TREATING SMOOTHNESS AND BALANCE DURING DATA EXCHANGE IN EXPLICIT SIMULATOR COUPLING OR COSIMULATION

DIRK SCHARFF, THILO MOSHAGEN, JAROSLAV VONDŘEJČ

Abstract. Cosimulation methods allow combination of simulation tools of physical systems running in parallel to act as a single simulation environment for a big system. As data is passed across subsystem boundaries instead of solving the system as one single equation system, it is not ensured that systemwide balances are fulfilled. If the exchanged data is a flow of a conserved quantity, approximation errors can accumulate and make simulation results inaccurate. The problem of approximation errors is typically addressed with extrapolation of exchanged data. Nevertheless balance errors occur as extrapolation is approximation. This problem can be handled with balance correction methods which compensate these errors by adding corrections for the balances to the signal in next coupling time step. This work aims at combining extrapolation of exchanged data and balance correction in a way that the exchanged signal not only remains smooth, meaning the existence of continuous derivatives, but even in a way reducing the derivatives, in order to avoid unphysical dynamics caused by the coupling. To this end, suitable switch and hat functions are constructed and applied to the problem.

Cosimulation, balance correction, extrapolation of signals, smoothing, error control, approximation error

1. Introduction

Engineers are increasingly relying on numerical simulation techniques. Models and simulation tools for various physical problems have come into existence in the past decades. The desire to simulate a system that consists of well described and treated subsystems by using appropriate solvers for each subsystem and letting them exchange the data that forms the mutual influence is immanent. The situation usually is described by two coupled differential-algebraic
systems $S_1$ and $S_2$ that together form a system $S$:

$$S_1:$$
$$\dot{x}_1 = f_1(x_1, x_2, z_1, z_2)$$
$$0 = g_1(x_1, x_2, z_1, z_2)$$

$$S_2:$$
$$\dot{x}_2 = f_2(x_1, x_2, z_1, z_2)$$
$$0 = g_2(x_1, x_2, z_1, z_2).$$

The $(x_1, x_2)$ are the differential states of $S$, their splitting into $x_i$ determines the subsystems $S_i$ together with the choices of the $z_i$. In Co-Simulation the immediate mutual influence of subsystems is replaced by exchanging data at fixed time points and subsystems are solved separately and parallelly but using the received parameter:

$$S_1:$$
$$\dot{x}_1 = f_1(x_1, z_1, u_2)$$
$$0 = g_1(x_1, z_1, u_2)$$

$$S_2:$$
$$\dot{x}_2 = f_2(x_2, z_2, u_1)$$
$$0 = g_2(x_2, z_2, u_1)$$

where $u_i$ are given by coupling conditions that have to be fulfilled at exchange times $T_k$

$$0 = h_1(x_1, z_1, u_1)$$
$$0 = h_2(x_2, z_2, u_2)$$

and are not dependent on subsystem $i$’s states any more, so are mere parameters between exchange time steps.

Full row rank of $dz_i, g_i$ can be assumed, such that the differential-algebraic systems are of index 1. This description of the setting is widespread ([1]). With the $h_i$ being solved for $u_i$ inside the $S_i$ (let solvability be given), for systems with more than two subsystems is more convenient to write output variable $y_i$ instead of $u_i$ and now redefine $u_j$ as the input of $S_j$, consisting of some components of the outputs $(y_i)_{i=1,n}$ [2]. This structure is defined as kind of a standard for connecting simulators for cosimulation by the Functional Mockup Interface Standard [3]. It defines clearly what information a subsystems implementation provides. From chapter 2 on, we use this notation. Mind input $u_j$ is then indexed with its subsystems index.

In Co-Simulation the variables establishing the mutual influence of subsystems are exchanged at fixed time points. This results in continuous variables being approximated by piecewise constant extrapolation, as shown in the following picture:
If one does not want to iterate on those inputs by restarting the simulations using the newly calculated inputs, one just proceeds to the next timestep. This gives the calculations an explicit component, the mutual influence is now not immediate any more, inducing the typical stability problems, besides the approximation errors.

But for good reasons, co-simulation is a widely used method: It allows to put separate submodels, for each of which a solver exists, together into one system and simulate that system by simulating each subsystem with its specialised solver - examination of mutual influence becomes possible without rewriting everything into one system, and simulation speed benefits from the parallel calculation of the submodels.

The following fields of work on explicit co-simulation can be named to be the ones of most interest:

1. Improvement of the approximation of the exchanged data will most often improve simulation results [4]. This is usually done by higher-order extrapolation of exchanged data, as shown in this plot, where the function plotted with dots is linearly extrapolated:

   ![Figure 1. Constant extrapolation of an input signal](image)

   **Figure 1.** Constant extrapolation of an input signal

2. When the mutual influence between subsystems consists of flow of conserved quantities like mass or energy, it turns out that the improvement of the approximation of this influence by extrapolation of past data is not sufficient to establish the conservation of those quantities with the necessary accuracy. The error that arises from the error in exchange adds up over time and becomes obvious (and lethal to simulation results many times). In a cooling cycle example ([5, Section 6.3]), a gain of 1.25% in coolant mass occurs when simulating a common situation. It
has been tried to meet this challenge by passing the amount of exchanged quantity for the past timestep along with the actual flow on to the receiving system, where then the error that has just been committed is calculated and added to the current flow to compensate the past error. For well damped example problems in fluid circles this method has fulfilled the expectations \cite{6}. It has been labelled balance correction.

(3) There is good reason to prevent jumps in exchanged data by smoothing. Higher order extrapolation polynomials cannot make extrapolated data at the end of the exchange timestep match the newly given value.

While under all circumstances it is desirable to guess the influence information at the subsystem boundaries from past data as well as possible, balance correction techniques bear the profound problem that for establishing conservation \textit{a posteriori} over the whole time an instantaneous error in the exchanged data has to be accepted. More precisely, balance correction means making an error in the amount of a quantity \( \int_{t_i}^{t_{i+1}} u(\tau) d\tau \) that is exchanged between subsystems during the interval \([t_i, t_{i+1})\), \(u\) being the exchanged signal, for the purpose of lowering the error in states \(x\) that would be caused by this accumulated and now persistent error in amount of \(\int u d\tau\). This paper reveals the extend of this problem and proposes a way to reduce the time delay.

Of course, also an error in the derivatives of the exchanged data is made, which might cause dynamics in the receiving system and its neighbours. The presented work prepares cosimulation to control the error in the derivatives by construction and use of suitable functions for smoothing during switching and adding of correction terms.

2. Preliminaries

2.1. Notation. Let information exchange times start with \(t_0\) and end with \(t_n\) and time intervals of exchange \([t_{i-1}, t_i)\) be indexed from 1 to \(n\), such that \(\Delta_i x := x(t_i) - x(t_{i-1})\) for any variable \(x\).

Input variables in general are denoted by \(u\), algebraic and differential states by \(z\) and \(x\), respectively. Lower indices indicate vector indices. \(^{3}\)Exchange at calculations end could be relevant as data determines algebraic variables at end of simulation, so \(t_n\) is also the end of simulation.
here the subsystem the components belong to (a slight change from indexing $u$ in eq. (9)-(10)), upper indices indicate the restriction of the function to the time interval, e.g. $u^j_i$ denotes the function $u_i : [t_{j-1}, t_j] \rightarrow \mathbb{R}^{n_i}$.

Within this paper, concepts are extended from flows of conserved quantities to general input treatment. Conserved variables are not distinguished in notation, conservation is mentioned in text if given. Input variables $u$ extrapolation is denoted by $\text{Ext}u$, sometimes referred to as hold function, thus a guess of $u^j_i$ calculated from the known $u^j_j$, $j < i$, while $\overline{u^j}$ be the signal as it is used for calculation constructed on $[t_{j-1}, t_j)$, here a smooth combination of $\text{Ext}u^j_{i-1}$ and $\text{Ext}u^j_i$, referred to as the smooth extrapolation. Let $\Delta E_j = \int_{t_{j-1}}^{t_j} u_i dt - \int_{t_{j-1}}^{t_j} \overline{u_i} dt$ be the error of the exchanged quantity itself.

2.2. Choice of exchanged Variables. The direction in which the variables are exchanged cannot be chosen arbitrarily. Careful consideration is needed to not place contradictorily constraints on the models. If i.e. at one of the model boundaries $S_1$ exchanges the value of a differential state, it must only export and not receive this value – otherwise this state is turned into a parameter and extrapolated instead of solved, or one would overwrite an integration result and reinitialize the solver at each exchange. The choice is to receive arguments of the derivative of the state and send the state value itself [6, 5]. Often, the physics of coupling is described by a flow driven by a potential, where the flux depends on the potential by $f \sim \nabla x P$, or in non-spatial context $f = \text{const} [P_{S_2} - P_{S_1}]$. Then exchange is commonly implemented in the following way: System $S_1$ passes the value of the flow-determining potential $P$ on to $S_2$, whereas the latter calculates the flux $f$ and passes it to $S_1$.

2.3. Detailed description of Techniques in coupling Simulation Software. After the brief overview in the introduction on the issues that appear when coupling simulation software and means to treat them was given. Those fields are described more precisely here, giving citations.

2.3.1. Determination of consistent Inputs. As $u_i$ fulfils $0 = h_i(x_i, z_i, u_j)$, where $z_i$ results from $0 = g_i(x_i, z_i, u_j)$ it in general depends on $u_j$.
So the coupling equations \[0 = h_i(x_i, z_i, u_j)\] (9)-(10) $(h_i)_{i=1..n} = 0$ are a coupled system in fact. The $0 = h_i$ can be solved with respect to the $u_i$ one after the other only if the directed graph of influence of outputs on each other contains neither bidirectional dependencies nor loops. Otherwise, the coupling conditions remain not fully fulfilled, or one solves them iteratively. Such methods are the Interface-Jacobian based methods, e.g. given in [7], which determine the $u_i$ with respect to $h_i = 0$ and $g_i = 0$ with the help of a newton solver at exchange times, which means full
consistency at those times. But not all tools simulating the subsystems provide the residual of $g_i = 0$.

2.3.2. **Iterative coupled Methods.** If restarting the timestep is an option, iterations can be performed and instead of using $\text{Ext}(u_i)$, the time-dependant numerical solution for $u_i$ from the old iteration can be used. Schemes are distinguished by the exchange directions of input, as the waveform iteration or the Gauss-Seidel scheme, which can exploit the directions of influences between the subsystems ([8], [1], [4]). Stability and accuracy, also in terms of balance, can be augmented, of course at computational cost due to the restarts.

In [9], a Newton method is used to solve the coupling equations (9)-(10), but different from [7] not only $(h_i(x_i(T_k), z_i, u_i))_{i=1..n}{\mathbf{0}}$ at the end of the timestep, but $(h_i(x_i(u_i), z_i, u_i))_{i=1..n} = 0$. This requires repeated solving of the subsystems alone for Jacobian evaluation.

2.3.3. **Increasing the extrapolation order of inputs.** To improve approximation $\text{Ext} u_i$ has been calculated as an extrapolation polynomial of past $u_j$, $j < i$. In his dissertation [4] Busch examines Lagrangian and Hermite extrapolation polynomials for exchanged quantities using a system concatenated from two coupled spring-mass oscillators, thus a linear second order ODE with four degrees of freedom, as model problem.

2.3.4. **Investigations of convergence.** Busch examines the effect of the approximation order in exchanged data on global convergence and judges the stability of the method for his problem, assuming exact integration of each subsystem. Convergence to exact solution in this case is limited to $O(h^2)$ for piecewise constant extrapolation, while it is limited to $O(h^3)$ for degree one Hermite polynomial extrapolation. As Busch examines a system of ordinary differential equations without algebraic parts and given explicitly, he feels no need to consider smoothing. Balance errors that occur in his problem remain undiscussed. Besides the aforementioned examination [4] which is limited to linear problems, in [1] there is a more general examination, treating iterative coupling schemes. Mainly, a fixed point argument is used here.

2.3.5. **Smooth switching between old and new input.** The smoothing of exchanged data is originally motivated by the need to provide a starting value close to the solution during searching for solutions of the algebraic part of the differential-algebraic equation system. Another reason for smoothing the inputs can be that it enables the calculation of derivatives resp. difference quotients from them. Although this is unfortunate, such needs sometimes occur in practice.

In [5], two halves of a cooling cycle are co-simulated, the mass flow $f_m$ being exchanged, so a component of $u$. The smooth input $\overline{u}_i$ is
concatenated as a convex sum of $\text{Ext} \ u_{i}^{j-1}$ and $\text{Ext} \ u_{j}^{j}$ weighted with a sufficiently smooth, here degree 5 polynomial, function switching from 0 at $T_{j-1}$ to 1 at $T_{j}$. Besides from stabilizing the Newton solver as intended, Kossels work gives an impression on balance errors: During some simulations balance error of about 1.25% in coolant mass is observed.

It is obvious that by this smoothed switching the balance error $\Delta E_{i}^{j}$ possibly increases compared to a unsmoothed switching procedure. Another work that considers smoothing is the paper [10].

2.3.6. Balance Correction. Classical balance errors of extrapolations is

$$\Delta E_{i}^{k} = \int_{T_{k-1}}^{T_{k}} u_{i} - \overline{(u)_{i}}(t)dt,$$

where $u_{i}$ is a flux of a conserved quantity, but it is defined for arbitrary quantities. Negative and positive contributions from the intervals partly compensate each other, and as in a typical simulation the system often ends up in a stationary state similar to the one at begin, thus graphs of quantities and their derivatives tend to end at values where they started, the conserved quantities balance error at the end of a simulation may be small. This does not imply it is small during all intervals of the simulation. As a typical such simulation situation think of an automotive driving cycle ([5]):

Using piecewise constant extrapolation of $f$, there is a gap in mass after the phase of rising system velocity has passed - this loss remains uncompensated during the (significant) phase of elevated speed. Scharff, motivated by the loss in balance stated in [5] as a collateral result, proposed in [6] that the errors

$$\Delta E_{i}^{j} = \int_{T_{j-1}}^{T_{j}} u_{i} dt - \int_{T_{j-1}}^{T_{j}} \overline{u_{i}} dt$$

are added in the time step $j + 1$. The correction that is applied in the $j$-th interval is then $\Delta E_{i}^{j-1} \phi_{j}(t)$, where the function $\phi_{j}$ is scaled such that its integral is 1 and may be constant but when one wants to preserve smoothness, it is a smooth function smoothly vanishing at the boundaries of the time interval $j$. In spite of the lack of strictness and the increase of errors in derivatives of the exchanged quantities, this method enables coupling of simulations that would be impossible without - Kossels example was recalculated successfully using balance correction.

3. Smooth Balancing and Stability

As a first generalisation, this paper uses that also nonconserved quantities can be balanced over short times. It makes sense to apply for example an amount of force that has been omitted due to extrapolation error in the next timestep.
Figure 4. An input signal and ways it might be treated by the receiving system: green: linearly interpolated, red: smoothed with s-shaped $\psi$, black: error from previous time step added to smooth signal. In the lower plot, the contributions themselves are plotted: Black dotted: error of linear extrapolation from previous timestep, blue: balance error due to smooth (slow) switching in actual timestep, black: sum of balance error from previous and switching error from previous step. It is obvious that switching error cannot be neglected, and that derivatives of the input signal oscillate.

If the algebraic equations $g_i = 0$ are solvable w.r.t. $z_i$, the systems can be solved by state space method, this is determining $z_i$ with a Newton solver at each evaluation of the ODE. Such system is then solved as an ODE, why in the following we consider only ODE. The right scheme below gives the cosimulation scheme for the coupled problem on the left:
Each method described in the preceding section provides relief for a certain problem that arises when applying the standard cosimulation scheme: In a now widened field of physical problems, this paper names simulation settings in which the problems balance and stability, smoothness and subsequent oscillations, and both of them arise, and combines balance correction and smoothing for solving them, examines the results and suggests an early refeed of early-known balance error contributions to tackle the delaying effect of smooth switching. It makes suggestions for further reducing the exchange-induced oscillations in receiving systems.

Smooth switching as in [5] is done using an S-shaped function \( \psi(t) \) for switching between extrapolation polynomials \( \text{Ext} u_{j-1} ^i \) and \( \text{Ext} u_j ^i \) such that the smoothed input is

\[
(12) \quad u_i ^j(t) = (1 - \psi(t)) \text{Ext} u_{j-1} ^i(t) + \psi(t) \text{Ext} u_j ^i(t).
\]

The ODE in above scheme then is \( \dot{x} = f_1(x_1, \overline{u}_i) \). If balance correction contributions \( \Delta E_i ^{j-1} \phi_j(t) \) as in Section 2.3.6 are added to such smooth \( \overline{u}_i \), the function \( \phi(t) \) should not be constant but a hat such that it is smooth on the boundaries of the interval, i.e. that \( \dot{\phi}^{(n)}(t) = 0 \) at \( t_j \) and \( t_{j-1} \) for some, better for all \( n \), in order not to spoil the smoothness of the \( \overline{u}_i \). Implementations of \( \phi \) and \( \psi \) are presented in Section 5. A sketch of the input signals used during combination of methods is given in figure 4.

For solving the ODE inside each subsystem, in this contribution pythons vode and zvode are used, methods relying on implicit Adams method and backward differentiation formulas (BDF). For both methods, a stepwidth can be found that conserves stability of linear problem, and the internal step size control does so, so subsystems methods behave like stable methods.
3.1. **Spring and Mass Example and its Implications.** Consider the linear mechanical oscillator given by

\[
\dot{x} = Ax = \begin{pmatrix} 0 & 1 \\ -\frac{c}{m} & -\frac{d}{m} \end{pmatrix} x, \quad x = \begin{pmatrix} x \\ \dot{x} \end{pmatrix}
\]

with mass \( m = 1 \), spring constant \( c = 1 \), and in which the damping constant \( d \) shall vanish.

None of its states is a conserved quantity in the view of this system.

3.1.1. **Cosimulation.** The system is treated as subsystems in the cosimulation scheme given in Table 1, an implementation of the standard scheme given in the beginning of Section 3.

| Spring System States | Mass System States |
|-----------------------|---------------------|
| \( x_1 := s = x \)   | \( x_2 := v = \dot{x} \) |
| Outputs               |                      |
| \( y_1 := F = -cx \) | \( y_2 := v = \dot{x} \) |
| Inputs                |                      |
| \( u_1 := y_2 \)     | \( u_2 := y_1 \)    |
| Equations             |                      |
| \( \dot{x}_1 = \text{Ext}(u_1) \) | \( \dot{x}_2 = -\frac{\text{Ext}(u_2)}{m} \) |
|                       | \( = v \)             | \( = \frac{F}{m} \) |

**Table 1.** Cosimulation Schemes for the spring-mass system, left constant, right linear extrapolation

Output of the spring is the force \( F = -cx \), that of the mass is the displacement \( s = x \), and these quantities are again non-conserved. Note this problem is the simplest splittable problem possible. It has the least number of states that can be splitted, the splitted systems depend only on input, and in the most simple way, which is linear.

Whereas the system is easily simulated in whole using an A-stable method, when simulated with the cosimulation scheme as above, the system picks up energy (figures 5 and 6) as the exchanged quantities force and displacement are factors of it and the approximation error of those factors cause an error (here: rise) in the systems energy as will be treated in section 3.2.2.

The balance correction algorithm (see Section 2.3.6) delivers a better result (figures 7 and 8), but as most of the quantity missing in one exchange step is refeeded with delay, the energy flux from the spring...
Figure 5. Simulation of the system (13) in the cosimulation scheme with constant extrapolation, varying the exchange step size $H$. Upper row, left: $H = 0.2$, right: $H = 0.1$, lower row: left: $H = 0.05$, right: $H = 0.025$. Convergence of order $H$ is given, but there is no stability for any step size in sight.

Figure 6. Simulation of the system (13) in the cosimulation scheme with linear extrapolation, exchange step size $H = 0.2$. Simulation is obviously instable. Right plot: The cosimulation method using constant extrapolation converges of order $H$, using linear extrapolation, with order $H^2$. 
Figure 7. Simulation of the system (13) in the cosimulation scheme with constant extrapolation and balance correction, varying the exchange step size \( H \). Left: \( H = 0.2 \), right: \( H = 0.1 \). The step size \( H = 0.1 \) is now stable, but the instability of the \( H = 0.2 \) calculation reveals that the scheme is not made stable by the balance correction.

Figure 8. Simulation of the system (13) in the cosimulation scheme with linear extrapolation and balance correction, varying the exchange step size \( H \). Left: \( H = 0.2 \), right: \( H = 0.1 \). The step size \( H = 0.1 \) is now stable, as well as the \( H = 0.2 \)-calculation, but no conservation is established for the latter.

to the mass in cosimulation is still higher than in the exact solution. As a conclusion, balancing the flows improves, but does not enforce stability, and it motivates a reclassification of the balance errors.

3.2. Conclusion: Reclassification of Balance Errors.

3.2.1. Errors in balances while exchanging conserved Quantities. These errors are defined as those that can be calculated as \( \Delta E^i_j = \int_{t_{j-1}}^{t_j} u_i dt - \int_{t_{j-1}}^{t_j} \bar{u}_i dt \), thus those resulting directly from extrapolation errors in flow of the conserved quantity. They for example occur the mass of a fluid in a cycle like in [5] and are treated in [6] as the original balance errors,
3.2.2. Errors in balances while exchanging factors of conserved quantities. If an exchanged quantity influences a conserved quantity, the balance of that quantity is disturbed by the approximation error of an input as described above, even if the input quantity is nonconserved. Furthermore, as the conserved quantity depends on other factors, its imbalance may persist even if the balance of the exchanged quantity is reestablished, e.g. by errors compensating each other. For all we know, this problem has not been described so far. Yet, in our case studies, it becomes apparent and it will be subject of our future work.

4. Effects of discontinuous Inputs: Method-induced Oscillations

4.1. Spring-Mass system on moving ground. To examine the effect of discontinuity of input due to extrapolation on subsequent subsystems, consider a spring on moving soil.
To relate it to the cosimulation context, this system can be seen as two coupled spring and mass systems in its limiting case, in which the first mass is very big, and so is the stiffness of the first spring, but still such that the frequency \( f = \frac{\omega}{2\pi}, \omega_{1,2} = \lambda_{1,2} \sim \pm \sqrt{-\frac{c}{m}} \) is well bigger than the frequency of the second system. So the system

\[
\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{c_1 + c_2}{m_1} & -\frac{d_1 + d_2}{m_1} - \frac{c_2}{m_1} & \frac{d_2}{m_1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{1}{m_1}(c_1 l_{0,1} - c_2 l_{0,2}) \end{pmatrix},
\]

is mimicked, for the case it is nearly decoupled by \( m_1 \gg m_2, \frac{c_2}{m_2} > \frac{c_1}{m_1} \), by the system

\[
\dot{x} = A x = \begin{pmatrix} 0 & \frac{1}{m} \\ -\frac{c}{m} & -\frac{d}{m} \end{pmatrix} \begin{pmatrix} x(t) \\ \dot{x} \end{pmatrix}, \quad x(t) = x_{t_0} \cos(\omega_1 t)
\]

The real behavior of such a system, given that the spring is relaxed at \( t = 0 \), is that the mass quietly follows the ground movement, i.e. the position of the mass is approximately \( x(t) \sim x_{t_0} \cos(\omega_1 t) + h_0 \) without own oscillations.
The position of the very heavy systems mass \( x_{t_0}(t) \), moving slowly, is then taken as the input \( u \), and the exchange step width is chosen in orders of magnitude of the small spring-mass systems eigen frequency. Now the effect of the discontinuities between piecewise extrapolation can be studied (the simplifications of course might hide instability).
Figure 9. The effect of extrapolation on the receiving spring mass system, using constant and linear extrapolation, respectively. Here and in the following simulations $m_2 = 0.0005$, $k_2 = 5.0$, and $\omega_1 = 2\pi$, corresponding to e.g. $m_1 = 1.0$, $k_1 = 4\pi^2$, see eq. (14). Red and blue curve are the extrapolants of moving soils displacement and velocity. The light green curve is the velocity of the mass, showing strong oscillations due to discontinuous input. Additionally, the numerical integral of the velocity is plotted, using only the values at exchange times (upper, light green dotted line) and values inside exchange interval (dark green dots on the displacement graph. This reveals a further pitfall in cosimulation.

Figure 10. First half of the previous plot 9 for details.

4.1.1. Simulations without and with balance correction. The figures 9 and 11 show well these effects on the subsystem. In the constant extrapolation case, the velocity inside a step shows typical bellies, which are caused by the extrapolant being too high in the beginning of each interval and lower as the real one in the end, or vice versa. In the linear extrapolation case, the velocity also shows bellies, but they are due to the initial error caused by the constant extrapolation in first step: When $H$ is modified, that oscillations stay. Nevertheless quadratic errors in $x_0$ and thus in the force amplify these oscillations: They keep
The effect of extrapolation on the receiving spring mass system, using constant and linear extrapolation with Balance Correction, respectively. Plot colouring and conclusions as in fig. 9.

rising in the standard case (fig. 9 right). It is clear that the oscillations in velocity cause oscillations in the displacement. Both are unphysical. Figure 11 shows identical setting of the simulation, except that now balance correction is applied - it can be seen that this has the desirable effect of stabilizing the initial oscillation in the linear extrapolation calculation, but the oscillations induced by input discontinuity remain. Balance correction does not affect smoothness.

Another treacherous property of the cosimulation scheme is revealed here: If velocity values are only stored at exchange times (dots in the plot), in the constant extrapolation case the impression is made that the velocity is low. Then the displacement is inconsistent with the velocity, see figures 10, left, and 11, left image.

These plots of simulation behavior clarify how desirable it is to provide a method that smoothens the input and the correction contributions. Considering DAE problems with disturbance index greater zero (see [11, chap. 3.1]), it is clear that unsmoothed cosimulation must fail in total.

4.1.2. Conclusion. Exchange step induced oscillations matter in some situations, and the methods so far discussed, even smoothed linear extrapolation with balance correction, need further improvements for settings vulnerable to exchange induced oscillations.

4.1.3. Simulations while smoothing the input. Now the effect of smoothed input on oscillations of subsequent systems is examined using the moving ground example. Figure 12 shows that if input is smoothed, input-induced oscillations are reasonably reduced compared to nonsmoothed input situation. In the linear extrapolation case, the smoothing even damps the high-frequency oscillations.

Figure 11.
Figure 12. The effect of extrapolation on the receiving spring mass system, using constant and linear smooth extrapolation, respectively. Plot colouring as in fig. 9. The oscillations are reasonably reduced (leaving aside the initial numerical excitation) compared to nonsmoothed input (9).

Figure 13. The effect of extrapolation on the receiving spring mass system, using constant and linear smooth extrapolation, respectively, and balance correction. Plot colouring as in fig. 9. The oscillations are amplified by the balance correction contributions, compare to (9).

If balance correction is used, again big oscillations in data exchange frequency can be observed, quite similar as in the nonsmoothed input (12) case and much bigger than without those contributions. Obviously, the balancing contributions excite the spring-mass in a similar way as the discontinuities. This can be explained by the fact that switch errors contribute to the corrections, and their amplitude may increase, even if fed back with a smooth hat function.

4.1.4. Conclusion. The problem of exchange step induced oscillations can be relieved by smooth switching, but balance correction contributions cause high derivatives in input too, a matter that should be addressed.
4.2. **Stability effect of smooth switching and balance correction.** It is left to examine the effect of smooth switching onto the stability of the cosimulation. We have discussed that the exchange of factors of a conserved quantity leads to disbalances (see section 3.2.2) and further, that balance correction relieves this effect but does not solve the problem, as corrections come with delay. If smooth switching is used, the error in input signals value is bigger than without. So the balance error of the extrapolation will also be bigger, as more quantity comes with delay, and so will be the error in the energy (the conserved quantity). This is the effect that shows up in picture 17, and it is stated as a result of the examination that smoothing in general worsens stability.

To examine the effect of smooth switching on stability in a case where smoothness matters, simulations with the system (14) are considered. The constant extrapolation here leads to a damped coupled system. Smoothing alone amplifies the damping (fig. 17, top left), which as before is caused by the bigger error. Balance correction turns this damping into an (undesired) amplification - fig. 17, lower left. For the linear extrapolation case, the logics is analogous, but the different sign of the balance correction contributions let the corrections have an additional damping effect. The effect is, as one expects, slightly increased by smooth switching if compared to nonsmooth switching, as fig. 17, top, shows, same for the smooth balance corrected scheme, below.
4.3. Splitting off balance error’s early known parts. At data exchange time \( t_j \) the amount \( \int_{t_{j-1}}^{t_j} u_2 dt \) is passed on to \( S_2 \) along with the value \( u_2(t_j) \) for extrapolation. The balance error is \( \Delta E^j_i = \int_{t_{j-1}}^{t_j} u_i dt - \int_{t_{j-1}}^{t_j} \bar{u}_i dt \). Mind that correction contributions alive in \( \Delta_j \) concern past errors and are not part of \( \bar{u}_i \).

The error that arises from smooth switching between \( \text{Ext} \ u^{j-1}_i \) and \( \text{Ext} \ u^j_i \) obviously consists of contributions which are already known at \( t_{i-1} \). Thus it is advantageous to start the correction of this error in interval \( [t_{i-1}, t_i] \), as soon as it is known. For this aim, we split the error

\[ \Delta_j = \int_{t_{j-1}}^{t_j} u_i dt - \int_{t_{j-1}}^{t_j} \bar{u}_i dt \]

\[ = \Delta_j \text{correction} + \Delta_j \text{smoothing} \]
into an error between true $S_1$ output and extrapolation and extrapolation and smoothly switched extrapolation.

\[
\Delta E^j_i = \int_{t_{j-1}}^{t_j} \mathbf{u}_i dt - \int_{t_{j-1}}^{t_j} \mathbf{\tilde{u}}_i dt
\]

\[
= \int_{t_{j-1}}^{t_j} \mathbf{u}_i dt - \int_{t_{j-1}}^{t_j} \mathbf{\tilde{u}}_i dt + \int_{t_{j-1}}^{t_j} \mathbf{\tilde{u}}_i dt - \int_{t_{j-1}}^{t_j} \mathbf{\tilde{u}}_i dt
\]

\[
= \left( \int_{t_{j-1}}^{t_j} \mathbf{u}_i dt - \int_{t_{j-1}}^{t_j} \mathbf{\tilde{u}}_i dt \right) + \int_{t_{j-1}}^{t_j} \mathbf{\tilde{u}}_i dt - \int_{t_{j-1}}^{t_j} \mathbf{\tilde{u}}_i dt
\]

and the balance correction part of this is added beginning with $t_j$, while the switching error is added right away beginning at $t_{j-1}$ together with the BC contributions from step $j - 1$.

4.3.1. Numerical example. If one compares figures 18 and 19 to their nonquick correspondants 15 and 17, lower rows, one finds a positive, conservation-enhancing effect for the spring-mass example, as one would expect, but hardly a difference in the conservation behavior of the double-spring-mass example.
4.4. Smoothing and Balance Correction together with special interest on oscillation reduction. It is the aim of this work to provide a calculation of $\overline{u}_i$ and balance correction contributions to $\tilde{u}_i$, resp. $\overline{u}_i$ such that $\overline{u}_i$ and balance correction are smooth not only in the sense of mere existence of derivatives, but the number of additional changes in curvature, this is the sign of the second time derivative, is reasonably related to number of information exchange times. The restriction to correct all known balance errors in the next time step, as it is done in \cite{6}, is dropped: The balance correction contribution of $[t_j, t_{j+1})$ is added during the next $n > 1$ timesteps, thus the intervals in which the balance correction contributions are performed are now overlapping. Adding correction contributions immediately in the next timestep means that the hat function $\phi$ is such that supp$\phi \subset \Delta_i t$, which if smooth has at least one extremum and two turning points. These then possibly cause local extrema and turning points on the corrected flow input signal, making it having unnecessarily high derivatives and look bumpy, see figure 4.

The unphysical derivative contributions may cause dynamics of exchange-rate frequency in the receiving subsystem and those subsequent in causality, and hide the observations one wanted to make, or worse,
even influence system behavior. For example, unphysical noise may occur. All this could be seen in section 4.1.

Doing cosimulation with smooth extrapolation of inputs, one has to do everything to minimize errors in derivatives of signals.

Thus motivated, a more sensible $\phi$ and support is chosen. There are mainly two ways to realize the new objective to reasonably connect the number of changes of curvature to the number of new information:

1. Let the hat function $\phi$ have at most one convex interval and at most one concave interval per time interval, which means per change of approximation of $u_i$. This means
   - using an S-shaped function $\psi$ as in [5] for switching between extrapolation polynomials $\tilde{u}^{-1}_i$ and $\tilde{u}^1_i$:
     \[
     \tilde{u}_i^j = (1 - \psi)\tilde{u}^{-1}_i + \psi\tilde{u}^1_i.
     \]
   - and a function $\phi$ to add balance corrections as $\Delta E^j_i \phi$ with the property that it is S-shaped in each time interval too.
Figure 19. Simulation of the system \([14]\) in the cosimulation scheme with constant (left) and linear (right) extrapolation \(H = 0.02\) and quick refeed of the switch error. Stability properties similar to smooth, not quick refeed case (fig. \([17]\), no real improvement in this setting. - Viell. hochfrequenter Energieübergang? - Kann man von einem steifen System reden, was an der Steifheitsgrenze aufgetrennt wird? - mal Eigenwertanalyse machen.

Straightforwardly this is done by using a hat function consisting of a S-shaped rising branch and a mirrored falling branch. The support of \(\phi\) thus stretches over two time intervals. Such a hat function is constructed in \(5.2\).

(2) Let the hat function for the balance correction contribution \(\phi\) have only one sign in curvature per time interval, which means that the support of it stretches over 4 time intervals, being convex in the first, concave in the second and third, having its maximum between them, and being convex again in the forth. This method has the advantage of damped derivatives of the corrections and because usually four \(\phi_i\) are overlapping inside the time domain, they may compensate each other, bringing \(u_i^{j-1}\) and its approximation closer together. Stretching \(\phi\) over even more time intervals can be thought of. Switching between extrapolation polynomials is done as above,
Figure 20. Input signal and ways to re-contribute errors by the receiving system: Early refeed of known parts: Contributions from smooth switching are re-contributed earlier. As before: Green: linearly interpolated, red: smoothed with s-shaped $\psi$, black: error from previous time step added to smooth signal. In the lower plot, the contributions themselves are plotted: Black: balance correction from previous timestep, blue: balance error due to smooth (slow) switching in actual timestep, black dotted: sum of both. It is obvious that switching error cannot be neglected, and that derivatives of the input signal oscillate.

so in this case changes of curvature arising from this switching stay the same. However, especially the choice of a four-time-interval support of $\phi$ is charming because the change of curvature of the correction contributions once per time interval can be seen as one reaction of the correction procedure per information exchange.

The suggested method has a drawback in terms of time delay compared to classical balance correction as in [2,3]. In the former approach, all $\Delta E^t_j$ is recontributed in $[t_j, t_{j+1})$. In the suggested way to distribute correction contributions over more time intervals - let us choose a support of two time intervals for the hat functions - the switching part of $\Delta E^t_j$ as in equation (16) is already recontributed in $[t_{j-1}, t_{i+1}$, thus starting earlier and ending at the same time. But the standard BC part arrives in $[t_j, t_{j+2})$.

As the (early) switching part amounts to roughly half of the (late) typical balance error, a time delay remains.
Early refeed of known parts, use of one-turning-point-per-information-policy. Contributions from smooth switching are recontributed earlier. As before: green: linearly interpolated, red: smoothed with s-shaped $\psi$, black dotted: error from previous time step added to smooth signal. The lower plot (rescaled!) shows the correction contributions themselves: Dotted are contributions including actual switch error, using an one-interval hat function, black the same using the two-interval test function. In magenta the sum of the two-interval-test-function contributions: It shows no unnecessary extrema. This property is inherited to the extrapolated signal: While the signal with the one-interval hat functions (black dotted) returns to the uncorrected signal at end of time step, the magenta signal has no such bends.

If support of hat functions is chosen as four time intervals, the delay and its effect grows.

4.4.1. Numerical example. The aim of reducing method induced oscillations is achieved well, see figure 22. But the figures 23 and 24 show that the delayed refeed of the balance error which is inherent to the spreading of the refeed over more intervals has a negative effect on the stability, as foreseen above.
5. Definition of smooth switching and hat functions

We refer as hat function to a real function with a convex support in which it is positive and has one maximum. An example is

\[
 d(t) = \begin{cases} 
  e^{-\frac{t+1}{\sigma^2}} & t \in [-1, 1] \\
  0 & \text{else}
\end{cases}
\]

(e.g. \cite{12}), reminding of the Gaussian standard distribution. It is not used in this context as polynomials are cheaper to evaluate. All hat functions \( \bar{\phi} \) are constructed here on the reference interval \([-1, 1]\) with the property that its integral is 1. On a general interval \([t_n, t_{n+k}]\), \( n \in \{1, 2, 4\} \), we use

\[
 \phi_n(t) = \left( \frac{2}{t_{n+k} - t_k} \right) \bar{\phi} \left( \frac{2}{t_{n+k} - t_k} \left( t - \frac{t_{n+k} + t_k}{2} \right) \right),
\]

a transformation of \( \bar{\phi} \) which itself is a hat function on \([t_n, t_{n+k}]\). The property that its integral is 1 is conserved by the transformation such that balance correction can be added to signal \( \bar{u} \) as \( \Delta E_{f,m,i} \phi_i(t) \), starting at time \( t_i \).
Figure 23. Simulation of the system (13) in the cosimulation scheme with constant (left) and linear (right) smooth extrapolation and BC with quick refeed of the switch error and refeed distributed over two steps, $H = 0.2$. Stability properties suffer from the time delay in refeed if compared to fig. 18, but are quite similar to smooth calculations without quick refeed (lower row in fig. 15).

All switch functions $\overline{\psi}$ are constructed analogously as $\overline{\phi}$ on the interval $[-1,1]$ with the property that its integral is 1, and again the transformation (19) conserves this property.

5.1. **Definition of a polynomial smooth hat function.** The function we search should be easy to calculate and implement. We concentrate our search to the realm of polynomials. The function should be symmetrical to its middleaxes, therefore the degree of the polynomial has to be even. Furthermore, all hat functions for our purposes should continuously vanish at the boundaries of their support, and so should at least the first and second derivative. Thus a polynomial of at least sixth degree is needed, and its first derivative should have double roots at -1 and 1.
Figure 24. Simulation of the double spring mass system (14) in the cosimulation scheme with constant (left) and linear (right) extrapolation, $H = 0.02$. Stability properties suffer from the delayed refeed, compare to fig. 23.

(and of course a root at 0):

\begin{equation}
\ddot{\phi} = ax \left[ (x - 1)^2 (x + 1)^2 \right] = ax(x^2 - 1)^2 = a(x^5 - 2x^3 + x),
\end{equation}

the integral of which is

\begin{equation}
\overline{\phi} = a \left[ \frac{1}{6} x^6 - \frac{1}{2} x^4 + \frac{1}{2} x^2 + c \right].
\end{equation}

We determine $c = -\frac{1}{6}$ by using $\overline{\phi}(1) = 0$, and

\begin{equation}
\int_{-1}^{1} \overline{\phi}(t) dt = a \left[ \frac{1}{42} x^7 - \frac{1}{10} x^5 + \frac{1}{6} x^3 - \frac{1}{6} x \right]_{-1}^{1} = 1
\end{equation}

yields $a = -\frac{105}{16}$.

No attention to turning points inside $(-1, 1)$ was paid, and one easily verifies, second derivative being $\overline{\phi}'' = (x-1)^2(x+1)^2 [(x-1)(x+1) + 2x(x+1) + 2x(x-1)]$, that $\pm \sqrt{\frac{1}{5}}$ are roots of it. Turning points of this function are not equidistant, which makes this function unsuitable for use in the second
A polynomial of 8th order would be necessary. We do not carry out these calculations because in the following a function with the desired property will be constructed.

5.2. Construction of switch functions as integrals of hat functions and of integral-of-two-hats type. If $\sum \phi_i = \text{const}$ in some interval, all derivatives of $\sum \phi_i$ vanish, which means that if corrections $\Delta E_{m,i}$ do not change, the sum $\sum \Delta E_{m,i} \phi_i$ does not change and thus only a constant is added to $\tilde{f}_m$. So if there shall be no contributions from the $\phi_i$ alone to derivatives, it is necessary (yet not sufficient) that the two resp. four test functions sum up to a constant on the common of their support $\cap \text{supp} \phi_i$.

Nevertheless, to this objective, let the real function $h(t)$ be one of the above defined hats, all of which have the property that they are symmetric w.r.t. $t = 0$.

We define a function

$$\psi := \int_{-1}^{t} h(\tau) d\tau.$$  

Its maximum is 1, $\psi(1) = 1$, as all hats from the former section are already normed such that $\int_{-1}^{1} h(\tau) d\tau = 1$. Clearly $\psi$ is a s-shaped switch function (so its name is well-chosen) with a turning point at (even rotational symmetry with respect to) $(0, \int_{-1}^{1} d(\tau)d\tau/2)$. Moreover, as for the integral of any function $h$ symmetric to $t_{\text{symm}}$ it holds

**Figure 25.** Left: Example for $\bar{\phi}$: Polynomial (21) and its integral as an example for $\psi$. Right: Example for $\phi$ as integral-of-two-hats: Polynomial (21) integrated, stretched and shifted and mirrored.
that $\int_{t_i}^{t_{symm}} h(\tau) d\tau = \int_{t_{symm}}^t h(\tau) d\tau$ is valid, $\psi(-t) = 1 - \psi(t)$ follows, see fig. 25.

$$\psi(-t) = 1 - \psi(t)$$

5.3. **Construction a of hat function of integral-of-two-hats type.** We make use of this property by defining a hat again

$$\overline{\phi} = \begin{cases} 
\psi \left( 2 \left( t + \frac{1}{2} \right) \right) & t < 0 \\
\psi \left( 2 \left( -t + \frac{1}{2} \right) \right) & t \geq 0,
\end{cases}$$

from a transformed $\psi$ (compare eq. (19)) and its mirrored counterpart. According to the integral of hat (22) and the transformation (19) the function

$$\overline{\phi} = \frac{-105}{16} \begin{cases} 
\frac{(2x + 1)^7}{42} & - \frac{(2x + 1)^5}{10} + \frac{(2x + 1)^3}{6} - \frac{2x + 1}{6} - \frac{8}{105} & t < 0 \\
\frac{(-2x + 1)^7}{42} & - \frac{(-2x + 1)^5}{10} + \frac{(-2x + 1)^3}{6} - \frac{-2x + 1}{6} - \frac{8}{105} & t \geq 0
\end{cases}$$

is a realization of an integral-of-hat type hat function, namely the one with the lowest possible degree.

6. **Discussion and Conclusion**

Unchanged by results of this examinations, problems without strong conservation properties can be well treated with cosimulation schemes, and problems with vulnerability to imbalance of exchanged conserved quantity in terms of biased system properties due to wrong quantity amount over a long time interval can be well treated with balance correction.

Considering systems with conservation properties in which factors of the conserved quantity are exchanged, the smoothing and recombination methods presented can improve stability and smoothness at the same time, but not solve the issues.
Improving the extrapolation order is always beneficial. Besides this, all attempts to cure the drawbacks of methods that are explicit in some variables encounter new drawbacks:

If smoothing and balance correction, then of course also smooth refeed is needed. Although smooth, the recontributions may cause quite high derivatives. The smoothing error contributes to the balance error - so to the derivatives that are caused by recontributions.

If they matter, so if excitation of subsystems can be expected, and so one tries to tackle them by distributing the recontributions over more time intervals, balance errors depending on the exchanged quantities (energy) are worsened.

Balance correction methods in general remain unable to fully cure balance violations and their consequences, even less so those that distribute recontrtribution over more than one time interval. Yet they are able to prolong the time that the simulation can deliver meaningful results.

6.1. **Convergence and stability considerations.** In \[4\] convergence results and stability criteria has been derived for cosimulation of linear ODE systems. To the best of our knowledge, nobody has written down such results for cosimulation of nonlinear ODE or DAE systems. Although simple considerations make it appear likely that cosimulation methods are convergent for such systems, it remains to prove that this is the case. Once such a proof is available, it has to be shown that balance correction does not spoil convergence.

6.2. **Conclusion.** Using smooth contributions significantly reduces unphysical dynamics in the frequency of exchange time stepwidth. This allows to draw conclusions concerning higher-frequency dynamics from cosimulation results while using balance correction, which would be impossible without careful choice of hat functions.

Finally, establishing balance of a quantity *a posteriori* cannot establish consistency and thus balance in all quantities of the system. Thus, the techniques discussed in this contribution sometimes can improve stability, but it cannot make the cosimulation scheme stable. It is hardly surprising that an explicit method as the cosimulation cannot easily be turned into a stable method.

The view on cosimulation as an explicit method finally puts the results into the right context: Nobody would expect an explicit method to work with a stepwidth close to the systems one eigenfrequency like in the example of a slow, heavy coupled to a quick, light spring-mass system. Why should cosimulation? Dense knowledge of the simulated system remains necessary. But still, the suggested measures have the capacity to transform cosimulations Euler-Forward-like simplicity and
limitedness into applicability for a big range of coupled problems - they turn cosimulation into a method with reliable results.

Anyhow, the future task will be to find methods that determine input signals such that effects of past errors, rather than the errors themselves, on the receiving system are compensated.

REFERENCES

[1] M. Arnold and M. Günther, Preconditioned dynamic iteration for coupled differential-algebraic systems, BIT Numerical Mathematics 41 (2001), pp. 1–25, Available at http://dx.doi.org/10.1023/A:101021909032551.

[2] M. Arnold, C. Clauss, and T. Schierz, Error Analysis and Error Estimates for Co-Simulation in FMI for Model Exchange and Co-Simulation v2.0 60.1 (2013), pp. 75–94, Available at doi:10.2478/meceng-2013-0005.

[3] M. Arnold, C. Bausch, T. Blochwitz, C. Clauß, M. Monteiro, T. Niedhold, J.V. Peetz, and S. Wolf, Functional Mock-up Interface for Co-Simulation (2010), Available at https://svn.modelica.org/fmi/branches/public/specifications/v1.0/FMI_for_CoSimulation_v1.0.pdf.

[4] M. Busch, Zur effizienten Kopplung von Simulationsprogrammen, Ph.D. thesis, 2012, Available at http://books.google.de/books?id=OqBpXp-f2gQC.

[5] R. Kossel, Hybride Simulation thermischer Systeme am Beispiel eines Reisebusses, Ph.D. thesis, Braunschweig, Techn. Univ., 2012.

[6] D. Scharff, C. Kaiser, W. Tegethoff, and M. Huhn, Ein einfaches Verfahren zur Bilanzkorrektur in Kosimulationsumgebungen, in SIMVEC - Berechnung, Simulation und - Erprobung im Fahrzeugbau, 2012.

[7] R. Kübler and W. Schiehlen, Modular Simulation in Multibody System Dynamics, Multibody System Dynamics 4 (2000), pp. 107–127, Available at http://dx.doi.org/10.1023/A:1009810318420.

[8] U. Miekkala and O. Nevanlinna, Convergence of dynamic iteration methods for initial value problems, SIAM J. Sci. and Stat. Comput 8 (1987), pp. 459–482.

[9] S.B.E. Elmqvist, Interface Jacobian-based Co-Simulation, International Journal for Numerical Methods in Engineering 98 (2014), pp. 418–444, Available at http://dx.doi.org/10.1002/nme.4637.

[10] M. Wells J.; Hasan and C. Lucas, Predictive Hold with Error Correction Techniques that Maintain Signal Continuity in Co-Simulation Environments, SAE Int. J. Aerosp (2012), pp. 481–493.

[11] P. Deuflhard and F.A. Bornemann, Numerische Mathematik II, de Gruyter, 1994.

[12] W. Walter, Einführung in die Theorie der Distributionen, BI-Wiss.-Verlag, 1994, Available at http://books.google.de/books?id=ATvvAAAAMAAJ.