For odd n, it is found that $\epsilon_{\min}$ is dependent on the width of the receive band, but is not significantly affected by the choice of center frequency. Fig. 5 shows $\epsilon_{\min}$ plotted against filter bandwidth for $n=5$, $d=1,2$, and $n=7, d=1,2$. It can be seen from this figure that the use of higher order discriminator filters ($d=2$) can greatly increase $\epsilon_{\min}$. Hence, improvements in amplitude response obtained by increasing d, may, in some cases be at the expense of increasing the amplitude of the group delay ripple. It may be noted that when the ripple amplitude is of the order of $\epsilon_{\min}$, a more rapid convergence may be obtained by replacing $(-1)^{n+1}$ in (8) with $(-1)^n$. This has the effect of making the average group delay of $T_{l,h}(\omega)$ at ω_c equal to $T_0-\epsilon$. In this case a small negative ϕ, which may be corrected quickly, can be expected.

IV. Conclusions

A method for designing equiripple group delay, all-pole filter sets, for use in frequency-discriminating FSK modems has been presented. The design algorithm involves the simultaneous minimization of two objective functions. In practice, the algorithm has always been found to converge rapidly to the required solution, provided that the restrictions regarding minimum ripple amplitude are observed. From a theoretical viewpoint, however, it is not clear that such an optimum will always exist. In addition, it has been shown that the minimum ripple amplitude obtainable is dependent upon the order of the receive filter, the order of the discriminators and the width of the receive band. Results have been presented to illustrate the relationship between these parameters.

The algorithm has been successfully used to design the receive-discriminator filter set for an all-digital implementation of a (V-21 type) 300 baud FSK modem. Details of this implementation are reported in [2].

Acknowledgment

Acknowledgement is given to the Director of Research of British Telecom for permission to publish this paper. The author also wishes to thank P. Challener, P. C. Millar, and J. M. Rollett of BTRL for their helpful advice and discussions, and for providing the modem simulation software.

References

[1] A. G. Deczky, "Recursive digital filters having equiripple group delay," *IEEE Trans. Circuits Syst.*, vol. CAS-21, pp. 131–134, Jan. 1974.

[2] P. M. Hughes, "Digital filters for use in an FSK modem," *Proc. Conf. on Digital Processing of Signals in Communications*, Loughborough, England, 1985, to be published.

Ill-Posed Sampling Theorems

KWAN FAI CHEUNG AND ROBERT J. MARKS, II

Abstract —**There are a number of innocent appearing sampling theorems that are ill-posed, i.e., a small amount of noise superimposed on the data can render the interpolation unstable. Using Papoulis' Generalized Sampling Theorem, we show that a sufficient condition for a sampling theorem to be ill-posed is that an interpolation function has infinite energy. Specific examples include the case where (a) the signal and the $(2n)$th derivative of the signal are both simultaneously sampled at half the Nyquist rate and (b) the signal's and the $(2n+1)$th derivative's samples are interlaced at Nyquist intervals.**

Manuscript received January 23, 1984; revised June 22, 1984.

The authors are with the Department of Electrical Engineering, University of Washington, Seattle, WA 98195.

I. Introduction

There have been a large number of generalizations of the sampling theorem. A band-limited signal can be regained from samples of the output of an all-pass filter, bunched samples, or signal-derivative samples [1]–[6]. In each case, the average sampling rate must equal or exceed the Nyquist rate. One could infer that any such set of independent data taken at the Nyquist rate might suffice to uniquely specify the signal. Indeed, such statements have been made in textbooks. Although possibly true in the absence of noise, there are certain cases where a small perturbation on sample values yields unbounded interpolation noise levels.

An example is signal and derivative sampling. Shannon [7] was the first to note that one could sample at half the Nyquist rate if at each sample location two samples were taken: one of the signal and one of the signal's derivative. This sampling theorem was generalized to m derivatives by Linden [1] and has found its way into a number of tutorials and texts [4]–[6]. Consider the seemingly innocent alteration of sampling at the Nyquist rate with interlaced signal and first derivative samples taken at each Nyquist interval. As we will demonstrate, restoration here is ill-posed. Indeed, subjecting the samples to sample-wise white noise renders the restoration unstable. Hence, one would wish to sample an odometer and speedometer simultaneously, rather than sequentially, to determine position.

Our purpose herein is to discuss a class of ill-posed sampling theorems as generated by Papoulis' Generalized Sampling Theorem [4], [8]. Sufficient conditions for ill-posedness will be give along with some specific examples. Clearly, such sampling theorems should be avoided.

II. Preliminaries

A finite energy signal is said to be σ-band limited if

$$f(t)=\frac{1}{2\pi}\int_{-\sigma}^{\sigma}F(\omega)\,e^{j\omega t}\,d\omega$$

where

$$F(\omega)=\mathscr{F}f(t)=\int_{-\infty}^{\infty}f(t)\,e^{-j\omega t}\,dt$$

and $\mathscr{F}$ denotes the Fourier transform operator.

Much work has been done on finding $f(t)$ from either partial knowledge or filtered versions of $f(t)$. Define $p_\tau(t)$ as unity for $|t|<\tau$ and zero, otherwise. Then, regaining $f(t)$ from $f(t)p_\tau(t)$ and $f(t)[1-p_\tau(t)]$ are, respectively, the classic extrapolation and interpolation problems. Restoring $f(t)$ from its samples $\{f(nT)|-\infty<n<\infty\}$ results in the classic Shannon sampling theorem [3]:

$$f(t)=\sum_{n=-\infty}^{\infty}f(nT)\frac{\sin\sigma(t-nT)}{\sigma(t-nT)},\qquad T=\pi/\sigma.$$

In any practical restoration procedure, the known data are accompanied by noise. If the noise is additive (and signal independent) and the restoration algorithm is linear, then the restoration algorithm yields $f(t)+\eta(t)$ as its result where $\eta(t)$ is the algorithm response to the data noise alone. The restoration-noise level is then

$$\overline{\eta^2(t)} = E\left[|\eta(t)|^2\right]$$

where E denotes the expectation operator. If the input noise level is bounded and $\overline{\eta^2(t)}$ is not, then the algorithm is ill-posed.

The extrapolation problem is ill-posed [9]–[12]. The sampling theorem and interpolation problems are well-posed [9], [13]. There are cases where the restoration noise level can be bounded over finite intervals rendering a globally ill-posed problem locally well posed [14].

III. Generalized Sampling Theorem

Many of the generalizations of the sampling theorem discussed in the introduction were eloquently brought under the umbrella of a single theorem by Papoulis [4], [8]. Briefly stated, let $\{g_k(t)|k=1,2,\cdots,m\}$ be the outputs from m specified filters with transfer functions $\{H_k(\omega)\}$ and common input $f(t)$. We sample each $g_k(t)$ at $1/m$th the Nyquist rate. The input can then be restored by the interpolation formula:

$$f(t) = \sum_{k=1}^{m} \sum_{n=-\infty}^{\infty} g_k(nT)\, y_k(t-nT), \qquad T=\frac{m\pi}{\sigma}. \quad (1)$$

The interpolation functions are found by

$$y_k(t) = \frac{1}{c}\int_{-\sigma}^{-\sigma+c} Y_k(\omega,t)\, e^{j\omega t}\, d\omega, \qquad k=1,2,\cdots,m \quad (2)$$

where $c=2\pi/T=2\sigma/m$ and the $Y_k(\omega,t)$'s are solutions of the set of equations:

$$\begin{bmatrix} H_1(\omega) & H_2(\omega) & \cdots & H_m(\omega) \\ H_1(\omega+c) & H_2(\omega+c) & \cdots & H_m(\omega+c) \\ \vdots & \vdots & & \vdots \\ H_1[\omega+(m-1)c] & H_2[\omega+(m-1)c] & \cdots & H_m[\omega+(m-1)c] \end{bmatrix} \begin{bmatrix} Y_1(\omega,t) \\ Y_2(\omega,t) \\ \vdots \\ Y_m(\omega,t) \end{bmatrix} = \begin{bmatrix} 1 \\ e^{jct} \\ \vdots \\ e^{j(m-1)ct} \end{bmatrix}. \quad (3)$$

Here, t is arbitrary and $-\sigma<\omega<-\sigma+c$. Clearly there is no solution if the H matrix is identically zero over any finite subinterval.

IV. Noise Sensitivity

In this section, we explore the sensitivity of the Generalized Sampling Theorem to sample wise white noise. We demonstrate that a sufficient condition for a sampling theorem to be ill-posed is that the energy of any one of the m interpolation functions is infinite.

Let $\{\xi_k(nT)|k=1,2,\cdots,m;\ -\infty<n<\infty\}$ denote a zero mean discrete stochastic noise sequence. If $g_k(nT)+\xi_k(nT)$ is used in (3) instead of $g_k(nT)$, the output is $f(t)+\eta(t)$ where

$$\eta(t) = \sum_{k=1}^{m} \sum_{n=-\infty}^{\infty} \xi_k(nT)\, y_k(t-nT).$$

We will assume that the noise is stationary and sample-wise white:

$$E\left[\xi_k(nT)\xi_q^*(lT)\right] = \overline{\xi^2}\,\delta_{n-l}\delta_{k-q}$$

where δ_n is the Kronecker delta, $\overline{\xi^2}=E[|\xi_k(nT)|^2]$ is the data noise level, and the asterisk denotes complex conjugate. The interpolation noise level then follows as

$$\overline{\eta^2(t)} = \overline{\xi^2} \sum_{k=1}^{m} \sum_{n=-\infty}^{\infty} |y_k(t-nT)|^2.$$

Clearly, $\overline{\eta^2(t)}$ is periodic with period T. Application of the Poisson sum formula yields

$$\overline{\eta^2(t)} = \frac{\overline{\xi^2}}{T} \sum_{k=1}^{m} \sum_{n=-\infty}^{\infty} W_k(nc)\, e^{jnct} \quad (4)$$

where

$$W_k(\omega) = \mathscr{F}|y_k(t)|^2 = \frac{1}{2\pi}\int_{-\sigma}^{\sigma} Y_k(\beta)\, Y_k^*(\beta-\omega)\, d\beta$$

and

$$Y_k(\omega) = \mathscr{F} y_k(t). \quad (5)$$

Note that (4) is simply a Fourier series with coefficients

$$c_n = \frac{\overline{\xi^2}}{T} \sum_{k=1}^{m} W_k(nc).$$

We, accordingly, define the average interpolation noise level by

$$c_0 = \frac{\overline{\xi^2}}{T} \sum_{k=1}^{m} W_k(0) = \frac{\overline{\xi^2}}{2\pi T} \sum_{k=1}^{m} \int_{-\sigma}^{\sigma} |Y_k(\omega)|^2\, d\omega. \quad (6)$$

or, using Parseval's Theorem

$$c_0 = \frac{\overline{\xi^2}}{T} \sum_{k=1}^{m} \int_{-\infty}^{\infty} |y_k(t)|^2\, dt.$$

Thus the average interpolation noise level is infinite if any one of the m interpolation functions has unbounded energy.

V. Examples

1. Derivative Sampling

Consider the $m=1$ case corresponding to pth-order derivative sampling

$$H_1(\omega) = (j\omega)^p.$$

We can, in principle, regain all frequency components other than zero. Note, however, that (6) becomes

$$c_0 = \frac{\overline{\xi^2}}{2\pi T}\int_{-\sigma}^{\sigma} \omega^{-2p}\, d\omega = \infty.$$

Thus the corresponding sampling theorem is ill-posed.

2. Interlaced Signal-Derivative Sampling

A less obvious ill-posed sampling theorem arises when we nonuniformly interlace pth order derivative samples with signal

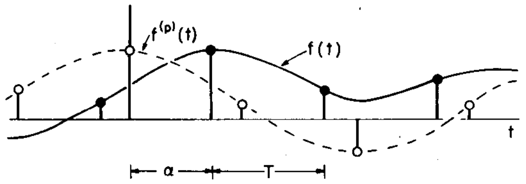

Fig. 1. Interlaced Signal—Derivative Sampling. The hollow dots represent samples of the pth derivative and the solid dots are signal samples. The restoration problem is ill-posed if (a) $\alpha = 0$ and p is even, or (b) $\alpha = T/2$ and p is odd. The Nyquist interval is $T/2$.

samples. For this $m = 2$ sampling theorem, the corresponding filters are

$$H_1(\omega) = (j\omega)^p$$
$$H_2(\omega) = e^{j\alpha\omega}.$$

The filter outputs are thus

$$g_1(t) = f^{(p)}(t)$$
$$g_2(t) = f(t + \alpha).$$

The sampling geometry is illustrated in Fig. 1.

Solving (3) and using (2) and (5) gives

$$Y_1(\omega) = \frac{T}{j^p}\left[\frac{e^{j\alpha\sigma} p_{\sigma/2}\left(\omega + \frac{\sigma}{2}\right)}{\Delta(\omega)} - \frac{p_{\sigma/2}\left(\omega - \frac{\sigma}{2}\right)}{\Delta(\omega - \sigma)}\right] \tag{7a}$$

$$Y_2(\omega) = T\left[\frac{-(\omega + \sigma)e^{j\alpha\sigma} p_{\sigma/2}\left(\omega + \frac{\sigma}{2}\right)}{\Delta(\omega)} - \frac{(\omega - \sigma)^p e^{j(\omega - \sigma)\alpha} p_{\sigma/2}\left(\omega - \frac{\sigma}{2}\right)}{\Delta(\omega - \sigma)}\right] \tag{7b}$$

where

$$\Delta(\omega) = \omega^p e^{j\alpha\sigma} - (\omega + \sigma)^p.$$

Clearly, both $Y_1(\omega)$ and $Y_2(\omega)$ have poles either when $\Delta(\omega) = 0$ or $\Delta(\omega - \sigma) = 0$. The former occurs when

$$\omega^p e^{j(\alpha\sigma + 2\pi n)} = (\omega + \sigma)^p$$

or

$$\omega = \frac{-\sigma}{2} - j\frac{\sigma}{2}\cot\left(\frac{\alpha\sigma + 2\pi n}{2p}\right), \qquad n = 0, 1, \cdots, p-1.$$

One of these roots is real when (a) $\alpha = 0$ and p is even, or (b) $\alpha = T/2$ and p is odd (corresponding to $n = p/2$ and $n = (p-1)/2$, respectively.) In either case, the real pole generated by $\Delta(\omega)$ is at $-\sigma/2$ and that generated by $\Delta(\omega - \sigma)$ is at $\sigma/2$. Clearly, application of (6) exposes this class of sampling theorems as ill-posed.

VI. Notes

1. *Sample Contributions in the Ill-Posed Sampling Theorems*

Insight into the ill-posedness of the sampling theorems can be gained by inspection of the interpolation functions. Consider, for example, $m = 1$ derivative sampling with $p = 1$. It follows that

$$y_1(t) = \frac{1}{2\sigma}\int_{-\sigma}^{\sigma} \frac{e^{j\omega t}\,d\omega}{j\omega} = \frac{1}{\sigma} Si(\sigma t)$$

where the sine integral is

$$Si(t) = \int_0^t \frac{\sin\tau}{\tau}\,d\tau.$$

Since $Si(t) \xrightarrow[t \to \pm\infty]{} \pm\pi/2$, interpolation at any point is affected significantly by every sample value, no matter how distant.

A similar contribution occurs for the ill-posed cases of interlaced signal-derivative sampling. We can, in general, invert (7) using contour integration [15]. For $p = 2$ and $\alpha = 0$, the results are

$$y_1(t) = \frac{1}{2}\left[\sin\left(\frac{\sigma t}{2}\right) Si\left(\frac{\sigma t}{2}\right) + \cos\left(\frac{\sigma t}{2}\right)\operatorname{sinc}\left(\frac{\sigma t}{2\pi}\right) + \operatorname{sinc}^2\left(\frac{\sigma t}{2\pi}\right)\right]$$

$$y_2(t) = \frac{2}{\sigma^2}\sin\left(\frac{\sigma t}{2}\right) Si\left(\frac{\sigma t}{2}\right)$$

where $\operatorname{sinc} x = \sin(\pi x)/(\pi x)$. Again, the occurrence of the sine integrals makes possible equally significant contributions from all sample values, no matter how far removed from the point of interpolation. The weighted noise levels from each sample value thus add to a random variable with unbounded variance.

2. *Effects of Oversampling*

Suppose $f(t)$ is $r\sigma$-band limited, where $r < 1$. Then it is also σ band limited. Thus the generalized sampling theorem expression is applicable. The transform of (1) is

$$F(\omega) = \left[\sum_{k=1}^{m} Y_k(\omega) \sum_{n=-\infty}^{\infty} g_k(nT) e^{-jn\omega T}\right] p_\sigma(\omega).$$

We can pass $F(\omega)$ through a low-pass filter unity for $|\omega| < r\sigma$ and zero elsewhere. The result is

$$F(\omega) = \left[\sum_{k=1}^{m} Y_k(\omega) \sum_{n=-\infty}^{\infty} g_k(nT) e^{-jn\omega T}\right] p_{r\sigma}(\omega).$$

In the time domain, this is equivalent to using the interpolation function set $\{\hat{y}_k(t)\}$ in place of $\{y_k(t)\}$ in (1) where

$$\hat{y}_k(t) = \frac{1}{2\pi}\int_{-r\sigma}^{r\sigma} Y_k(\omega) e^{j\omega t}\,d\omega.$$

Going through the same analysis as before, we find the average interpolation noise level for the oversampled case is

$$\hat{c}_0 = \frac{\overline{\xi^2}}{2\pi T}\sum_{k=1}^{m}\int_{-r\sigma}^{r\sigma} |Y_k(\omega)|^2\,d\omega. \tag{7}$$

Comparing with (6), we conclude that $\hat{c}_0 \leqslant c_0$. Oversampling, in general, thus buys us a lower average interpolation noise level [13], [15].

Consider, then, the ill-posed interlaced derivative signal sampling theorem. If we sample at a rate greater than twice the Nyquist rate, the integral in (7) will not include the poles at $\omega = \pm\sigma/2$ and the resulting sampling theorem becomes well-posed.

At exactly twice the Nyquist rate, the integration limits in (7) are at the pole locations. Thus $\hat{c}_0 = \infty$. We can, however, discard the derivative samples and use the conventional (well-posed) sampling theorem to restore the signal. Thus we are confronted with the curious task of discarding the derivative samples to improve the interpolation noise level.

Acknowledgment

The authors wish to thank Loretta Tollefson for her help in preparing the manuscript and one of the reviewers for his/her helpful comments.

References

[1] D. A. Linden, "A discussion of sampling theorems," *Proc. IRE*, vol. 47, pp. 1219–1226, 1959.
[2] A. Nathan, "On sampling a function and its derivatives," *Inform. Contr.*, vol. 22, pp. 172–182, 1973.
[3] A. J. Jerri, "The Shannon sampling theorem—Its various extensions and applications: A tutorial review," *Proc. IEEE*, vol. 65, p. 1565, 1977.
[4] A. Papoulis, *Signal Analysis*. New York: McGraw-Hill, 1977.
[5] R. Bracewell, *The Fourier Transform and Its Applications*. New York: McGraw-Hill, 1965.
[6] J. D. Gaskill, *Linear Systems, Fourier Transforms and Optics*. New York: Wiley, 1978, pp. 266–285.
[7] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, p. 12, 1948.
[8] A. Papoulis "Generalized Sampling Expansion" *IEEE Trans. Circuits Syst.*, vol. CAS-24, pp. 652–654, 1977.
[9] R. J. Marks II and D. K. Smith, "Gerchberg-type linear deconvolution and extrapolation algorithms," in *Transformations in Optical Signal Processing*, W. T. Rhodes, J. R. Fienup, and B. E. A. Saleh, Eds., Society of Photo-Optics Instrumentation Engineers, Bellingham, Washington, DC, 1983.
[10] H. Stark, D. Cahana, and H. Webb, "Restoration of arbitrary finite energy optical objects from limited spatial and spectral information," *J. Opt. Soc. Am.*, vol. 71, p. 635, 1981.
[11] T. K. Sarkar, D. D. Wiener, and V. K. Jain, "Some mathematical considerations in dealing with the inverse problem," *IEEE Trans. Antennas Propagat.*, vol. AP-29, p. 373, 1981.
[12] D. C. Youla, "Generalized image restoration by method of alternating orthogonal projections," *IEEE Trans. Circuits Syst.*, vol. CAS-25, p. 694, 1978.
[13] R. J. Marks II, "Noise sensitivity of band-limited signal derivative interpolation," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-31, pp. 1028–1032, 1983.
[14] R. J. Marks II, "Posedness of a bandlimited image-extension problem in tomography," *Opt. Let.*, 7, pp. 376–377, 1982.
[15] K. F. Cheung, *The generalized sampling expansion: its stability, posedness, and discrete equivalent*, thesis, Univ. Washington, Dep. Elect. Eng., Seattle, WA, 1983.

Stability Assessment of Two-Dimensional State-Space Systems

K. V. FERNANDO AND H. NICHOLSON

Abstract—The two-dimensional Fornasini–Marchesini model [1], [2] is one of the most general state-space models available. Although conditions for stability for this model are theoretically simple, actual numerical verification is not a trivial exercise. One way to overcome this problem is to compute the matrix norm $\max \|A_1 + e^{jw}A_2\|_1$ for real w, and the system is stable if this value is less than unity. We compute this value by transforming the two-dimensional system into a canonical form based on the generalized eigenstructure of the state matrices A_1 and A_2.

I. Introduction

The two-dimensional Fornasini–Marchesini model [1], [2] is one of the most general state-space models available as it imbeds other known models [1]. Thus by studying the stability of the Fornasini–Marchesini model, it is possible to obtain results also for less general models.

Manuscript received May 6, 1984; revised August 22, 1984.

The authors are with Department of Control Engineering, University of Sheffield, Sheffield S1 3JD, England.

Although conditions for stability of the Fornasini–Marchesini model are theoretically simple and elegant, actual numerical verification is difficult and sometimes impossible due to the very high burden of computation. This difficulty is essentially due to the use of one-dimensional conditions for stability in verification of two-dimensional stability.

It is known that the Fornasini–Marchesini model is stable if and only if the matrix $(A_1 + e^{jw}A_2)$ is stable for all real w, or equivalently the maximum spectral radius of that matrix with respect to w should be less than unity. Thus by knowing the maximum value of the norm $\|A_1 + e^{jw}A_2\|_1$ with respect to w, the stability of the system can be determined. In this article we propose a method for evaluating this norm by transforming the system into a canonical form based on the generalized eigenstructure of the state matrices A_1 and A_2.

II. Preliminaries

We consider the linear, stationary, finite dimensional, double-indexed dynamical system $S(A_1, A_2, B_1, B_2, C)$ defined by the first-order partial difference equation [1], [2]

$$\begin{aligned} x(h+1,k+1) &= A_1 x(h,k+1) + A_2 x(h+1,k) \\ &\quad + B_1 u(h,k+1) + B_2 u(h+1,k) \\ y(h,k) &= Cx(h,k) \end{aligned}$$

where $u(h,k)$ is the input and $y(h,k)$ is the output at "time" (h,k). We assume that $u(h,k)$ and $y(h,k)$ are defined in the field of real numbers and (h,k) takes integer values. We further assume that the local state-space is n-dimensional and thus

$$\begin{array}{llll} x \in R^{n\times 1} & A_i \in R^{n\times n} & B_i \in R^{n\times m}, & i=1,2, \\ u \in R^{m\times 1} & y \in R^{r\times 1} & C \in R^{r\times n} & \end{array}$$

The two-dimensional z-transform of the system S is given by

$$\tilde{y}(z_1,z_2)/\tilde{u}(z_1,z_2) = C(I - z_1A_1 - z_2A_2)^{-1}(z_1B + z_2B)$$

where the forward shift operators z_1 and z_2 can be associated with the indices h and k, respectively.

The state transitions of the system are based on the "shuffle product" [1], [2] of the matrices A_1 and A_2 which is defined as

$$A_1{}^i \sqcup\!\sqcup^j A_2 \in R^{n\times n}$$

where

$$\begin{aligned} &A_1{}^i \sqcup\!\sqcup^0 A_2 = A_1^i \\ &A_1{}^0 \sqcup\!\sqcup^j A_2 = A_2^j \\ &A_1{}^i \sqcup\!\sqcup^j A_2 = A_1\left(A_1{}^{i-1} \sqcup\!\sqcup^j A_2\right) + A_2\left(A_1{}^i \sqcup\!\sqcup^{j-1} A_2\right) \end{aligned}$$

For example,

$$\begin{aligned} A_1{}^3 \sqcup\!\sqcup^2 A_2 &= A_1^3A_2^2 + A_1^2A_2A_1A_2 + A_1A_2A_1^2A_2 + A_2A_1^3A_2 \\ &\quad + A_2^3A_1^3 + A_2A_1^2A_2A_1 + A_2A_1A_2A_1^2 \\ &\quad + A_1^2A_2^2A_1 + A_1A_2^2A_1^2 + A_1A_2A_1A_2A_1. \end{aligned}$$

We note that the shuffle product $A_1{}^i \sqcup\!\sqcup^j A_2$ is composed of ${}^{i+j}C_i$ matrix product terms and computations of these products for large i, j is not an easy task. For example for $i=3$ and $j=2$, the binomial coefficient ${}^5C_2 = 10$ and for $i=10$, $j=5$, ${}^{15}C_5 = 3003$.