eigenvalues are of different signs, the steady state point is called a *saddle* point. It is *unstable*, attracting along some axes and repelling along others. If there are also complex components, the nature of the fixed point doesn't change (it's still



**FIGURE 6 | Barplot showing the values of the complexity indices maximal Lyaopunov exponent (A), fractal dimension (B), time lag (C), and embedding dimension (D) estimated from the time series of each molecular species in the system. A red line marks the average value.**

a sink, source, or saddle point) but with a twist. If the eigenvalues are purely complex, then there are closed orbits around the fixed point.

The eigenvectors of the Jacobian matrix give the axes along which the behaviors indicated by the eigenvalues are centered. So, the eigenvector associated with a negative eigenvalue is a vector along which the fixed point attracts. The eigenvector associated



with a positive eigenvalue is an axis along which the fixed point repels.

#### 4. RESULTS

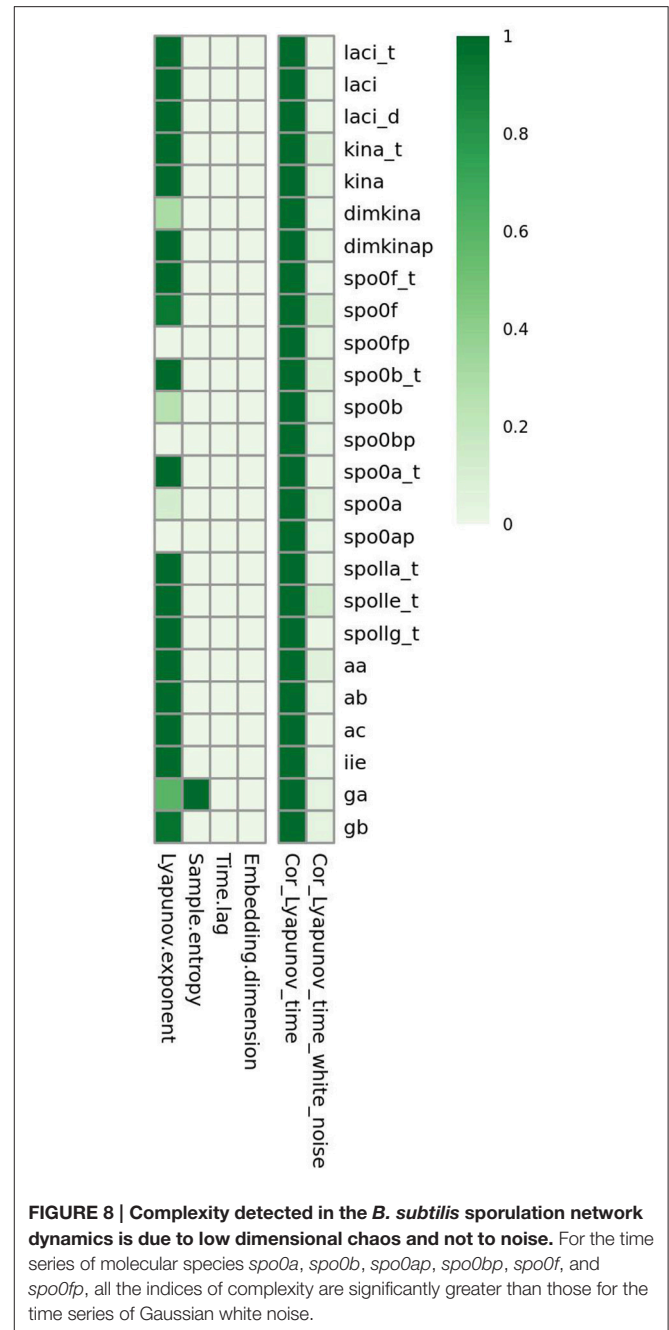
In this section we collect the results of three different analyses: (i) the parameter sensitivity analysis, (ii) the model time series analysis, and (iii) the phase space analysis. The first two analyses capture different aspects and manifestations of the presence of chaos and their outputs are sets of molecular species whose behavior is a likely candidate for chaotic dynamics. The final result is an intersecting set of molecular species that represents the consensus set of molecular species whose dynamics is affected by chaos. The third analysis aims at determining how the topology and the parameters of the network of interactions among the molecular species evolves with time. This last analysis allows the determination of the time variation of the active degrees of freedom in the system (Hilborn, 2000).

##### 4.1. Kinetic Rate Constants Controlling the Dynamics

We explored the parameters' space in which the systems of ordinary differential equations that represent the model has a solution. We found that the largest range of variation for the parameters at which the systems still admits a unique solution is defined by

$$I(p_h) = \left[ \frac{p_h}{10}, 10 \times p_h \right], \quad h = 1, 2, \dots, N_p \quad (13)$$

where  $p_h$  is the value of the  $h$ -th parameter assigned from experimental data. For each parameter we randomly sampled 50 values from a uniform distribution positively defined in  $I(p_h)$ . In turn these were used in simulations to give sensitivity indices according to Equation (2). In **Figure 5** a heatmap shows the value of the sensitivity index collapsed into intervals. Moreover, **Table 1** lists the variables and the parameters which affect them most.



Appreciable parameter sensitivity is only apparent in an interval ranging from a tenth to ten times the parameter value obtained from data. The dynamics of the majority of the molecular species is robust with respect to the variations of the parameters' values on smaller intervals (see Figure S2). The molecular species *spo0a*, *spo0b*, *spo0ap*, and *spo0bp* are the most sensitive to perturbation of parameters even on small intervals.

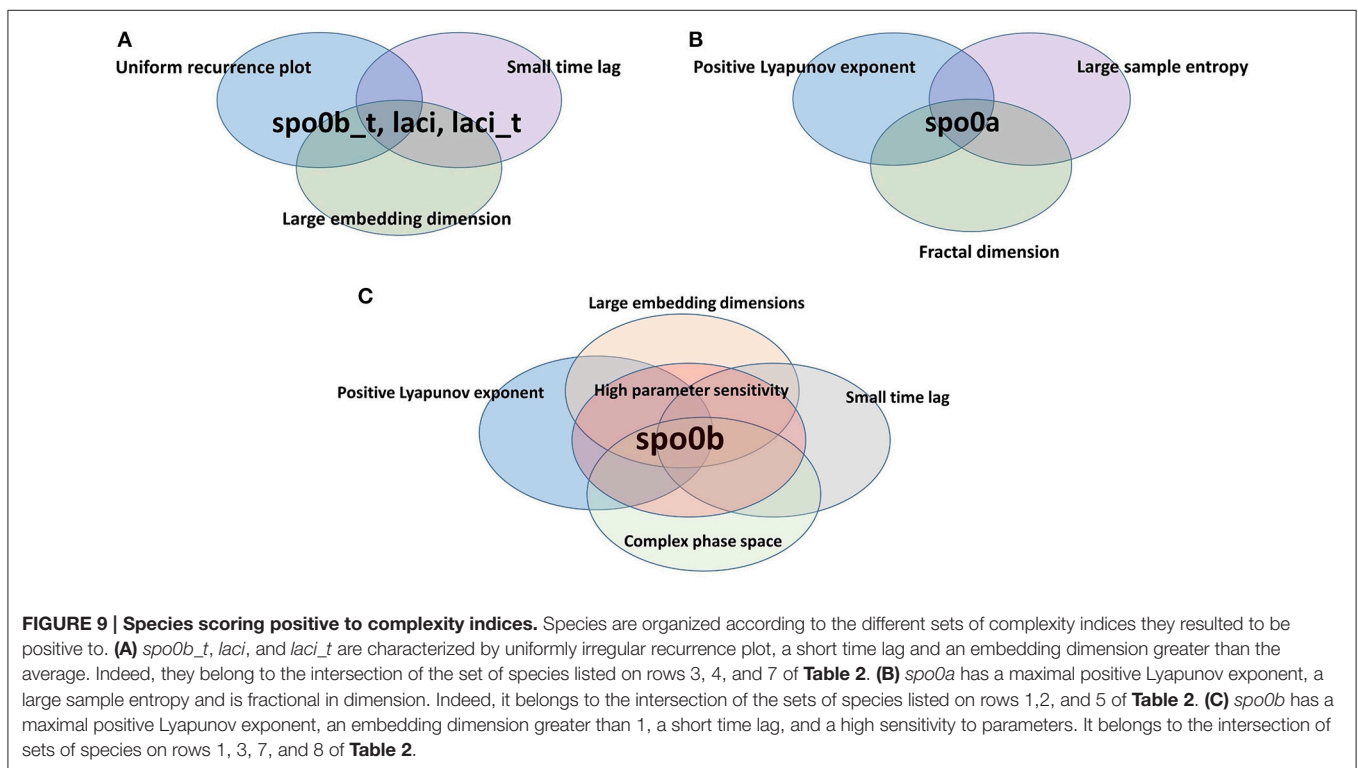
##### 4.2. Complexity Indices Identify the Drivers of Chaotic Dynamics

**Figure 6** gives a graphical summary of the complexity indices estimated for the time series of each molecular species. The

**TABLE 2 | Sets of molecular species with values of the complexity indices indicative of presence of chaos.**

Index	Variables	% of total nr. of variables
Positive Lyapunov exponent	dimkina, dimkinap, spo0f, spo0fp, spo0b, spo0bp, spo0a, spo0ap	32
Fractal dimension (above the average)	kina_t, kina, dimkina, dimkinap, spo0fp, spo0a, spo0ap, spollg_t	32
Embedding dimension (above the average)	kina_t, kina, dimkina, dimkinap, spo0fp, spo0a, spo0ap, spollg_t	36
Time lag (below the average)	laci_t, laci, laci_d, spo0f_t, spo0f, spo0b_t, spo0b, spo0bp, spo0a_t spo0a, spolla_t, spolle_t, aa, ab, ac, iie	64
Sample entropy (above the average)	spo0f, spo0bp, spo0a	12
Complex trajectory in phase space	ga, gb, spo0a, spo0a_t, spo0b, spo0f_t, spollg_t	28
Recurrence plot reveals chaos	spo0b, spo0bp, spo0a, spo0ap, aa, ab	24
High parameter sensitivity (ordered by sensitivity)	aa, spo0a, dimkina, spo0ap, ac, iie, spo0bp, ab, laci_d, ga, spo0b, gb, spo0f	52

The third column reports the percentage of species scoring positive to the test of chaos according to a set of indices.



majority of the molecular species have positive Lyapunov exponents, fractional dimension, time lag between 0 and 400 (that is the about 3% of the time range used in the simulation), and embedding dimension greater than 1. **Figure 7** shows the sample entropy values and reports that the highest vale of sample entropy is assumed by *spo0bp*.

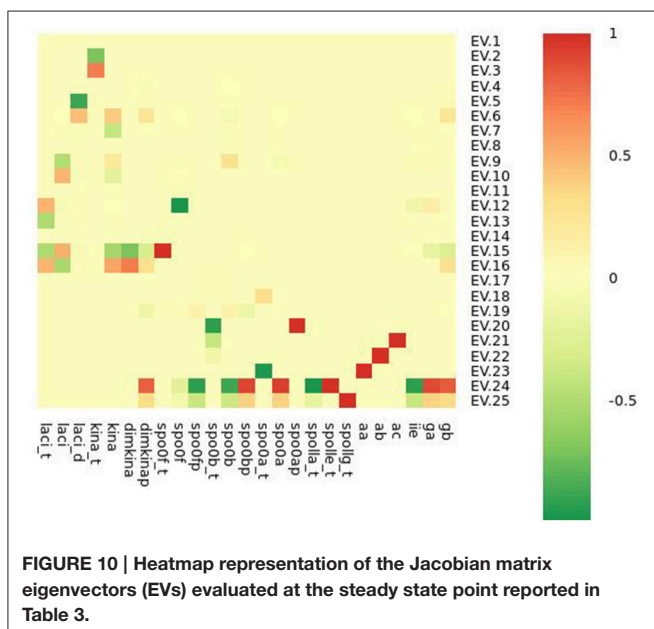
To distinguish chaotic from noisy dynamics, we compared the complexity indices of the time series of each variable with the mean and the standard deviation of the complexity indices estimated from 50 time series of white Gaussian noise of mean  $\mu = 0$ , variance  $\sigma^2 = 1$  generated for each variable and having amplitude equal to the range of variability of the

variable. The heatmap in **Figure 8** shows, on the left side, the frequency at which the noisy time series has an index value higher than that observed for the time series from the real model. Comparison of indices under chaotic and noisy conditions is performed for each index (shown by column) and each variable (shown by row). The heatmap on the right side shows correlation between Lyapunov index and time points in the time series. Column “Cor\_Lyapunov\_time” displays statistical significance for Pearson’s correlation relating Lyapunov exponent and time. Column “Cor\_Lyapunov\_time\_white\_noise” displays similar information for the time series of white Gaussian noise generated for each variable. We found that sample entropy,



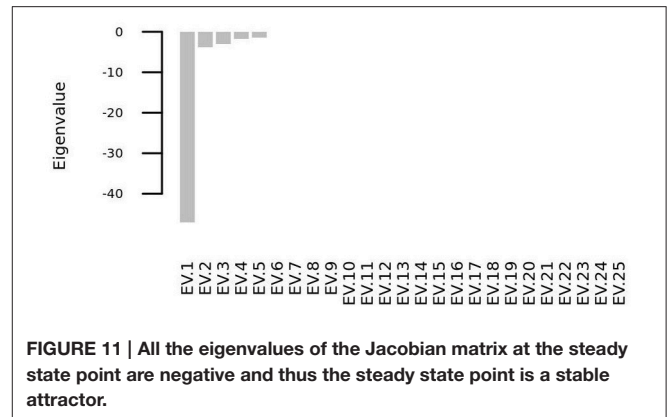
**TABLE 3 | Coordinates of the stable steady state point of the model solved with initial conditions  $s_i(t = 0) = 0 \forall i = 1, 2, \dots, 25$ .**

Molecular species	Coordinate (nM)
<i>laci_t</i>	17.24
<i>laci</i>	18.27
<i>laci_d</i>	2855.29
<i>kina_t</i>	77.89
<i>kina</i>	674.32
<i>dimkina</i>	1719.72
<i>dimkinap</i>	81.81
<i>spo0f_t</i>	20.68
<i>spo0f</i>	1219.14
<i>spo0fp</i>	26.7
<i>spo0b_t</i>	41.1
<i>spo0b</i>	3683.52
<i>spo0bp</i>	2.09
<i>spo0a_t</i>	23.92
<i>spo0a</i>	2350.24
<i>spo0ap</i>	1921.15
<i>spolla_t</i>	76.6
<i>spolle_t</i>	57.45
<i>spollg_t</i>	81.74
<i>aa</i>	7979.31
<i>ab</i>	3542.82
<i>ac</i>	880.92
<i>iie</i>	660.73
<i>ga</i>	2316.09
<i>gb</i>	940.06



**FIGURE 10 | Heatmap representation of the Jacobian matrix eigenvectors (EVs) evaluated at the steady state point reported in Table 3.**

time lag and embedding dimension are significantly higher in the model time series than in the white noise time series. The Lyapunov exponents are significantly greater for the white noise time series compared with the model's time series, except for



**FIGURE 11 | All the eigenvalues of the Jacobian matrix at the steady state point are negative and thus the steady state point is a stable attractor.**

*spo0a*, *spo0ap*, *spo0bp*, *spo0f*, and *spo0fp*. This result suggests that these molecular species exhibit remarkable chaotic dynamics. The left part of the heatmap confirms a non linear time behavior of the Lyapunov exponents of the white noise time series, and suggests a linear time behavior of the Lyapunov exponents of the model's time series. Again, this result distinguishes between chaotic dynamics and random noisy dynamics (Gao and Zheng, 1994). In the Supplementary Material (Figures S4–S6), we report similar results obtained in the comparison of the complexity indices of the model's time series with the ones for the colored and power law noise.

For each index of complexity, **Table 2** (and a graphical summary of it in **Figure 9**) reports the set of molecular species where results indicate the presence of chaos. In this Table, we also included two qualitative indicators of the presence of chaos, such as a complex phase space (i.e., convoluted trajectories) and recurrence plot. These plots are provided in Supplementary Material (Figure S7), and visualize a square matrix, whose elements are the times at which a state of a dynamical system recurs (columns and rows correspond then to a certain pair of times) (Marwan et al., 2016, 2007). We refer the reader to the Supplementary Material for a comprehensive description of the recurrence plots analysis.

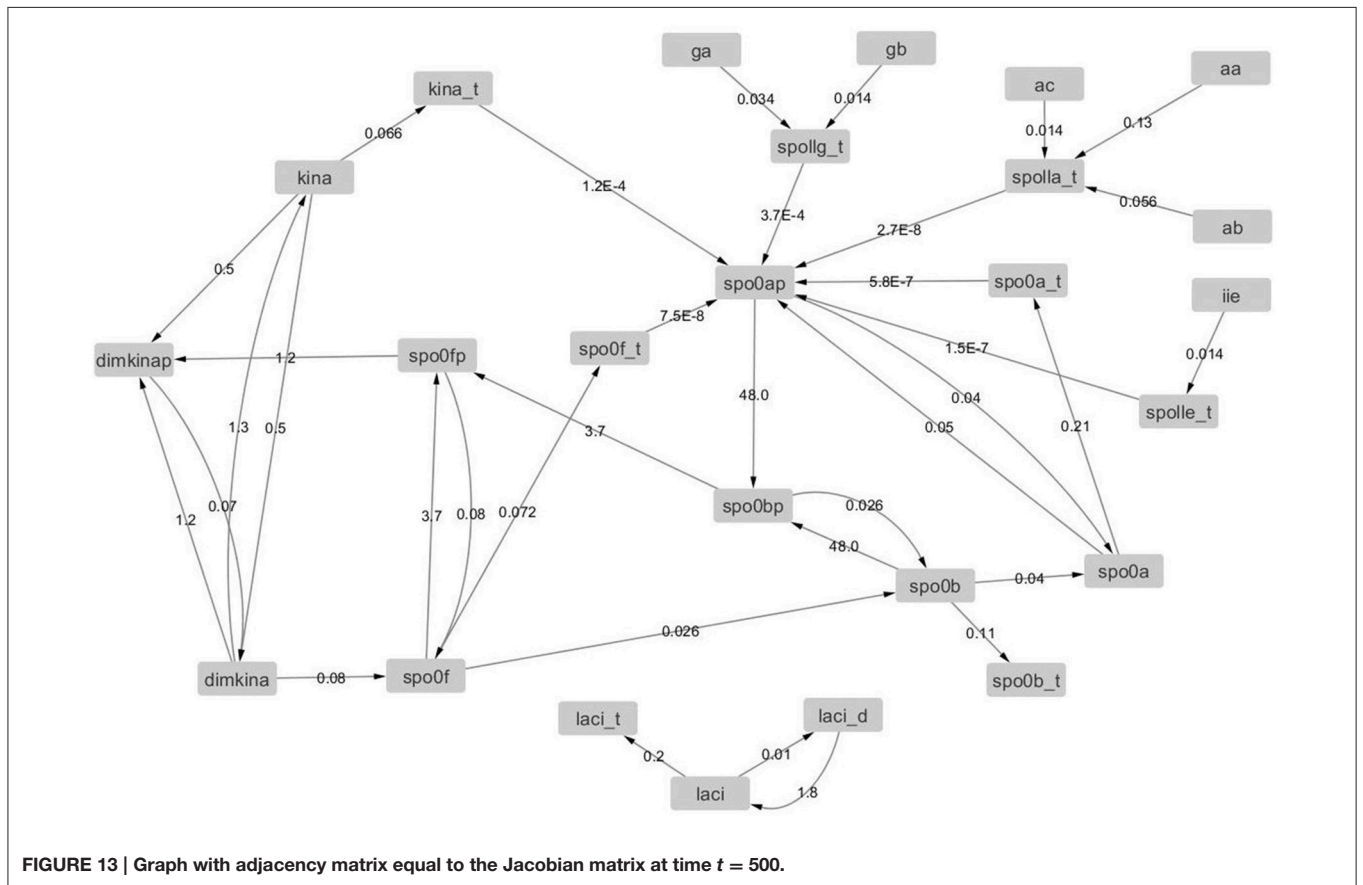
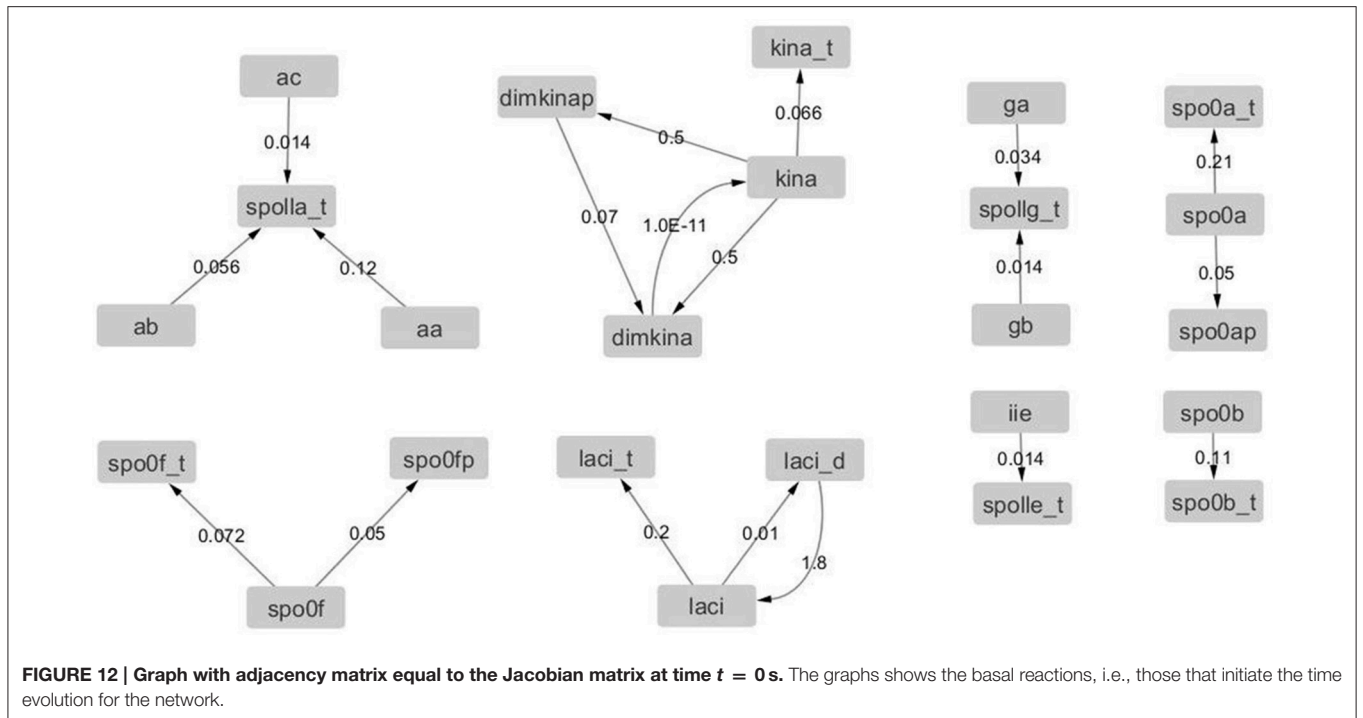
The variable *spo0b* is the one with the largest set of complexity indices whose values point to the presence of chaos.

### 4.3. Phase Space Analysis

We solved the systems of differential equations setting to zero the initial concentration of all the molecular species (i. e.  $s_i(t = 0) = 0 \forall i = 1, 2, \dots, 25$ ), and we found that the system has one steady state point, whose coordinates are shown in **Table 3**. This point is a stable equilibrium, as the eigenvalues of the Jacobian matrix (**Figure 10**) of the system are all negative (**Figure 11**).

We also calculated the value of the elements of the Jacobian matrix at different time points to determine its time evolution. Since the entries of the Jacobian matrix are the partial derivatives of the rate equations with respect to the variables [i.e.,  $J_{ij} = \frac{\partial f_i}{\partial s_j}$  (Equation 12)], the Jacobian matrix can be represented by a weighted graph, where the nodes represent the variables (i.e., the molecular species) and the edge weights are the elements of the matrix  $J_{ij}$ . We introduced the estimate of the error on





$J_{ij}$  defined as  $\Delta J_{ij} = \text{prec} \times \frac{\partial^2 f_i}{\partial S_j^2}$ , where  $\text{prec} = 10^{-8}$  is the precision of the numerical solution of the model, and set the threshold of 20% on the relative error  $ER = \frac{\Delta J_{ij}}{J_{ij}}$ . Edges with  $ER < 20\%$  were retained and allowed for the estimation of the number of active degrees of freedom in the system, i.e., the number of variables involved in active interactions (Hilborn, 2000). Hence, the graphs derived from the Jacobian matrices estimated at different time points are temporal snapshots of the molecular interaction network.

**TABLE 4 | Minimum, first quartile, median, mean, third quartile, and maximum of the distributions of the Jacobian matrix values (i.e., edge weights of the networks) at times  $t = (0, 500, 1000, \dots, 7500)$ .**

Time (sec)	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0	0.00	0.02	0.07	0.20	0.16	1.80
500	0.00	0.01	0.07	3.09	0.50	48.00
1000	0.00	0.01	0.07	3.09	0.50	48.00
1500	0.00	0.01	0.07	3.04	0.50	47.00
2000	0.00	0.01	0.07	3.04	0.50	47.00
2500	0.00	0.01	0.07	3.03	0.50	47.00
3000	0.00	0.01	0.07	3.04	0.50	47.00
3500	0.00	0.01	0.07	3.04	0.50	47.00
4000	0.00	0.01	0.05	0.90	0.46	11.00
4500	0.00	0.01	0.05	2.76	0.50	43.00
5000	0.00	0.01	0.06	3.30	0.50	52.00
5500	0.00	0.01	0.07	3.36	0.50	53.00
6000	0.00	0.01	0.07	3.25	0.50	51.00
6500	0.00	0.01	0.07	3.20	0.50	50.00
7000	0.00	0.01	0.07	3.14	0.50	49.00
7500	0.00	0.01	0.07	3.09	0.50	48.00

Steep changes of the values of the quartiles reveal an irregular behavior of the dynamics, and reflect the presence of chaos.

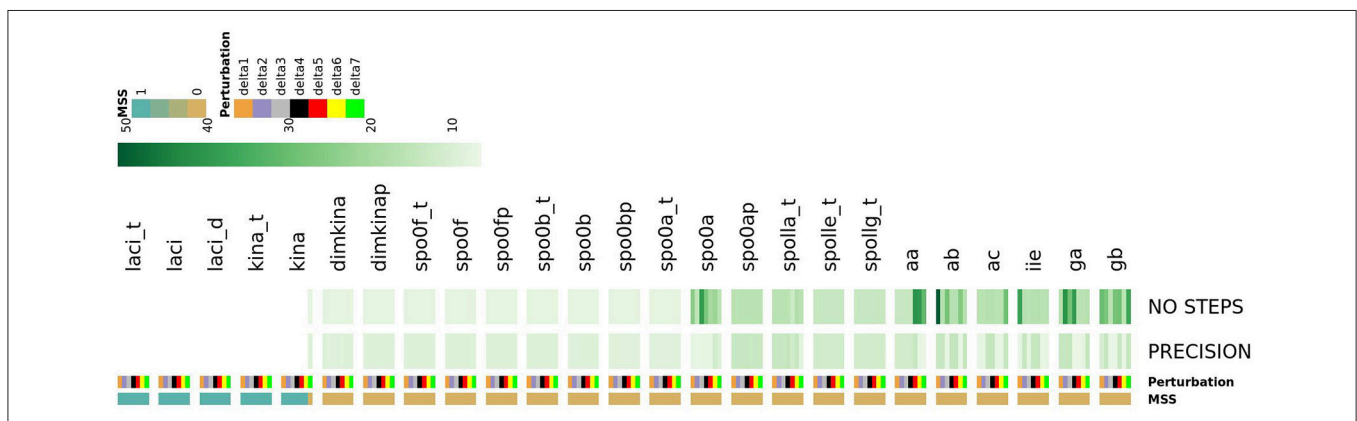
The graphs derived from the Jacobian matrix estimated at different time points are time snapshots of the interaction network of the molecular species. These graphs visualize the interactions that are active (i.e., with an edge that has a weight significantly different from zero) at a given time, and thus provide an approximate estimation and representation of the number of active degree of freedom of the system. We report in **Figures 12, 13**, the graphs obtained from the Jacobian matrices evaluated at times  $t = \{0, 500\}$  s. For  $t > 500$  s the variations of the Jacobian matrix are minimal and thus not shown here (in the Supplementary Material we provide the graphs for  $t > 500$  s in GraphML format).

In **Table 4**, we observe a rapid increment of the edge weights from  $t = 0$  to  $t = 500$ , than a plateau till  $t = 4,000$ , and then a decrement at  $t = 4,000$  s. These changes reflect the changes of the topology of the network of *B. subtilis* sporulation initiation. These steep decrements and increments are expressions of a stiff and highly non linear dynamics, and in turn confirms the presence of chaos (unpredictability) in (of) it.

### Sensitivity to Initial Conditions

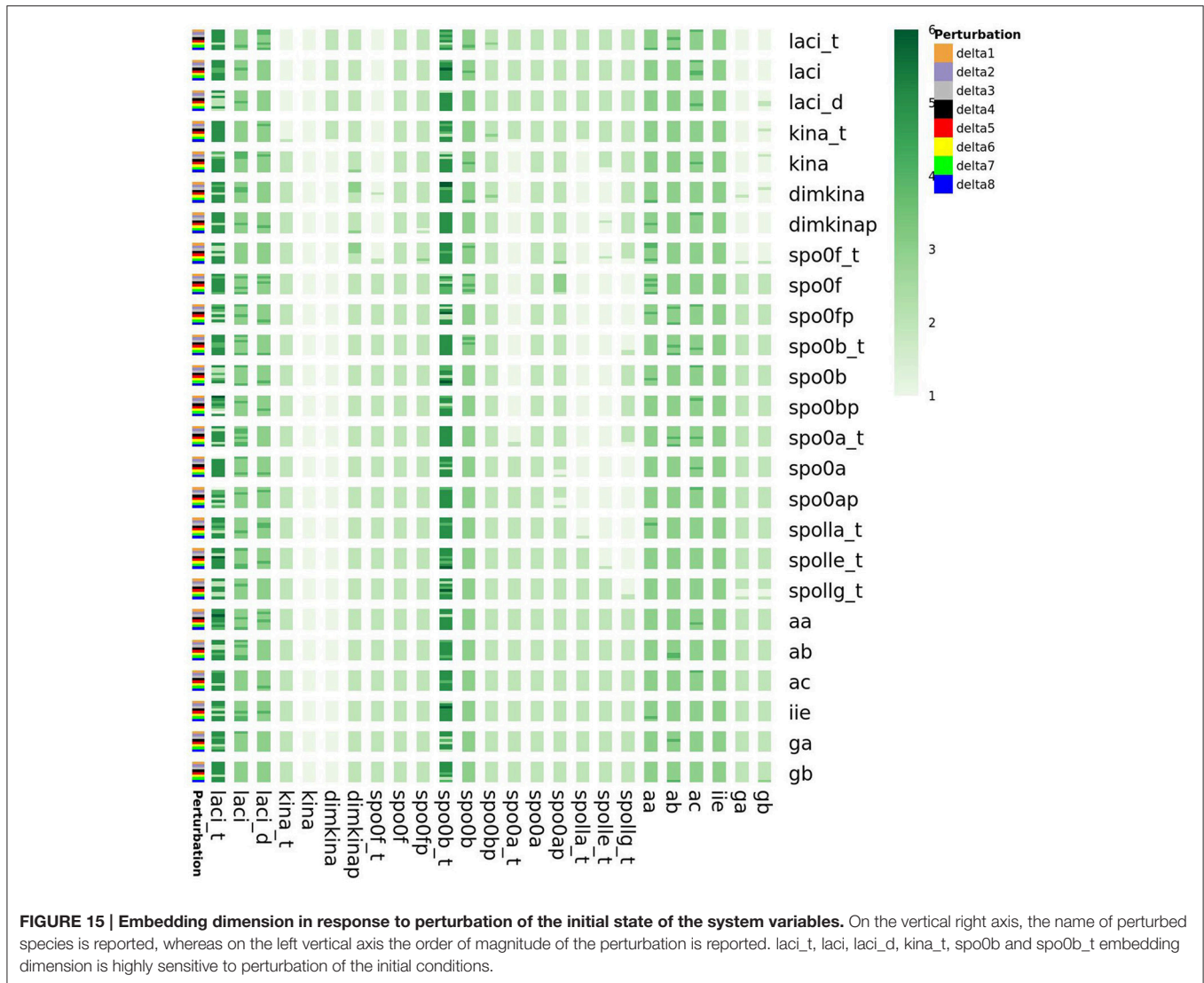
In order to assess the existence and, eventually, the sensitivity of the steady state point in response to perturbations of the initial conditions, we introduced a set  $\Delta$  of perturbations of different magnitudes:  $\Delta = \{\Delta_h, h = 1, 2, \dots, 7\}$ , where  $\Delta_1 = 10^{-5}$  and  $\Delta_{i+1} = 10\Delta_i, i = 1, 2, \dots, 6$ . We perturbed the initial state of one variable at a time by  $\Delta_h (h = 1, 2, \dots, 7)$ , and calculated the steady state point with the Newton-Raphson method (Deuffhard, 2004; Bressoud, 2007), which is one of the most consolidated and efficient algorithms for finding the zeros of a function (Hoppensteadt, 1993; Ortega and Rheinboldt, 2000).

As shown in **Figure 14**, this analysis allowed detecting a number of species which, once perturbed, can cause different types of complex evolutionary trajectories from the initial state of the system. Perturbing the initial state of the



**FIGURE 14 | The heatmap describes the systems ability to reach the steady state in response to seven different perturbation values of each species.**

System's response is described by three features: MSS, PRECISION and NO STEPS. Depending on the perturbation magnitude of the initial state, the Newton-Raphson methods warns about the non-convergence to a steady state. This occurs when, depending on the initial conditions, the system has multiple steady states (MSS), and the Newton-Raphson solution methods flips unpredictability among them (Bloomfield, 2014). MSS is a binary feature assuming 1 if the system reaches multiple steady states and 0 otherwise; PRECISION is the precision in computation of a single steady state (reported in negative log scale.); NO STEPS is the number of steps required converge to a single steady state. Perturbations extent is shown for each systems species by a color bar. The map shows that the perturbation of laci\_t, laci, laci\_d, and kina causes an unresolved multi-stationarity.



*laci\_t*, *laci*, *laci\_d*, *kina\_t* and *kina* species caused the most noticeable consequence, preventing the system from reaching the steady state, irrespectively of the applied perturbation extent. This finding reflects the role of the species as starters of sporulation dynamics. For another subset of species, such as *spo0a*, *aa*, *ab*, *ga*, and *gb*, perturbing the initial state caused an increase in the number of iterative steps required by the Newton-Raphson method to reach the steady state, compared to the majority of the species in the system. This effect is indicative of increased complexity in the evolving system.

Furthermore, as shown in the Supplementary Material, we also found that the perturbations of the initial state of all the species cause variations in the minimum embedding dimension of *spo0b*, and *spo0b\_t*, as well as of *laci\_t*, *laci*, *laci\_d*, *kina\_t* (Figure 15). Quantification of the increase in minimum embedding dimension upon perturbation of the initial state is a further indication of the high sensitivity to perturbations, for the species highlighted by the initial analysis of complexity indices,

such as *spo0b*, *spo0b\_t*, and for the species emerging from the previous characterization of perturbation experiments.

Perturbations of initial conditions also cause variations of the Kaplan-Yorke ratio (i.e.,  $(\sum_{i=1}^j \lambda_i)/\lambda_{i+j}$ , where  $j$  is the largest integer such that  $\sum_{i=1}^j \lambda_i > 0$ ) of *spo0b\_t*, *kina\_t*, and *gb* and of the Kolmogorov-Sinai entropy (i.e., the sum of positive Lyapunov exponents) of *kina\_t* (Figures S8–S10). Hence, by multiple analyses the system showed a highly unpredictable behavior to perturbation of the initial state, as reflected by exponentially growing separation of the trajectories as well as by the topological complexity of the time series.

## CONCLUSIONS

We have presented a detailed, and original analysis of the *B. subtilis* sporulation initiation network dynamics. This analysis aimed to detect the presence of low-dimensional chaos in the dynamics of the system. Unlike more common approaches to chaos detection, this analysis includes and, is based on,

a mathematical model of the *B. subtilis* sporulation initiation dynamics. This approach allows a comprehensive explanation of the mechanisms through which those molecular species with chaotic dynamics interact with the others and propagate the effects of chaos throughout the system.

Our analysis has: (i) assessed sensitivity of the dynamics by varying the kinetic parameters and the initial state of the model to determine the dynamic control parameters and identify the most crucial molecular species; (ii) calculated the complexity indices of the time series obtained from the model, and used these to identify the drivers of chaotic dynamics, and finally, (iii) calculated the Jacobian matrix of the system of equations as a function of time to find the steady state points and their nature to give an estimate of the number of active degrees of freedom of system as function of time.

We found that the dynamics of the *B. subtilis* sporulation initiation network is affected by low-dimensional chaos and identified Spo0B as the principal driver of the chaotic dynamics in this system. Spo0B scored positive for the majority of the chaos indices. This result suggests a new role for this molecular species, which so far has received little attention, and highlights the importance of its dynamics and interactions within the network model structure. Our analysis also indicates additional experimental work that could be conducted to improve our understanding of the sporulation network and to determine the role of Spo0B. On one hand, it would be important to conduct phosphoproteomics experiments to measure the amount of phosphorylated species, for which no experimental data is yet available. On the other hand, the study of Spo0B mutated (overexpressed/silenced) strains would refine our knowledge about the dynamics of the 3-level *B. subtilis* phosphorelay and the so far elusive mechanisms that could be regulating its expression. Other molecular species that also showed positive results for the test of chaos were *spo0a*, *spo0b\_t*, *laci*, and *laci\_t*. These molecular species also represent the degrees of freedom that are active during most of the range of the simulated time. These species have been identified as the drivers of the chaotic dynamics of

*B. subtilis* sporulation initiation network model, and have an active role in determining its predictability.

One of the most challenging goal of studying a complex biological system is to control it. In this vein, our work proposes a method to identify the drivers in our *B. subtilis* sporulation initiation network on which such methods of chaos control might be applied. At the same time, our work highlights the need to investigate on these drivers and their mechanism of interaction in order to successfully implement chaos control.

## AUTHOR CONTRIBUTIONS

All authors have equally contributed to the design and implementation of the mathematical analytical procedure, the discussion of the results and the writing of this paper.

## FUNDING

This work was partially supported by the Biotechnology and Biological Sciences Research Council (BBSRC) Institute Strategic Programme [BB/J004529/1]: The Gut Health and Food Safety ISP.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the Biotechnology and Biological Sciences Research Council (BBSRC). We thank everyone at quantitative microbial risk assessment group and the Clostridium botulinum bacterial pathogens group at the IFR.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2016.01760/full#supplementary-material>

## REFERENCES

Abarbanel, H. D. (1996). *Analysis of Observed Chaotic Data*. New York, NY: Springer-Verlag. doi: 10.1007/978-1-4612-0763-4

Azar, A. T., and Vaidyanathan, S. (2016). *Advances in Chaos Theory and Intelligent Control*. Springer International Publishing. doi: 10.1007/978-3-319-30340-6

Bloomfield, V. A. (2014). *Using R for Numerical Analysis in Science and Engineering*. Boca Raton, FL: CRC Press, Taylor and Francis Group.

Bressoud, D. M. (2007). *A Radical Approach to Real Analysis*. Washington, DC: The Mathematical Association of America.

Cao, L. (1997). Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D* 110, 43–50. doi: 10.1016/S0167-2789(97)00118-8

Carlin, F. (2011). Origin of bacterial spores contaminating foods. *Food Microbiol.* 28, 177–182. doi: 10.1016/j.fm.2010.07.008

Cencini, M., Cecconi, F., and Vulpiani, A. (2009). *Chaos. From Simple Models to Complex Systems*. Singapore: World Scientific. doi: 10.1142/7351

de Jong, I. G., Veening, J. W., and Kuipers, O. P. (2010). Heterochronic phosphorelay gene expression as a source of heterogeneity in *Bacillus subtilis* spore formation. *J. Bacteriol.* 192, 2053–2067. doi: 10.1128/JB.01484-09

Deuffhard, P. (2004). *Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms*. Springer Series in Computational Mathematics. Berlin: Springer.

Ding, M., Grebogi, C., Ott, E., Sauer, T., and Yorke, J. A. (1993). Estimating correlation dimension from a chaotic time series: when does plateau onset occur? *Physica D* 69, 404–424. doi: 10.1016/0167-2789(93)90103-8

Eswaramoorthy, P., Guo, T., and Fujita, M. (2009). *In vivo* domain-based functional analysis of the major sporulation sensor kinase, KinA, in *Bacillus subtilis*. *J. Bacteriol.* 191, 5358–5368. doi: 10.1128/JB.00503-09

Fawcett, P., Eichenberger, P., Losick, R., and Youngman, P. (2000). The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8063–8068. doi: 10.1073/pnas.140209597

Fujita, M., and Losick, R. (2005). Evidence that entry into sporulation in *Bacillus subtilis* is governed by a gradual increase in the level and activity of the master regulator Spo0A. *Genes Dev.* 19, 2236–2244. doi: 10.1101/gad.1335705

Gao, J., and Zheng, Z. (1994). Direct dynamical test for deterministic chaos. *Europhys. Lett.* 25:485. doi: 10.1209/0295-5075/25/7/002

Gao, J. B., Hu, J., Tung, W. W., and Cao, Y. H. (2006). Distinguishing chaos from noise by scale-dependent Lyapunov exponent. *Phys. Rev. E* 74:066204. doi: 10.1103/PhysRevE.74.066204



- Higuchi, T. (1988). Approach to an irregular time series on basis of fractal dimension. *Physica D* 21, 277–283. doi: 10.1016/0167-2789(88)90081-4
- Hilborn, R. (2000). *Chaos and Nonlinear Dynamics, An Introduction for Scientists and Engineers*. Oxford: Oxford University Press.
- Hoppensteadt, F. C. (1993). *Analysis and Simulation of Chaotic Systems*. New York, NY: Springer-Verlag.
- Ihekwa, A., Mura, I., and Barker, G. C. (2014). Computational modelling and analysis of the molecular network regulating sporulation initiation in *Bacillus subtilis*. *BMC Syst. Biol.* 8:119. doi: 10.1186/s12918-014-0119-x
- Jiang, M., Shao, W., Perego, M., and Hoch, J. A. (2000). Multiple histidine kinases regulate entry into stationary phase and sporulation in *Bacillus subtilis*. *Mol. Microbiol.* 38, 535–542. doi: 10.1046/j.1365-2958.2000.02148.x
- Kaffashia, F., Foglyanob, R., Wilsonb, C. G., and Loparao, K. A. (2008). The effect of time delay on approximate & sample entropy calculations. *Physica D* 237, 3069–3074. doi: 10.1016/j.physd.2008.06.005
- Kalitin, B. S. (2004). Lyapunov stability and orbital stability of dynamical systems. *Diff. Equat.* 40, 1096–1105. doi: 10.1023/B:DIEQ.0000049826.73745.c1
- Kaneko, K. (2006). *Life: An Introduction to Complex Systems Biology*. Berlin; Heidelberg: Springer-Verlag.
- Kaneko, K., and Tsuda, I. (2001). *Complex Systems: Chaos and Beyond. A Constructive Approach with Applications in Life Sciences*. Berlin; Heidelberg: Springer-Verlag.
- Kantz, H. (2004). *Nonlinear Time Series Analysis*. Cambridge: Cambridge University Press.
- Kellert, S. H. (1993). *In the Wake of Chaos: Unpredictable Order in Dynamical Systems*. Chicago, IL: University of Chicago Press.
- Kothamachu, V. B., Feliu, E., Wiuf, C., Cardelli, L., and Soyer, O. S. (2013). Phosphorelays provide tunable signal processing capabilities for the cell. *PLoS Comput. Biol.* 9:e1003322. doi: 10.1371/journal.pcbi.1003322
- Kuchina, A., Espinar, L., Garcia-Ojalvo, J., and Süel, G. M. (2011). Reversible and noisy progression towards a commitment point enables adaptable and reliable cellular decision-making. *PLoS Comput. Biol.* 7:e1002273. doi: 10.1371/journal.pcbi.1002273
- Lai, Y.-C. (2014). Controlling complex, non-linear dynamical networks. *Nat. Sci. Rev.* 1, 339–341. doi: 10.1093/nsr/nwu023
- Layek, G. C. (2015). *An Introduction to Dynamical Systems and Chaos*. Springer.
- Lecca, P., Re, A., Ihekwa, A., Mura, I., and Nguyen, T.-P. (2016). *Computational Systems Biology. Inference and Modelling*. Oxford: Elsevier.
- Mao, D. (2011). *Biological Time Series Classification Via Reproducing Kernels and Sample Entropy*. Oxford: Proquest, Umi Dissertation Publishing.
- Marwan, N., Romano, M., and Thiel, M. (2016). *Recurrence Plots and Cross Recurrence Plots*. Available online at: <http://www.recurrence-plot.tk/>
- Marwan, N., Romano, M. C., Thiel, M., and Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Phys. Rep.* 438, 237–329. doi: 10.1016/j.physrep.2006.11.001
- Mera, M. E., and Morán, M. (2002). Degrees of freedom of a time series. *J. Stat. Phys.* 106, 125–145. doi: 10.1023/A:1013172129075
- Muchová, K., Lewis, R. J., Perecko, D., Brannigan, J. A., Ladds, J. C., Leech, A., et al. (2004). Dimer-induced signal propagation in Spo0A. *Mol. Microbiol.* 53, 829–842. doi: 10.1111/j.1365-2958.2004.04171.x
- Narula, J., Devi, S. N., Fujita, M., and Igoshin, O. A. (2012). Ultrasensitivity of the *Bacillus subtilis* sporulation decision. *Proc. Natl. Acad. Sci. U.S.A.* 109, E3513–E3522. doi: 10.1073/pnas.1213974109
- Ortega, J. M., and Rheinboldt, W. C. (2000). *Iterative Solution of Nonlinear Equations in Several Variables*. Philadelphia, PA: Society for Industrial and Applied Mathematics.
- Ravetti, M. G., Carpi, L. C., Goncalves, B. A., Frery, A. C., and Rosso, O. A. (2014). Distinguishing noise from chaos: objective versus subjective criteria using horizontal visibility graph. *PLoS ONE* 9:e108004. doi: 10.1371/journal.pone.0108004
- Richman, J. S., and Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.* 278, H2039–H2049.
- Rohde, G. K. (2008). Chaotic signal detection and estimation based on attractor sets: applications to secure communications. *Chaos* 18:013114. doi: 10.1063/1.2838853
- Rosenstein, M. T., Collins, J. J., and Luca, C. J. D. (1993). A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D* 65, 117–134. doi: 10.1016/0167-2789(93)90009-P
- Rosso, O. A., Larrondo, H. A., Martin, M. T., Plastino, A., and Fuentes, M. A. (2007). Distinguishing noise from chaos. *Phys. Rev. Lett.* 99:154102. doi: 10.1103/PhysRevLett.99.154102
- Satola, S., Kirchman, P. A., and Moran, C. P. (1991). Spo0A binds to a promoter used by sigma A RNA polymerase during sporulation in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* 88, 4533–4537. doi: 10.1073/pnas.88.10.4533
- Satola, S. W., Baldus, J. M., and Moran, C. P. (1992). Binding of Spo0A stimulates spoIIG promoter activity in *Bacillus subtilis*. *J. Bacteriol.* 174, 1448–1453.
- Sen, S., Garcia-Ojalvo, J., and Elowitz, M. B. (2011). Dynamical consequences of bandpass feedback loops in a bacterial phosphorelay. *PLoS ONE* 6:e25102. doi: 10.1371/journal.pone.0025102
- Sinha, S. (1997). Controlling chaos in biology. *Curr. Sci. Biol.* 75, 977–983.
- Skiadas, C. H., and Skiadas, C. (2016). *Handbook of Applications of Chaos Theory*. Boca Raton, FL: Taylor and Francis Group, CRC Press.
- Skinner, J. E. (1994). Low-dimensional chaos in biological systems. *Biotechnology* 12, 596–600. doi: 10.1038/nbt0694-596
- Sokunbi, M. O. (2014). Sample entropy reveals high discriminative power between young and elderly adults in short fMRI data sets. *Front. Neuroinformatics* 8:69. doi: 10.3389/fninf.2014.00069
- Sprott, J. C. (2003). *Chaos and Time-Series Analysis*. Oxford: Oxford University Press.
- Stragier, P., and Losick, R. (1996). Molecular genetics of sporulation in *Bacillus subtilis*. *Ann. Rev. Genet.* 30, 297–241. doi: 10.1146/annurev.genet.30.1.297
- Strauch, M. A., Wu, J. J., Jonas, R. H., and Hoch, J. A. (1993). A positive feedback loop controls transcription of the spoOF gene, a component of the sporulation phosphorelay in *Bacillus subtilis*. *Mol. Microbiol.* 7, 967–974. doi: 10.1111/j.1365-2958.1993.tb01188.x
- Tamma, A., and Khubchandani, B. L. (2016). Accurate determination of time delay and embedding dimension for state space reconstruction from a scalar time series. *Chaotic Dyn.* arxiv:1605.01571v1.
- Tél, T., and Gruiz, M. (2005). *Chaotic Dynamics: An Introduction Based on Classical Mechanics*. Cambridge: Cambridge University Press.
- Theiler, J. (1990). Estimating the fractal dimension of chaotic time series. *Lincoln Lab. J.* 3, 63–86.
- Tojo, S., Hirooka, K., and Fujita, Y. (2013). Expression of kinA and kinB of *Bacillus subtilis*, necessary for sporulation initiation, is under positive stringent transcription control. *J. Bacteriol.* 195, 1656–1665. doi: 10.1128/JB.02131-12
- Vasseur, D. A. (2015). *Biological Chaos and Complex Dynamics*. Available online at: [www.oxfordbibliographies.com/document/obo-9780199830060/obo-9780199830060-0024.xml](http://www.oxfordbibliographies.com/document/obo-9780199830060/obo-9780199830060-0024.xml)
- Vishnoi, M., Narula, J., Devi, S. N., Dao, H.-A., Igoshin, O. A., and Fujita, M. (2013). Triggering sporulation in *Bacillus subtilis* with artificial two-component systems reveals the importance of proper Spo0A activation dynamics. *Mol. Microbiol.* 90, 181–194. doi: 10.1111/mmi.12357
- Vulpiani, A. (2004). *Determinismo e caos*. Roma: Carocci.
- Wang, L., Fabret, C., Kanamaru, K., Stephenson, K., Dartois, V., Perego, M., et al. (2001). Dissection of the functional and structural domains of phosphorelay histidine kinase a of *Bacillus subtilis*. *J. Bacteriol.* 183, 2795–2802.
- Weiss, J. N., Garfinkel, A., Spano, M. L., and Ditto, W. L. (1994). Chaos and chaos control in biology. *J. Clin. Invest.* 93, 1355–1360. doi: 10.1172/JCI117111
- York, K., Kenney, T. J., Satola, S., Moran, C. P., Poth, H., and Youngman, P. (1992). Spo0A controls the sigma A-dependent activation of *Bacillus subtilis* sporulation-specific transcription unit spoIIE. *J. Bacteriol.* 174, 2648–2658.
- Zeraouia, E. (2011). *Models and Application of Chaos Theory in Modern Science*. St. Helier: Science Publisher, CRC Press.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Lecca, Mura, Re, Barker and Ihekwa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.