CHOPtrey: contextual online polynomial extrapolation for enhanced multi-core co-simulation of complex systems*

Abir Ben Khaled-El Feki[1], Laurent Duval[1,2], Cyril Faure[3], Daniel Simon[4], and Mongi Ben Gaid[1]

[1]IFP Energies nouvelles, 1 et 4 avenue de Bois-Préau, 92852 Rueil-Malmaison, France
[2]University Paris-Est, LIGM, ESIEE Paris, 93162 Noisy-le-Grand, France
[3]CEA List, Nano-INNOV, 8 avenue de la Vauve, 91120 Palaiseau, France
[4]INRIA and LIRMM - CAMIN team, 860 Rue Saint Priest, 34095 Montpellier Cedex 5, France

March 2, 2017

Abstract

The growing complexity of Cyber-Physical Systems (CPS), together with increasingly available parallelism provided by multi-core chips, fosters the parallelization of simulation. Simulation speed-ups are expected from co-simulation and parallelization based on model splitting into weak-coupled sub-models, as for instance in the framework of Functional Mockup Interface (FMI). However, slackened synchronization between sub-models and their associated solvers running in parallel introduces integration errors, which must be kept inside acceptable bounds.

CHOPtrey denotes a forecasting framework enhancing the performance of complex system co-simulation, with a trivalent articulation. First, we consider the framework of a Computationally Hasty Online Prediction system (CHOPred). It allows to improve the trade-off between integration speed-ups, needing large communication steps, and simulation precision, needing frequent updates for model inputs. Second, smoothed adaptive forward prediction improves co-simulation accuracy. It is obtained by past-weighted extrapolation based on Causal Hopping Oblivious Polynomials (CHOPoly). And third, signal behavior is segmented to handle the discontinuities of the exchanged signals: the segmentation is performed in a Contextual & Hierarchical Ontology of Patterns (CHOPatt).

Implementation strategies and simulation results demonstrate the framework ability to adaptively relax data communication constraints beyond synchronization points which sensibly accelerate simulation. The CHOPtrey framework extends the range of applications of standard Lagrange-type methods, often deemed unstable. The embedding of predictions in lag-dependent smoothing and discontinuity handling demonstrates its practical efficiency.

Keywords: parallel simulation; Functional Mockup Interface; smoothed online prediction; causal polynomial extrapolation; context-based decision; internal combustion engine.

*Published in Simulation: Transactions of the Society for Modeling and Simulation International, 2017, http://dx.doi.org/10.1177/0037549716684026

1 Introduction

Intricacy in engineered systems increases simulator complexity. However, most existing simulation softwares are currently unable to exploit multi-core platforms, as they rely on sequential Ordinary Differential Equations (ODE) and Differential Algebraic Equations (DAE) solvers. Co-simulation approaches can provide significant improvements by allowing to jointly simulate models coming from different areas, and to validate both individual behaviors and their interactions [1]. Different simulators may be exported from original authoring tools, for instance as Functional Mock-up Units (FMUs), and then imported in a co-simulation environment. Hence, they cooperate at run-time, thanks to Functional Mockup Interface (FMI) definitions [2] for their interfaces, and to the master algorithms of these environments.

Co-simulation has shown an important potential for parallelization (see [3] for a review). Meanwhile, synchronization between the different sub-models is required due to their mutual dependencies. It avoids the propagation of numerical errors in simulation results and guarantees their correctness. Unfortunately, synchronization constraints also lead some processors into waiting periods and idle time. This decreases the potential efficiency of the threaded parallelism existing in multi-core platforms.

To overcome this limitation and exploit the available parallelism more efficiently, the dependencies constraints between parallel sub-models should be relaxed as far as possible while preserving an acceptable accuracy of the simulation results. This can be performed by a well-grounded system decomposition, tailored to data dependency reduction between the sub-models. For instance, the method proposed in [4] for distributed simulation uses transmission line modeling (TLM) based on bilateral delay lines: decoupling points are chosen when variables change slowly and the time-step of the solver is relatively small.

Unfortunately, perfect decoupling cannot often be reached and data dependencies still exist between the different blocks. Thus, tight synchronization is required between blocks using small communication steps. This greatly limits the possibilities to accelerate the simulation and eventually reach the real-time.

We propose in this work a Computationally Hasty Online Prediction framework (CHOPred) to stretch out synchronization steps with negligible precision changes in the simulation, at low-complexity. It is based on a Contextual & Hierarchical Ontology of Patterns (CHOPatt) that selects appropriate Causal Hopping Oblivious Polynomials (CHOPoly) for signal forecasting, allowing data exchange between sub-models beyond synchronization points.

This paper is organized as follows. We first review challenges as well as related work and summarize contributions in Section 2. Section 3 presents a formal model of a hybrid dynamical system and the motivations for twined parallel simulation and extrapolation. The background on prediction and the proposed Causal Hopping Oblivious Polynomials (CHOPoly) are developed in Section 4, with details in A. Then, the principles of the Contextual & Hierarchical Ontology of Patterns (CHOPatt) for the management of hybrid models are exposed in Section 5. Finally, the methodology's performance is assessed in Section 7 using an internal combustion engine model described in Section 6.

2 Related work and contributions

The continuous systems usually attain high integration speeds with variable-step solvers. The major challenge for hybrid systems resides in their numerous discontinuities that prevent similar performance. Since we are especially interested in modular co-simulation [3], it is shown in [5] that integrating each sub-system with its own solver allows to avoid interrupts coming from unrelated events. Moreover event detection and location inside a sub-system can be processed faster because they involve a smaller set of variables. However, partitioning the original complex model into several lesser complex models may add virtual algebraic loops, therefore involving delayed outputs, even with an efficient execution order.

To take advantage of the model splitting without adding useless delays, we propose in [6] a new co-simulation method based on a refined scheduling approach. This technique, denoted "RCosim", retains the speed-up advantage of modular co-simulation thanks to the parallel execution of the system's components. Besides, it improves the accuracy of the simulation results through an offline scheduling of operations that takes care of model input/output dynamics.

However, in practical applications, current co-simulation set-ups use a constant communication grid shared by all the models. In fact, the size of communication steps has a direct impact on simulation errors, and effective communication step control should rely on online estimations of the errors induced by slackened exchange rates. Schierz *et al.* [7] propose the use of adaptive communication step sizes to better handle the various changes in model dynamics. Meanwhile, the stability of multi-rate simulators with adaptive steps needs to be carefully assessed, for example based on errors propagation inside modular co-simulations [8].

Data extrapolation over steps is also expected to enhance the simulation precision over large communication steps. Nevertheless, actual complex systems (CPS) usually present non-linearities and discontinuities, making hard to predict their future behavior from past observations only. Moreover, the considered models are generated using the Simulink Coder target or the FMI for Model Exchange framework, which does not provide input derivatives (in contrast with the FMI for Co-Simulation architecture). Hence polynomial prediction cannot always correctly extrapolate. For example, [9] uses a constant, linear or quadratic extrapolation and a linear interpolation to improve the accuracy of the modular time integration. This method is successful for non-stiff systems but fails in the stiff case.

In the purpose of defining a method for the parallel simulation of hybrid dynamical systems, our previous work on a single-level context-based extrapolation [10] accounts for steps, stiffness, discontinuities or weird behaviors. It uses adapted extrapolation to limit excessively wrong prediction. It shows that properly-chosen context-based extrapolation, combined with model splitting and parallel integration, can potentially improve the speed vs. precision trade-off needed to reach real-time simulation.

In [10], context-based extrapolation is exclusively intended for FMU models and extrapolation is performed on integration steps only. This paper improves upon the preceding results with a better-grounded methodology, by:

- adding oblivion to past samples through a weighting factor in polynomial prediction;
- increasing the computational efficiency via a novel matrix formulation;
- adding an online error evaluation at each communication step to select the best context and weighting factor at the forthcoming step, further minimizing extrapolation errors;
- adding a hierarchy of decisional and functional contexts to improve prediction strategy.

3 Problem formalization and motivation

3.1 Model definition

Complex physical systems are generally modeled by hybrid non-linear ODEs or DAEs. The hybrid behavior is due to the discontinuities, raised by events triggered off by the crossing of a given threshold (zero-crossing). It plays a key role in the simulation complexity and speed. Indeed, more events slow down numerical integration.

Let us provide a formal model, considering a hybrid dynamic system Σ whose continuous state evolution is governed by a set of non-linear differential equations:

$$\begin{aligned} \dot{\boldsymbol{X}} &= \boldsymbol{f}(t, \boldsymbol{X}, \boldsymbol{D}, \boldsymbol{U}_{\text{ext}}) \quad \text{for} \quad t_n \leq t < t_{n+1}\,, \\ \boldsymbol{Y}_{\text{ext}} &= \boldsymbol{g}(t, \boldsymbol{X}, \boldsymbol{D}, \boldsymbol{U}_{\text{ext}})\,, \end{aligned}$$

where $\boldsymbol{X} \in \mathbb{R}^{n_X}$ is the continuous state vector, $\boldsymbol{D} \in \mathbb{R}^{n_D}$ is the discrete state vector, $\boldsymbol{U}_{\text{ext}} \in \mathbb{R}^{n_{U_{\text{ext}}}}$ is the external input vector, $\boldsymbol{Y}_{\text{ext}} \in \mathbb{R}^{n_Y}$ is the external output vector and $t \in \mathbb{R}^+$ denotes the time. The sequence $(t_n)_{n\geq 0}$ of strictly increasing time instants represents discontinuity points called *state events*, which are the roots of the equation:

$$\boldsymbol{h}(t, \boldsymbol{X}, \boldsymbol{D}, \boldsymbol{U}_{\text{ext}}) = 0\,.$$

The function $\boldsymbol{h}$ is usually called *zero-crossing* function or *event indicator*, used for *event detection* and *location* [11]. At each time instant t_n, a new continuous state vector can be computed as a result of the *event handling*:

$$\boldsymbol{X}(t_n) = \boldsymbol{I}(t_n, \boldsymbol{X}, \boldsymbol{D}, \boldsymbol{U}_{\text{ext}})\,,$$

and a new discrete state vector can be computed as a result of a discrete state update:

$$\boldsymbol{D}(t_n) = \boldsymbol{J}(t_{n-1}, \boldsymbol{X}, \boldsymbol{D}, \boldsymbol{U}_{\text{ext}})\,.$$

If no discontinuity affects a component of $\boldsymbol{X}(t_n)$, the right limit of this component will be equal to its value at t_n. This hybrid system model is adopted by several modeling and simulation environments and is underlying the FMI specification [2].

We assume that Σ is well-posed, in the sense that a unique solution exists for each admissible initial conditions $\boldsymbol{X}(t_0)$ and $\boldsymbol{D}(t_0)$ and that consequently $\boldsymbol{X}$, $\boldsymbol{D}$, $\boldsymbol{U}_{\text{ext}}$, and $\boldsymbol{Y}_{\text{ext}}$ are piece-wise continuous functions, i.e. continuous on each sub-interval $]t_n, t_{n+1}[$.

3.2 Model parallelization with modular co-simulation

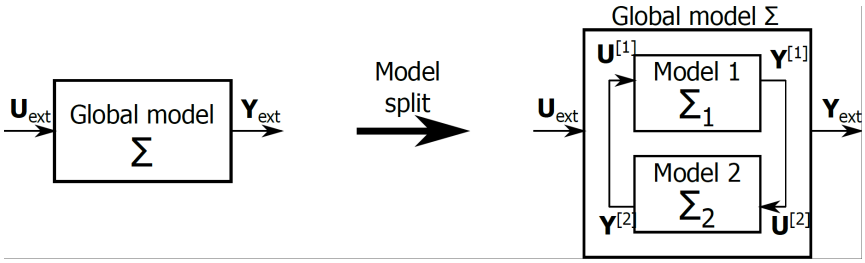


Figure 1: System splitting for parallelization.

To execute the system in parallel, the model must be split into several sub-models. For simplicity, assume that the system is decomposed into two separate blocks denoted Model 1 and Model 2, in Figure 1 with $\boldsymbol{X} = [\boldsymbol{X}^{[1]}\ \boldsymbol{X}^{[2]}]^T$ and $\boldsymbol{D} = [\boldsymbol{D}^{[1]}\ \boldsymbol{D}^{[2]}]^T$, where ^{T} denotes the matrix transpose. Therefore, the

sub-systems can be written as:

$$\begin{cases} \dot{\boldsymbol{X}}^{[1]} = \boldsymbol{f}^{[1]}(t, \boldsymbol{X}^{[1]}, \boldsymbol{D}^{[1]}, \boldsymbol{U}^{[1]}, \boldsymbol{U}_{\text{ext}}), \\ \boldsymbol{Y}^{[1]} = \boldsymbol{g}^{[1]}(t, \boldsymbol{X}^{[1]}, \boldsymbol{D}^{[1]}, \boldsymbol{U}^{[1]}, \boldsymbol{U}_{\text{ext}}), \end{cases}$$

$$\begin{cases} \dot{\boldsymbol{X}}^{[2]} = \boldsymbol{f}^{[2]}(t, \boldsymbol{X}^{[2]}, \boldsymbol{D}^{[2]}, \boldsymbol{U}^{[2]}, \boldsymbol{U}_{\text{ext}}), \\ \boldsymbol{Y}^{[2]} = \boldsymbol{g}^{[2]}(t, \boldsymbol{X}^{[2]}, \boldsymbol{D}^{[2]}, \boldsymbol{U}^{[2]}, \boldsymbol{U}_{\text{ext}}). \end{cases}$$

Here, $\boldsymbol{U}^{[1]}$ are the inputs needed for Model 1 (Σ_1), directly provided by the outputs $\boldsymbol{Y}^{[2]}$ produced by Model 2 (Σ_2). Similarly, $\boldsymbol{U}^{[2]}$ are the inputs needed for Model 2 directly provided by the outputs $\boldsymbol{Y}^{[1]}$ produced by Model 1. Our approach generalizes to any decomposition of the system Σ into B blocks, where each block is indexed by $b \in \{1, \ldots, B\}$.

3.3 Model of computation

To perform the numerical integration of the whole multi-variable system, each of these simulators needs to exchange, at communication (or synchronization) points t_{s_b} ($b = 1$ or $b = 2$), the data needed by the other (see Figure 2). To speed up integration, the parallel branches must be as independent as possible, so that they are synchronized at a rate $H^{[b]} = t_{s_b+1} - t_{s_b}$, by far slower than their internal integration step $h_{n_b}^{[b]}$ ($H^{[b]} \gg h_{n_b}^{[b]}$). Therefore, between communication points, each simulator integrates at its own rate (assuming a variable-step solver), and considers that data incoming from others simulators is held constant.

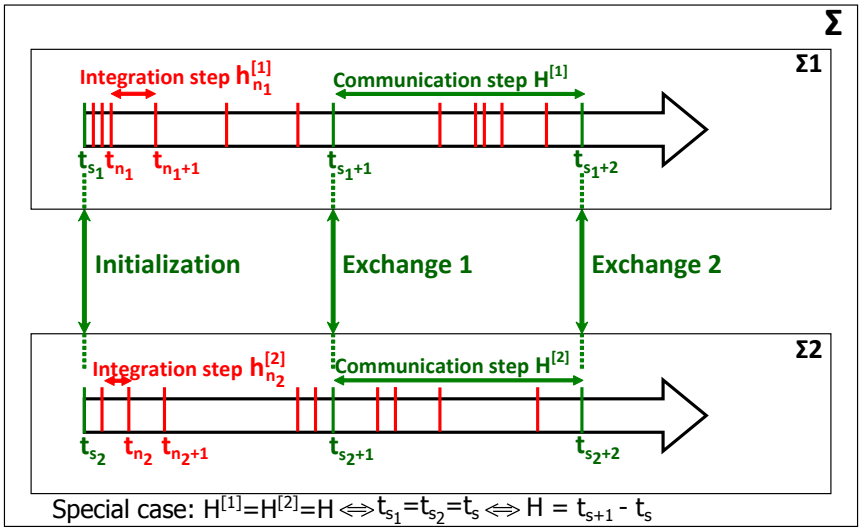


Figure 2: Σ split into Σ_1 and Σ_2 for parallel simulation.

Besides, when $H^{[b_1]} \neq H^{[b_2]}$, data incoming from other slower simulators are held constant. For instance in Figure 3, Model 1 needs to communicate with an external model two times faster than Model 2). Data incoming from Model 2 is held constant during $2H^{[1]}$, potentially causing aliasing.

It is likely that large and multi-rate communication intervals allow to speed up the numerical integration, but may result in integration errors and poor confidence in the final result.

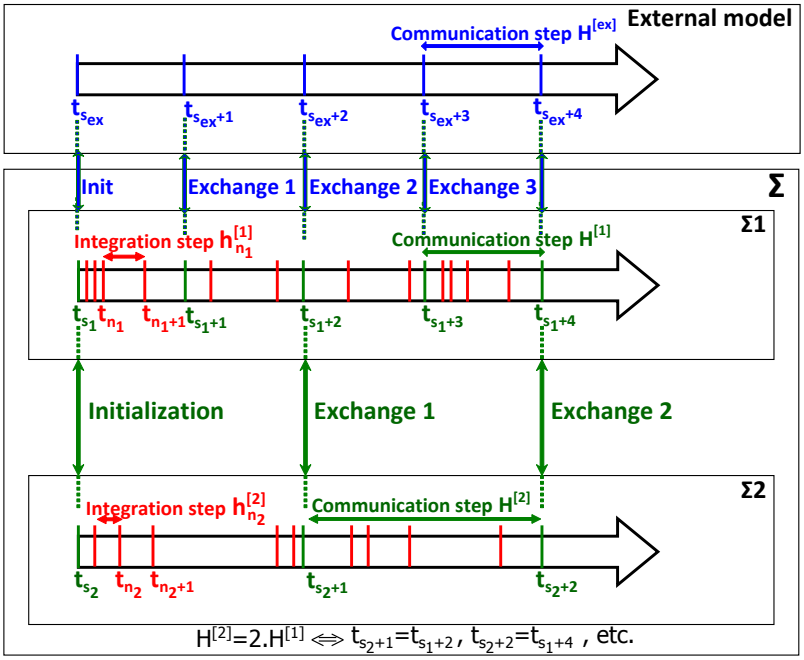


Figure 3: Parallel multi-rate simulation.

3.4 Motivation for extrapolation

Modeling the errors induced by slackened synchronization is a first direction in improving the trade-off between integration speed and accuracy.

3.4.1 Integration errors and parallelism

Error evaluation and convergence analysis were performed in [3, Chapter 9] for different schemes (sequential and parallel modular co-simulation, models with real and artificial loops, mono-rate and multi-rate co-simulation, etc.). Assuming that the b^{th} sub-model is connected to $B_c - 1$ other sub-models, its global error on states (1) and outputs (2) are bounded in Landau-Bachmann $\mathcal{O}()$ notation. Errors are functions is function of the communication step (or synchronization interval) $H = \max H^{[b]}$, $b \in \{1, \ldots, B_c\}$, the integration time-step h and the order of accuracy p of the used numerical solver [3, Chapter 4]:

$$\left\| \boldsymbol{X}^{[b]}(t_{n+1}) - \boldsymbol{X}^{[b]}_{n+1} \right\| \leq \mathcal{O}(h^p) + \mathcal{O}(H) \ , \tag{1}$$

$$\left\| \boldsymbol{Y}^{[b]}(t_n) - \boldsymbol{Y}^{[b]}_{n} \right\| \leq \mathcal{O}(h^p) + \mathcal{O}(H) \ . \tag{2}$$

Both global errors are clearly bounded by two terms. The first term is related to the applied numerical solver, more specifically the time-step and the order. The worst case scenario on the bounding would be the maximum used integration step h and order p. The second term is related to the size of the communication step H. However, the weight of each term on the error is deeply related to the size of the communication step relatively to the integration step. Based on the same approach as in [9], it is clear that the communication step H dominates the error when $H \gg h$.

Therefore, considering a split model and a parallel execution, a trade-off must be found between acceptable simulation errors, thanks to tight enough synchronization, and simulation speed-up thanks to decoupling between sub-models.

3.4.2 Contribution of the extrapolation

To add a degree of freedom to this trade-off achievement, we propose to extrapolate model inputs to compensate the stretching out of communication steps between sub-models. In fact, it was proven previously that numerical solutions are first order accurate, $O(H)$, when choosing larger communication steps H. Holding inputs constant between two synchronization intervals plays the role of a zeroth-order hold (ZOH, constant extrapolation). To generalize the error bound, the $O(H)$ term can then be replaced with $O\left(H^{\delta+1}\right)$, where δ is the extrapolation degree. Using for example linear ($\delta = 1$) or quadratic ($\delta = 2$, A.1) extrapolation instead of constant update ($\delta = 0$) has the potential to reduce the bound of simulation errors, as proven in [9, p. 238 *sq.*]. However, the difficulty with extrapolation is that it could be sensitive for different reasons:

- there exists no universal prediction scheme, efficient for every signal;
- prediction should be efficient: causal, sufficiently fast and reliable;
- standard polynomial prediction may fail in stiff cases [12] (cf. Section 5 for details).

We choose to base our extrapolation on polynomial prediction, which allows fast and causal calculations. In this situation, the rationale is that the computing cost of a low-order polynomial predictor would be by far smaller than the extra model computations needed by shorter communication steps. Since such predictions would be accurate neither for any signal (for instance, blocky versus smooth signals) nor for any signal behavior (slow variations versus steep onsets), we borrow the context-based approach from lossless image encoders [13], such as GIF (Graphics interchange format) or PNG (Portable Network Graphics) compression formats. The general aim of these image coders is to predict a pixel value based on a pattern of causal neighboring pixels. Compression is achieved when the prediction residues possess smaller intensity values, and more generally a sparser distribution (concentrated around close-to-zero values) than that of pixels in the original image. They may therefore be coded on smaller "bytes", using entropy coding techniques. In images, one distinguishes basic "objects" (smooth-intensity varying regions, edges with different orientations). Based on simple calculations over prediction patterns, different behaviors are inferred (e.g. flat, smooth, $+45^o$ or -45^o edges, etc.). Look-up table predictors are then used, depending on the context.

In the proposed approach, we build a heuristic table of contexts (Section 5) based on a short frame of past samples, and affect pre-computed polynomial predictors to obtain context-dependent extrapolated values. We now review the principles of extrapolation.

4 CHOP$_{\text{oly}}$: Causal Hopping Oblivious Polynomials for low-complexity extrapolation

4.1 Background on prediction, real-time constraints and extrapolation strategies

The CHOPtrey framework requires a dedicated instance of forecasting for discrete time series. The neighboring topics of prediction, interpolation or extrapolation represent a large body of knowledge in signal processing [14, 15], econometrics [16], control [17, 18] and numerical analysis [19]. They are paramount in complex systems that live, sample and communicate at different rates. Their use in simulation might be

milder. There exists a natural and common belief that signal extrapolation may introduce additional error terms and cause numerical instability [20, 21]. Building upon our previous work [10], two additions diminish the importance of extrapolation caveats: past sample weighting (oblivious polynomials) and a hierarchy of contexts defined by a pattern ontology.

As we operate under real-time conditions, implying strong causality, only a tiny fraction of time series extrapolation methods are practically applicable. For instance, the theory of multi-rate filter banks bears formal similarities with distributed simulation or co-simulation, and have long been deployed for interpolation, extrapolation, and reduction of computational needs [22]. Such systems allow optimized noise handling at variable redundancy rates [23]. The overall propagation delay is harmless in traditional signal and image processing (for compression or signal restoration). However, it prevents their use in distributed simulation. Additionally, the decomposition of signals into frequency bands is not adapted to simulation data, that sometimes exhibit stiff behavior, preventing the use of band-limited extrapolation [24]. We alternatively propose to feed a bank of weighted smoothing polynomials, whose outputs are selected upon both simulation and real-time behaviors, organized in a Contextual & Hierarchical Ontology of Patterns (CHOPatt, Section5).

ZOH (zeroth-order hold) or nearest-neighbor extrapolation is probably the most natural, the less hypothetical, and the less computationally expensive forecasting method. It consists in using the latest known sample as the predicted value. It possesses small (cumulative) errors when the time series is relatively flat or when its sampling rate is sufficiently high, with respect to the signal dynamics. In other words, it is efficient when the time series is sampled fast enough to ensure small variations between two consecutive sampling times. However, it indirectly leads to under-sampling or aliasing related disturbances [25]. They affect the signal information content, including its derivatives, and consequently its processing, for instance through differential equation systems. They appear as quantization-like noise [26, p. 609 sq.], delay induction, offset or bump flattening. As a remedy, higher order extrapolation schemes can be used. They consist in fitting a δ-degree polynomial through $\delta + 1$ supporting points [27, p. 15 sq.]. The most commons are the predictive first-order [28, 29] or the second-order hold (FOH or SOH) [25]. They result in linear or parabolic extrapolation. Polynomial parameters are obtained in a standard way with Lagrange polynomials. Hermite polynomials can be used when one is interested in extrapolating derivatives as well, and their relative stability has been studied [31, p. 61 sq.]. Both, although being exact at synchronization points, tend to oscillate amidst them. To avoid introducing discontinuities, extrapolation can be smoothed with splines [32] or additional interpolation [33].

In our co-simulation framework, communication intervals are not chosen arbitrarily small for computational efficiency. Thus, the slow variation of inputs and outputs cannot be ensured in practice. Provided a cost vs. error trade-off is met, borrowing additional samples from past data and using higher-order extrapolation schemes could be beneficial. Different forecasting methods of various accuracy and complexity may be efficiently evaluated. We focus here on online extrapolation with causal polynomials, for simplicity and ease of implementation, following initial work in [34, Chap. 16]. Their main feature is that they are obtained in a least-squares fashion: we do not impose that they pass exactly through supporting points. However, they do so when the degree δ and the number of supporting points $\delta + 1$ is set as above. We improve on [10] by adding a time-depending oblivion faculty to polynomial extrapolation with an independent power weighting on past samples. It accounts for memory depth changes required to adapt to sudden variations. Computations are performed on frames of samples hopping at synchronization steps, allowing a low-complexity matrix formulation. The resulting Causal Hopping Oblivious Polynomials (CHOPoly) are described next, evaluated on the case study described in Section 6 and tested in Section 7.

4.2 Notations

The first convention abstracts a sampling framework independent of the actual sampling period H and the running time index. This amounts to considering a unit communication step for any signal $u(t)$, i.e. $H = 1$, and to using a zero-reindexing convention: the last available sample is indexed by 0 (u_0), and the previous samples are thus indexed backwards: u_{-1}, u_{-2}, u_{-3},... This simplification possesses two main traits:

1. local indices hop at each communication step and become independent of the actual timing and communication rate;

2. some intermediate computations may thus be performed only once, reducing the risk of cumulative numerical errors [35].

We can either use a recursive formulation with infinite memory, or a finite prediction frame. The first option is used for instance in adaptive filtering or Kalman estimation. It generally includes oblivion, implemented for instance using a scalar forgetting factor w_l, $l \in]-\infty, \ldots, -1, 0]$, assigning a decreasing weight, often exponential [36], to older error samples. We instead consider the second option with a finite frame of the λ last consecutive past samples $\{u_{1-\lambda}, u_{2-\lambda}, \ldots, u_0\}$. Limited-sized buffers are more natural for polynomial forecasting, and they also appear beneficial to reset computations in the case of sudden changes in the data, as discussed in Section 5.2. To emulate a variable memory depth weight, we choose a piece-wise power weighting with order $\omega \geq 0$. It can be expressed as follows, without λ and ω indices to lighten notations:

$$w_l = \begin{cases} 0 & \text{if } l < 1-\lambda\,, \\ \left(\dfrac{\lambda + l}{\lambda}\right)^{\omega} & 1-\lambda \leq l \leq 0\,. \end{cases} \tag{3}$$

Their behavior for different powers ω is illustrated in Figure 4. For $\omega = 0$, all samples in the frame are assigned the same importance. The higher the power, the smaller the influence of older samples.

4.3 CHOPoly: Type I, symbolic and Type II implementations

Although the forthcoming derivations are elementary, they are rarely exposed with their complexity evaluated. We consider polynomial predictors, performed in the least-squares sense [26, p. 227 sq.]. We denote by $P_{\delta,\lambda,\omega}$ the least-squares polynomial predictor of degree $\delta \in \mathbb{N}$, frame length $\lambda \in \mathbb{N}^*$ and weighting factor ω. Its stable estimation requires that $\lambda > \delta$. The polynomial $P_{\delta,\lambda,\omega}$ is defined by the vector of length $\delta + 1$, with coefficients $\boldsymbol{a}_\delta$; hence $u(t) = a_\delta + a_{\delta-1}t + \cdots + a_0 t^\delta$. A causal prediction consists in the estimation of unknown data at a future and relative time $\tau \geq 0$. Generally $\tau \in]0, 1[$, i.e., inside the time interval between the last known sample u_0 and the forthcoming communication step. The predicted value at time τ is loosely denoted by $u_{P_{\delta,\lambda,\omega}}(\tau)$, or simply $u(\tau)$ when the context is self-explanatory. We refer to A.1 for an introductory example on constant weighting, degree-two or parabolic prediction ($P_{2,\lambda,0}$).

In the more generic context, we minimize the least-squares prediction error:

$$e(\boldsymbol{a}_\delta) = \sum_{l=1-\lambda}^{0} \left(\frac{\lambda + l}{\lambda}\right)^{\omega} \times \left(u_l - \left(\sum_{d=0}^{\delta} a_d l^{\delta-d}\right)\right)^2 . \tag{4}$$

If one chooses $\omega = 0$, and $\lambda = \delta + 1$, one recovers the Lagrange polynomials, and the error $e(\boldsymbol{a}_\delta)$ is exactly zero. Consequently, CHOPoly encompasses standard extrapolation used in co-simulation (*e.g.* ZOH, FOH, SOH or higher-order). Choosing $\lambda > \delta + 1$ induces a form of smoothing, that reduces oscillations.

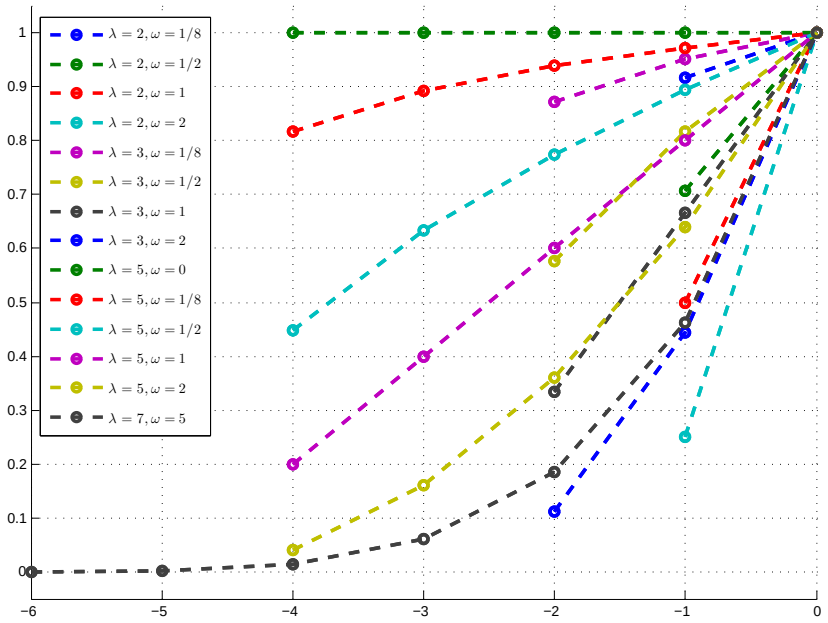


Figure 4: Effect of the weighting power ω with different choices of λ on the memory depth.

Choosing $\omega > 0$ generally promotes a better extrapolation near the last know sample, making smoothing lag-dependent. Both parameters take into account the possibility that the signal values at communication steps could be imperfect, due to jitter in the actual sampling or round-off errors for instance.

As λ^ω is a non-null constant, the derivation of (4) with respect to a_i yields $\delta + 1$ equations, for each $i \in \{0, \dots, \delta\}$:

$$\sum_{l=1-\lambda}^{0} (\lambda + l)^\omega \, l^{\delta - i} u_l = \sum_{d=0}^{\delta} \left[\sum_{l=1-\lambda}^{0} (\lambda + l)^\omega \, l^{2\delta - d - i} \right] a_d \, .$$

We define the sums of powers $z_{d,\lambda} = \sum_{l=0}^{\lambda-1} l^d$ (see A.1) and the weighted sums of powers $\bar{z}_{d,\lambda,\omega} = \sum_{l=0}^{\lambda-1} (\lambda - l)^\omega \, l^d$. They can be rewritten, if $\omega \in \mathbb{N}$:

$$\bar{z}_{d,\lambda,\omega} = \sum_{o=0}^{\omega} (-1)^o \binom{o}{\omega} \lambda^{\omega - o} z_{o+d,\lambda} \, ,$$

where $\binom{o}{\omega}$ denotes the binomial coefficient. The associated weighted moments write, accordingly, $\overline{m}_{d,\lambda,\omega} = \sum_{l=0}^{\lambda-1} (\lambda - l)^\omega \, l^d u_{-l}$. We refer to A.1 for additional information on power sums $z_{d,\lambda} = \sum_{l=0}^{\lambda-1} l^d$ and moments.

The generic extrapolation pattern takes the first following matrix form (Type I):

$$u(\tau) = \begin{bmatrix} 1 & \tau & \cdots & \tau^\delta \end{bmatrix} \begin{bmatrix} \overline{z}_{0,\lambda,\omega} & -\overline{z}_{1,\lambda,\omega} & \cdots & (-1)^\delta \overline{z}_{\delta,\lambda,\omega} \\ -\overline{z}_{1,\lambda,\omega} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ (-1)^\delta \overline{z}_{\delta,\lambda,\omega} & \cdots & \cdots & \overline{z}_{2\delta,\lambda,\omega} \end{bmatrix}^{-1} \begin{bmatrix} \overline{m}_{0,\lambda,\omega} \\ -\overline{m}_{1,\lambda,\omega} \\ \vdots \\ (-1)^\delta \overline{m}_{\delta,\lambda,\omega} \end{bmatrix} . \tag{5}$$

We note $\boldsymbol{\tau}_\delta = [1, \tau, \cdots, \tau^\delta]^T$ the vector of τ powers, and

$$\overline{\boldsymbol{m}}_{\delta,\lambda,\omega} = \left[\overline{m}_{0,\lambda,\omega}, -\overline{m}_{1,\lambda,\omega}, \cdots, (-1)^\delta \overline{m}_{\delta,\lambda,\omega}\right]^T .$$

The inverse of the $(\delta + 1) \times (\delta + 1)$ Hankel matrix $\overline{\boldsymbol{Z}}_{\delta,\lambda,\omega}$ in (5):

$$\begin{bmatrix} \overline{z}_{0,\lambda,\omega} & -\overline{z}_{1,\lambda,\omega} & \cdots & (-1)^\delta \overline{z}_{\delta,\lambda,\omega} \\ -\overline{z}_{1,\lambda,\omega} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ (-1)^\delta \overline{z}_{\delta,\lambda,\omega} & \cdots & \cdots & \overline{z}_{2\delta,\lambda,\omega} \end{bmatrix}$$

is denoted by $\overline{\boldsymbol{Z}}_{-\delta,\lambda,\omega}$. Hence, the Type-I formula for CHOPoly writes:

$$u(\tau) = \boldsymbol{\tau}_\delta^T \overline{\boldsymbol{Z}}_{-\delta,\lambda,\omega} \overline{\boldsymbol{m}}_{\delta,\lambda,\omega} . \tag{6}$$

The generic matrix formulation in (6) is compact. The matrix $\overline{\boldsymbol{Z}}_{-\delta,\lambda,\omega}$ may be computed beforehand, as a look-up table. A polynomial of degree δ involves terms of degree 2δ. We note that this could lead to huge values computed from large sample indices for long simulation signals. This situation is avoided by the frame-hopping zero-reindexing convention, which thus limits round-off errors and subsequent issues in the inversion of matrices that could become defective [37].

The moment-based formulation conceals more direct symbolic formulae, linear in past samples and polynomial in τ. Some examples are given in A.2. They can be implemented with Ruffini-Horner's rule [39] for higher extrapolation orders. However, symbolic polynomial evaluations are not always handy to implement.

We note $\boldsymbol{u}_\lambda = [u_0, u_{-1}, \ldots, u_{2-\lambda}, u_{1-\lambda}]^T$. Then (6) can be rewritten in a Type II form. For instance, since $\boldsymbol{\tau}_0^T = 1$, extrapolation with polynomial $P_{0,\lambda,1}$ (17) rewrites as a matrix product with $\boldsymbol{\Pi}_{0,\lambda,1}$:

$$u(\tau) = \frac{2}{\lambda(\lambda+1)} \begin{bmatrix} \lambda & \lambda - 1 & \cdots & 1 \end{bmatrix} \boldsymbol{u}_\lambda = \boldsymbol{\tau}_0^T \boldsymbol{\Pi}_{0,\lambda,1} \boldsymbol{u}_\lambda , \tag{7}$$

The general Type-II formula for predictor matrices $\boldsymbol{\Pi}_{\delta,\lambda,\omega}$ of size $(\delta + 1) \times \lambda$ implements (6) as:

$$u(\tau) = \boldsymbol{\tau}_\delta^T \boldsymbol{\Pi}_{\delta,\lambda,\omega} \boldsymbol{u}_\lambda . \tag{8}$$

Such a formulation allows an efficient storage of matrices $\boldsymbol{\Pi}_{\lambda,\delta,\omega}$ with different weighting factors. Examples are provided in Table 5 in A.4. Estimates of the generic number of elementary operations required for each extrapolated τ are compared in Table 4 in A.3. Type I and Type II are roughly quadratic in (δ, λ). Although the rectangular matrix $\boldsymbol{\Pi}_{\delta,\lambda,\omega}$ is larger than $\overline{\boldsymbol{Z}}_{\delta,\lambda,\omega}$, since $\lambda > \delta$, the economy in lazy matrix addressing combined with the number of elementary operations make Type II implementation more practical and efficient.

5 CHOP$_{att}$: Contextual & Hierarchical Ontology of Patterns

5.1 Pattern representation and two-level context hierarchy definitions

We now introduce the pattern-based approach, borrowed from common lossless image encoders (Section 3.4.2), by providing a contextual and hierarchical framework for CHOPoly extrapolation. To this aim, the role of contexts is to cover all possible scenarios (an ontology) of signal's evolution for a hybrid dynamical system. For instance, slow and steep variations must be included, the same goes for blocky and smooth signals.

On the first level, we define a functional context that differentiates a set of typical patterns, represented in Figure 5. We are interested in the dynamics of the most recent samples, depicted with red links, with respect to the past behavior (black links). It defines six mutually exclusive entities, illustrated by a short name: "flat", "calm", "move", "rest", "take" and "jump".

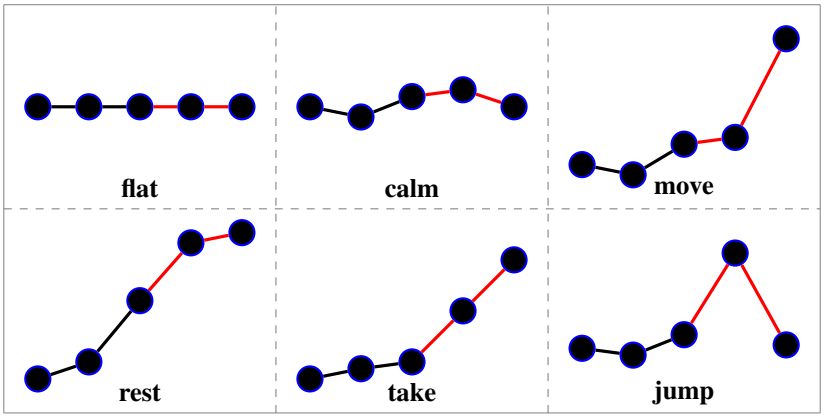


Figure 5: Pattern representation for the functional context in Table 1.

"Flat" addresses steady signals. "Calm" represents a sufficiently sampled situation, where value increments over time remain below fixed thresholds. "Move" defines a formerly "calm" signal whose novel value changes rapidly, above a pre-defined threshold. "Rest" handles signals previously varying above a threshold and becoming "calm". The "take" context addresses signals with constantly high variations. The "jump" pattern is a "take" with a sign change as for bumps and dips. They are detected in practice by comparing consecutive differences on past samples (d_i, *cf.* (9)), indexed by i, to thresholds (γ_i, *cf.* (10)). Their formal definition is provided in Table 1 and their adaptive selection in Section 5.2.

On the second level, we define a meta- or decisional context that determines if extrapolation would be beneficial or detrimental. Indeed, a major difficulty for hybrid complex systems resides in sharp and fast variations in signal patterns induced by stiffness and discontinuities. We thus detect if the signal can be characterized by a seventh "cliff" pattern, for which functional CHOPoly prediction would fail, by taking into account the importance of the signal's amplitude. This decision context is detected by comparing a ratio ρ of differences based on past samples (12) to a threshold (denoted by Γ, *cf.* (13)). The above seven patterns define a hierarchical context selection, composed of the decisional context embedding the functional context on a second level, summarized in Figure 6.

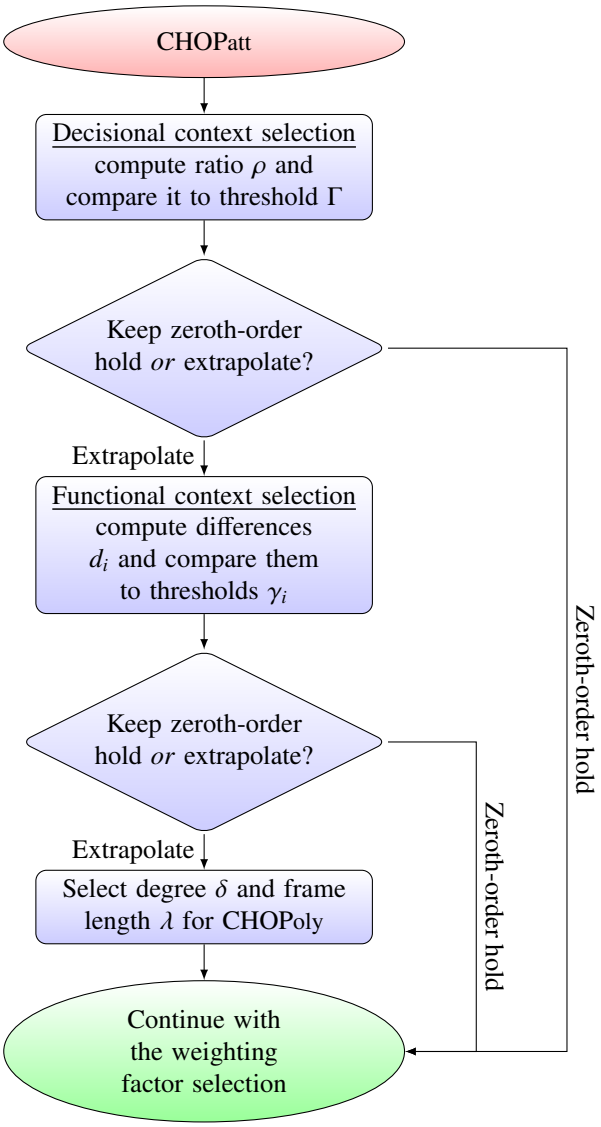


Figure 6: Hierarchical and functional context selection flowchart.

5.2 Functional context selection

As a concrete example, used in the remaining of our work, we propose two simple measures of variation based on the last three samples only: u_0, u_{-1} and u_{-2} (with the zero-reindexing convention from Section 4.2), the last and previous differences:

$$d_0 = u_0 - u_{-1} \quad \text{and} \quad d_{-1} = u_{-1} - u_{-2}\,. \tag{9}$$

Their absolute values are compared with thresholds γ_0 and γ_{-1}, respectively, defined in (10). To build

the different contexts, three complementary conditions are defined:

- O if $|d_i| = 0$;
- C_i if $0 < |d_i| \leq \gamma_i$;
- $\overline{C_i}$ if $|d_i| > \gamma_i$.

Table 1 formally defines the six entities from the functional context and presents examples of "default" ω-parametrized CHOPoly families. Since the "flat" context addresses steady signals, a mere ZOH suffices,

n(ame)	#	$\lvert d_{-1}\rvert$	$\lvert d_0\rvert$	$d_{-1}.d_0$	$(\delta, \lambda, \omega)$
f(lat)	0	O	O	O	$(0, 1, .)$
c(alm)	1	C_1	C_2	any	$(2, 5, .)$
m(ove)	2	C_1	$\overline{C_2}$	any	$(0, 1, .)$
r(est)	3	$\overline{C_1}$	C_2	any	$(0, 2, .)$
t(ake)	4	$\overline{C_1}$	$\overline{C_2}$	> 0	$(1, 3, .)$
j(ump)	5	$\overline{C_1}$	$\overline{C_2}$	< 0	$(0, 1, .)$

Table 1: Functional context: entities definition and associated prediction.

hence $P_{0,1,\omega}$. The "calm" context represents smooth signals, in this case, it could be approximated by a quadratic polynomial, for instance $P_{2,5,\omega}$. For the "flat" and "jump" contexts, an additional procedure consists in resetting the extrapolation to prevent inaccurate prediction. For example, when context 1 is chosen just after context 4, the quadratic extrapolation $P_{2,5,\omega}$ requires 5 valid samples, whereas the last 3 only are relevant, justifying our finite-length frames option.

The choice of thresholds γ_0 and γ_{-1} is potentially crucial. For instance, fixed values may reveal inefficient under important amplitude or scale variation in signals. Hence, we have chosen to update them adaptively, based on the statistics of a past frame $\{u_{1-\Lambda}, \ldots, u_{-3}\}$ (Λ denotes the maximum frame size).

With excessively low thresholds, high-order extrapolations would rarely be chosen, losing the benefits of prediction. Overly high thresholds would in contrast suffer from any unexpected jump or noise. As contexts are based on backward derivatives, we have used in the simulations presented here the mid-range statistical estimator of their absolute values. This amounts to set:

$$\gamma_0 = \gamma_{-1} = \frac{1}{2} \max_{i \in [1-\Lambda, \ldots, -3]} (|u_i - u_{i+1}|)\,, \tag{10}$$

which appeared to be sufficiently robust in our test-cases.

5.3 Next communication period: weighting factor and decisional context selection

To further improve our algorithm in [10] and to decrease even more prediction induced integration errors, the oblivious weighting factor from (3) makes the algorithm more subtly aware of data freshness. Decisions for the next communication period are taken with respect to the new updated input value u_0, compared to past predictions. It is gauged with two error measures, illustrated in Figure 7. They assess, a posteriori, what

would have been the best prediction. First, Δ_{worst} denotes the worst case scenario without extrapolation, the predicted value being held constant and equal to u_{-1}:

$$\Delta_{\text{worst}} = |u_0 - u_{-1}|.$$

Second, we estimate the best prediction pattern in retrospect, obtained by optimizing the free parameter ω in the prediction polynomial defined by the last selected context. We choose here a subset of weighting factors in $\Omega = \left\{0, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1, 2\right\}$ (see A.4 for details):

$$\Delta_{\text{best}} = \min_{\omega \in \Omega} |u_0 - \hat{u}_{-1}^{\omega}|.$$

The best weighting factor ω_{best} defined in (11) is then selected during the next communication interval:

$$\Delta_{\text{best}} = |u_0 - \hat{u}_{-1}^{\omega_{\text{best}}}|. \tag{11}$$

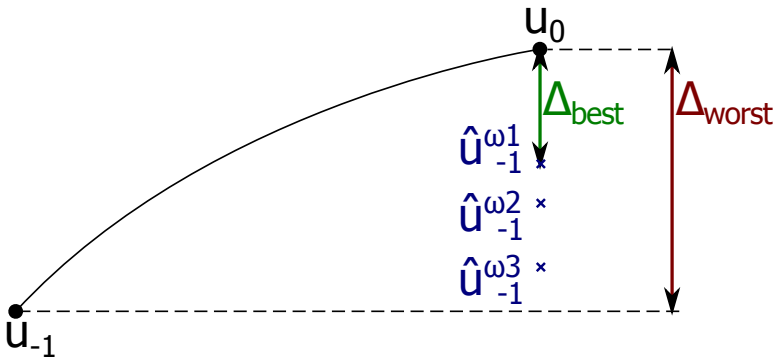


Figure 7: Illustration of two error measures: Δ_{worst} when there is no extrapolation and Δ_{best} for the best prediction pattern.

Regarding the selection of the decisional context called "cliff", we first define a ratio:

$$\rho = \frac{\Delta_{\text{best}}}{\Delta_{\text{worst}}}. \tag{12}$$

It specifies if there is a sharp fast variation or not by comparing it to a pre-defined threshold and then decides to extrapolate or not. The threshold denoted Γ is defined as follow:

$$0.7 \leq \Gamma < 1. \tag{13}$$

In fact, when $\rho > \Gamma$ (e.g. with $\Gamma = 90\,\%$), it means that the input value is only enhanced by $1 - \Gamma$ (e.g. $10\,\%$) regarding the "true" value. This is the case when there is a sharp and fast variation or a weird behavior. The decisional context "cliff" is then selected and activated with its associated heuristic polynomial predictor $P_{\delta,\lambda,\omega} = P_{0,1,\omega}$. On the other hand, when $\rho \leq \Gamma$, the conventional functional context table introduced in Section 5.2 is used.

6 Case study

6.1 Engine simulator

In this study, a Spark Ignition (SI) RENAULT F4RT engine has been modeled with 3 gases (air, fuel and burned gas). It is a four-cylinder, in-line Port Fuel Injector (PFI), engine in which the engine displacement is 2000 cm^3. The combustion is considered as homogeneous. The air path (AP) consists in a turbocharger with a mono-scroll turbine controlled by a waste-gate, an intake throttle and a downstream-compressor heat exchanger. This engine is equipped with two Variable Valve Timing (VVT) devices, for intake and exhaust valves, to improve the engine efficiency (performance, fuel and emissions). The maximum power is about 136 kW at 5000 rpm.

The F4RT engine model was developed using the ModEngine library [40]. ModEngine is a Modelica [41] library that allows the modeling of a complete engine with diesel and gasoline combustion models.

Requirements for the ModEngine library were defined and based on the already existing IFP-Engine library. The development of the IFP-Engine library was performed several years ago at "IFP Energies nouvelles" and it is currently used in the AMESim[1] tool. ModEngine contains more than 250 sub-models. It has been developed to allow the simulation of a complete virtual engine using a characteristic time-scale based on the crankshaft angle. A variety of elements are available to build representative models for engine components, such as turbocharger, wastegate, gasoline or diesel injectors, valve, air path, Exhaust Gas Recirculation (EGR) loop etc. ModEngine is currently functional in Dymola[2].

The engine model and the split parts were imported into xMOD model integration and virtual experimentation tool [42], using the FMI export features of Dymola. This cyber-physical system has 118 state variables and 312 event indicators (of discontinuities).

6.2 Decomposition approach

The partitioning of the engine model is performed by separating the four-cylinder from the air path, then by isolating the cylinders (C_i, for $i \in \{1, \ldots, 4\}$) from each other. This kind of splitting allows for the reduction of the number of events acting on each sub-system. In fact, the combustion phase raises most of the events, which are located in the firing cylinder. The solver can process them locally during the combustion cycle of the isolated cylinder, and then enlarge its integration time-step until the next cycle.

From a thermodynamic point of view, the cylinders are weakly coupled, but a mutual data exchange does still exist between them and the air path.

The dynamics of the air path is slow (it produces slowly varying outputs to the cylinders, e.g. temperature) compared to those of the cylinders (they produce fast outputs to the air path, e.g. torque). Besides, unlike cylinders outputs, most air path outputs are not a direct function of the air path inputs (they are called Non Direct Feedthrough (NDF) outputs). This results in choosing the execution order of the split model from the air path to the cylinders (in accordance with the analysis of the behavior of NDF to Direct Feedthrough (DF) in [3, Chapter 9]).

The model is split into 5 components and governed by a basic controller denoted CTRL. It gathers 91 inputs and 98 outputs.

[1] www.lmsintl.com/imagine-amesim-1-d-multi-domain-system-simulation
[2] www.3ds.com/products/catia/portfolio/dymola

7 Tests and results

Tests are performed on a platform with 16 GB RAM and an "Intel Core i7" 64-bit processor, running 4 cores (8 threads) at 2.70 GHz.

7.1 Reference simulation

The model validation is based on the observation of some quantities of interest as the pressure, the gas mass fraction, the enthalpy flow rate, the torque, etc. These outputs are computed using LSODAR [3], a variable time-step solver with a root-finding capability that detects the events occurring during the simulation. It also has the ability to adapt the integration method depending on the observed system stiffness.

The simulation reference Y_{ref} is built from the integration of the entire engine model, the solver tolerance (tol) being decreased until reaching stable results, which is reached for tol = 10^{-7} (at the cost of an unacceptable slow simulation speed).

Then, to explore the trade-offs between the simulation speed and precision, simulations are run with increasing values of the solver tolerance until reaching a desired relative integration error Er, defined by (14)

$$\text{Er}(\%) = \frac{100}{N} \cdot \sum_{i=0}^{N-1} \left| \frac{Y_{\text{ref}}(i) - Y(i)}{Y_{\text{ref}}(i)} \right|, \tag{14}$$

with N the number of saved points during 1 s of simulation. Iterative runs showed that the relative error converge to a desired error (Er ≤ 1 %) for tol = 10^{-4}. The single-thread simulation of the whole engine with LSODAR and tol = 10^{-4} provides the simulation execution time reference, to which the parallel version is compared. When using the split model, each of its 5 components is assigned to a dedicated core and integrated by LSODAR with tol = 10^{-4}.

7.2 Automatic detection of fast and sharp variations

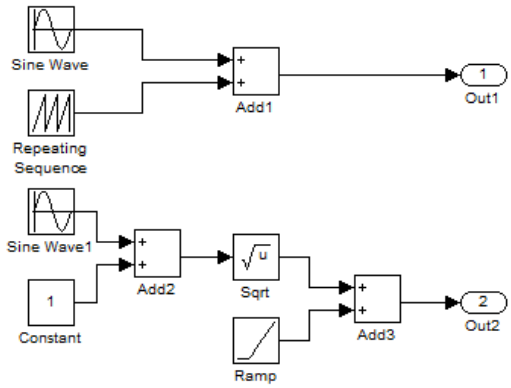
Adding a hierarchy of decisional and functional contexts overcomes previous limitation in [10]. We illustrate with two signals denoted "Out1" and "Out2". They are built in Matlab/Simulink as shown in Figure 8a. They exhibit different variations over time, as illustrated in Figure 8b. From Figure 9, extrapolation of "Out1" fails around $t = 8$ s at the sharp variation. Here extrapolation is detrimental since it increases errors instead of minimizing them.

To fix this limitation on high jumps, we apply hierarchical contexts' selection to detect the "cliff" context. Using the ratio ρ (12) and comparing it to a threshold Γ (13) equal to 90 %, the improvement of extrapolation at this step is found lesser than 10 %. The "cliff" context is then activated to avoid extrapolation. As a result, the decisional context prevents additional errors of prediction and the result of the simulation is improved as it is shown in Figure 10.

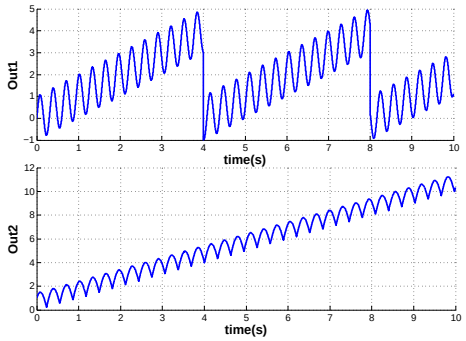
7.3 Automatic selection of the weighting factor

To determine the best weighting factor regarding error minimization, we first test them separately on "Out1" and "Out2". This first test is quite simple since there is no interaction between blocks, which means that

[3] Short for Livermore Solver for Ordinary Differential equations with Automatic method switching for stiff and nonstiff problems, and with Root-finding [43].



(a) Construction of Out1 and Out2 in Matlab/Simulink.



(b) Illustration of Out1 and Out2.

Figure 8: A test sample.

there is no effect of the action/reaction of extrapolated signals on each other. The purpose here is to show the difference between all weighting factors.

Figure 11 shows the absolute error, which is the absolute value of the difference between the reference in Figure 8b built with a small communication step $H = 10\,\mu s$ and signals simulated with a larger communication step $H = 100\,\mu s$, extrapolated or not. It can be inferred that the higher the weighting factor, the smaller the error.

Table 2 shows the cumulative absolute error on a long simulation run (during 10 s). It confirms that the weighting factor $\omega = 2$ is the best regarding error reduction. In fact, the extrapolation is enhanced from $\omega = 0$ to $\omega = 2$ by reducing the error of prediction by 20.10 % for "Out1" and by 11.39 % for "Out2".

The same experience is now applied on the F4RT engine model and extrapolation with different weighting factors is applied separately on all engine inputs. Figure 12 represents one of the cylinder 1 outputs, "the enthalpy flow rate", for the different extrapolation modes.

We notice more clearly in Figure 13 that for each communication step, there is a different best weighting factor that minimizes the absolute error. Besides, the computation of the cumulative integration error, during a long simulation run, shows that there is no unique best weighting factor.

The weighting factor ω is then chosen dynamically during the simulation as described in Section 5.3.

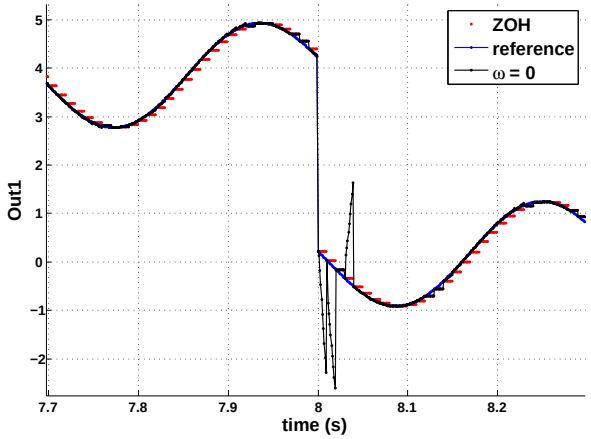


Figure 9: Failure of the old extrapolation.

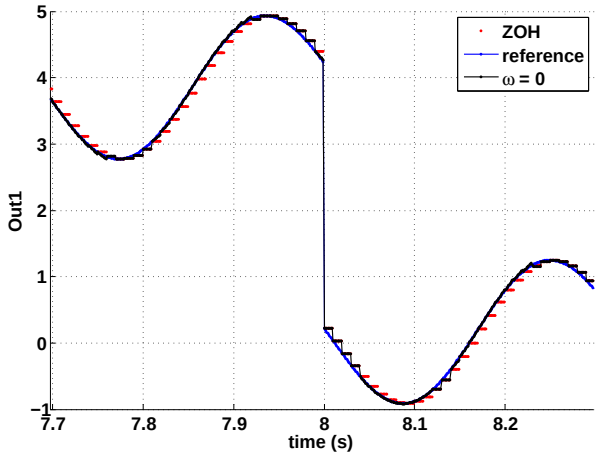


Figure 10: Success of the new extrapolation.

At each communication step, the weighting factor that minimizes the last error is selected and used for the current step.

Thanks to dynamic error evaluation and weighting factor selection, the cumulative integration error is almost always the lowest one for the different outputs as shown in Figures 14a and 14b. However, the worst enhancement of the error depends on the output, for instance it is obtained with $\omega = \frac{1}{2}$ for the air mass fraction of the cylinder 1 and with $\omega = 2$ for the fuel mass fraction of the cylinder 1. This confirms that for complex coupled systems, there is no unique best weighting factor, hence the necessity and the usefulness of combining different ones.

Besides, for the "burned gas mass fraction" of cylinder 1 (see Figure 15), the dynamic weighting factor

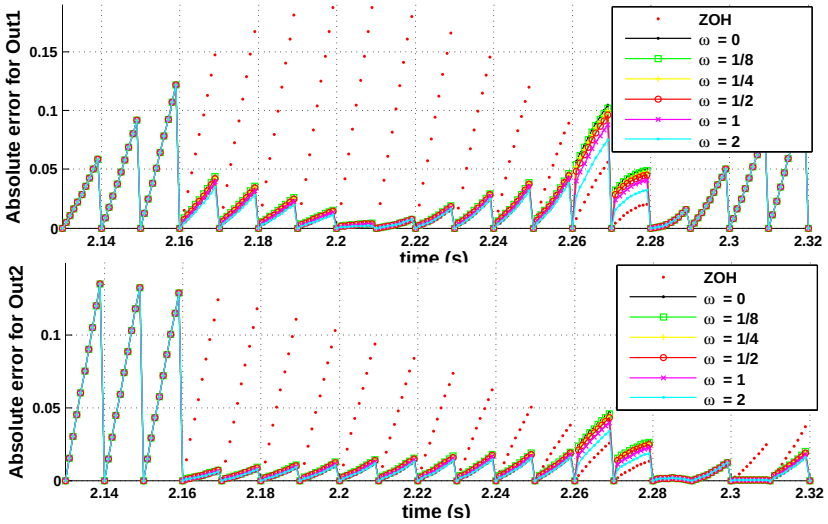


Figure 11: Behavior of the absolute error at each communication step.

Type	Error		Improvement	
	Out1	Out2	Out1	Out2
ZOH	0.572	0.399	—	—
$\omega = 0$	0.204	0.158	64.34 %	60.40 %
$\omega = 1/8$	0.201	0.157	64.86 %	60.65 %
$\omega = 1/4$	0.198	0.155	65.38 %	61.15 %
$\omega = 1/2$	0.193	0.153	66.27 %	61.65 %
$\omega = 1$	0.182	0.148	68.18 %	62.91 %
$\omega = 2$	0.163	0.140	71.50 %	64.91 %

Table 2: Cumulative absolute error during 10 s and relative improvement.

selection decreases the error of prediction by 32 % compared to the previous work (with $\omega = 0$, in [10]) as well as the simulation error by 40 % compared to the non-extrapolated signal.

Regarding now the achievement on the simulation speed-up, Table 3 shows the acceleration compared with the single-threaded reference. Firstly, the speed-up is supra-linear w.r.t. the number of cores when the model is split into 5 threads integrated in parallel on 5 cores. Indeed, the containment of events detection and handling inside small sub-systems allows for solvers accelerations, enough to over-compensate the multi-threading costs. Secondly, model splitting combined with enlarged communication steps, from $H = 100\,\mu s$ to $H = 250\,\mu s$, allows around +12.50 % extra speed-up. Unfortunately this extra speed-up is obtained at the

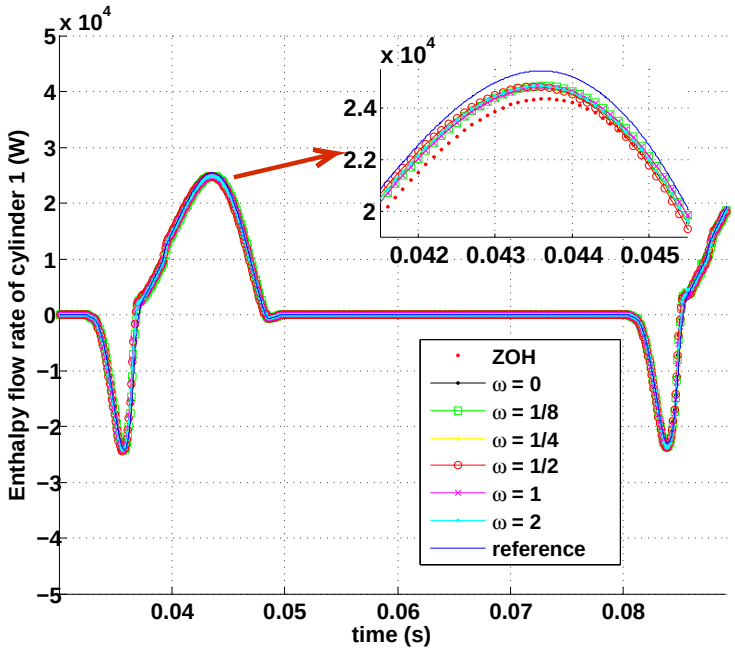


Figure 12: Enthalpy flow rate output of cylinder 1.

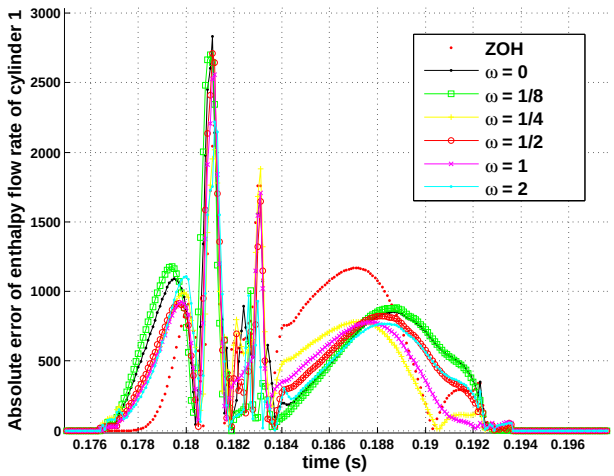
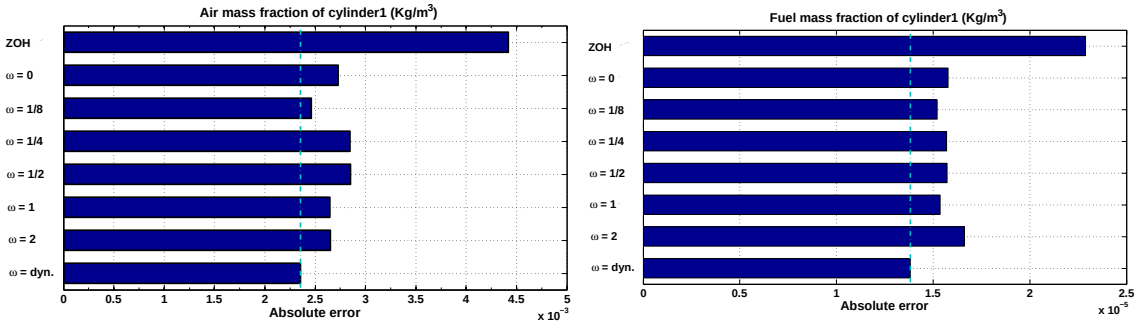


Figure 13: Absolute error of enthalpy flow rate output of cylinder 1.



(a) Air mass fraction.

(b) Fuel mass fraction.

Figure 14: Absolute error in cylinder 1.

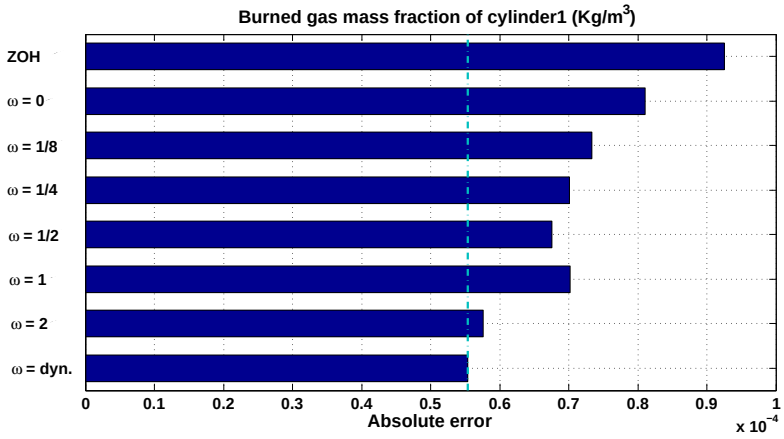


Figure 15: Absolute error of the burned gas mass fraction of cylinder 1.

cost of relative error increase (e.g. +341 % for the fuel density). Thirdly, the combination of model splitting with expanded communication steps (use of $H = 250\,\mu s$) as well as CHOPtrey allows to keep the same extra speed-up while decreasing the relative error to values close to, or below, those measured with $H = 100\,\mu s$ and ZOH. We can conclude that the enhancement brought with CHOPtrey allows for improved performance on both sides: simulation time and results' accuracy.

8 Summary and perspectives

The main objective of CHOPtrey is to provide a framework for hybrid dynamical systems co-simulation speed-up [44] based on extrapolation. Cheaper slackened synchronization intervals are allowed by a combination of prediction and multi-level context selection. It aims at reaching real-time simulation while pre-

Communication time	Prediction	Speed-up factor	Relative error variation	
			Burned gas density	Fuel density
100 µs	ZOH	8.9	–	–
250 µs	ZOH	10.01	+7 %	+341 %
	CHOPtrey	10.07	−26 %	+21 %

Table 3: CHOPtrey performance: speed-up vs. accuracy. The speed-up factor is compared with single-threaded reference. The relative error variation is compared with ZOH at 100 µs.

serving result accuracy.

It is implemented in combination with model splitting and parallel simulation on a hybrid dynamical engine model. It results in effective simulation speed-up with negligible computational overheads. In addition, sustained or even improved simulation precision is obtained without noticeable instability.

This work can be extended in different directions. Keeping with data extrapolation, simulated signals can be cleaned from long range trends [45], to better detect subtle behavioral modifications, and subsequently adapt detection thresholds. They can be processed on different time grids [46] with multi-scale techniques [47]. Acting as local pseudo-derivatives, the latter can decompose signals into morphological components such as polynomials trends, singularities and oscillations. This would allow improvements in context assignment by measuring sharp variations and spurious events with data-relative sparsity metrics [48].

Moreover, the discrimination of cliff behaviors could be further improved by using knowledge of the plant model, allowing to discard out-of-bound values, e.g., non-negative variables.

Finally, simulation results suggest that widening the communication steps is an important source for integration acceleration. Beyond equidistant communication grids, adaptive, context and/or error-based closed-loop control of communication steps [49, 50] is a promising research objective.

9 Acknowledgments

This work was supported by the ITEA2 project MODRIO[4], and funded in part by the "Direction Générale des Entreprises" of the French Ministry of Industry.

A Complements on extrapolation

A.1 Toy parabolic CHOP$_{oly}$ extrapolation formulae

We estimate the best fitting parabola (i.e. $\delta = 2$) with a uniform weighting ($\omega = 0$): $u(t) = a_\delta + a_{\delta-1}t + a_{\delta-2}t^2$ to approximate the set of discrete samples $\{u_{1-\lambda}, u_{2-\lambda}, \ldots, u_0\}$. We consider here "uniform" weighting $w_l = 1$ for $1 - \lambda \leq l \leq 0$ (i.e. with weighting factor $\omega = 0$). The prediction polynomial $P_{2,\lambda,0}$ is defined by the

[4]Model Driven Physical Systems Operation

vector of polynomial coefficients: $\boldsymbol{a}_2 = [a_2, a_1, a_0]^T$.These coefficients are determined, in the least-squares sense [51], by minimizing the squared or quadratic prediction error (4):

$$e(\boldsymbol{a}_2) = \sum_{l=1-\lambda}^{0} 1 \times \Big(u_l - (a_2 + a_1 l + a_0 l^2)\Big)^2 \ .$$

Here indices l are non-positive, i.e. between $1-\lambda$ and 0. The minimum error is obtained by solving the following system of equations (zeroing the derivatives with respect to each of the free variables a_i):

$$\forall i \in \{0, 1, 2\}, \quad \frac{\partial e(\boldsymbol{a}_2)}{\partial a_i} = 0 \,,$$

namely:

$$\left\{ \begin{array}{l} \displaystyle\sum_{l=1-\lambda}^{0} l^0 \Big(u_l - (a_2 l^0 + a_1 l^1 + a_0 l^2)\Big) = 0 \,, \\ \displaystyle\sum_{l=1-\lambda}^{0} l^1 \Big(u_l - (a_2 l^0 + a_1 l^1 + a_0 l^2)\Big) = 0 \,, \\ \displaystyle\sum_{l=1-\lambda}^{0} l^2 \Big(u_l - (a_2 l^0 + a_1 l^1 + a_0 l^2)\Big) = 0 \,. \end{array} \right. \tag{15}$$

System (15) may be rewritten as:

$$\left\{ \begin{array}{l} \displaystyle\sum_{1-\lambda}^{0} u_l = a_2 \sum_{1-\lambda}^{0} l^0 + a_1 \sum_{1-\lambda}^{0} l^1 + a_0 \sum_{1-\lambda}^{0} l^2 \,, \\ \displaystyle\sum_{1-\lambda}^{0} l u_l = a_2 \sum_{1-\lambda}^{0} l^1 + a_1 \sum_{1-\lambda}^{0} l^2 + a_0 \sum_{1-\lambda}^{0} l^3 \,, \\ \displaystyle\sum_{1-\lambda}^{0} l^2 u_l = a_2 \sum_{1-\lambda}^{0} l^2 + a_1 \sum_{1-\lambda}^{0} l^3 + a_0 \sum_{1-\lambda}^{0} l^4 \,. \end{array} \right.$$

Closed-form expressions exist for the sum of powers $z_{d,\lambda}$, involving Bernoulli sequences [52]. For instance, up to the 4$^{\text{th}}$ power:

- $z_{0,\lambda} = \lambda$;
- $z_{1,\lambda} = (\lambda - 1)\lambda/2$;
- $z_{2,\lambda} = (\lambda - 1)\lambda(2\lambda - 1)/6$;
- $z_{3,\lambda} = (\lambda - 1)^2\lambda^2/4$;
- $z_{4,\lambda} = (\lambda - 1)\lambda(2\lambda - 1)(3\lambda^2 - 3\lambda - 1)/30$.

Let $m_{d,\lambda} = \overline{m}_{d,\lambda,0} = \sum_{l=0}^{\lambda-1} l^d u_{-l}$ (here indices l are positive) denote the d^{th} moment[5] of the samples u_i, and $\boldsymbol{m}_{2,\lambda}$ the vector of moments $[m_{0,\lambda}, -m_{1,\lambda}, m_{2,\lambda}]^T$. We now form the Hankel matrix $\boldsymbol{Z}_{2,\lambda}$ of sums of powers

[5]Definition: the d^{th} moment of a real function $f(t)$ about a constant c is usually defined as: $M_d = \int_{-\infty}^{\infty}(t-c)^d f(t)dt$. The m_ds may be interpreted as discrete versions of one-sided moments about $t = 0$ of the discrete function $u(t)$; alternatively — *cf.* definitions for $z_{d,\lambda}$ — the moments are sort of weighted (by u_l) sum-of-powers.

(depending on $\delta = 2$ and λ):

$$\boldsymbol{Z}_{2,\lambda} = \begin{bmatrix} z_{0,\lambda} & -z_{1,\lambda} & z_{2,\lambda} \\ -z_{1,\lambda} & z_{2,\lambda} & -z_{3,\lambda} \\ z_{2,\lambda} & -z_{3,\lambda} & z_{4,\lambda} \end{bmatrix}.$$

The system in (15) rewrites:

$$\begin{bmatrix} m_{0,\lambda} \\ -m_{1,\lambda} \\ m_{2,\lambda} \end{bmatrix} = \begin{bmatrix} z_{0,\lambda} & -z_{1,\lambda} & z_{2,\lambda} \\ -z_{1,\lambda} & z_{2,\lambda} & -z_{3,\lambda} \\ z_{2,\lambda} & -z_{3,\lambda} & z_{4,\lambda} \end{bmatrix} \times \begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix}$$

or

$$\boldsymbol{m} = \boldsymbol{Z}_{2,\lambda} \times \boldsymbol{a}\,.$$

We now want to find the value predicted by $P_{2,\lambda,0}$ at time τ. Let $\boldsymbol{\tau}_2 = [1, \tau, \tau^2]^T$ be the vector of τ powers. Then $u(\tau)$ is equal to $a_2 + a_1\tau + a_0\tau^2 = \boldsymbol{\tau}_2^T \times \boldsymbol{a}_2$. This system might be solved with standard pseudo-inverse techniques [53], by premultiplying by the transpose of $\boldsymbol{Z}_{2,\lambda}$. This is not required, as $\boldsymbol{Z}_{2,\lambda}$ is always invertible, provided that $\lambda > \delta$. Its inverse is denoted $\boldsymbol{Z}_{-2,\lambda}$. It thus does not need to be updated in real-time. It may be computed offline, numerically or even symbolically. Hence:

$$u(\tau) = \left(\boldsymbol{\tau}_2^T \times \boldsymbol{Z}_{-2,\lambda}\right) \times \boldsymbol{m}_{2,\lambda}\,.$$

The vector $\boldsymbol{\tau}_2$ and $\boldsymbol{Z}_{-2,\lambda}$ are fixed, and the product $\boldsymbol{\tau}_2^T \times \boldsymbol{Z}_{-2,\lambda}$ may be stored at once. Thus, for each prediction, the only computations required are the update of vector $\boldsymbol{m}_{2,\lambda}$ and its product with the aforementioned stored matrix.

A.2 CHOPoly symbolic formulation

When only one polynomial predictor is required, actual computations do not require genuine matrix calculus, especially for small degrees δ. With $\delta = 0$ and $\omega = 0$ (or $P_{0,\lambda,0}$), one easily sees that:

$$u(\tau) = \frac{m_0}{z_{0,\lambda}} = \frac{u_0 + \cdots + u_{1-\lambda}}{\lambda}, \tag{16}$$

that is, the running average of past frame values. It reduces to standard ZOH, $u(\tau) = u_0$, when $\lambda = 1$. For $\delta = 0$ and $\omega = 1$, one gets a weighted average giving more importance to the most recent samples:

$$u(\tau) = 2\frac{\lambda u_0 + \cdots + 2u_{2-\lambda} + u_{1-\lambda}}{\lambda(\lambda+1)}. \tag{17}$$

With $\lambda = 2$, $\delta = 1$ and $\omega = 0$, $P_{1,2,1}(\tau) = u_0 + (u_0 - u_{-1})\tau$ yields the simplest 2-point linear prediction or standard FOH. $P_{1,3,0}$ yields the simple estimator form:

$$u(\tau) = \frac{1}{6}(5u_0 + 2u_{-1} - u_{-2}) + (u_0 - u_{-2})\frac{\tau}{2}\,. \tag{18}$$

For $P_{2,5,1}$, we tediously get:

$$\begin{aligned} u(\tau) =& \frac{1}{70}(65u_0 + 12u_{-1} - 6u_{-2} - 4u_{-3} + 3u_{-4}) + (25u_0 - 12u_{-1} - 16u_{-2} - 4u_{-3} + 7u_{-4})\frac{\tau}{28} \\ &+ (5u_0 - 4u_{-1} - 4u_{-2} + 3u_{-4})\frac{\tau^2}{28} \end{aligned} \tag{19}$$

or with the Ruffini-Horner's method for polynomial evaluation, to slightly reduce the number of operations:

$$u(\tau) = (((25u_0 - 20u_{-1} - 20u_{-2} + 15u_{-4})\,\tau + (25u_0 - 12u_{-1} - 16u_{-2} - 4u_{-3} + 7u_{-4}))\,\tau + (130u_0 + 24u_{-1} - 12u_{-2} - 8u_{-3} + 6u_{-4}))\,/140\,. \tag{20}$$

One easily remarks that, when the weighting exponent ω is an integer, prediction polynomials have rational coefficients, which limits floating-point round-off errors, especially when prediction times and variables (for instance quantized ones) are integer or rational as well.

A.3 CHOP$_{\text{oly}}$ Type I and II computational complexity

The computational complexity of a single extrapolation is given in Table 4. They are evaluated by a number of elementary operations. They are only meant to provide rough intuitions and guidelines on the actual implemented complexity. Their expressions for Type I (6) and Type II (8) Causal Hopping Oblivious Polynomials implementations are given terms of δ, λ (and ω). Type I direct implementation is not efficient: unoptimized moments computations yield a cubic complexity in $(\delta, \lambda, \omega)$. Recurring results can be evaluated using call-by-need or lazy evaluation. For $\overline{m}_{d,\lambda,\omega}$, the factors $(\lambda - l)^\omega$ yield $\lambda - 2$ powers (the $\lambda - 1$ adds are unnecessary), since 0^ω and 1^ω are direct. For $2 \leq l \leq \lambda$, the l^d are gathered in a $(\delta + 1) \times \lambda$ array, with $\sum_{d=2}^{\delta} d = \frac{\delta(\delta+1)}{2} - 1$ products. The $(\lambda - l)^\omega l^d$ terms can be stored in a $(\delta + 1) \times \lambda$ matrix, involving $(\lambda - 2)(\delta - 1)$ products. These estimates are stored in the top-half of Table 4, above the dashed line.

Upon hopping frame update and τ determination, using the simplification $\tau^{d+1} = \tau\tau^d$, $\boldsymbol{\tau}_\delta$ requires $\delta - 1$ products. The evaluation of the weighted moments entails only $\delta(\lambda - 1)$ adds, and $(\lambda - 2)\delta + 1$ products, since the lazy matrix storing $(\lambda - l)^\omega l^d$ contains some zeroes and ones (Table 4, bottom-half). Both Type I and II are thus roughly quadratic in (δ, λ). Both are competitive with respect to Lagrange extrapolation, which requires $O\left(\delta^2\right)$ or $O(\delta)$ with $\delta = \lambda + 1$ depending on the implementation.

If we precisely compute the operation excess from Type I to II, we obtain $\Xi(\delta, \lambda) = 2\delta^2 + \delta + 3 - 2\lambda$. For the parameters given in A.4, we have for instance $\Xi(0, 2) = -1$, $\Xi(1, 3) = 0$, $\Xi(2, 5) = 3$.

Operations	+	×	power	+	×
$(\lambda - l)^\omega$			$\lambda - 2$		
l^d		$\delta(\delta-1)/2$			
$(\lambda - l)^\omega l^d$		$(\lambda - 2)(\delta - 1)$			
$\boldsymbol{\tau}_\delta$		$\delta - 1$			$\delta - 1$
$\overline{\boldsymbol{m}}_{\delta,\lambda,\omega}$	$\delta(\lambda - 1)$	$(\lambda - 2)\delta + 1$			
$\overline{\boldsymbol{Z}}_{-\delta,\lambda,\omega}\overline{\boldsymbol{m}}_{\delta,\lambda,\omega}/\boldsymbol{\Pi}_{\delta,\lambda,\omega}\boldsymbol{u}_\lambda$	$\delta(\delta+1)$	$(\delta+1)^2$		$(\delta+1)(\lambda-1)$	$(\delta+1)\lambda$
$\boldsymbol{\tau}_\delta \times \cdots$	δ	δ		δ	δ
Leading orders	Type I: $2(\delta\lambda + \delta^2)$			Type II: $2\delta\lambda$	

Table 4: Elementary operations required for $u(\tau)$ in Type I and II implementations. Top: lazy evaluation (computed once). Bottom: required for each hopping frame.

A.4 Examples for Type II CHOP$_{\text{oly}}$ implementation

Table 5 provides examples of predictor matrices $\boldsymbol{\Pi}_{\delta,\lambda,\omega}$ (*cf.* (8)) of fixed degree δ and length λ, for different integers and rational powers $\omega \in \left\{0, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1, 2\right\}$.

ω	$\mathbf{\Pi}_{0,2,\omega}$	$\mathbf{\Pi}_{1,3,\omega}$	$\mathbf{\Pi}_{2,5,\omega}$
0	$\frac{1}{2}\begin{bmatrix} 1 & 1 \end{bmatrix}$	$\frac{1}{6}\begin{bmatrix} 5 & 2 & -1 \\ 3 & 0 & -3 \end{bmatrix}$	$\frac{1}{70}\begin{bmatrix} 62 & 18 & -6 & -10 & 6 \\ 54 & -13 & -40 & -27 & 26 \\ 10 & -5 & -10 & -5 & 10 \end{bmatrix}$
$\frac{1}{8}$	$\begin{bmatrix} 0.52 & 0.48 \end{bmatrix}$	$\begin{bmatrix} 0.84 & 0.31 & -0.16 \\ 0.51 & -0.02 & -0.49 \end{bmatrix}$	$\begin{bmatrix} 0.89 & 0.25 & -0.09 & -0.13 & 0.08 \\ 0.79 & -0.21 & -0.58 & -0.35 & 0.36 \\ 0.15 & -0.08 & -0.14 & -0.06 & 0.14 \end{bmatrix}$
$\frac{1}{4}$	$\begin{bmatrix} 0.54 & 0.46 \end{bmatrix}$	$\begin{bmatrix} 0.85 & 0.30 & -0.15 \\ 0.52 & -0.05 & -0.48 \end{bmatrix}$	$\begin{bmatrix} 0.90 & 0.23 & -0.09 & -0.12 & 0.07 \\ 0.80 & -0.24 & -0.58 & -0.32 & 0.34 \\ 0.15 & -0.09 & -0.14 & -0.05 & 0.13 \end{bmatrix}$
$\frac{1}{2}$	$\begin{bmatrix} 0.59 & 0.41 \end{bmatrix}$	$\begin{bmatrix} 0.87 & 0.26 & -0.13 \\ 0.55 & -0.10 & -0.45 \end{bmatrix}$	$\begin{bmatrix} 0.91 & 0.21 & -0.09 & -0.09 & 0.06 \\ 0.83 & -0.30 & -0.58 & -0.26 & 0.31 \\ 0.16 & -0.10 & -0.15 & -0.04 & 0.13 \end{bmatrix}$
1	$\frac{1}{3}\begin{bmatrix} 2 & 1 \end{bmatrix}$	$\frac{1}{10}\begin{bmatrix} 9 & 2 & -1 \\ 6 & -2 & -4 \end{bmatrix}$	$\frac{1}{140}\begin{bmatrix} 130 & 24 & -12 & -8 & 6 \\ 125 & -60 & -80 & -20 & 35 \\ 25 & -20 & -20 & 0 & 15 \end{bmatrix}$
2	$\frac{1}{5}\begin{bmatrix} 4 & 1 \end{bmatrix}$	$\frac{1}{38}\begin{bmatrix} 36 & 4 & -2 \\ 27 & -16 & -11 \end{bmatrix}$	$\frac{1}{10164}\begin{bmatrix} 9750 & 1056 & -684 & -144 & 186 \\ 10375 & -7216 & -5028 & 368 & 1501 \\ 2275 & -2464 & -1176 & 644 & 721 \end{bmatrix}$

Table 5: Type II matrices $\mathbf{\Pi}_{0,2,\omega}$, $\mathbf{\Pi}_{1,3,\omega}$ and $\mathbf{\Pi}_{2,5,\omega}$, with integer and rational weighting powers $\omega \in \left\{0, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, 1, 2\right\}$.

References

[1] M. Valášek, Modeling, simulation and control of mechatronical systems, in: Arnold and Schiehlen [12], pp. 75–140. 2

[2] T. Blochwitz et al., The Functional Mockup Interface for tool independent exchange of simulation models, in: C. Clauß (Ed.), Proc. Int. Modelica Conf., Linköping Electronic Conference Proceedings, Linköping Univ. Electronic Press, Dresden, Germany, 2011. doi:10.3384/ecp11063105. 2, 4

[3] A. Ben Khaled-El Feki, Distributed real-time simulation of numerical models: application to powertrain, Ph.D. thesis, Université de Grenoble (May 2014).
URL https://tel.archives-ouvertes.fr/tel-01144469 2, 6, 16

[4] M. Sjölund et al., Towards efficient distributed simulation in Modelica using Transmission Line Modeling, in: 3rd Int. Workshop on Equation-Based Object-Oriented Languages and Tools EOOLT, Linköping Univ. Electronic Press, Oslo, Norway, 2010, pp. 71–80. 2

[5] A. Ben Khaled et al., Multicore simulation of powertrains using weakly synchronized model partitioning, in: IFAC Workshop on Engine and Powertrain Control Simulation and Modeling ECOSM, Rueil-Malmaison, France, 2012, pp. 448–455. doi:10.3182/20121023-3-FR-4025.00018. 2

[6] A. Ben Khaled et al., Fast multi-core co-simulation of cyber-physical systems: Application to internal combustion engines, Simul. Model. Pract. Theory 47 (2014) (2014) 79–91. doi:http://dx.doi.org/10.1016/j.simpat.2014.05.002.
URL http://www.sciencedirect.com/science/article/pii/S1569190X14000665 3

[7] T. Schierz, M. Arnold, C. Clauß, Co-simulation with communication step size control in an FMI compatible master algorithm, in: Proc. Int. Modelica Conf., Linköping Electronic Conference Proceedings, Linköping University Electronic Press, Munich, Germany, 2012, pp. 205–214. doi:10.3384/ecp12076205. 3

[8] M. Arnold, Stability of sequential modular time integration methods for coupled multibody system models, J. Comput. Nonlinear Dynam. 5 (3) (2010). doi:10.1115/1.4001389. 3

[9] M. Arnold, Numerical methods for simulation in applied dynamics, in: Arnold and Schiehlen [12], pp. 191–246. 3, 6, 7

[10] A. Ben Khaled et al., Context-based polynomial extrapolation and slackened synchronization for fast multi-core simulation using FMI, in: H. Tummescheit, K.-E. Årzén (Eds.), Proc. Int. Modelica Conf., Linköping Electronic Conference Proceedings, Linköping University Electronic Press, Lund, Sweden, 2014, pp. 225–234. 3, 8, 14, 17, 20

[11] F. Zhang, M. Yeddanapudi, P. Mosterman, Zero-crossing location and detection algorithms for hybrid system simulation, in: M. J. Chung, P. Misra (Eds.), 17th IFAC World Congress, Seoul, South Korea, 2008, pp. 7967–7972. doi:10.3182/20080706-5-KR-1001.01346. 4

[12] M. Arnold, W. Schiehlen (Eds.), Simulation Techniques for Applied Dynamics, Vol. 507 of CISM Courses and Lectures, Springer, 2008. 7, 28

[13] M. J. Weinberger, G. Seroussi, G. Sapiro, The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS, IEEE Trans. Image Process. 9 (8) (Aug. 2000) (2000) 1309–1324. doi:10.1109/83.855427.
URL http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=855427 7

[14] N. Wiener, Extrapolation, interpolation, and smoothing of stationary time series, MIT Press, 1949. 7

[15] E. Meijering, A chronology of interpolation: from ancient astronomy to modern signal and image processing, Proc. IEEE 90 (3) (Mar. 2002) (2002) 319–342. doi:10.1109/5.993400. 7

[16] R. G. Brown, Smoothing, Forecasting and Prediction of Discrete Time Series, Prentice-Hall, 1962. 7

[17] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, Time Series Analysis: Forecasting and Control, Probability and Statistics, Wiley, 2008. 7

[18] S. Väliviita, S. Ovaska, O. Vainio, Polynomial predictive filtering in control instrumentation: a review, IEEE Trans Ind. Electron. 46 (5) (Oct. 1999) (1999) 876–888. doi:10.1109/41.793335. 7

[19] L. F. Richardson, The approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam, Phil. Trans. R. Soc. A 210 (459-470) (Jan. 1911) (1911) 307–357. doi:10.1098/rsta.1911.0009.
URL http://dx.doi.org/10.1098/rsta.1911.0009 7

[20] M. Arnold, C. Clauß, T. Schierz, Error analysis and error estimates for co-simulation in FMI for model exchange and co-simulation V2.0, Arch. Mech. Eng. LX (1) (Jan. 2013). doi:10.2478/meceng-2013-0005.
URL http://dx.doi.org/10.2478/meceng-2013-0005 8

[21] G. Stettinger et al., A model-based approach for prediction-based interconnection of dynamic systems, in: Proc. IEEE Conf. Decision Control, Los Angeles, CA, USA, 2014, pp. 3286–3291. doi:10.1109/cdc.2014.7039897.
URL http://dx.doi.org/10.1109/CDC.2014.7039897 8

[22] M. G. Bellanger, J. L. Daguet, G. P. Lepagnol, Interpolation, extrapolation, and reduction of computation speed in digital filters, IEEE Trans. Acous., Speech Signal Process. 22 (4) (Aug. 1974) (1974) 231–235. doi:10.1109/TASSP.1974.1162581.
URL http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1162581 8

[23] J. Gauthier, L. Duval, J.-C. Pesquet, Optimization of synthesis oversampled complex filter banks, IEEE Trans. Signal Process. 57 (10) (Oct. 2009) (2009) 3827–3843. doi:10.1109/TSP.2009.2023947. 8

[24] V. C. Liu, P. P. Vaidyanathan, Finite length band-limited extrapolation of discrete signals, in: Proc. Int. Symp. Circuits Syst., Vol. 2, Portland, OR, USA, 1989, pp. 1037–1040. doi:10.1109/iscas.1989.100529.
URL http://dx.doi.org/10.1109/ISCAS.1989.100529 8

[25] M. Benedikt, D. Watzenig, A. Hofer, Modelling and analysis of the non-iterative coupling process for co-simulation, Math. Comp. Model. Dyn. Syst. 19 (5) (Oct. 2013) (2013) 451–470. doi:10.1080/13873954.2013.784340.
URL http://dx.doi.org/10.1080/13873954.2013.784340 8

[26] R. W. Hamming, Numerical methods for scientists and engineers, Dover publications, 1973. 8, 9

[27] M. Friedrich, Parallel co-simulation for mechatronic systems, Ph.D. thesis, Technischen Universität München (2011). 8

[28] M. Arnold, Multi-rate time integration for large scale multibody system models, in: IUTAM Symposium on Multiscale Problems in Multibody System Contacts, 2007, pp. 1–10. doi:10.1007/978-1-4020-5981-0_1.
URL http://dx.doi.org/10.1007/978-1-4020-5981-0_1 8

[29] S. Hoher, S. Röck, A contribution to the real-time simulation of coupled finite element models of machine tools — a numerical comparison, Simul. Model. Pract. Theory 19 (7) (aug 2011) (2011) 1627–1639. doi:10.1016/j.simpat.2011.03.002.
URL http://dx.doi.org/10.1016/j.simpat.2011.03.002 8

[30] J. Nutaro et al., The split system approach to managing time in simulations of hybrid systems having continuous and discrete event components, Simul. T. Soc. Mod. Sim. 88 (3) (May 2012) (2012) 281–298. doi:10.1177/0037549711401000.
URL http://dx.doi.org/10.1177/0037549711401000

[31] M. Busch, Zur effizienten Kopplung von Simulationsprogrammen, Ph.D. thesis, Universität Kassel (2012). 8

[32] S. Oh, S. Chae, A co-simulation framework for power system analysis, Energies 9 (3) (feb 2016) (2016) 131. doi:10.3390/en9030131.
URL http://dx.doi.org/10.3390/en9030131 8

[33] M. Busch, Continuous approximation techniques for co-simulation methods: Analysis of numerical stability and local error, ZAMM (2016). doi:10.1002/zamm.201500196.
URL http://dx.doi.org/10.1002/zamm.201500196 8

[34] C. Faure, Real-time simulation of physical models toward hardware-in-the-loop validation, Ph.D. thesis, Université Paris-Est, France (Oct. 17, 2011). 8

[35] B. Beckermann, E. B. Saff, The sensitivity of least squares polynomial approximation, in: Applications and Computation of Orthogonal Polynomials, Vol. 131 of Int. Ser. of Num. Math., Birkhäuser, 1999, pp. 1–19, conference at the Mathematical Research Institute Oberwolfach, Germany. 9

[36] E. S. Gardner, Jr., Exponential smoothing: The state of the art — part II, Int. J. Forecast. 22 (4) (2006) (2006) 637–666. doi:DOI:10.1016/j.ijforecast.2006.03.005.
URL http://www.sciencedirect.com/science/article/pii/S0169207006000392 9

[37] E. L. Kaltofen, W.-S. Lee, Z. Yang, Fast estimates of Hankel matrix condition numbers and numeric sparse interpolation, in: Proc. Int. Workshop Symbolic-Numeric Computation, San Jose, California, USA, 2011, pp. 130–136. doi:10.1145/2331684.2331704.
URL http://dx.doi.org/10.1145/2331684.2331704 11

[38] H. Wang, L. Chen, Y. Hu, A state event detecting algorithm for hybrid dynamic systems, Simul. T. Soc. Mod. Sim. 91 (11) (Nov. 2015) (2015) 959–969. doi:10.1177/0037549715606968.
URL http://dx.doi.org/10.1177/0037549715606968

[39] J. M. Peña, T. Sauer, On the multivariate Horner scheme, SIAM J. Numer. Anal. 37 (4) (2000) (2000) 1186–1197. doi:10.1137/s0036142997324150.
URL http://dx.doi.org/10.1137/S0036142997324150 11

[40] Z. Benjelloun-Touimi et al., From physical modeling to real-time simulation: Feedback on the use of Modelica in the engine control development toolchain, in: Proc. Int. Modelica Conf., Linköping Electronic Conference Proceedings, Linköping Univ. Electronic Press, Dresden, Germany, 2011. doi:10.3384/ecp11063. 16

[41] P. Fritzson, Principles of object-oriented modeling and simulation with Modelica 2.1, Wiley, 2010. 16

[42] M. Ben Gaïd et al., Heterogeneous model integration and virtual experimentation using xMOD: Application to hybrid powertrain design and validation, in: 7th EUROSIM Congress on Modeling and Simulation, Prague, Czech Republic, 2010. 16

[43] A. C. Hindmarsh, L. R. Petzold, Algorithms and software for Ordinary Differential Equations and Differential-Algebraic Equations, part II: Higher-order methods and software packages, Comput. Phys. 9 (1995) (1995) 148–155. 17

[44] D. Broman et al., Requirements for hybrid cosimulation, Tech. Rep. UCB/EECS-2014-157, Electrical Engineering and Computer Sciences, University of California at Berkeley (Aug. 16, 2014).
URL http://www.eecs.berkeley.edu/Pubs/TechRpts/2014/EECS-2014-157.html 22

[45] X. Ning, I. W. Selesnick, L. Duval, Chromatogram baseline estimation and denoising using sparsity (BEADS), Chemometr. Intell. Lab. Syst. 139 (Dec. 2014) (2014) 156–167. doi:10.1016/j.chemolab.2014.09.014. 23

[46] F. González et al., On the effect of multirate co-simulation techniques in the efficiency and accuracy of multibody system dynamics, Multibody Syst. Dyn. 25 (4) (Dec. 2010) (2010) 461–483. doi:10.1007/s11044-010-9234-7.
URL http://dx.doi.org/10.1007/s11044-010-9234-7 23

[47] C. Chaux, J.-C. Pesquet, L. Duval, Noise covariance properties in dual-tree wavelet decompositions, IEEE Trans. Inform. Theory 53 (12) (Dec. 2007) (2007) 4680–4700. doi:10.1109/TIT.2007.909104. 23

[48] A. Repetti et al., Euclid in a taxicab: Sparse blind deconvolution with smoothed ℓ_1/ℓ_2 regularization, IEEE Signal Process. Lett. 22 (5) (May 2015) (2015) 539–543. arXiv:1407.5465, doi:10.1109/LSP.2014.2362861.
URL http://dx.doi.org/10.1109/LSP.2014.2362861 23

[49] T. Schierz, M. Arnold, C. Clauß, Co-simulation with communication step size control in an FMI compatible master algorithm, in: Proc. Int. Modelica Conf., Linköping Electronic Conference Proceedings, Munich, Germany, 2012, pp. 205–214. doi:10.3384/ecp12076205. 23

[50] S. Sadjina et al., Energy conservation and power bonds in co-simulations: Non-iterative adaptive step size control and error estimation, PREPRINT (Feb. 2016). 23

[51] S. M. Stigler, Gauss and the invention of least squares, Ann. Statist. 9 (3) (1981) (1981) 465–474.
URL http://projecteuclid.org/euclid.aos/1176345451 24

[52] G. F. C. de Bruyn, J. M. de Villiers, Formulas for $1 + 2^p + 3^p + \ldots + n^p$, Fibonacci Q. 32 (3) (1994) (1994) 271–276. 24

[53] T. N. E. Greville, Some applications of the pseudoinverse of a matrix, SIAM Rev. 2 (1) (Jan. 1960) (1960) 15–22. doi:10.1137/1002004.
URL http://dx.doi.org/10.1137/1002004 25