End-to-End Statistical Model Checking for Parameterization and Stability Analysis of ODE Models*

DAVID JULIEN, Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, France
GILLES ARDOUREL, Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, France
GUILLAUME CANTIN, Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, France
BENOÎT DELAHAYE, Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, France

ACM Reference Format:
David JULIEN, Gilles ARDOUREL, Guillaume CANTIN, and Benoît DELAHAYE. 2024. End-to-End Statistical Model Checking for Parameterization and Stability Analysis of ODE Models. *ACM Trans. Model. Comput. Simul.* 1, 1, Article 1 (January 2024), 4 pages. https://doi.org/10.1145/3649438

Note: These appendices are mostly a rewriting of Appendix A and B in the context of initial conditions. The arguments provided to prove Lemma 4.3 and Theorem 4.4 are very similar to those provided to prove Lemma 3.3 and Theorem 3.4, respectively. We will highlight the differences, which are mostly constants, in blue.

1 PROOF OF LEMMA 4.3

Again, we recall the definition of *stability* of an approximation method:

Definition 1.1 (Method stability). We say that the approximation method determined by Equation (8) is stable if there exists a constant $\mathcal{K} > 0$, called *stability constant*, such that, for any two sequences $(y_k)_{0\leq k\leq J}$ and $(\widetilde{y}_k)_{0\leq k\leq J}$ defined as $y_{k+1} = y_k + h\,\Phi(\tau_k, y_k, h)$ and $\widetilde{y}_{k+1} = \widetilde{y}_k + h\,\Phi(\tau_k, \widetilde{y}_k, h) + \eta_k$ respectively, $(0 \leq k < J)$, with $\eta_k \in \mathbb{R}$, we have

$$\max_{0\leq k\leq J} \|y_k - \widetilde{y}_k\|_{\mathbb{R}^n} \leq \mathcal{K}\big(\, \|y_0 - \widetilde{y}_0\|_{\mathbb{R}^n} + \sum_{0\leq k\leq J} |\eta_k| \,\big). \tag{1}$$

Note that this stability is relative to the method: it does not imply stability of the studied system itself.

It is well-known that if Φ is κ-Lipschitz w.r.t. y, that is if $\forall t \in [0,T]$, $\forall y_1, y_2 \in \mathbb{R}$, and $\forall h \in \mathbb{R}$, $\|\Phi(t, y_1, h) - \Phi(t, y_2, h)\|_{\mathbb{R}^n} \leq \kappa \|y_1 - y_2\|_{\mathbb{R}^n}$, then stability is ensured (see for instance [1] or [2]).

Now, we fix $x_0, \widetilde{x}_0 \in \mathcal{B}_{x_0^*}$, and we consider the approximate solutions $(y_k^{x_0})_{0\leq k\leq J}$, $(y_k^{\widetilde{x}_0})_{0\leq k\leq J}$ to Equation (22) relative to x_0 and $\widetilde{x}_0$

$$\begin{cases} y_0^{x_0} &= x_0, \\ y_{k+1}^{x_0} &= y_k + h\,\Phi(t_k, y_k, h) \end{cases}, \qquad \begin{cases} y_0^{\widetilde{x}_0} &= \widetilde{x}_0, \\ y_{k+1}^{\widetilde{x}_0} &= y_k^{\widetilde{x}_0} + h\,\Phi(t_k, y_k^{\widetilde{x}_0}, h). \end{cases}$$

*This work is partially supported by ANR BisoUS (ANR-22-CE48-0012) and research project VERHYDYN (Verification of Hybrid Dynamical Systems) funded by INS2I-CNRS.

Authors' addresses: David JULIEN, david.julien@univ-nantes.fr, Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, Nantes, France, F-44000; Gilles ARDOUREL, Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, Nantes, France, F-44000; Guillaume CANTIN, Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, Nantes, France, F-44000; Benoît DELAHAYE, Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, Nantes, France, F-44000.

We recall that the exact solutions to Equation (22) relative to x_0 and $\widetilde{x}_0$ are denoted x^{x_0} and $x^{\widetilde{x}_0}$ respectively. For $x_0 \in \mathcal{B}_{x_0^*}$ and $0 \leq k \leq J$, we adapt the consistency error defined in Equation (36) to the initial condition:

$$\varepsilon_{h,k}(x_0) = \|x^{x_0}(\tau_k) - y^{x_0}(\tau_k)\|_{\mathbb{R}^n} \tag{2}$$

The consistency errors satisfy $\varepsilon_h(x_0) = \max_{0\leq k\leq J} \varepsilon_{h,k}(x_0)$, where $\varepsilon_h(x_0)$ is the global approximation error (defined in Equation (26)). The proof of Lemma 4.3 can be derived from the following theorem:

Theorem 1.2. *Assume that the function Φ defined in Equation (8) is κ_1-Lipschitz w.r.t. y. Then, the approximation method is stable w.r.t. the consistency error, i.e. there exists $\mathcal{K} > 0$ such that*

$$\forall x_0, \widetilde{x}_0 \in \mathcal{B}_{x_0^*}, \max_{O\leq k\leq J} \left|\varepsilon_{h,k}(x_0) - \varepsilon_{h,k}(\widetilde{x}_0)\right| \leq \mathcal{K} \|x_0 - \widetilde{x}_0\|_{\mathbb{R}^n}.$$

Proof. By assumption, Φ is κ_1-Lipschitz w.r.t. y, so we have, $\forall 0 \leq k \leq J$,

$$\left\|\Phi(t_k, y_k^{x_0}, h) - \Phi(t_k, y_k^{\widetilde{x}_0}, h)\right\|_{\mathbb{R}^n} \leq \kappa_1 \left\|y_k^{x_0} - y_k^{\widetilde{x}_0}\right\|_{\mathbb{R}^n}$$

It follows that

$$\begin{aligned}
\left\|y_{k+1}^{x_0} - y_{k+1}^{\widetilde{x}_0}\right\|_{\mathbb{R}^n} &\leq \left\|y_k^{x_0} - y_k^{\widetilde{x}_0}\right\|_{\mathbb{R}^n} + h\left\|\Phi(t_k, y_k^{x_0}, h) - \Phi(t_k, y_k^{\widetilde{x}_0}, h)\right\|_{\mathbb{R}^n} \\
&\leq \left\|y_k^{x_0} - y_k^{\widetilde{x}_0}\right\|_{\mathbb{R}^n} + h\kappa_1 \left\|y_k^{x_0} - y_k^{\widetilde{x}_0}\right\|_{\mathbb{R}^n} \\
&\leq (1 + h\kappa_1) \left\|y_k^{x_0} - y_k^{\widetilde{x}_0}\right\|_{\mathbb{R}^n} \\
&\leq (1 + h\kappa_1)^{k+1} \left\|y_0^{x_0} - y_0^{\widetilde{x}_0}\right\|_{\mathbb{R}^n}, \quad \text{by immediate recursion}
\end{aligned}$$

This leads to

$$\max_{0\leq k\leq J} \left\|y_k^{x_0} - y_k^{\widetilde{x}_0}\right\|_{\mathbb{R}^n} \leq (1 + h\kappa_1)^J \left\|y_0^{x_0} - y_0^{\widetilde{x}_0}\right\|_{\mathbb{R}^n} \tag{3}$$

Further, it is proved in [2] that if Φ is Lipschitz continuous w.r.t. the initial condition x_0, then the exact solution x^{x_0} is also Lipschitz continuous w.r.t. x_0, that is, there exists $\kappa_2 > 0$ such that

$$\forall x_0, \widetilde{x}_0 \in \mathcal{B}_{x_0^*}, \forall T_1 \leq t \leq T_2, \left\|x^{x_0}(t) - x^{\widetilde{x}_0}(t)\right\|_{\mathbb{R}^n} \leq \kappa_2 \|x_0 - \widetilde{x}_0\|_{\mathbb{R}^n} \tag{4}$$

Thus, we have

$$\begin{aligned}
\left|\varepsilon_{h,k}(x_0) - \varepsilon_{h,k}(\widetilde{x}_0)\right| &\leq \left\|x^{x_0}(\tau_k) - y^{x_0}(\tau_k) - x^{\widetilde{x}_0}(\tau_k) - y^{\widetilde{x}_0}(\tau_k)\right\|_{\mathbb{R}^n} \\
&\leq \left\|x^{x_0}(\tau_k) - x^{\widetilde{x}_0}(\tau_k)\right\|_{\mathbb{R}^n} + \left\|y^{x_0}(\tau_k) - y^{\widetilde{x}_0}(\tau_k)\right\|_{\mathbb{R}^n} \\
&\leq \mathcal{K}\|x_0 - \widetilde{x}_0\|_{\mathbb{R}^n},
\end{aligned}$$

where $\mathcal{K} = (1 + \kappa_1)^J + \kappa_2$, which completes the proof for Equation (1.2). □

It remains to show that Theorem 1.2 implies Lemma 4.3.

Proof of Lemma 4.3. Let $\varepsilon > 0$, then $(h_i)_{i\geq 0}$ be a sequence of discretization steps such that $\lim_{i\to\infty} = 0$. Since the approximation method given by Equation (8) is assumed to be convergent, each function $\varepsilon_{h_i}(\cdot)$ defined in Equation (26) is pointwise convergent to 0. Furthermore, we recall that Φ is Lipschitz continuous w.r.t. the initial condition $x_0 \in \mathcal{B}_{x_0^*}$. Hence, Theorem 1.2 implies that the functions $\left(\varepsilon_{h_i}(\cdot)\right)_{i\geq 0}$ are also Lipschitz continuous, with uniform Lipschitz constant $\mathcal{K}$:

$$\left|\varepsilon_{h_i}(x_0) - \varepsilon_{h_i}(\widetilde{x}_0)\right| \leq \mathcal{K}\|x_0 - \widetilde{x}_0\|_{\mathbb{R}^n}, \quad \forall x_0, \widetilde{x}_0 \in \mathcal{B}_{x_0^*}, \quad \forall i \in \mathbb{N}.$$

Consequently, the function $\left(\varepsilon_{h_i}(\cdot)\right)_{i\geq 0}$ are uniformly equicontinuous. Then, Arzelà-Ascoli Theorem [3] implies that the sequence $\left(\varepsilon_{h_i}(\cdot)\right)_{i\geq 0}$ converges uniformly to 0 on $\mathcal{B}_{x_0^*}$: there exists $i^* \in \mathbb{N}$ such that

$$\varepsilon_{h_i}(x_0) < \varepsilon, \forall i \geq i^*, \forall x_0 \in \mathcal{B}_{x_0^*}.$$

and Lemma 4.3 is proved. □

2 PROOF OF THEOREM 4.4

First step. We begin the proof of Theorem 4.4 by showing how to compute an estimator $\hat{p}_-^\varepsilon$ of the probability p_-^ε defined in Equation (28).

Let $(x_{0,i})_{i\in\mathbb{N}}$ be a sequence of values in the ball $\mathcal{B}_{x_0^*}$. We write B_i the random variable corresponding to the test "$\Gamma_-^\varepsilon(x_{0,i})$ holds": all the B_i are i.i.d. variables and follow a Bernoulli's law of parameter p_-^ε. We write b_i the evaluation of B_i. We introduce the transfer function $g_- : \mathcal{B}_{x_0^*} \to \{0,1\}$ corresponding to the test regarding $\Gamma_-^\varepsilon(x_{0,i})$, defined by $g_-(x_{0,i}) = 1$ if $\Gamma_-^\varepsilon(x_{0,i})$ holds, 0 otherwise. Next, we consider

$$G = \mathbb{E}(g_-(X)) = \int_{\mathcal{B}_{x_0^*}} g_-(s) f_X(s) \mathrm{d}s, \tag{5}$$

where f_X is defined by a uniform distribution, that is, $f_X(s) = \frac{1}{|\mathcal{B}_{x_0^*}|}$, $s \in \mathcal{B}_{x_0^*}$. We produce a sample $(x_{0,1}, x_{0,2}, \ldots, x_{0,N})$ of the variable X in $\mathcal{B}_{x_0^*}$, and use it to compute the Monte-Carlo estimator G. By virtue of the Law of Large Numbers, the sample mean satisfies: $\overline{g}_N = \frac{1}{N}\sum_{i=1}^N g_-(s_{0,i})$. The Central Limit Theorem states that the variable $Z = \frac{\overline{g}_N - G}{\sigma_{\overline{g}_N}}$ approximately follows a Standard Normal Distribution $\mathcal{N}(0,1)$; hence, for a risk θ, we can bound the error $|\alpha_N|$ of swapping G with $\overline{g}_N$ by building confidence intervals:

$$\mathbb{P}\left(|\alpha_N| \leq \chi_{1-\frac{\theta}{2}} \frac{\sigma_{g_-}}{\sqrt{N}}\right) = 1 - \theta, \tag{6}$$

where $\chi_{1-\frac{\theta}{2}}$ is the quantile of the Standard Normal Distribution $\mathcal{N}(0,1)$ and σ_{g_-} is the variance of g_-.

Since we are interested in finding p_-^ε with a certain confidence, we can perform this process after setting the desired target error α and risk θ, knowing how many simulations must be ran using Hoeffding's inequality [4]:

$$\theta = \mathbb{P}(\overline{g}_N \notin [p_-^\varepsilon - \alpha, p_-^\varepsilon + \alpha]) \leq 2\exp(-2\alpha^2 N),$$

or equivalently $N \geq \frac{\log(2/\theta)}{2\alpha^2}$. Here, it is worth emphasizing that N can be chosen independently of ε.

Further, the variance of $\overline{g}_N$ can be expressed with the variance of $g_-(X)$:

$$\sigma_{g_-}^2 = \mathbb{E}\left([g_-(X) - \mathbb{E}(g_-(X))]^2\right) = \int_{\mathcal{B}_{x_0^*}} (g_-(s))^2 f_X(s)\mathrm{d}s - G^2.$$

We consider i.i.d. samples, hence $\sigma_{g_-}^2$ can be estimated with the variance $S_{g_-}^2$:

$$\sigma_{g_-}^2 \simeq S_{g_-}^2 = \frac{1}{N}\sum_{i=1}^N (g_-(x_{0,i})^2 - \overline{g}_N^2).$$

It follows that σ_{g_-} can be estimated with its empirical counterpart $\hat{\sigma}_{g_-} = \sqrt{S_{g_-}^2}$, which shows that the error displays a $1/\sqrt{N}$ convergence.

Finally, after estimating σ_{g_-}, we can find $\hat{p}^\varepsilon_-$ using the variance of Bernoulli's law $\hat{\sigma}^2_{g_-} = \hat{p}^\varepsilon_- \times (1 - \hat{p}^\varepsilon_-)$. We conclude that the probability that $\Gamma^\varepsilon_-(\boldsymbol{\lambda})$ holds is estimated by $\hat{p}^\varepsilon_- = \frac{1}{2}\left(1 \pm \sqrt{1 - 4\hat{\sigma}^2_{g_-}}\right)$, with an error α and a risk θ, provided we perform $N \geq \frac{\log(2/\theta)}{2\alpha^2}$ simulations. It follows that

$$\mathbb{P}\big(p^\varepsilon_- \in [\hat{p}^\varepsilon_- - \alpha, \hat{p}^\varepsilon_- + \alpha]\big) \geq 1 - \theta. \tag{7}$$

Similarly, we determine an estimator $\hat{p}^\varepsilon_+$ of p^ε_+ by running $N \geq \frac{\log(2/\theta)}{2\alpha^2}$ additional simulations, and obtain a confidence interval satisfying

$$\mathbb{P}\big(p^\varepsilon_+ \in [\hat{p}^\varepsilon_+ - \alpha, \hat{p}^\varepsilon_+ + \alpha]\big) \geq 1 - \theta. \tag{8}$$

Second step. Now, let us show how a confidence interval for the probability p can be derived from the confidence intervals given in (7), (8), involving the estimators $\hat{p}^\varepsilon_-$ and $\hat{p}^\varepsilon_+$ respectively. The independence of the samples used to determine the estimators $\hat{p}^\varepsilon_-$, $\hat{p}^\varepsilon_+$ guarantees that

$$\mathbb{P}(p \in [\hat{p}^\varepsilon_- - \alpha, \hat{p}^\varepsilon_+ + \alpha]) = \mathbb{P}\big(\{p \geq \hat{p}^\varepsilon_- - \alpha\}\big) \times \mathbb{P}\big(\{p \leq \hat{p}^\varepsilon_+ + \alpha\}\big).$$

By virtue of (7), we have $\mathbb{P}(p^\varepsilon_- \geq \hat{p}^\varepsilon_- - \alpha) \geq 1 - \theta$. Next, the estimate (30) implies $\mathbb{P}(p \geq \hat{p}^\varepsilon_- - \alpha) \geq \mathbb{P}(p^\varepsilon_- \geq \hat{p}^\varepsilon_- - \alpha) \geq 1 - \theta$. Similarly, we have $\mathbb{P}(p \leq \hat{p}^\varepsilon_+ + \alpha) \geq 1 - \theta$, and finally $\mathbb{P}(p \in [\hat{p}^\varepsilon_- - \alpha, \hat{p}^\varepsilon_+ + \alpha]) \geq (1 - \theta)^2 = 1 - \xi$, since $\theta = 1 - \sqrt{1 - \xi}$.

Third step. Finally, let us prove how Lemma 4.3 guarantees that proper values of h and ε can be found, in order to control the distance between $\hat{p}_-$ and $\hat{p}_+$.

Indeed, the continuity of the probability measure $\mathbb{P}$ ensures that there exists $\varepsilon_0 > 0$ such that $\left|p^\varepsilon_- - p^\varepsilon_+\right| \leq \alpha$, for $\varepsilon < \varepsilon_0$. Next, we write

$$\left|\hat{p}^\varepsilon_- - \hat{p}^\varepsilon_+\right| \leq \left|\hat{p}^\varepsilon_- - p^\varepsilon_-\right| + \left|\hat{p}^\varepsilon_+ - p^\varepsilon_+\right| + \left|p^\varepsilon_- - p^\varepsilon_+\right|,$$

hence we have, for $\varepsilon < \varepsilon_0$:

$$\begin{aligned}\mathbb{P}\big(\left|\hat{p}^\varepsilon_- - \hat{p}^\varepsilon_+\right| \leq 3\alpha\big) &\geq \mathbb{P}(\left|\hat{p}^\varepsilon_- - p^\varepsilon_-\right| \leq \alpha) \times \mathbb{P}(\left|\hat{p}^\varepsilon_+ - p^\varepsilon_+\right| \leq \alpha) \times \mathbb{P}(\left|p^\varepsilon_- - p^\varepsilon_+\right| \leq \alpha) \\ &\geq (1 - \theta)^2 \times 1 = 1 - \xi.\end{aligned}$$

In parallel, Lemma 4.4 guarantees that for h sufficiently small, the global stability error can be uniformly bounded on $\mathcal{B}_{x_0^*}$ by ε_0. The proof is complete. □

REFERENCES

[1] Butcher, J. C. *Numerical methods for ordinary differential equations.* John Wiley & Sons, 2016. Third Edition.
[2] Crouzeix, M., and Mignot, A. L. *Analyse numérique des équations différentielles*, vol. 1. Masson, 1984.
[3] Dunford, N., and Schwartz, J. T. *Linear operators. 1. General theory.* A Wiley Interscience Publication. Interscience Publishers, 1967.
[4] Hoeffding, W. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association 58*, 301 (1963), 13–30.