On the Realization of Impulse Invariant Low-Rank Volterra Kernels

Phillip M. S. Burt, *Member, IEEE*, and José Henrique de Morais Goulart

***Abstract*—Volterra models can accurately model numerous non-linear systems of practical interest, but often at an unacceptable computational cost. If the Volterra kernels of a system have low-rank structure (like, e.g., kernels of bilinear systems), this major drawback can in principle be mitigated. Yet, when one seeks an exact discrete-time model of a mixed-signal chain involving that system, the existing formula that generalizes the impulse invariance principle to Volterra kernels yields discrete-time kernels that do not share the same low rank. At first sight this would seem to seriously complicate the otherwise simple discrete-time realization of low-rank kernels. We show here that this not the case. By defining a cascade operator, the structure of generalized impulse invariance can be unveiled, leading to a realization without an inordinate increase in computational complexity. Finally, we give a numerical example involving a physical system that shows the relevance of our proposal.**

***Index Terms*—Nonlinear systems, bilinear systems, Volterra model, impulse invariance.**

I. Introduction

VOLTERRA models are a popular choice for the modeling of non-linear systems of various kinds [1]–[7]. In particular, the most commonly used variant in digital signal processing applications, known as a Volterra filter (VF), is essentially a feed-forward polynomial model whose output is linear in the model parameters—a desirable feature for system identification—and whose wide applicability has been established by Boyd and Chua in the eighties [8]. Unfortunately, though, the amount of parameters of a VF grows quite rapidly with the system memory length and the order (nonlinearity degree) of the model, whose practical realization thus often becomes too costly.

Motivated by this drawback, a whole line of research has been devoted to devising more practical alternatives which trade some generality of the VF—and often also the linearity in the parameters—by a lower parametric complexity [9]–[18]. One of the most effective and elegant proposals is based on the simple assumption that the pth-order Volterra kernel $h_p(n_1, \ldots, n_p)$ (high-order analogues of the impulse response, see Section III for a definition) approximately decomposes as a sum of a few *separable* functions, that is, $h_p(n_1, \ldots, n_p) \approx \sum_{r=1}^{R_p} h_r^{(1)}(n_1) \ldots h_r^{(p)}(n_p)$ with a sufficiently small R_p. This amounts to a low-rank (specifically, rank-R_p) approximation of the kernels, viewed as tensors [13]. Besides the potential dramatic reduction in computational complexity, one major advantage of this approach is its straightforward realization by a combination of linear blocks and instantaneous nonlinearities or multipliers.

Manuscript received 26 April 2022; accepted 15 June 2022. Date of publication 24 June 2022; date of current version 7 July 2022. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Demetrio Labate.

Phillip M. S. Burt is with the Escola Politécnica, University of São Paulo, 05508-900 São Paulo, Brazil (e-mail: phillip@lcs.poli.usp.br).

José Henrique de Morais Goulart is with the IRIT, University of Toulouse, Toulouse INP, CNRS, 31071 Toulouse Cedex 7, France (e-mail: josehenrique.demoraisgoulart@toulouse-inp.fr).

Digital Object Identifier 10.1109/LSP.2022.3185962

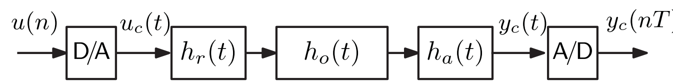

Fig. 1. Signal chain starting in discrete-time, passing through continuous-time and then returning to discrete-time.

However, in applications where an exact discrete-time modeling of a mixed-signal chain consisting of discrete-time and continuous-time parts is desirable (for instance, in the cancellation of signals originating from discrete-time, such as in acoustic echo cancellation or nonlinearity mitigation [10], [19]), the low-rank realization is not as simple anymore. The reason is that the well-known impulse invariance [20] between discrete-time and continuous-time linear time-invariant (LTI) systems does not generalize "cleanly" to nonlinear systems, but has to be somewhat modified so as to incorporate a factor which depends on the pattern of repeated kernel arguments, as pointed out in [21], [22]. We show that, as a consequence, the computational cost of a naive realization of generalized impulse invariance, while still much smaller than that of a VF realization (whose cost is not affected by generalized invariance), would increase by as much as 2^{p-1} times. We then deduce a much more efficient realization, preserving the great attractiveness of the low-rank approach.

To the best of our knowledge this problem has not been previously addressed. Apart from its practical relevance, as discussed above, the result we present is of interest in itself as an addition to the theory of signal processing.

II. Impulse Invariance of LTI Systems

A continuous-time LTI system bandlimited to $1/2T$ Hz, with impulse response $h_c(t)$, can be implemented [20, p. 173] with a mixed-signal chain containing the *impulse invariant* discrete-time system with impulse response[1]

$$h(n) = h_c(nT). \tag{1}$$

Impulse invariance also comes into question when modeling the mixed-signal chain depicted in Fig. 1, which is of greater

[1] Hereafter, the subscript c is used to distinguish a continuous-time signal from its discrete-time version. For convenience, we have dropped the factor T from the definition $h(n) = Th_c(nT)$ of impulse invariance of [20].

concern here and is now described. From the input sequence $u(n)$, an impulsive D/A converter[2] with sampling period T provides

$$u_c(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT)u(k). \tag{2}$$

After a reconstruction filter $h_r(t)$, an application-specific system $h_o(t)$, and an anti-aliasing filter $h_a(t)$, all LTI, it follows from (2) that

$$y_c(t) = \sum_{k=-\infty}^{\infty} h_c\,(t - kT)\,u(k), \tag{3}$$

where $h_c(t)$ is the overall impulse response given by the convolutions $h_r(t) * h_o(t) * h_a(t)$. Finally, an A/D sampler gives the output $y_c(nT)$. It follows that a system with impulse response $h(n) = h_c(nT)$, input $u(n)$ and output $y(n) = \sum_{k=-\infty}^{\infty} h(n-k)u(k)$ is an *exact* discrete-time model of the signal chain[3] in the sense that

$$y(n) = y_c(nT), \tag{4}$$

as desired, for instance, in acoustic echo cancellation [23].

Remark The sampling of $y_c(t)$ in (4) and of $h_c(t)$ in (1) must be consistent at discontinuities. For instance, let $h_c(t) = e^{-at}$ if $t > 0$ and $h_c(t) = 0$ if $t < 0$. From (3) then, $y_c(t)$ is discontinuous at $t = nT$ if $u(n) \neq 0$. Assuming the A/D sampler always provides the right-side limit $y_c(nT_+)$ (respectively, the left-side limit $y_c(nT_-)$ or $[y_c(nT_+) + y_c(nT_-)]/2$), it follows from (3) that, to achieve (4), $h(0)$ must be given by $h_c(0_+) = 1$ (respectively, $h_c(0_-) = 0$ or $[h_c(0)_+ + h_c(0_-)]/2 = 1/2$).

It should be noted that when $h_c(t)$ represents an actual physical system, the impulse response $h(n) = h_c(nT)$ will, in general, have infinite duration. As long, though, as the system $h_c(t)$ is rational, the exact realization (with a finite number of operations) of a discrete-time system with impulse response $h(n)$ is straightforward [24].

III. Generalization of Impulse Invariance to Volterra Kernels

Let the analog portion of the chain in Fig. 1 be now nonlinear. (This can arise, for instance, from a nonlinear loudspeaker in acoustic echo cancellation [19].) We assume then that its input/output relation is given by the (causal) Volterra series $y_c(t) = \sum_{p=1}^{\infty} y_{c,p}(t)$, with homogeneous outputs given by

$$y_{c,p}(t) = \int_0^{\infty}\!\cdots\int_0^{\infty} h_{c,p}(\tau_1,\ldots,\tau_p)\prod_{i=1}^{p} u_c(t - \bar{\tau}_i)d\tau_1\ldots d\tau_p, \tag{5}$$

where $\bar{\tau}_i = \sum_{j=i}^{p} \tau_j$ and $h_{c,p}(\tau_1,\ldots,\tau_p)$ is a *regular* Volterra kernel of order p [21, p. 15], continuous for $\tau_1,\ldots,\tau_p > 0$. Although the existence of the realization problem addressed in this paper is independent of employing conventional[4] or regular Volterra kernels, the latter are more convenient for the required algebraic manipulation.

To provide a discrete-time model of the signal chain, let $\bar{n}_i = \sum_{j=i}^{p} n_j$, where n_j always represents discrete time, and, for some causal discrete-time kernel $v_p : \mathbb{Z}^p \to \mathbb{R}$ let

$$y_p(n) = \sum_{n_p=0}^{\infty}\ldots\sum_{n_1=0}^{\infty} v_p(n_1,\ldots,n_p)\prod_{i=1}^{p} u(n - \bar{n}_i), \tag{6}$$

with $u(n)$ and $u_c(t)$ still being related by (2). It follows from [22] that achieving $y_p(n) = y_{c,p}(nT)$, $p > 1$, assuming right-side sampling at discontinuities, requires generalizing (1) as

$$v_p(n_1,\ldots,n_p) = \frac{h_{c,p}(n_1 T,\ldots,n_p T)}{m_1!\ldots m_q!}, \tag{7}$$

where q is the number of groups of consecutive null indices among $n_1,\ldots,n_{p-1}$ and $m_1 - 1,\ldots,m_q - 1$ are the numbers of indices in each group.[5] For instance, if $p=5$, $n_1=n_2=0\neq n_3$ and $n_4=0$, then $q=2$, $m_1=3$ and $m_2=2$.

This generalized impulse invariance results from the impulsive terms of the integrand in (5) and the possible discontinuity of the kernel on the border of the domain $\tau_1,\ldots,\tau_{p-1} \geq 0$. In the interior of this domain we retrieve a direct extension of the invariance condition (1), that is, we have $v_p(n_1,\ldots,n_p) = h_{c,p}(n_1 T,\ldots,n_p T)$ for $n_1,\ldots,n_{p-1} > 0$. For more details on the steps leading to (7), the reader is referred to [22].

Similarly to the discretized $h(n)$ of (1), (7) gives in general an infinite duration $v_p(n_1,\ldots,n_p)$. This raises the issue of its realization with a finite number of operations, not addressed in [22]. We do this now, for the class of low-rank kernels.

IV. Realization of Low-Rank Kernels

We consider systems with *low-rank* Volterra kernels

$$h_{c,p}(\tau_1,\ldots,\tau_p) = \sum_{r=1}^{R_p} \mathbf{H}_{c,r}^{(p)}(\tau_p)\ldots\mathbf{H}_{c,r}^{(1)}(\tau_1), \tag{8}$$

for any set of (vector- and) matrix-valued functions $\mathbf{H}_{c,r}^{(i)}(\tau_i)$ of compatible dimensions,[6] and a given $R_p \in \mathbb{N}^*$, termed the *rank* of $h_{c,p}$. Of particular practical interest are bilinear systems [21], [25], further discussed in Section V, for which $R_p = 1$. We consider $R_p = 1$ from here onward, and thus look into the realization of one of the parallel branches of (8).

A. Cascade Structure and Operator

From (8) and (5) it follows directly that low-rank kernels can be realized quite simply by a cascade of linear blocks and multipliers. This is depicted in Fig. 2 ($p = 3$, $R_3 = 1$), where

$$\begin{aligned}\mathbf{z}_1(t) &= [\textstyle\int_0^{\infty} \mathbf{H}_c^{(1)}(\tau)u_c(t-\tau)d\tau]u_c(t)\\ &= [\,\mathbf{H}_c^{(1)} * u_c(t)]\,u_c(t) = \mathbf{H}_c^{(1)} \circ u_c(t),\end{aligned}$$

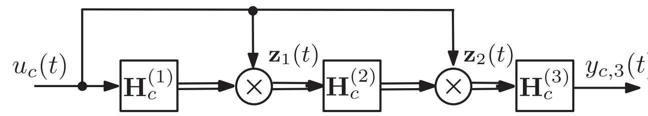

Fig. 2. Cascade realization of a separable kernel, $p = 3$, $R_3 = 1$. Double arrows denote vector-valued signals.

[2] $\delta(t)$ is the Dirac delta function. The assumption of ideal impulsive excitation is not restrictive, since the reconstruction filter $h_r(t)$ can absorb the rectangular impulse response of a real-world zero-order hold D/A.

[3] In this case, is not required that system $h_c(t)$ be bandlimited.

[4] With a conventional kernel $h_{c,p}^{(\mathrm{conv})}(\tau_1,\tau_2,\ldots,\tau_p) = h_{c,p}^{(\mathrm{reg})}(\tau_1-\tau_2,\ldots,\tau_{p-1}-\tau_p,\tau_p)$, (5) assumes the more familiar form with τ_i instead of $\bar{\tau}_i$.

[5] A related result is stated without proof in [21, p. 254]. For simplicity, the dependence of m_i on $n_1,\ldots,n_{p-1}$ is omitted. If $y_p(n) = y_{c,p}(nT_-)$ or $y_p(n) = [y_{c,p}(nT_-) + y_{c,p}(nT_+)]/2$ at discontinuities, the result is similar.

[6] For uniformity of notation, $\mathbf{H}_{c,r}^{(p)}$ (resp., $\mathbf{H}_{c,r}^{(1)}$) is denoted as a matrix, though being a row (resp., column) vector. If $R_p = 1$, we drop subscripts r.

$$\mathbf{z}_2(t) = [\mathbf{H}_c^{(2)} * \mathbf{z}_1(t)]\, u_c(t) = \mathbf{H}_c^{(2)} \circ \mathbf{z}_1(t),$$

$$y_{c,3}(t) = \mathbf{H}_c^{(3)} * \mathbf{z}_2(t).$$

Here, $*$ stands for convolution and $\circ$ stands for the *cascade* operator defined, given $u_c(t)$, by $h \circ x(t) \triangleq [h * x(t)]u_c(t)$, which is linear in h and in x. Also, with it we can write

$$y_{c,3}(t) = \mathbf{H}_c^{(3)} * \{\, \mathbf{H}_c^{(2)} \circ [\, \mathbf{H}_c^{(1)} \circ u_c(t)]\,\}.$$

Hence, the realization can be expressed as a sequential calculation with $p-1$ applications of the cascade operator, followed by a convolution at the final stage.

B. Parallel-Cascade Realization of Impulse Invariance

Consider the sampled kernel factors $\mathbf{H}^{(i)}(n_i) = \mathbf{H}_c^{(i)}(n_i T)$. We can readily verify that, assuming the input has the form (2), the cascade structure has this very particular property:

Property Replacing $\mathbf{H}_c^{(i)}(\tau_i)$ with $\mathbf{H}^{(i)}(n_i)$ and $u_c(t)$ with $u(n)$ in the cascade structure that realizes $h_{c,p}(\tau_1, \ldots, \tau_p) = \mathbf{H}_c^{(p)}(\tau_p) \ldots \mathbf{H}_c^{(1)}(\tau_1)$, gives a realization of[7]

$$\tilde{v}_p(n_1, \ldots, n_p) = \mathbf{H}^{(p)}(n_p) \ldots \mathbf{H}^{(1)}(n_1). \tag{9}$$

This discrete-time kernel, however, is not impulse invariant in relation to $h_{c,p}(\tau_1, \ldots, \tau_p)$, since the term $1/m_1! \ldots m_q!$ in (7) is missing. To include the missing term, let us initially rewrite the invariance condition (7), assuming $R_p = 1$, as

$$v_p(n_1, \ldots, n_p) = c_p(n_1, \ldots, n_{p-1})\mathbf{H}^{(p)}(n_p) \ldots \mathbf{H}^{(1)}(n_1), \tag{10}$$

where $c_p(n_1, \ldots, n_{p-1}) \triangleq 1/m_1! \ldots m_q!$. We note then that c_p can be decomposed as

$$c_p(n_1, \ldots, n_{p-1}) = \sum_{r=1}^{2^{p-1}} a_r c_r^{(1)}(n_1) \ldots c_r^{(p-1)}(n_{p-1}), \tag{11}$$

where $c_r^{(i)}(n_i)$ is either the unit impulse[8] $\delta(n)$ or its complement $\bar{\delta}(n) = [1 - \delta(n)]$.

Example Since $c_3(n_1, n_2) = 1$ when $n_1, n_2 > 0$, $c_3(n_1, n_2) = 1/2$ if either $n_1 = 0$ or $n_2 = 0$ (but not both), and $c_3(n_1, n_2) = 1/3!$ if $n_1 = n_2 = 0$, it decomposes as

$$c_3(n_1, n_2) = \bar{\delta}(n_1)\bar{\delta}(n_2) + \frac{1}{2}\, \delta(n_1)\bar{\delta}(n_2) + \tfrac{1}{2}\, \bar{\delta}(n_1)\delta(n_2) + \tfrac{1}{3!}\, \delta(n_1)\delta(n_2). \tag{12}$$

Now, from (10) and (11) we get

$$v_p(n_1, \ldots, n_p) = \sum_{r=1}^{2^{p-1}} a_r \mathbf{G}_r^{(p)}(n_p) \ldots \mathbf{G}_r^{(1)}(n_1), \tag{13}$$

where $\mathbf{G}_r^{(i)}(n_i)$ is either given by $\bar{\delta}(n_i)\mathbf{H}^{(i)}(n_i)$ or by $\delta(n_i)\mathbf{H}^{(i)}(n_i), 1 \le i < p$, and $\mathbf{G}_r^{(p)}(n_p) = \mathbf{H}^{(p)}(n_p)$. It follows that v_p can be realized by summing the output of 2^{p-1} parallel cascade structures. Nevertheless, because all blocks $\mathbf{G}_r^{(i)}(n_i)$ in (13) come from the decomposition $(\delta(n_i) + \bar{\delta}(n_i))\, \mathbf{H}^{(i)}(n_i)$, we can largely mitigate this increase in computational complexity by "sharing" computations among the branches of the parallel structure. We describe next a systematic way of implementing this strategy.

[7]This is not true in general for interconnections of linear blocks. For instance, two discrete-time linear systems $f(n) = f_c(nT)$ and $g(n) = g_c(nT)$ in series have impulse response $\sum_k f(k)g(n-k) \neq \int f_c(\tau)g_c(nT - \tau)d\tau$.

[8]$\delta(n) = 1$ if $n = 0$ and $\delta(n) = 0$ if $n \neq 0$.

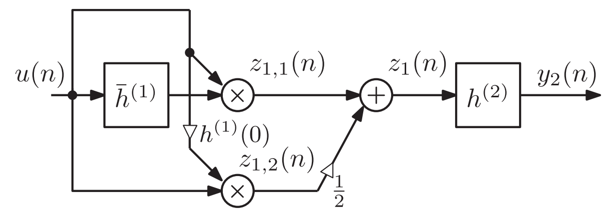

Fig. 3. Realization of impulse invariant kernel of order $p = 2$.

C. Efficient Realization of Impulse Invariance

For simplicity, we consider henceforth low-rank kernels $h_p(\tau_1, \ldots, \tau_p) = h_c^{(1)}(\tau_1) \ldots h_c^{(p)}(\tau_p)$ with scalar (instead of matrix) factors, the extension to matrices being straightforward. With $h^{(i)}(n_i) = h_c^{(i)}(n_i T)$, we define:

$$\bar{h}^{(i)}(n_i) \triangleq \bar{\delta}(n_i)h^{(i)}(n_i), \quad h_0^{(i)}(n_i) \triangleq \delta(n_i)h^{(i)}(0). \tag{14}$$

With $p = 2$, initially, from (10) and (14) we have $c_2(n_1) = \bar{\delta}(n_1) + 1/2\, \delta(n_1)$, and, therefore,

$$v_2(n_1, n_2) = \left[\bar{\delta}(n_1) + \tfrac{1}{2}\, \delta(n_1)\right] h^{(1)}(n_1)h^{(2)}(n_2) = \left[\bar{h}^{(1)}(n_1) + \tfrac{1}{2}\, h_0^{(1)}(n_1)\right] h^{(2)}(n_2).$$

Using this in (6) and a discrete-time version of the cascade operator $h \circ x(n) \triangleq [h * x(n)]u(n)$, we can write the output as

$$y_2(n) = h^{(2)} * \left[\left(\bar{h}^{(1)} + \tfrac{1}{2} h_0^{(1)}\right) \circ u(n)\right]. \tag{15}$$

Defining now $z_1(n) \triangleq (\bar{h}^{(1)} + 1/2\, h_0^{(1)}) \circ u(n)$, it follows from the linearity in h of the cascade operator that

$$z_1(n) = z_{1,1}(n) + \tfrac{1}{2}\, z_{1,2}(n), \tag{16}$$

where

$$z_{1,1}(n) = \bar{h}^{(1)} \circ u(n), \tag{17}$$

$$z_{1,2}(n) = h_0^{(1)} \circ u(n) = h^{(1)}(0)u^2(n), \tag{18}$$

and the output is $y_2(n) = h^{(2)} * z_1(n)$. This realization is depicted in Fig. 3.

Let, now, $p = 3$. From (10) and (12) we can write, initially,

$$v_3(n_1, n_2, n_3) = \left\{\left[\bar{\delta}(n_1) + \tfrac{1}{2}\, \delta(n_1)\right] \bar{\delta}(n_2) + \tfrac{1}{2}\, \bar{\delta}(n_1)\delta(n_2) + \tfrac{1}{3!}\, \delta(n_1)\delta(n_2)\right\} h^{(1)}(n_1)h^{(2)}(n_2)h^{(3)}(n_3).$$

Moving $h^{(1)}(n_1)h^{(2)}(n_2)$ into the brackets, we get then

$$v_3(n_1, n_2, n_3) = \left\{\left[\bar{h}^{(1)}(n_1) + \tfrac{1}{2}\, h_0^{(1)}(n_1)\right] \bar{h}^{(2)}(n_2) + \tfrac{1}{2}\, \bar{h}^{(1)}(n_1)h_0^{(2)}(n_2) + \tfrac{1}{3!}\, h_0^{(1)}(n_1)h_0^{(2)}(n_2)\right\} h^{(3)}(n_3),$$

so that, using this in (6), introducing another cascade operator for the second stage, and with (16)–(18), we can write

$$y_3(n) = h^{(3)} * \left[\bar{h}^{(2)} \circ z_1(n) + \tfrac{1}{2}\left(h_0^{(2)} \circ z_{1,1}(n)\right) + \tfrac{1}{3!}\left(h_0^{(2)} \circ z_{1,2}(n)\right)\right].$$

Organizing as for $p = 2$ then, $y_3(n) = h^{(3)} * z_2(n)$, where

$$z_2(n) = z_{2,1}(n) + \tfrac{1}{2}\, z_{2,2}(n) + \tfrac{1}{3!}\, z_{2,3}(n),$$

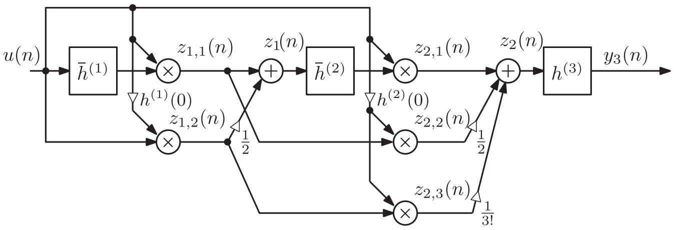

Fig. 4. Realization of impulse invariant kernel of order $p = 3$.

with

$$z_{2,1}(n) = \bar{h}^{(2)} \circ z_1(n),$$

$$z_{2,2}(n) = h_0^{(2)} \circ z_{1,1}(n) = h^{(2)}(0) z_{1,1}(n) u(n)$$

$$z_{2,3}(n) = h_0^{(2)} \circ z_{1,2}(n) = h^{(2)}(0) z_{1,2}(n) u(n),$$

which is depicted in Fig. 4. Generalizing, we should compute:

- For $i = 1, \ldots, p-1$ and with $z_0(n) = z_{0,1}(n) = u(n)$,
$$z_{i,1}(n) = \left[\bar{h}^{(i)} * z_{i-1}(n)\right] u(n),$$
$$z_{i,j}(n) = h^{(i)}(0) z_{i-1,j-1}(n) u(n),\ j = 2, \ldots, i+1,$$
$$z_i(n) = \sum_{j=1}^{i+1} \frac{1}{j!} z_{i,j}(n).$$
- $y_p(n) = h^{(p)} * z_{p-1}(n)$.

V. Example: Bilinear Systems

Bilinear systems have state-space equations of the form

$$\mathbf{x}'_c(t) = \mathbf{F}\mathbf{x}_c(t) + \mathbf{G}\mathbf{x}_c(t) u_c(t) + \mathbf{b} u_c(t)$$

$$y_c(t) = \mathbf{c}^\top \mathbf{x}_c(t),$$

and can approximate, up to any kernel order p, the large class known as linear-analytic systems [21], [26], [27]. Their kernels read $h_{c,p}(\tau_1, \ldots, \tau_p) = \mathbf{c}^\top e^{\mathbf{F}\tau_p} \mathbf{G} e^{\mathbf{F}\tau_{p-1}} \mathbf{G} \ldots \mathbf{G} e^{\mathbf{F}\tau_1} \mathbf{b}$, $\tau_i \geq 0$, thus having the low-rank form of (8) with $R_p = 1$ (that is, rank one), and $\mathbf{H}_c^{(1)}(\tau_1) = e^{\mathbf{F}\tau_1}\mathbf{b}$, $\mathbf{H}_c^{(i)}(\tau_i) = e^{\mathbf{F}\tau_i}\mathbf{G}$, $1 < i < p$, and $\mathbf{H}_c^{(p)}(\tau_p) = \mathbf{c}^\top e^{\mathbf{F}\tau_p}\mathbf{G}$.

As an example, consider the bilinear model of a bass loudspeaker [22], sampled at a rate of 1.5 kHz. An infinite-memory discrete-time realization of its fourth-order impulse invariant Volterra kernel was derived as in Section IV-C, and the corresponding output $y_4(n)$ obtained for an unit-power AWGN input $u(n)$. For validation, the output $\widehat{y}_4(n)$ of the time-truncated kernel was directly calculated using (7) in (6) (in other words, a VF realization) with $\bar{n}_i \leq 120$, aiming at a small discrepancy $\epsilon(n) = y_4(n) - \widehat{y}_4(n)$. Indeed, as seen in Fig. 5, $\epsilon(n)$ is of the order of 10^{-16} (mainly due to computing with 64 b precision), validating the proposed procedure. Also displayed is the output $\tilde{y}_4(n)$ of the kernel given by (9). Its large discrepancy in relation to $y_4(n)$ shows that the invariance principle of (7) can be very relevant in practice.

To compare computational costs now, a VF filter realization of the truncated kernel, with $\bar{n}_i < N$ in (6), requires at least $\binom{N+p-1}{p}$ multiplications [28, p. 36].[9] In the previous loudspeaker example, even allowing for less precision such that

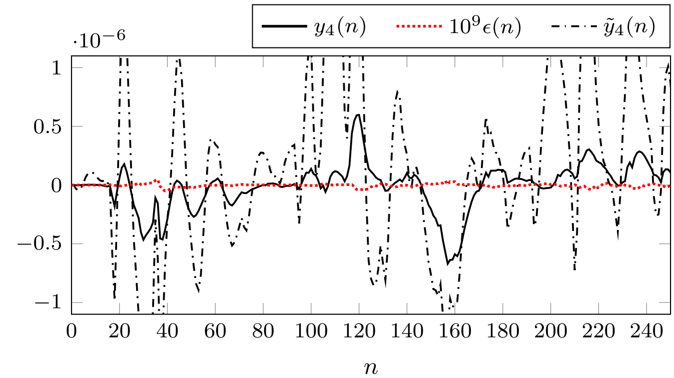

Fig. 5. Output $y_4(n)$ of the realization of an impulse invariant fourth-order kernel and its discrepancy $\epsilon(n)$ to a time-truncated VF realization. For comparison, we also depict the output $\tilde{y}_4(n)$ of the non-impulse invariant kernel of (9).

$\{E[\epsilon^2(n)]/E[y_4^2(n)]\}^{1/2} = 10^{-3}$, still requires $N = 48$ and, thus, 249900 multiplications.

Consider next the last stage in Fig. 4, now with matrix factors. For a bilinear kernel, $\mathbf{H}^{(p)}(n) = \mathbf{c}^\top e^{\mathbf{F}Tn}\mathbf{G}$, $n \geq 0$, so $y_p(n) = \mathbf{H}^{(p)} * \mathbf{z}_{p-1}(n) = \sum_{k=0}^{\infty} \mathbf{c}^\top e^{\mathbf{F}Tk}\mathbf{G}\mathbf{z}_{p-1}(n-k)$, which we readily see is realized by

$$\mathbf{x}_p(n+1) = \mathbf{A}\mathbf{x}_p(n) + \mathbf{B}\mathbf{z}_{p-1}(n) \tag{19}$$

$$y_p(n) = \mathbf{c}^\top \mathbf{x}_p(n) + \mathbf{d}^\top \mathbf{z}_{p-1}(n), \tag{20}$$

where $\dim[\mathbf{x}_p(n)] = M \triangleq \dim[\mathbf{x}_c(t)]$, $\mathbf{A} = e^{\mathbf{F}T}$, $\mathbf{B} = e^{\mathbf{F}T}\mathbf{G}$ and $\mathbf{d} = \mathbf{G}\mathbf{c}$. Assuming $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{d}$ are pre-calculated and have no structure to be exploited for reducing computational cost, (19) and (20) require $2(M^2 + M)$ multiplications. Proceeding similarly for the outputs of the remaining linear blocks gives then a sub-total cost of $C_O = (2p-1)M^2 + 3M$. Finally, the computation of the inputs $\mathbf{z}_i(n)$ of the linear blocks requires $[(p-3)(p/2+1)+1]M^2 + [(p-1)(p/2+1)+3]M$ multiplications [29].[10] In the loudspeaker example, $p=4$ requires $M=34$ [22], giving a total of 13226 multiplications, much less than the at least 249900 required by the VF (and with no loss in precision) and the slightly over $2^{p-1}[C_O + (p-1)M] = 66368$ required by the parallel-cascade of Section IV-B.

VI. Conclusion

By defining a cascade operator, we have shown how to construct a realization of discrete-time kernels obtained from continuous-time low-rank regular kernels by the generalized impulse invariance principle. This construction is required because such discrete-time kernels are not themselves of the same low-rank and thus cannot be realized by the same cascade structures that realize their continuous-time counterparts. The proposed structure requires additional multipliers, not incurring however in an inordinate increase of computational complexity. The low-rank property is found in kernels with practical relevance, and holds in particular for kernels of bilinear systems.

[9] Excluding the computation of $\prod_{i=1}^{p} u(n - \bar{n}_i)$, for simplicity. All multiplication figures refer to the computation of one output sample. In [28] a triangular kernel equivalent to v_p is considered.

[10] This results from eq. (27) in [29], with all $M_i = M$, taking the upper bound $M_i M_{i-1}$ for all μ_i, and adding the $(p-1)M$ multiplications by $u(n)$ required for $\mathbf{z}_{i,1}$, $1 \leq i < p$.

REFERENCES

[1] T. Hélie and D. Roze, "Sound synthesis of a nonlinear string using Volterra series," *J. Sound Vib.*, vol. 314, no. 1–2, pp. 275–306, 2008.

[2] T. Hélie, "Volterra series and state transformation for real-time simulations of audio circuits including saturations: Application to the Moog ladder filter," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 4, pp. 747–759, May 2010.

[3] G. D. Mitsis, M. G. Markakis, and V. Z. Marmarelis, "Nonlinear modeling of the dynamic effects of infused insulin on glucose: Comparison of compartmental with Volterra models," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 10, pp. 2347–2358, Oct. 2009.

[4] B. F. Beidas, "Intermodulation distortion in multicarrier satellite systems: Analysis and turbo Volterra equalization," *IEEE Trans. Commun.*, vol. 59, no. 6, pp. 1580–1590, Jun. 2011.

[5] L. A. Azpicueta-Ruiz, M. Zeller, A. R. Figueiras-Vidal, J. Arenas-García, and W. Kellermann, "Adaptive combination of Volterra kernels and its application to nonlinear acoustic echo cancellation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 1, pp. 97–110, Jan. 2011.

[6] G. Gowtham *et al.*, "A family of adaptive Volterra filters based on maximum correntropy criterion for improved active control of impulsive noise," *Circuits, Syst., Signal Process.*, vol. 41, no. 2, pp. 1019–1037, 2022.

[7] G. Stepniak, M. Marzecki, and J. Bojarczuk, "Volterra predistorter for the dynamic nonlinearity of LED," *Opt. Lett.*, vol. 47, no. 5, pp. 1161–1164, Mar. 2022. [Online]. Available: http://opg.optica.org/ol/abstract.cfm?URI=ol- 47--5-1161

[8] S. Boyd and L. Chua, "Fading memory and the problem of approximating nonlinear operators with Volterra series," *IEEE Trans. Circuits Syst.*, vol. CAS- 32, no. 11, pp. 1150–1161, Nov. 1985.

[9] R. D. Nowak and B. D. Van Veen, "Tensor product basis approximations for Volterra filters," *IEEE Trans. Signal Process.*, vol. 44, no. 1, pp. 36–50, Jan. 1996.

[10] C. Crespo-Cadenas, J. Reina-Tosina, M. J. Madero-Ayora, and J. Muñoz Cruzado, "A new approach to pruning Volterra models for power amplifiers," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2113–2120, Apr. 2010.

[11] E. L. O. Batista, O. J. Tobias, and R. Seara, "A sparse-interpolated scheme for implementing adaptive Volterra filters," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2022–2035, Apr. 2010.

[12] M. Zeller and W. Kellermann, "Fast and robust adaptation of DFT-domain Volterra filters in diagonal coordinates using iterated coefficient updates," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1589–1604, Mar. 2010.

[13] G. Favier, A. Y. Kibangou, and T. Bouilloc, "Nonlinear system modeling and identification using Volterra-PARAFAC models," *Int. J. Adapt. Control Signal Process.*, vol. 26, no. 1, pp. 30–53, Jan. 2012.

[14] E. L. O. Batista and R. Seara, "A fully LMS/NLMS adaptive scheme applied to sparse-interpolated Volterra filters with removed boundary effect," *Signal Process.*, vol. 92, no. 10, pp. 2381–2393, 2012.

[15] P. M. S. Burt and J. H. de M. Goulart, "Evaluating the potential of Volterra-PARAFAC IIR models," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2013, pp. 5745–5749.

[16] E. L. O. Batista and R. Seara, "A reduced-rank approach for implementing higher-order Volterra filters," *EURASIP J. Adv. Signal Process.*, vol. 2016, no. 1, pp. 1–8, 2016.

[17] K. Batselier, Z. Chen, and N. Wong, "Tensor network alternating linear scheme for MIMO Volterra system identification," *Automatica*, vol. 84, pp. 26–35, 2017.

[18] J. H. de. M. Goulart and P. M. S. Burt, "Volterra kernels of bilinear systems have tensor train structure," in *Proc. 29th Eur. Signal Process. Conf.*, 2021, pp. 1070–1074.

[19] M. I. Mossi, C. Yemdji, N. Evans, C. Beaugeant, and P. Degry, "Robust and low-cost cascaded non-linear acoustic echo cancellation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2011, pp. 89–92.

[20] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2009.

[21] W. J. Rugh, *Nonlinear System Theory: The Volterra/Wiener Approach*. Baltimore, MD, USA: Johns Hopkins Univ. Press, 1981. [Online]. Available: sites.google.com/site/wilsonjrugh/

[22] P. M. S. Burt and J. H. de M. Goulart, "Efficient computation of bilinear approximations and Volterra models of nonlinear systems," *IEEE Trans. Signal Process.*, vol. 66, no. 3, pp. 804–816, Feb. 2018.

[23] C. Breining *et al.*, "Acoustic echo control. An application of very-high-order adaptive filters," *IEEE Signal Process. Mag.*, vol. 16, no. 4, pp. 42–69, Jul. 1999.

[24] C. Chen, *Linear System Theory and Design*. London, U.K.: Oxford Univ. Press, 2013.

[25] C. Bruni, G. Dipillo, and G. Koch, "Bilinear systems: An appealing class of "nearly linear" systems in theory and applications," *IEEE Trans. Autom. Control*, vol. AC-19, no. 4, pp. 334–348, Aug. 1974.

[26] H. Hermes and G. W. Haynes, "On the nonlinear control problem with control appearing linearly," *J. Soc. Ind. Appl. Math., Ser. A: Control*, vol. 1, no. 2, pp. 85–108, 1963.

[27] R. W. Brockett, "Volterra series and geometric control theory," *Automatica*, vol. 12, no. 2, pp. 167–176, Mar. 1976.

[28] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*. Hoboken, NJ, USA: Wiley, 2000.

[29] P. M. S. Burt and J. H. de M. Goulart, "On the realization of impulse invariant bilinear Volterra kernels," 2021, *arXiv:2107.06144*.