49th IEEE Conference on Decision and Control
December 15-17, 2010
Hilton Atlanta Hotel, Atlanta, GA, USA

Approximative Classification of Regions in Parameter Spaces of Nonlinear ODEs Yielding Different Qualitative Behavior

Jan Hasenauer, Christian Breindl, Steffen Waldherr, and Frank Allgöwer

Abstract— **Nonlinear dynamical systems can show a variety of different qualitative behaviors depending on the actual parameter values. As in many situations of practical relevance the parameter values are not precisely known it is crucial to determine the region in parameter space where the system exhibits the desired behavior.**

In this paper we propose a method to compute an approximative, analytical description of this region. Employing Markov-chain Monte-Carlo sampling, nonlinear support vector machines, and the novel notion of margin functions, an approximative classification function is determined.

The properties of the method are illustrated by studying the dynamic behavior of the Higgins-Sel'kov oscillator.

I. Introduction

Nonlinear dynamical systems can show a large variety of different types of dynamic behavior [2]. For many applications it is of interest to check whether a dynamical system has a certain property and, more interestingly, how this property depends on the system parameters. A classical example for such a property is the stability of an equilibrium point, but also more complex questions, such as whether a system state can reach a particular threshold from a given initial condition, are relevant for example in safety considerations.

While for a single given parameter vector it may be computationally inexpensive to check whether the system has a certain property or not, determining the global characteristics of the parameter space with respect to the property of interest is in general hard. However, the problem of determining this characteristic is highly relevant for analysis or design aspects, as it can give insight into to the system's robustness properties [6]. Also for identification purposes classification of parameter regions is helpful as they contain information about the relevance of certain parameters.

Efficient methods for solving such general classification problems for nonlinear systems only exist for very restricted problem setups. One example is the computation of codimension 1 bifurcation surfaces in two and three dimensional parameter spaces [11], for which continuation methods can be employed [9]. For higher dimensional parameter spaces no efficient methods are available.

In this paper we pursue the goal of constructing an approximative analytic classification function to compute the system property for new parameter vectors. This method will also be applicable in cases where no analytic rule for determining the parameter dependent property exists. Furthermore, this approximative classification function can be used to compute an approximation of the hypersurface separating regions in the parameter space which lead to qualitatively different dynamic properties. The resulting approximation can be used for a first robustness analysis or serve as starting point for computationally more demanding methods.

For computing the approximative classification function we propose to combine Markov-chain Monte-Carlo (MCMC) sampling and nonlinear support vector machines (SVM) with the concept of margin functions. Given a certain property and a parameter value, a margin function quantifies the extent to which this property is present. We will show that using MCMC sampling together with an appropriate margin function allows to obtain a high sample density close to the separating boundary. This fact facilitates the computation of a good approximative classification function using SVMs.

The paper is structured as follows. In Section II the problem of deriving a classification function is described more precisely. In Section III the definition of margin functions is given and the application of nonlinear SVM, and MCMC sampling is presented. As a specific example of a margin function the concept of loop breaking is presented in Section IV. In Section V the method is then applied to study two properties of a model of the Higgins-Sel'kov oscillator. The paper concludes with a summary in Section VI.

II. Problem Description

In the following we consider systems of nonlinear differential equations,

$$\Sigma(\theta): \quad \dot{x} = f(x,\theta), \qquad x(0) = x_0(\theta), \tag{1}$$

in which $x(t,\theta) \in \mathbb{R}^n$ is the state vector and $\theta \in \Omega \subset \mathbb{R}^q$ is the parameter vector. The set Ω of admissible parameter values is open and bounded. In order to guarantee existence and uniqueness of the solution $x(t,\theta)$ we assume that the vector field $f : \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}^n$ is locally Lipschitz. Additionally, $x_0 : \mathbb{R}^q \to \mathbb{R}^n$ is considered to be smooth.

Besides the system class, the properties of interest are restricted as well. We only consider properties for which there exists a $(q-1)$ – dimensional manifold that separates Ω into regions where the property is present and regions where it is absent. In this case, we can subdivide the parameter set Ω in three disjunct subsets,

$$\Omega = \Omega_1 \cup \Omega_0 \cup \Omega_{-1}, \tag{2}$$

where Ω_1 is an open set such that $\Sigma(\theta)$ has the property of interest for $\theta \in \Omega_1$, Ω_{-1} is an open set such that $\Sigma(\theta)$ does not have the property of interest for $\theta \in \Omega_{-1}$, and Ω_0 is a $(q-1)$ – dimensionally manifold.

In case the property of interest is local exponential stability of an equilibrium point $x_s(\theta)$, Ω_1 is the set of all parameters for which $x_s(\theta)$ is locally exponentially stable, Ω_{-1} the set of parameters for which $x_s(\theta)$ is unstable, and Ω_0 the set of parameters for which $x_s(\theta)$ is marginally stable.

III. COMPUTATION OF APPROXIMATIVE CLASSIFICATION FUNCTION

In this section it is shown how nonlinear SVM, MCMC sampling, and suitable margin and classification functions can be used to find an explicit analytic approximative classification function for the qualitative system properties.

A. Margin Function

An essential element of the approach developed in this work is the construction of a suitable margin function for the subdivision of Ω into the subsets Ω_1, Ω_{-1}, and Ω_0.

Definition 1: A continuous function $\mathrm{MF} : \mathbb{R}^q \to \mathbb{R}$ is called a *margin function* for the system property of interest, if $\mathrm{MF}(\theta) > 0$ for $\theta \in \Omega_1$, $\mathrm{MF}(\theta) < 0$ for $\theta \in \Omega_{-1}$, and $\mathrm{MF}(\theta) = 0$ for $\theta \in \Omega_0$.

Based on the margin function the classification function

$$\mathrm{CF} : \mathbb{R}^q \to \{-1, 0, 1\} : \theta \mapsto \mathrm{sgn}(\mathrm{MF}(\theta)), \qquad (3)$$

is defined, such that each parameter vector θ can be characterized by the relation $\theta \in \Omega_{\mathrm{CF}(\theta)}$.

As an example, consider the property that the norm of the trajectory starting at $x_0(\theta)$ passes a certain threshold, i.e. $\exists t \in [0, t_{\max}] : ||x(t,\theta)||_2 > \gamma$. For this problem, a possible margin function is given by: $\mathrm{MF}(\theta) = \max_{t\in[0,t_{\max}]} ||x(t,\theta)||_2 - \gamma$.

In the following, margin and classification functions are employed to sample and classify points in the parameter space.

B. Nonlinear Support Vector Machine

Let us assume that a set of S points in the parameter space, $\{\theta^1, \ldots, \theta^S\}$, and a classification function CF are given. Then a set of classified points,

$$T = \left\{ \left(\theta^1, c^1\right), \ldots, \left(\theta^S, c^S\right) \right\}, \qquad (4)$$

with class $c^i = \mathrm{CF}(\theta^i)$ can be derived. This set is called training set. Based on the generated training set T a nonlinear SVM [4] is learned. The process of learning the nonlinear SVM consists hereby of two main steps [4].

The first step is a mapping of the training set T from the input space into a feature space of higher dimension,

$$\Phi(T) = \left\{ \left(\Phi(\theta^1), c^1\right), \ldots, \left(\Phi(\theta^S), c^S\right) \right\}, \qquad (5)$$

illustrated in Figure 1. This feature space is a Hilbert space of finite or infinite dimension [4] with coordinates defined by $\Phi : \mathbb{R}^q \times \mathbb{R} \mapsto \mathbb{R}^{q^*} \times \mathbb{R}$, where q^* is the dimension of the feature space. After the transformation Φ of the data into the feature space a linear separation of the data is performed. Therefore, the optimization problem,

$$\begin{aligned} \min_{w,b,\xi} \quad & \frac{1}{2} w^{\mathrm{T}} w + C \sum_{i=1}^{S} \xi^i \\ \text{s.t.} \quad & c^i \left(w^{\mathrm{T}} \Phi(\theta^i) + b\right) \geq 1 - \xi^i, \quad i = 1, \ldots, S, \\ & \xi^i \geq 0, \qquad i = 1, \ldots, S, \end{aligned} \qquad (6)$$

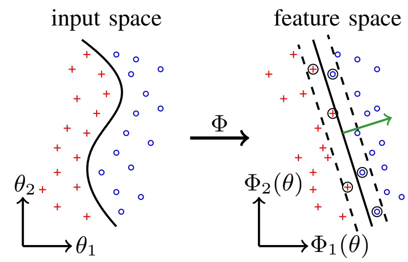

Fig. 1. Visualization of the mapping Φ from input space to feature space (class = 1: +, class = -1: o) . Left: Nonlinearly distributed data in the input space which do not allow a linear separation. Right: Samples transformed in the feature space where they can be separated linearly, and normal vector w ($\rightarrow$) of the separating hyperplane. Support vectors of the separating hyperplane are encircled (o). Only data points close to separating hyperplane influence the classification.

is solved. Thereby, w denotes the normal vector of the separating hyperplane and b the offset of the plane, as depicted in Figure 1. The first term of the objective function, $\frac{1}{2} w^{\mathrm{T}} w$, yields a maximization of the margin between the separating hyperplane and the data points, whereas the second term, $\sum_{i=1}^{S} \xi^i$, penalizes mis-classifications. The weighting of the different terms can be influenced via C. The constraints are that all data points $(\Phi(\theta^i), c^i)$ are correctly classified within a certain error ξ^i.

To solve the constraint optimization problem (6) its dual problem is derived [4],

$$\begin{aligned} \min_{\lambda} \quad & \sum_{i=1}^{S} \lambda_i - \frac{1}{2} \sum_{i=1}^{S} \sum_{j=1}^{S} \lambda_i \lambda_j c^i c^j \Phi^{\mathrm{T}}(\theta^i) \Phi(\theta^j) \\ \text{s.t.} \quad & \sum_{i=1}^{S} \lambda_i c^i = 0, \quad 0 \leq \lambda_i \leq C, \quad i = 1, \ldots, S, \end{aligned} \qquad (7)$$

in which $\lambda \in \mathbb{R}_+^S$ is the vector of Lagrange multipliers. Given the solution of (7), w and b can be determined (see [4]), and used to derive the approximative classification function

$$\begin{aligned} \mathrm{aCF}(\theta) &= \mathrm{sgn}\left(w^{\mathrm{T}} \Phi(\theta) + b\right) \\ &= \mathrm{sgn}\left(\sum_{i=1}^{S} \lambda_i c^i \Phi^{\mathrm{T}}(\theta^i) \Phi(\theta) + b \right). \end{aligned} \qquad (8)$$

Given this approximative classification function the approximations,

$$\hat{\Omega}_{\pm 1} = \left\{ \theta \in \Omega | \pm (w^{\mathrm{T}} \Phi(\theta) + b) > 0 \right\}, \qquad (9)$$

$$\hat{\Omega}_0 = \left\{ \theta \in \Omega | w^{\mathrm{T}} \Phi(\theta) + b = 0 \right\}, \qquad (10)$$

can be defined, which are analytical even if $\mathrm{CF}(\theta)$ was not an analytic function.

Remark 1: Note that from (8) one can see that only points θ^i in the training set T for which $\lambda_i \neq 0$ contribute to the classification of a new point θ^j. Points θ^i for which $\lambda_i \neq 0$ are called support vectors. These are the points which are close to the separating hypersurface [4].

From this one can conclude that points which are far away from the separating hypersurface are not of interest for the computation of the nonlinear SVM and a high sample density close to the separating hypersurface is desirable.

C. Markov-chain Monte-Carlo Sampling

As outlined in the previous section, obtaining a good approximation of the separating hypersurface requires a large portion of the sample $\{\theta^i\}$ to be close to the interface. These sample member can then function as support vectors. Sampled points far away from the interface are also needed but their number can be smaller.

Achieving a high sample density close to the separating hypersurface requires sophisticated sampling techniques. Classical Monte-Carlo sampling approaches or latin hypercube sampling would yield an equal spread of the sampled points in the whole parameter region Ω. Thus, the sample size necessary to obtain a good resolution of the interface would be very large. To achieve a sampling targeted to the separating hypersurface, we propose to employ Markov-chain Monte-Carlo sampling [8] in combination with an acceptance probability related to the margin function $\mathrm{MF}(\theta)$.

MCMC sampling approaches are two-step methods. In the first step a new parameter vector θ^{i+1} with

$$\theta^{i+1} \sim J(\theta^{i+1}|\theta^i) \quad \text{for } \theta^{i+1}, \theta^i \in \Omega \tag{11}$$

where θ^i is the current parameter vector and $J(\theta^{i+1}|\theta^i)$ a transition kernel. The kernel $J(\theta^{i+1}|\theta^i)$ is a probability density and for this work we chose a multi-variate Gaussian,

$$J(\theta^{i+1}|\theta^i) = \frac{1}{(2\pi)^{\frac{q}{2}} (\det(\Sigma))^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\Delta\theta^{\mathrm{T}}\Sigma^{-1}\,\Delta\theta\right)$$

with $\Delta\theta = \theta^{i+1} - \theta^i$ and positive definite covariance matrix $\Sigma \in \mathbb{R}^{q\times q}$. Also define the function $\mathrm{D}: \mathbb{R}^q \to \mathbb{R}$ which maps the margin function to a suitable measure of the distance to the separating hypersurface. In this work D is defined by

$$\mathrm{D}(\theta) = \begin{cases} \exp\left(-\frac{\mathrm{MF}^2(\theta)}{\sigma^2}\right) & \text{for } \theta \in \Omega \\ 0 & \text{otherwise,} \end{cases} \tag{12}$$

hence, takes its maximum on the separating hypersurface ($\mathrm{MF}(\theta) = 0$), and is zero outside Ω. The parameter σ is a design parameter.

In the second step, the proposed parameter vector θ^{i+1} is accepted with probability

$$p_{\text{acc}} = \min\left\{1, \frac{\mathrm{D}(\theta^{i+1})}{\mathrm{D}(\theta^i)}\right\}. \tag{13}$$

If the parameter θ^{i+1} is accepted i is updated. Otherwise, a new parameter vector θ^{i+1} is proposed. This procedure is repeated till the required samples size, S, is reached. The approach which is used to generate a sample from $\mathrm{D}(\theta)$ with $\theta \in \Omega$ is shown in Algorithm 1.

Employing this algorithm a series of parameter vectors θ^i and corresponding qualitative properties c^i is computed. This set can then be used as training set for the nonlinear SVM.

If the margin function MF and Ω_0 satisfy some additional conditions, Theorem 1 guarantees that a desired percentage of the MCMC sample will be contained in an ϵ neighborhood of Ω_0, defined as

$$\Omega_\epsilon := \left\{\theta \in \Omega | \exists\bar{\theta} \in \Omega_0 : ||\theta - \bar{\theta}||_2 \le \epsilon\right\}, \tag{14}$$

Algorithm 1 MCMC sampling of parameter space.

Require: distance $\mathrm{D}(\theta)$, initial point $\theta^{(1)} \in \Omega$.
Initialization of the Markov-chain with $\theta^{(1)}$.
while $i \le S$ **do**
  Given θ^i, propose $\theta^{i+1} \in \Omega$ using $J(\theta^{i+1}|\theta^i)$.
  Determine $\mathrm{D}(\theta^{i+1})$ and $c^{i+1} = \mathrm{CF}(\theta^{i+1})$.
  Generate uniform random number $r \in [0, 1]$.
  if $r \le \min\left\{1, \frac{\mathrm{D}(\theta^{i+1})}{\mathrm{D}(\theta^i)}\right\}$ **then**
    i = i + 1.
  end if
end while

(see Figure 2). Let $a(\theta) := \mathrm{sgn}(\mathrm{MF}(\theta)) \cdot \min_{\bar{\theta}\in\Omega_0} ||\theta - \bar{\theta}||_2$ be the signed distance of a point θ to the hypersurface Ω_0, then Ω_ϵ can also be written as

$$\Omega_\epsilon = \{\theta \in \Omega \mid ||a(\theta)|| < \epsilon\}. \tag{15}$$

Theorem 1: Assume that $\Omega_0 \subset \Omega$ is a regular manifold and that the margin function MF is bounded by

$$\begin{aligned} (\hat{m}a(\theta) - \mathrm{MF}(\theta))a(\theta) \ge 0, \qquad & \forall\theta \in \Omega_\epsilon \\ ||\mathrm{MF}(\theta)|| \ge \gamma, \qquad & \forall\theta \in \Omega \setminus \Omega_\epsilon, \end{aligned} \tag{16}$$

where the bounds $\hat{m}$, $\gamma > 0$ satisfy

$$\hat{m}^2 \le 3\frac{\sigma^2}{\epsilon^2}\left[1 - \frac{1-\delta}{\delta}\frac{||\Omega|| - ||\Omega_\epsilon||}{||\Omega_\epsilon||}\exp\left(-\frac{\gamma^2}{\sigma^2}\right)\right], \tag{17}$$

for some ϵ, $0 < \epsilon \ll 1$, and $\delta > 0$, with $||\Omega||$ the Lebesgue measure of Ω. Then, for $S \gg 1$ the expected number of sampled points contained in Ω_ϵ is greater than $(1 - \delta)S$.

The conditions on MF are illustrated in Figure 3.

Proof: Given a continuous function $\mathrm{D}(\theta)$ such that $\forall\theta \in \Omega:\ \mathrm{D}(\theta) > 0$, it is known [8] that for $S \gg 1$ and an appropriate transition kernel J the expected value of the probability of finding a sampled point θ^i in Ω_ϵ approaches

$$\Pr(\theta^i \in \Omega_\epsilon) = \int_{\Omega_\epsilon} \mathrm{D}(\theta)d\theta \bigg/ \int_{\Omega} \mathrm{D}(\theta)d\theta. \tag{18}$$

Thus, $\Pr(\theta^i \in \Omega_\epsilon) \ge (1 - \delta)$ is equivalent to

$$\int_{\Omega_\epsilon} \mathrm{D}(\theta)d\theta \ge \frac{1-\delta}{\delta}\int_{\Omega\setminus\Omega_\epsilon} \mathrm{D}(\theta)d\theta. \tag{19}$$

To proof that (19) holds supposed Theorem 1 is satisfied, the diffeomorphism $\Gamma : \mathcal{S} \times [-\epsilon, \epsilon] \to \Omega_\epsilon : (s, a) \mapsto \theta$ and its inverse Γ^{-1} are introduced, as depicted in Figure 2. Employing these and integrating by substitution we can write

$$\int_{\Omega_\epsilon} \mathrm{D}(\theta)d\theta = \int_{\mathcal{S}}\int_{-\epsilon}^{\epsilon} \mathrm{D}(\Gamma(s, a))\left|\det\left(\frac{\partial\Gamma}{\partial[s, a]}\right)\right| da\, ds, \tag{20}$$

with $\Gamma(S \times [-\epsilon, \epsilon]) = \Omega_\epsilon$. Note that Γ and Γ^{-1} exist as we have assumed Ω_0 to be a regular manifold. This fact further allows to choose Γ such that $\frac{\partial\Gamma}{\partial[s,a]}$ is an orthogonal matrix for small a. Therefore, with $\epsilon \ll 1$, on the integration domain we have $\left|\det\left(\frac{\partial\Gamma}{\partial[s,a]}\right)\right| = 1$, resulting in

$$\int_{\Omega_\epsilon} \mathrm{D}(\theta)d\theta = \int_{\mathcal{S}}\int_{-\epsilon}^{\epsilon} \exp\left(-\frac{\mathrm{MF}^2(\Gamma(s, a))}{\sigma^2}\right) da\, ds. \tag{21}$$

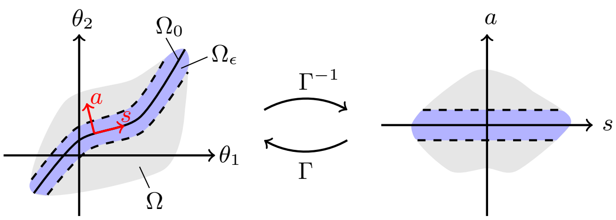

Fig. 2. Illustration of the mappings Γ and Γ^{-1} for the parameter - coordinate system (θ-plane) to the manifold - distance from manifold - coordinate system (s-a-plane). a describes the distance from Ω_0 and is positive for $\Gamma(s,a) \in \Omega_1$ and negative for $\Gamma(s,a) \in \Omega_{-1}$.

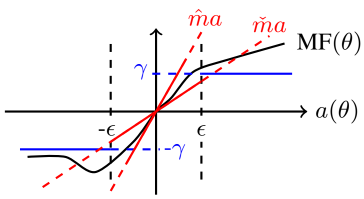

Fig. 3. Illustration of γ and the slopes $\hat{m}$ and $\check{m}$.

If the slope of MF(θ) along the a-direction is now upper bounded by $\hat{m}$ on the interval $[-\epsilon, \epsilon]$, and as $\exp(-x^2) \geq 1 - x^2 \; \forall x$, one obtains

$$\int_{\Omega_\epsilon} \mathrm{D}(\theta) d\theta \geq \int_{\mathcal{S}} \int_{-\epsilon}^{\epsilon} \exp\left(-\frac{\hat{m}^2 a^2}{\sigma^2}\right) da\, ds \tag{22}$$

$$\geq \int_{\mathcal{S}} \int_{-\epsilon}^{\epsilon} \left(1 - \frac{\hat{m}^2 a^2}{\sigma^2}\right) da\, ds \tag{23}$$

$$\geq ||\Omega_\epsilon|| \left(1 - \frac{1}{3}\frac{\hat{m}^2 \epsilon^2}{\sigma^2}\right). \tag{24}$$

Next, the right hand side of (19) is upper bounded. Therefore, the lower bound γ of the absolute value of MF(θ) in $\Omega \setminus \Omega_\epsilon$ is employed, resulting in

$$\int_{\Omega\setminus\Omega_\epsilon} \mathrm{D}(\theta) d\theta \leq \int_{\Omega\setminus\Omega_\epsilon} \exp\left(-\frac{\gamma^2}{\sigma^2}\right) d\theta \tag{25}$$

$$\leq (||\Omega|| - ||\Omega_\epsilon||) \exp\left(-\frac{\gamma^2}{\sigma^2}\right). \tag{26}$$

Plugging (24) and (26) into (19) yields (17). ∎

As it can be seen from Theorem 1, for a suitable choice of the MF it can be guaranteed that a certain percentage of the sample is contained in Ω_ϵ. Unfortunately, this does not ensure that the separating hypersurface is approximated well everywhere. Therefore, also the distribution of the sample along Ω_0 has to be considered. For this purpose, let us introduce the ϵ neighborhood of a point $\bar{\theta}^i \in \Omega_0$,

$$\Omega_\epsilon(\bar{\theta}^i) := \left\{\theta \in \Omega \,\middle|\, ||\Gamma^{-1}(\theta) - \Gamma^{-1}(\bar{\theta}^i)||_\infty \leq \epsilon\right\},$$

in transformed coordinates, where Γ is defined as before.

Theorem 2: Assume the conditions of Theorem 1 hold, then the expected number of samples in $\Omega_\epsilon(\bar{\theta}^1)$ and $\Omega_\epsilon(\bar{\theta}^2)$, $\bar{\theta}^1, \bar{\theta}^2 \in \Omega_0$, differs at most by a factor of $\beta \geq 0$, i.e.,

$$\forall \bar{\theta}^1, \bar{\theta}^2 \in \Omega_0 : \; \Pr(\theta^i \in \Omega_\epsilon(\bar{\theta}^1)) \leq (1+\beta)\Pr(\theta^i \in \Omega_\epsilon(\bar{\theta}^2)), \tag{27}$$

if

$$\hat{m}^2 - \check{m}^2 \leq \frac{\sigma^2}{\epsilon^2}\ln(1+\beta), \tag{28}$$

with $\hat{m}$ and $\check{m}$ being slope bounds of MF(θ) (see Figure 3).

Proof: To prove Theorem 2, note that from (28) it follows that

$$\forall a \in [-\epsilon, \epsilon] : \; \hat{m}^2 - \check{m}^2 \leq \frac{\sigma^2}{a^2}\ln(1+\beta). \tag{29}$$

By applying $\exp(\cdot)$ to (29) and integrating both sides over a and s we obtain

$$\begin{aligned}&\int_{\mathcal{S}_1} \int_{-\epsilon}^{\epsilon} \exp\left(-\frac{\check{m}^2}{\sigma^2}a^2\right) da\, ds \\ &\quad \leq (1+\beta) \int_{\mathcal{S}_2} \int_{-\epsilon}^{\epsilon} \exp\left(-\frac{\hat{m}^2}{\sigma^2}a^2\right) da\, ds,\end{aligned} \tag{30}$$

where $\mathcal{S}_i = \left\{s \,\middle|\, ||s - \Gamma^{-1}(\bar{\theta}^i)||_\infty \leq \epsilon\right\}$. With (12) it follows that

$$\int_{\mathcal{S}_1} \int_{-\epsilon}^{\epsilon} \mathrm{D}(\Gamma(s,a)) da\, ds \leq (1+\beta) \int_{\mathcal{S}_2} \int_{-\epsilon}^{\epsilon} \mathrm{D}(\Gamma(s,a)) da\, ds,$$

which with a change of variables and $\Pr(\theta^i \in \Omega_\epsilon(\bar{\theta}^j)) = \int_{\Omega_\epsilon(\bar{\theta}^j)} \mathrm{D}(\theta) d\theta$ finally yields (27). ∎

Theorems 1 and 2 provide conditions on the margin function MF which guarantee that the separating interface is well approximated by the MCMC sample $\{\theta^i\}$.

IV. MARGIN FUNCTION BASED ON FEEDBACK LOOP BREAKING

A suitable margin function for classifying stability and instability of steady states can be constructed via the feedback loop breaking approach proposed in [12]. Thereby, we associate with (1) a control system

$$\begin{aligned}\dot{x} &= g(x, u, \theta) \\ y &= h(x, \theta)\end{aligned} \tag{31}$$

such that $g(x, h(x,\theta), \theta) = f(x, \theta)$. Note that $x_s(\theta)$ is a steady state of (31) for the constant input $u_s(\theta) = h(x_s(\theta), \theta)$. According to the procedure outlined in [12], we compute a linear approximation of (31) around the steady state $x_s(\theta)$ with the transfer function

$$G(\theta, s) = C(\theta)(sI - A(\theta))^{-1}B(\theta), \tag{32}$$

with $A(\theta) = \frac{\partial g}{\partial x}(x_s(\theta), u_s(\theta), \theta)$, $B(\theta) = \frac{\partial g}{\partial u}(x_s(\theta), u_s(\theta), \theta)$, and $C(\theta) = \frac{\partial h}{\partial x}(x_s(\theta), u_s(\theta), \theta)$.

The realness locus of G is defined as

$$\mathcal{R}(\theta) = \{\omega \geq 0 \mid G(\theta, j\omega) \in \mathbb{R}\}. \tag{33}$$

Define $\alpha \in \mathbb{N}$ by

$$\alpha = |p_+ - p_- + z_- - z_+|, \tag{34}$$

where p_+ (p_-) is the number of poles of $G(\theta, \cdot)$ and z_+ (z_-) is the number of zeros of $G(\theta, \cdot)$ in the right (left) complex half plane. Following [12], we say that $\mathcal{R}(\theta)$ is minimal, if it has α elements in case α is odd, and $\alpha - 1$ elements otherwise.

We make the following assumptions on the open-loop system (31) and the associated transfer function G:

- $A(\theta)$ is asymptotically stable for all $\theta \in \Omega$.

- The realness locus $\mathcal{R}(\theta)$ of G is minimal for all $\theta \in \Omega$.
- The numerator and denominator polynomials of $G(\theta, s)$ are of constant degree with respect to s for all $\theta \in \Omega$.

The following result then follows immediately from Theorem 3.5 in [12].

Proposition 1: Under the above assumptions, the Jacobian $\frac{\partial f}{\partial x}(x_s(\theta), \theta)$ of (1) has eigenvalues with positive real part, if and only if

$$G(\theta, j\omega_c) > 1, \tag{35}$$

where $\omega_c = \arg\max_{\omega \in \mathcal{R}(\theta)} G(\theta, j\omega)$.

This result immediately suggests to use

$$\mathrm{MF}(\theta) = G(\theta, j\omega_c) - 1, \tag{36}$$

with $\omega_c = \arg\max_{\omega \in \mathcal{R}(\theta)} G(\theta, j\omega)$, as a margin function for the considered classification problem, with the resulting classification function

$$\mathrm{CF}(\theta) = \begin{cases} -1 & \text{for } G(\theta, j\omega_c) < 1 \\ 1 & \text{for } G(\theta, j\omega_c) > 1 \\ 0 & \text{for } G(\theta, j\omega_c) = 1. \end{cases} \tag{37}$$

Note that according to the results in [12], $G(\theta, j\omega_c) = 1$ is equivalent to $A_{cl}(\theta)$ having an eigenvalue $j\omega_c$ on the imaginary axis, while $G(\theta, j\omega_c) < 1$ is equivalent to asymptotic stability of $A_{cl}(\theta)$. Thus, the proposed classification function CF indeed structures the parameter space into regions of instability and stability, with bifurcations occurring on the boundary.

V. HIGGINS–SEL'KOV OSCILLATOR

To illustrate the properties of the proposed method the qualitative behavior of the Higgins-Sel'kov oscillator [10] is analyzed in this section. This system describes one elementary step in glycolysis [3]. The Higgins-Sel'kov oscillator system can exhibit a stable steady state or an unstable steady state coexisting with a stable limit cycle, depending on parameter values. The model we consider in this work is taken from [7] and a basal conversion rate of S to P is added yielding

$$\begin{aligned} \dot{S} &= \theta_1 - \theta_2 P^2 S - \theta_3 S \\ \dot{P} &= \theta_2 P^2 S - \theta_4 P + \theta_3 S. \end{aligned} \tag{38}$$

The parameters θ_3 and θ_4 are fixed to $\theta_3 = 0.04$ and $\theta_4 = 2$, whereas the parameters $[\theta_1, \theta_2]^T \in \Omega$, with $\Omega = (0, 5)^2$ are considered.

A. Stability Analysis and Existence of a Stable Limit Cycle

The first property we want to analyze is the stability of the unique steady state $x_s(\theta) = [S_s(\theta), P_s(\theta)]^T$, with

$$S_s(\theta) = (\theta_1 \theta_4^2)/(\theta_2 \theta_1^2 + \theta_3 \theta_4^2), \quad P_s(\theta) = \theta_1/\theta_4. \tag{39}$$

As system (38) is two dimensional, all trajectories remain bounded, and $x_s(\theta)$ is the unique steady state, according to the theorem of Poincaré-Bendixson it can be concluded that a limit cycle exists if the steady state $x_s(\theta)$ is unstable [5]. Therefore, a margin function MF(θ) is chosen employing the loop breaking method introduced in Section IV, with

$$g(x, u, \theta) = \begin{bmatrix} \theta_1 - \theta_3 S - \theta_2 S u^2 \\ \theta_3 S - \theta_4 P - \theta_2 S u^2 \end{bmatrix}, \quad h(x, \theta) = P. \tag{40}$$

For the MCMC sampling, the parameters $\Sigma = 0.15\, I_2$, with I_2 being the 2×2 identity matrix, and $\sigma^2 = 1/40$ were chosen. In total 5000 parameter vectors have been proposed, 1958 of which have been accepted and used for the calculation of the approximative classification function aCF(θ). The classification error of this training set was 0.4%.

To evaluate the proposed scheme, the obtained result is compared to the analytical solution. Therefore, determinant and trace of the Jacobian $J|_{(x_s(\theta),\theta)}$ of (38) at steady state $x_s(\theta)$ are considered. As $\det(J|_{\bar{x}})$ is always positive, the steady state $x_s(\theta)$ is stable iff $\mathrm{tr}(J|_{(x_s(\theta),\theta)})$ is negative, and unstable if it is positive. This yields the two separating hypersurfaces (bifurcation surfaces),

$$\begin{aligned} &\mathrm{SH1}: \theta_2 = ((\theta_4^6 - 8\theta_3\theta_4^5)^{\frac{1}{2}} - 2\theta_3\theta_4^2 + \theta_4^3)/(2\theta_1^2) \\ &\mathrm{SH2}: \theta_2 = -((\theta_4^6 - 8\theta_3\theta_4^5)^{\frac{1}{2}} + 2\theta_3\theta_4^2 + \theta_4^3)/(2\theta_1^2). \end{aligned} \tag{41}$$

As visible in Figure 4 (left) the margin function derived from loop breaking yields that a large percentage of the sample is close to the separating hypersurface Ω_0. Also sample densities in different areas along Ω_0 are comparable. Using these samples as training set for the nonlinear SVM [1] with Gaussian kernel functions yields an almost perfect agreement between the analytical and the approximated solution of the separating hypersurface. The quality of the approximative classification function is evaluated using 1000 uniformly distributed sample points $\theta^i \in \Omega$, which were not contained in the training set. Thereby, 99.60% of the samples were classified correctly and the computation times using the analytical and the approximative (SVM) classification functions were in the same order of magnitude.

Remark 2: Note that the proposed method is a global method as different branches of the separating hypersurface are found. This is not possible using continuation methods.

B. Amplitude of Stable Limit Cycle Oscillation

As a second example we study the amplitude of the oscillation of S on the limit cycle $\mathcal{L}$. This parameter dependent amplitude $A(\theta)$ is defined as

$$A(\theta) = \frac{1}{2}\left(\max_{\hat{t} \in \tau(\theta)} S_{\mathcal{L}}(\hat{t}, \theta) - \min_{\check{t} \in \tau(\theta)} S_{\mathcal{L}}(\check{t}, \theta)\right), \tag{42}$$

in which $S_{\mathcal{L}}(t, \theta)$ is the time course of S on the limit cycle and $\tau(\theta) = [0, t_{\mathcal{L}}(\theta)]$, where $t_{\mathcal{L}}(\theta)$ denotes the parameter dependent period of the limit cycle oscillation. For the amplitude of the limit cycle no analytic solution exists and therefore it is computed numerically. For parameter vectors θ resulting in a stable steady state we define $A(\theta) = 0$.

The property we are interested in for the remainder of this section is whether or not $A(\theta) \in [3.5, 5]$. For the MCMC sampling the margin function

$$\mathrm{MF}(\theta) = -(A(\theta) - 3.5)(A(\theta) - 5), \tag{43}$$

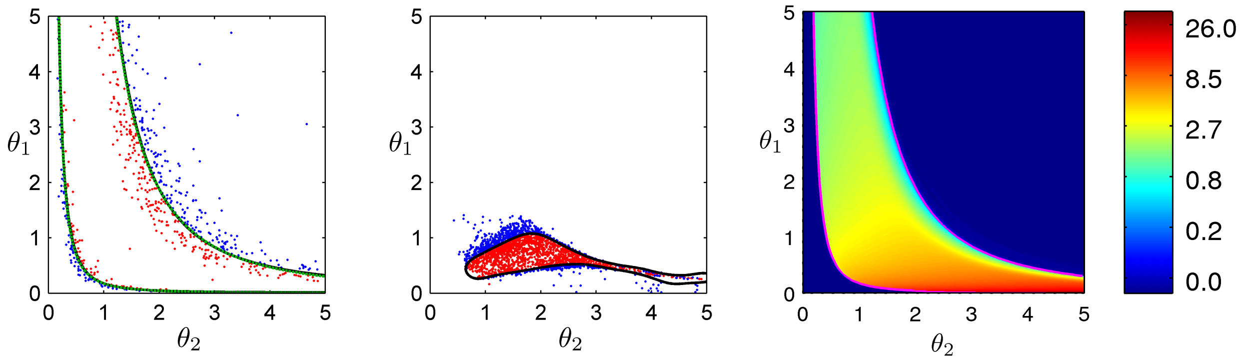

Fig. 4. Classification of parameters in the set $[\theta_1, \theta_2]^{\mathrm{T}} = (0,5)^2$ and fixed parameters $\theta_3 = 0.04$ and $\theta_4 = 2$ with respect to different properties of the Higgins-Sel'kov oscillator. Left: Sample (stable (•) and unstable (•) steady states) generated using the MCMC sampling and used for the estimation approximative classification function (—), and analytical solution of the bifurcation surfaces (– –). Middle: Sample ($A \notin [3.5, 5]$ (•) and $A \in [3.5, 5]$ (•)) generated using the MCMC sampling and used for the estimation approximative classification function (—). It can be seen that some samples members are not classified correctly and that in the lower left part blue samples are missing. Right: Illustration of amplitudes A of the limit cycle oscillation as function of θ_1 and θ_2. An in-depth analysis reveals that the margin function is not differentiable at the bifurcation manifold (Theorems 1 and 2 are not applicable) which complicates the sampling.

has been selected. This margin function is zero for $A(\theta) = 3.5$ and $A(\theta) = 5$, positive for $A(\theta) \in (3.5, 5)$, and negative otherwise. During the MCMC sampling 6000 parameter vectors were proposed, with $\Sigma = 0.15\, I_2$ and $\sigma^2 = 1/10$. Out of these 6000 samples 2449 parameter vectors were accepted and used to learn the nonlinear SVM. The classification error for this training set was 2.53%. These training points and the resulting separating hypersurface are shown in Figure 4 (middle). It can be seen that the sample concentrates in some parameter regions and are not equally distributed along the interface. The reason for this behavior can be seen in Figure 4 (right). For parameters $\theta_2 > 3$ and low values of θ_1, the norm of the gradient $\frac{\partial A}{\partial \theta}$ becomes very large and the set Ω_1 becomes very pointed. Therefore, only a small portion of the sample reachs this region and the quality of the approximation is worse in this area. A problem similar to this occurs in the region $\theta_1 < 0.5$ and $\theta_2 < 1$. Here, no parameter vectors in Ω_{-1} were drawn. The reason for this is again the large norm of the gradient of $A(\theta)$, and thus also in MF due to the disappearance of the limit cycle. Hence, the conditions of Theorems 1 and 2 to guarantee a good sampling of the interface are not fulfilled.

However, despite of these sampling problems the quality of the approximative classification function is satisfying. Out of 1000 uniformly distributed parameter vectors $\theta^i \in \Omega$ which are not contained in the training set, only six were mis-classified. Furthermore, as for this property no analytical relation exists, the computation time required to classify these 1000 samples is three orders of magnitude smaller than computing the property via simulation.

VI. CONCLUSIONS

In this work a novel method to determine a simple approximative classification function is presented. Employing this function, the classification of a point is simplified to an evaluation of an explicit analytic function. It could be shown that this function can be computed efficiently using margin function, MCMC sampling targeted to the separating hypersurface, and support vector machines.

Using the Higgins-Sel'kov oscillator the properties and benefits of the proposed approach are presented. It is shown that in cases were good margin functions are available, e.g. based on feedback loop breaking, the method yields a good approximative classification function already for small sample numbers. For cases where the margin functions are non-smooth, a larger samples are required but still a small classification error can be achieved.

ACKNOWLEDGMENTS

The authors acknowledge financial support from the German Federal Ministry of Education and Research (BMBF) within the FORSYS-Partner program (grant nr. 0315-280A), the Helmholtz Society, project "Control of Regulatory Networks with focus on noncoding RNA" (CoReNe), and the German Research Foundation within the Cluster of Excellence in Simulation Technology (EXC310/1) at the University of Stuttgart.

REFERENCES

[1] S. Canu, Y. Grandvalet, V. Guigue, and A. Rakotomamonj. SVM and kernel methods MATLAB toolbox. Perception Systèmes et Information, INSA de Rouen, Rouen, France, 2005.
[2] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, volume 42 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.
[3] J. Higgins. A chemical mechanism for oscillation of glycolytic intermediates in yeast cells. *Biochem.*, 51(6):989–994, 1964.
[4] O. Ivanciuc. *Reviews in computational chemistry*, volume 23, chapter Applications of support vector machines in chemistry, pages 291–400. Wiley-VCH, Weinheim, 2007.
[5] H.K. Khalil. *Nonlinear Systems*. Prentice Hall, 2002.
[6] J. Kim, D.G. Bates, I. Postlethwaite, L. Ma, and P.A. Iglesias. Robustness analysis of biochemical network models. *IEE Proc. of the Syst. Biol.*, 153(3):96—104, 2006.
[7] E. Klipp, R. Herwig, A. Kowald, C. Wierling, and H. Lehrach. *Systems Biology in Practice*. Wiley-VCH, Weinheim, 2005.
[8] D.J.C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2005.
[9] S.L. Richter and R.A. DeCarlo. Continuation methods: Theory and applications. *IEEE Trans. Circuits Syst.*, 30(6):347—352, 1983.
[10] E.E. Sel'kov. Self-oscillations in glycolysis. 1. a simple kinetic model. *Eur. J. Biochem.*, 4(1):79–86, 1968.
[11] D. Stiefs, T. Gross, R. Steuer, and U. Feudel. Computation and visualization of bifurcation surfaces. *Int. J. Bifurcation Chaos*, 18(8):2191–2206, 2008.
[12] S. Waldherr and F. Allgöwer. Searching bifurcations in high-dimensional parameter space via a feedback loop breaking approach. *Int. J. Syst. Sci.*, 70(4):769–782, 2009.