

# Análise do Comportamento Estocástico do Algoritmo KLMS para um Sinal de Entrada Correlacionado

Wemerson D. Parreira, Márcio H. Costa e José C. M. Bermudez

**Resumo**— O algoritmo Kernel Least-Mean-Square (KLMS) é um algoritmo popular para filtragem adaptativa não-linear devido à sua simplicidade e robustez. Em filtragem adaptativa baseada em kernel com modelos de ordem finita as estatísticas do sinal de entrada do filtro linear dependem do kernel e dos seus parâmetros, assim como da política de atualização do dicionário de vetores de entrada que define o espaço de Hilbert em que o filtro opera. A existência de relações altamente não-lineares entre os parâmetros do filtro adaptativo e os critérios de desempenho faz do projeto desses filtros uma tarefa quase impossível sem a ajuda de modelos analíticos que prevejam seu desempenho. Análises do comportamento estocástico do KLMS já foram realizadas considerando dicionários fixos e dicionários variáveis sob a consideração de que a sequência temporal dos vetores de entrada é estatisticamente independente. Este trabalho estuda o comportamento médio do vetor de coeficientes adaptativos do KLMS com kernel Gaussiano para dicionários variáveis e vetores de entrada correlacionados no tempo.

**Palavras-Chave**— Filtragem adaptativa, kernel least-mean-square (KLMS), sistema não-linear, entrada correlacionada.

**Abstract**— The Kernel Least-Mean-Square (KLMS) algorithm is a popular algorithm for adaptive nonlinear filtering due to its simplicity and robustness. In kernel-based adaptive filtering with finite order models, the statistics of the linear filter depend on the kernel and its parameters, as well as, on the input vector dictionary update rule that defines the Hilbert space in which the filter operates. The existence of the highly nonlinear relationships between the adaptive filter parameters and the performance criteria makes the design of these filters an almost impossible task without the help of analytical models that predict their performance. Theoretical analyses of the stochastic behavior of the KLMS have already been performed assuming fixed dictionaries and variable dictionaries under the consideration that the temporal sequence of the input vectors is statistically independent. This paper studies the KLMS mean-weight behavior with Gaussian kernel for variable dictionaries and time-correlated input vectors.

**Keywords**— Adaptive filtering, kernel least-mean-square (KLMS), nonlinear system, correlated input-signal.

## I. INTRODUÇÃO

Aplicações práticas em diversas áreas como comunicações e engenharia biomédica requerem um processamento não linear dos sinais. Métodos de estimação não linear baseados em kernels têm se popularizado nas últimas décadas [1], [2]. Recentemente, os filtros adaptativos baseados em kernel têm sido

Wemerson D. Parreira, Márcio H. Costa e José C. M. Bermudez, Departamento de Engenharia Elétrica e Eletrônica, Univ. Federal de Santa Catarina, Florianópolis-SC, Brasil, E-mails: wemerson.d.p@ufsc.br, costa@eel.ufsc.br e j.bermudez@ieec.org. Este trabalho foi parcialmente financiado pela CAPES (41001010005P1-PNPD) e pelo CNPq (304867/2015-2 e 307071/2013-8).

reconhecidos como uma das melhores soluções para problemas de estimação não linear em tempo real. Sua popularidade baseia-se no mapeamento do espaço de sinais de entrada para um espaço de Hilbert definido por kernel reproduzível (Reproducing Kernel Hilbert Space – RKHS), mapeamento esse que permite o uso de estruturas lineares na solução de problemas não lineares. Uma visão geral dos algoritmos de filtragem adaptativa baseada em kernel é apresentada em [3]. A Figura 1 mostra um diagrama em blocos típico de problemas de estimação não linear baseada em kernel. Nessa

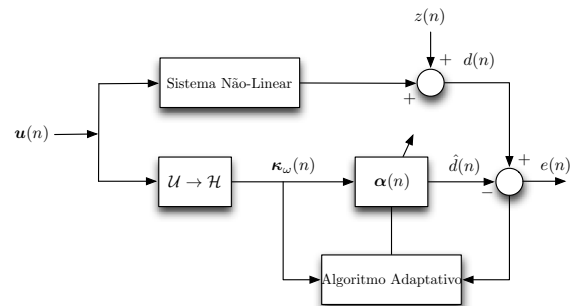


Fig. 1. Estimação de sistemas utilizando algoritmos KAF.

figura,  $\mathcal{U} \subseteq \mathbb{R}^q$  é o espaço dos vetores de entrada  $\mathbf{u}(n)$ ,  $\kappa: \mathcal{U} \times \mathcal{U} \rightarrow \mathbb{R}$  é o kernel que define o RKHS  $\mathcal{H}$  induzido pelo produto interno  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ ,  $d(n)$  é a soma do sinal a ser estimado com um ruído aditivo  $z(n)$  de média zero e decorrelacionado de qualquer outro sinal.

O teorema da representação [1] estabelece que a função  $\psi(\cdot)$  que minimiza o erro quadrático da estimação  $\sum_{n=1}^N [d(n) - \psi(\mathbf{u}(n))]^2$ , dados  $N$  vetores de entrada  $\mathbf{u}(n)$  e saídas desejadas  $d(n)$ , pode ser escrita como uma expansão em termos de funções kernel  $\psi(\cdot) = \sum_{n=1}^N \alpha_n \kappa(\cdot, \mathbf{u}(n))$ . O Problema se reduz então a determinar o vetor  $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_N]^T$  que minimiza  $\|\mathbf{d} - \mathbf{K}\boldsymbol{\alpha}\|^2$ , em que  $\mathbf{K}_{N \times N}$  é a matriz Gram com  $(n, \ell)$ -ésima entrada dada por  $\kappa(\mathbf{u}(n), \mathbf{u}(\ell))$  e  $\mathbf{d} = [d(1), \dots, d(N)]^T$ . Note que neste caso a ordem do problema, e conseqüentemente a ordem do filtro adaptativo, cresce com o número  $N$  de observações do sinal de entrada  $\mathbf{u}(\cdot)$ . Tal situação inviabiliza a aplicação dessa metodologia na resolução de problemas em tempo real. Uma forma de contornar essa limitação é utilizar modelos de ordem limitada da forma

$$\psi(\cdot) = \sum_{m=1}^M \alpha_m \kappa(\cdot, \mathbf{u}(\omega_m)) \quad (1)$$

em que  $M$  funções kernel  $\{\kappa(\cdot, \mathbf{u}(\omega_m))\}_{m=1}^M$  compõem o dicionário  $\mathcal{D}$  de dimensão  $M$ , em que  $\omega_j, j = 1, \dots, M$  formam um subconjunto de  $\{1, 2, \dots, N\}$  correspondendo aos índices temporais dos  $M < N$  vetores de entrada escolhidos para compor o modelo (1). Algoritmos adaptativos baseados nessa formulação incluem o *Kernel Least-Mean-Square* (KLMS) [4], objeto deste estudo, o *Kernel Recursive-Least-Squares* (KRLS) [5], o KLMS normalizado (KNLMS) e o *Kernel Affine Projection* (KAPA) [6], [7].

Além da escolha da ordem  $M$  do modelo, o dicionário pode ser estático, se um conjunto fixo de vetores for escolhido, ou dinâmico, se o dicionário for atualizado a intervalos temporais pré-definidos. O segundo caso é especialmente importante nos casos de dicionários pequenos ou de sinais não estacionários.

O comportamento estocástico do algoritmo KLMS com ordem finita foi analisado em [8] para sequências de entrada compostas por vetores  $\mathbf{u}(n)$  Gaussianos, identicamente distribuídos e independentes (i.i.d.), e com dicionário dinâmico. Em [9] o comportamento do KLMS foi estudado no caso de emprego de um dicionário pré-projetado e fixo durante a operação. Essa análise foi depois estendida para o caso de sinais e coeficientes complexos em [10].

Este trabalho avança no estudo do comportamento do algoritmo KLMS ao considerar a importante implementação com dicionário dinâmico e sequência de entrada composta por vetores  $\mathbf{u}(n)$  Gaussianos e correlacionados no tempo, situação ainda não estudada na literatura. São determinadas as expressões analíticas das matrizes de correlação necessárias para o estudo do comportamento médio dos coeficientes adaptativos. Resultados numéricos de simulação Monte Carlo ilustram a acurácia do novo modelo na previsão do comportamento do algoritmo durante o transitório e em regime permanente.

## II. CONSTRUÇÃO DO DICIONÁRIO

Diferentes métodos podem ser empregados para selecionar um conjunto esparsos dos vetores de entrada para compor o modelo (1). Algumas dessas técnicas estão descritas em [3], [11]. Uma técnica popular por ser eficaz e de baixa complexidade emprega a função coerência. Por essa regra, uma função kernel  $\kappa(\cdot, \mathbf{u}(\ell))$  é inserida no dicionário se

$$\max_j |\kappa(\mathbf{u}(\ell), \mathbf{u}(\omega_j))| \leq \varepsilon_0 \quad (2)$$

em que  $\varepsilon_0$  determina a coerência do dicionário<sup>1</sup>. Neste trabalho consideramos a utilização da coerência para a definição da dimensão  $M$  do dicionário e para sua inicialização. Durante a operação o dicionário é atualizado a cada iteração utilizando uma linha de retardo. Assim, na  $k$ -ésima iteração  $\mathcal{D}(k) = \{\kappa(\cdot, \mathbf{u}(k-1)), \dots, \kappa(\cdot, \mathbf{u}(k-M))\}$ .

## III. ANÁLISE DO ERRO QUADRÁTICO MÉDIO

Considere o problema descrito na Figura 1 de estimação do sinal de saída de um sistema não linear usando o modelo de ordem finita (1) baseado no kernel Gaussiano:

$$\kappa(\mathbf{u}, \mathbf{u}') = \exp[-\|\mathbf{u} - \mathbf{u}'\|_2^2 / 2\xi^2] \quad (3)$$

<sup>1</sup>Mostra-se em [7] que essa regra leva a um dicionário de dimensão finita.

em que  $\xi \in \mathbb{R}_+^*$  é o parâmetro do kernel,  $\mathbf{u}, \mathbf{u}' \in \mathcal{U} \subseteq \mathbb{R}^q$  e  $\|\cdot\|_2$  é a norma Euclidiana. O ambiente é assumido estacionário, significando que  $\psi(\mathbf{u}(n))$  é estacionário para  $\mathbf{u}(n)$  estacionário. Essa condição é satisfeita por vários sistemas não lineares usados para modelar situações práticas, tais como sistemas sem memória, Wiener e Hammerstein [8]. A entrada do sistema é uma sequência de vetores  $\mathbf{u}(n)$  Gaussianos, identicamente distribuídos e com matriz de correlação  $\mathbf{R}_{\mathbf{u}\mathbf{u}, \ell} = E\{\mathbf{u}(n) \mathbf{u}^\top(n-\ell)\}$ .

Seja  $\boldsymbol{\kappa}_\omega(n)$  o vetor de funções kernel no instante  $n$ , isto é,

$$\boldsymbol{\kappa}_\omega(n) = [\kappa(\mathbf{u}(n), \mathbf{u}(\omega_1)), \dots, \kappa(\mathbf{u}(n), \mathbf{u}(\omega_M))]^\top \quad (4)$$

em que  $\kappa(\cdot, \mathbf{u}(\omega_m))$  é a  $m$ -ésima função que forma o dicionário e  $\mathbf{u}(n) \neq \mathbf{u}(\omega_m)$ . Da Figura 1 e do modelo (1), a estimativa da saída desejada é

$$\hat{d}(n) = \boldsymbol{\alpha}^\top(n) \boldsymbol{\kappa}_\omega(n) \quad (5)$$

em que  $\boldsymbol{\alpha}(n) = [\alpha_1(n), \dots, \alpha_M(n)]^\top$ . O correspondente erro de estimação é definido como:

$$e(n) = d(n) - \hat{d}(n). \quad (6)$$

Elevando ao quadrado ambos os lados de (6) e calculando o valor esperado  $E\{\cdot\}$  condicionado em  $\boldsymbol{\alpha}(n) = \boldsymbol{\alpha}$  obtém-se a já conhecida expressão para o erro quadrático médio (*mean-square error* – MSE)

$$\begin{aligned} J_{\text{mse}}(n) &= E\{e^2(n)\} \\ &= E\{d^2(n)\} - 2\mathbf{p}_{\kappa d}^\top \boldsymbol{\alpha} + \boldsymbol{\alpha}^\top \mathbf{R}_{\kappa\kappa} \boldsymbol{\alpha} \end{aligned} \quad (7)$$

em que  $\mathbf{R}_{\kappa\kappa} = E\{\boldsymbol{\kappa}_\omega(n) \boldsymbol{\kappa}_\omega^\top(n)\}$  é a matriz de correlação do vetor de entrada do filtro adaptativo, e  $\mathbf{p}_{\kappa d} = E\{d(n) \boldsymbol{\kappa}_\omega(n)\}$  é o vetor de correlação cruzada entre  $\boldsymbol{\kappa}_\omega(n)$  e  $d(n)$ . Assumindo que  $\mathbf{R}_{\kappa\kappa}$  é definida positiva, o vetor de coeficientes ótimo é dado por

$$\boldsymbol{\alpha}_{\text{opt}} = \mathbf{R}_{\kappa\kappa}^{-1} \mathbf{p}_{\kappa d} \quad (8)$$

denominado solução de Wiener. O correspondente erro quadrático médio mínimo (MMSE) para uma entrada  $\boldsymbol{\kappa}_\omega(n)$  é

$$J_{\text{min}} = E\{d^2(n)\} - \mathbf{p}_{\kappa d}^\top \mathbf{R}_{\kappa\kappa}^{-1} \mathbf{p}_{\kappa d}. \quad (9)$$

Assim, para determinar a solução de Wiener  $\boldsymbol{\alpha}_{\text{opt}}$  é necessário determinar a matriz  $\mathbf{R}_{\kappa\kappa}$ , considerando as propriedades estatísticas de  $\mathbf{u}(n)$  e da função kernel utilizada  $\kappa(\cdot, \cdot)$ .

### A. Matriz de correlação – $\mathbf{R}_{\kappa\kappa}$

Diferentemente da análise em [8] para uma entrada i.i.d., no caso estudado aqui os elementos de  $\mathbf{R}_{\kappa\kappa}$  são afetados pela correlação entre a entrada  $\mathbf{u}(n)$  e cada elemento  $\{\mathbf{u}(\omega_m)\}_{m=1}^M$  do dicionário.

A matriz  $\mathbf{R}_{\kappa\kappa}$  tem componentes dadas por:

$$[\mathbf{R}_{\kappa\kappa}]_{ij} = \begin{cases} E\{[\boldsymbol{\kappa}_\omega(n)]_i^2\}, & i = j \\ E\{[\boldsymbol{\kappa}_\omega(n)]_i [\boldsymbol{\kappa}_\omega(n)]_j\}, & i \neq j \end{cases} \quad (10)$$

em que  $1 \leq i, j \leq M$ . Para o cálculo dos valores esperados em (10) usam-se as seguintes expressões na forma quadrática:

$$\begin{aligned} \|\mathbf{u}(n) - \mathbf{u}(\omega_i)\|_2^2 &= \mathbf{y}_2^\top \mathbf{Q}_2 \mathbf{y}_2 \\ \|\mathbf{u}(n) - \mathbf{u}(\omega_i)\|_2^2 + \|\mathbf{u}(n) - \mathbf{u}(\omega_j)\|_2^2 &= \mathbf{y}_3^\top \mathbf{Q}_3 \mathbf{y}_3, \text{ com } i \neq j \end{aligned} \quad (11)$$

em que

$$\begin{aligned} \mathbf{y}_2 &= (\mathbf{u}^\top(n) \mathbf{u}^\top(\omega_i))^\top \\ \mathbf{y}_3 &= (\mathbf{u}^\top(n) \mathbf{u}^\top(\omega_i) \mathbf{u}^\top(\omega_j))^\top \end{aligned} \quad (12)$$

e

$$\mathbf{Q}_2 = \begin{pmatrix} \mathbf{I} & -\mathbf{I} \\ -\mathbf{I} & \mathbf{I} \end{pmatrix} \quad \mathbf{Q}_3 = \begin{pmatrix} 2\mathbf{I} & -\mathbf{I} & -\mathbf{I} \\ -\mathbf{I} & \mathbf{I} & \mathbf{O} \\ -\mathbf{I} & \mathbf{O} & \mathbf{I} \end{pmatrix} \quad (13)$$

em que  $\mathbf{I}$  é a matriz identidade ( $q \times q$ ) e  $\mathbf{O}$  é a matriz nula ( $q \times q$ ). Usando a função geradora de momentos  $\psi_z(s)$  da forma quadrática  $z = \mathbf{y}^\top \mathbf{Q} \mathbf{y}$ , a partir de [12, p. 100], em que  $\mathbf{y}$  é um vetor de variáveis aleatórias com distribuição Gaussiana de média  $\boldsymbol{\mu}_y$  e matriz de covariância  $\boldsymbol{\Sigma}_{yy} = E\{(\mathbf{y} - \boldsymbol{\mu}_y)(\mathbf{y} - \boldsymbol{\mu}_y)^\top\}$ , obtém-se:

$$\begin{aligned} \psi_z(s) &= E\{e^{s^\top z}\} = \det\{\mathbf{I} - 2s \mathbf{Q} \boldsymbol{\Sigma}_{yy}\}^{-1/2} \\ &\times \exp\left[(-1/2)\boldsymbol{\mu}_{yy}^\top \boldsymbol{\Sigma}_{yy}^{-1}(\mathbf{I} - 2s \mathbf{Q} \boldsymbol{\Sigma}_{yy})^{-1} \boldsymbol{\mu}_y\right] \end{aligned} \quad (14)$$

Fazendo  $\mathbf{u}(\omega_\ell) = \mathbf{u}(n - \ell)$ , pode-se obter o  $\ell$ -ésimo elemento do vetor de valores esperados de  $\boldsymbol{\kappa}_\omega(n)$ ,

$$\begin{aligned} [\boldsymbol{\mu}_k(n)]_\ell &= [E\{\boldsymbol{\kappa}_\omega(n)\}]_\ell = E\{\boldsymbol{\kappa}(u(n), u(n - \ell))\} \\ &= \det\{\mathbf{G}_\ell\}^{-1/2} \exp\left[-\frac{1}{2}\boldsymbol{\mu}_{y_2}^\top (\mathbf{G}_\ell \boldsymbol{\Sigma}_{y_2 y_2, \ell})^{-1} \boldsymbol{\mu}_{y_2}\right] \end{aligned} \quad (15)$$

em que  $\mathbf{G}_\ell = \mathbf{I} - 2s \mathbf{Q}_2 \boldsymbol{\Sigma}_{y_2 y_2, \ell}$  e  $s = -1/2\xi^2$ . A  $\ell$ -ésima matriz de correlação correspondente é:

$$\mathbf{R}_{y_2 y_2, \ell} = \begin{pmatrix} \mathbf{R}_{uu,0} & \mathbf{R}_{uu,\ell} \\ \mathbf{R}_{uu,\ell}^\mathbf{H} & \mathbf{R}_{uu,0} \end{pmatrix} \quad (16)$$

em que  $\mathbf{R}_{uu,p} = E\{\mathbf{u}(n) \mathbf{u}^\top(n - p)\}$ , com  $p = 0, \dots, M$  e  $\mathbf{R}^\mathbf{H} = (\mathbf{R}^*)^\top$ . A partir deste ponto usa-se simplesmente  $\mathbf{R}_{uu}$  para denotar  $\mathbf{R}_{uu,0}$ .

Assim, para determinar  $\boldsymbol{\Sigma}_{y_2 y_2, \ell}$  é necessário obter a expressão de  $\mathbf{R}_{uu, \ell}$  ( $q \times q$ ) que é dada por:

$$\mathbf{R}_{uu, \ell} = \begin{pmatrix} r_{uu}(\ell) & \dots & r_{uu}(\ell + q - 1) \\ \vdots & \ddots & \vdots \\ r_{uu}(\ell - q + 1) & \dots & r_{uu}(\ell) \end{pmatrix}. \quad (17)$$

Note que para  $\ell \neq 0$  a matriz  $\mathbf{R}_{uu, \ell}$  não é simétrica.

Finalmente, com um procedimento análogo ao usado para obter (15) a partir de (14) e da simetria da matriz  $\mathbf{R}_{\kappa\kappa}$  definida em (10) pode-se obter a  $(i, j)$ -ésima componente da matriz  $\mathbf{R}_{\kappa\kappa}$ , que é:

$$[\mathbf{R}_{\kappa\kappa}]_{ij} = \begin{cases} \det\{\mathbf{G}'_i\}^{-1/2} \\ \times \exp\left[-\frac{1}{2}\boldsymbol{\mu}_{y_2}^\top (\mathbf{G}'_i \boldsymbol{\Sigma}_{y_2 y_2, i})^{-1} \boldsymbol{\mu}_{y_2}\right], & i = j \\ \det\{\mathbf{G}''_{ij}\}^{-1/2} \\ \times \exp\left[-\frac{1}{2}\boldsymbol{\mu}_{y_3}^\top (\mathbf{G}''_{ij} \boldsymbol{\Sigma}_{y_3 y_3, ij})^{-1} \boldsymbol{\mu}_{y_3}\right], & i < j \end{cases} \quad (18)$$

em que  $\mathbf{G}'_i = \mathbf{I}_2 - 4s \mathbf{Q}_2 \boldsymbol{\Sigma}_{y_2 y_2, i}$ ,  $\mathbf{G}''_{ij} = \mathbf{I}_3 - 2s \mathbf{Q}_3 \boldsymbol{\Sigma}_{y_3 y_3, ij}$  com

$$\mathbf{R}_{y_3 y_3, \ell p} = \begin{pmatrix} \mathbf{R}_{uu} & \mathbf{R}_{uu, \ell} & \mathbf{R}_{uu, p} \\ \mathbf{R}_{uu, \ell}^\mathbf{H} & \mathbf{R}_{uu} & \mathbf{R}_{uu, h} \\ \mathbf{R}_{uu, p}^\mathbf{H} & \mathbf{R}_{uu, h}^\mathbf{H} & \mathbf{R}_{uu} \end{pmatrix} \quad (19)$$

e  $h = |\ell - p|$ . Em (18), cada matriz  $\boldsymbol{\Sigma}_{y_\beta y_\beta}$  é uma matriz de covariância de dimensão  $(\beta q \times \beta q)$  do vetor  $\mathbf{y}_\beta$ ,  $\mathbf{I}_\beta$  é uma matriz identidade de dimensão  $(\beta q \times \beta q)$  e  $\det\{\cdot\}$  denota o determinante de uma matriz. Para vetores de entrada  $\{\mathbf{u}(n)\}_{n=0}^\infty$  com média nula, as expressões para se obter  $(i, j)$ -ésima componente da matriz de autocorrelação  $\mathbf{R}_{\kappa\kappa}$  se reduzem a:

$$[\mathbf{R}_{\kappa\kappa}]_{ij} = \begin{cases} \det\{\mathbf{I}_2 - 4s \mathbf{Q}_2 \mathbf{R}_{y_2 y_2, i}\}^{-1/2}, & i = j \\ \det\{\mathbf{I}_3 - 2s \mathbf{Q}_3 \mathbf{R}_{y_3 y_3, ij}\}^{-1/2}, & i < j. \end{cases} \quad (20)$$

#### IV. ANÁLISE DO COMPORTAMENTO TRANSITÓRIO DO VETOR DE ERRO NOS COEFICIENTES: MODELO DE PRIMEIRA ORDEM

A equação de atualização dos coeficientes do algoritmo KLMS para o sistema apresentado na Figura 1 é [3]

$$\boldsymbol{\alpha}(n+1) = \boldsymbol{\alpha}(n) + \eta e(n) \boldsymbol{\kappa}_\omega(n). \quad (21)$$

Definindo-se  $\mathbf{v}(n) = \boldsymbol{\alpha}(n) - \boldsymbol{\alpha}_{\text{opt}}$  obtém-se a equação para atualização do vetor de erro nos coeficientes

$$\mathbf{v}(n+1) = \mathbf{v}(n) + \eta e(n) \boldsymbol{\kappa}_\omega(n). \quad (22)$$

De (5), (6) e da definição de  $\mathbf{v}(n)$ , o erro é dado por

$$e(n) = d(n) - \boldsymbol{\kappa}_\omega^\top(n) \mathbf{v}(n) - \boldsymbol{\kappa}_\omega^\top(n) \boldsymbol{\alpha}_{\text{opt}} \quad (23)$$

e o erro da estimação ótima é

$$e_0(n) = d(n) - \boldsymbol{\kappa}_\omega^\top(n) \boldsymbol{\alpha}_{\text{opt}}. \quad (24)$$

Substituindo (23) em (22) obtém-se:

$$\begin{aligned} \mathbf{v}(n+1) &= \mathbf{v}(n) + \eta d(n) \boldsymbol{\kappa}_\omega(n) - \eta [\boldsymbol{\kappa}_\omega^\top(n) \mathbf{v}(n)] \boldsymbol{\kappa}_\omega(n) \\ &\quad - \eta [\boldsymbol{\kappa}_\omega^\top(n) \boldsymbol{\alpha}_{\text{opt}}] \boldsymbol{\kappa}_\omega(n). \end{aligned} \quad (25)$$

##### A. Hipóteses estatísticas simplificadoras

Hipóteses simplificadoras são necessárias para que seja matematicamente possível um estudo do comportamento estocástico do vetor de erro nos coeficientes. Uma importante hipótese usada nessa e em outras análises de comportamento estocástico de algoritmos adaptativos é que  $\boldsymbol{\kappa}_\omega(n) \boldsymbol{\kappa}_\omega^\top(n)$  é estatisticamente independente de  $\mathbf{v}(n)$ . Essa hipótese, detalhadamente justificada em [13] e denominada *modified independence assumption* (MIA), tende a ser menos restritiva que a hipótese de independência clássica [14].

Outras hipóteses estatísticas simplificadoras são apresentadas em diferentes partes deste trabalho. Essas premissas são declaradas e discutidas à medida que forem aplicadas para facilitar a compreensão.

##### B. Comportamento médio do vetor de erro nos coeficientes

Calculando o valor esperado em ambos os lados de (25) e usando MIA obtém-se o comportamento médio do vetor de erro nos coeficientes para o algoritmo KLMS:

$$E\{\mathbf{v}(n+1)\} = (\mathbf{I} - \eta \mathbf{R}_{\kappa\kappa}) E\{\mathbf{v}(n)\} \quad (26)$$

em que  $\mathbf{R}_{\kappa\kappa}$  é dada por (18) (ou (20) para o caso da sequência  $\{\mathbf{u}(n)\}_{n=0}^\infty$  ter média nula).

## V. RESULTADOS

Um exemplo para verificar a eficiência do modelo analítico do comportamento estocástico do algoritmo KLMS quando submetido a uma entrada correlacionada é apresentado nesta seção. Este problema foi adaptado de [15]. Aqui o sinal de entrada é uma sequência de vetores, gerados a partir de um processo autorregressivo (AR) de primeira ordem

$$u(n) = a u(n-1) + \eta_u(n) \quad (27)$$

em que  $a = 0,5$ ,  $\eta_u(n)$  é um sinal Gaussiano branco tal que  $\sigma_u^2 = 1$  e a sequência de entrada  $\{\mathbf{u}(n)\}_{n=0}^{\infty} \in \mathbb{R}^2$ ,  $\mathbf{u}(n) = [u(n) \ u(n-1)]^T$ . O sistema não linear que será estimado é dado por:

$$\varphi(\mathbf{u}(n)) = \sum_{i=0}^3 \left(\frac{1}{2}\right)^{i+1} \exp\left\{\frac{\|\mathbf{u}(n-i) - \mathbf{b}_i\|_2^2}{[\mathbf{s}]_{i+1}^2}\right\} \quad (28)$$

em que  $\mathbf{b}_0 = [-0,1454 \ -0,3862]^T$ ,  $\mathbf{b}_1 = [1,3162 \ -0,7965]^T$ ,  $\mathbf{b}_2 = [0,1354 \ 0,4178]^T$  e  $\mathbf{b}_3 = [0,8199 \ -0,8544]^T$ ,  $\mathbf{s} = [0,863 \ 0,9873 \ 0,2756 \ 0,7662]^T$ . O sinal desejado não linear  $d(n)$  foi corrompido por um ruído Gaussiano branco  $z(n)$  com média nula e variância  $\sigma_z^2 = 10^{-6}$ ,

$$d(n) = \varphi(n) + z(n). \quad (29)$$

Após alguns testes foram selecionados os parâmetros para ilustrar a acurácia do novo modelo. O nível de coerência  $\varepsilon_0 = 0,1$  para (2) e parâmetro do kernel  $\xi \in \{0,3, 0,5\}$ . O passo de adaptação foi escolhido para prover uma aprendizagem lenta. O dicionário possui dimensão fixa  $M$  obtido a partir da média de  $M_i$ s dicionários gerados em 500 realizações de Monte Carlo (para satisfazer o nível de coerência  $\varepsilon_0 = 0,1$ ) Os dicionários são atualizados a cada interação utilizando linha de retardo<sup>2</sup>, ou seja, na  $k$ -ésima interação  $\mathcal{D}(k) = \{\kappa(\cdot, \mathbf{u}(k-1)), \dots, \kappa(\cdot, \mathbf{u}(k-M))\}$ . O vetor de coeficientes  $\alpha$  foi inicializado em  $[0 \ 0 \ \dots \ 0]^T$ .

A Fig. 2 apresenta o comportamento da norma Euclidiana do vetor de erro nos coeficientes para 100 realizações de uma simulação de Monte Carlo, para o modelo do algoritmo KLMS derivado em [8] e para o novo modelo considerando a tripla de valores:  $\eta = 10^{-4}$ ,  $\xi = 0,5$  e  $M = 5$ . Neste exemplo, o vetor de coeficientes ótimos é  $\alpha_{\text{opt}} = [0,2041 \ 0,0911 \ 0,1375 \ 0,1324 \ 0,1671]^T$ . O valor do erro mínimo (9) é  $J_{\min}(0,5; 5) = -16,5 \text{ dB}^3$ . A Fig. 3 apresenta o comportamento do vetor de coeficientes para  $\{\alpha_i\}_{i=1}^4$ . É possível notar pelas Fig. 2 e 3 uma maior acurácia do modelo dado por (18) e (26), em comparação com o modelo apresentado em [8] para o comportamento médio dos coeficientes.

Usando  $\eta = 2 \times 10^{-4}$ ,  $\xi = 0,3$  e  $M = 8$ , o vetor de pesos ótimos obtido é  $\alpha_{\text{opt}} = [0,2270 \ 0,1439 \ 0,1933 \ 0,1909 \ 0,1959 \ 0,1988 \ 0,2182]^T$  e o MMSE  $J_{\min}(0,3; 8) = -15,9 \text{ dB}$ . O comportamento da norma Euclidiana do vetor erro nos coeficientes é apresentado

<sup>2</sup>Essa técnica já foi empregada em outros trabalhos tais como em [16], [15], [8] produzindo bons resultados.

<sup>3</sup>Para o cálculo do MMSE a partir de (9),  $E\{d^2(n)\}$  foi obtida usando (29) e um procedimento análogo ao usado para calcular  $\mathbf{R}_{\kappa\kappa}$  e  $\mathbf{p}_{\kappa d}$ . Em casos em que a não linearidade é desconhecida podem ser utilizadas as estimativas de  $\mathbf{p}_{\kappa d}$  e  $E\{d^2(n)\}$ .

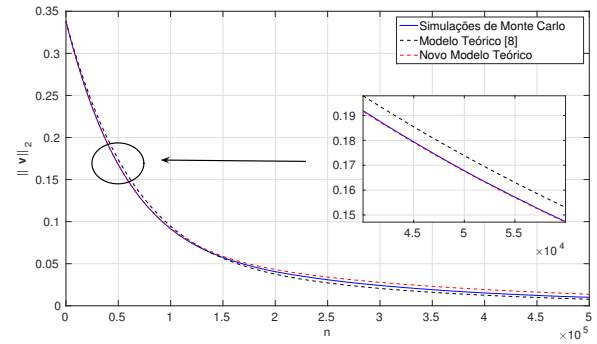


Fig. 2. Comportamento da norma do vetor de erro nos coeficientes. A curva contínua em azul representa as 100 realizações de Monte Carlo, a curva tracejada em preto representa o modelo KLMS dado por [8] e a curva tracejada em vermelho representa o novo modelo proposto.

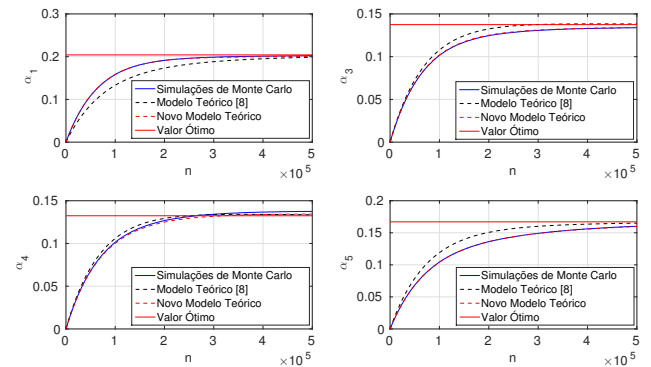


Fig. 3. Comportamento do vetor de coeficientes. A curva contínua em vermelho representa o comportamento das componentes do vetor  $\alpha_{\text{opt}}$ , a curva contínua em azul as 100 realizações de Monte Carlo, em preto tracejado o modelo KLMS dado por [8] e em tracejado vermelho o novo modelo proposto.

na Fig. 4 e o comportamento do vetor de coeficientes é apresentado pela Fig. 5. A simulação do comportamento estocástico do KLMS apresentado nas Fig. 4 e 5 foi obtida a partir de 100 realizações de uma simulação de Monte Carlo. Novamente, o modelo deduzido aqui apresentou uma acurácia maior que a apresentada em [8].

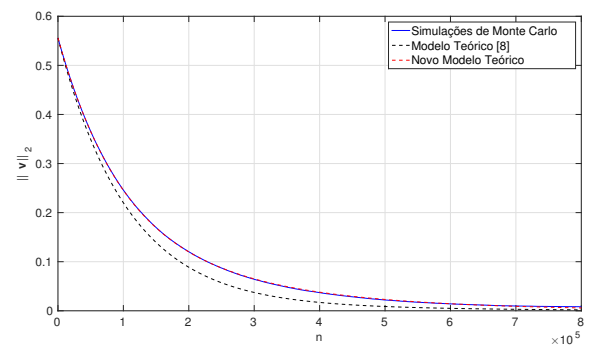


Fig. 4. Comportamento da norma do vetor de erro nos coeficientes. A curva contínua em azul representa as 100 realizações de Monte Carlo, a curva tracejada em preto representa o modelo KLMS dado por [8] e a curva tracejada em vermelho o novo modelo proposto.

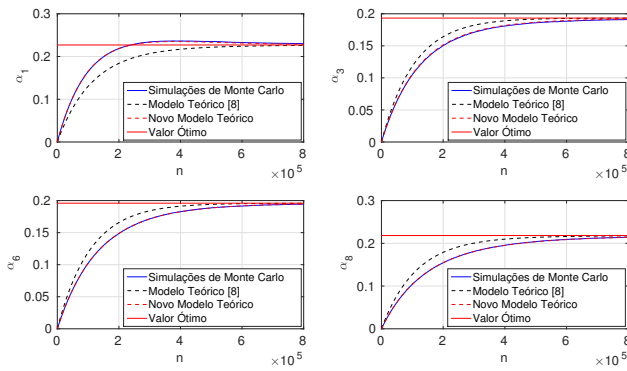


Fig. 5. Comportamento do vetor de coeficientes. A curva contínua em vermelho representa as componentes do vetor  $\alpha_{\text{opt}}$ , a curva contínua em azul as 100 realizações de Monte Carlo, em tracejado preto o modelo KLMS dado por [8] e em tracejado vermelho o novo modelo proposto.

## VI. CONCLUSÕES

A análise do transitório bem como do regime permanente para o comportamento médio do vetor de coeficientes adaptativos do KLMS com kernel Gaussiano para uma entrada correlacionada no tempo e dicionários variáveis foi apresentada neste trabalho. Uma expressão analítica para a matriz de correlação  $\mathbf{R}_{\kappa\kappa}$  quando o dicionário é atualizado utilizando linha de retardo permitiu uma melhor estimativa do comportamento do vetor peso nos coeficientes em relação ao derivado em [8]. As simulações de Monte Carlo ilustram a acurácia do modelo analítico proposto.

### APÊNDICE I

#### CÁLCULO DA CORRELAÇÃO CRUZADA - $\mathbf{p}_{\kappa d}$

Para uma não linearidade conhecida descrita por (28) e estatísticas de  $\mathbf{u}(\cdot)$  como descrita no Exemplo 1 é possível obter uma expressão analítica para a correlação cruzada entre  $\kappa_{\omega}(n)$  e  $d(n)$ , que é :

$$\mathbf{p}_{\kappa d} = E[\kappa_{\omega}(n) d(n)] = \begin{bmatrix} E[\kappa(\mathbf{u}(\omega_1), \mathbf{u}(n)) d(n)] \\ E[\kappa(\mathbf{u}(\omega_2), \mathbf{u}(n)) d(n)] \\ \vdots \\ E[\kappa(\mathbf{u}(\omega_M), \mathbf{u}(n)) d(n)] \end{bmatrix} \quad (30)$$

em que a  $\ell$ -ésima linha do vetor  $\mathbf{p}_{\kappa d}$ , com  $1 \leq \ell \leq M$  é dada por

$$[\mathbf{p}_{\kappa d}]_{\ell} = \sum_{i=0}^3 a^{i+1} E \left\{ \kappa(\mathbf{u}(\omega_{\ell}), \mathbf{u}(n)) \exp \left[ \frac{-\|\mathbf{u}(n-i) - \mathbf{b}_i\|_2^2}{[s'_i]^2} \right] \right\} \quad (31)$$

Assim, se  $\mathbf{u}(\omega_{\ell}) = \mathbf{u}(n-\ell)$  usando a forma quadrática com  $\mathbf{y}'_3 = [\mathbf{u}(n) \ \mathbf{u}(n-\ell) \ \mathbf{u}(n-i)]^T$  e matrizes

$$\mathbf{W}_i = \begin{pmatrix} s\mathbf{I} & -s\mathbf{I} & \mathbf{O} \\ -s\mathbf{I} & s\mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & s'_i\mathbf{I} \end{pmatrix}, \mathbf{r}_i = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ 2s'_i\mathbf{b}_i \end{pmatrix} \quad (32)$$

com  $s'_i = 1/[s_i]^2$  e  $\mathbf{0} = [0 \ 0]^T$ , tem-se por [12] a  $\ell$ -ésima

componente do vetor de correlação cruzada:

$$[\mathbf{p}_{\kappa d}]_{\ell} = \sum_{i=0}^3 a^{i+1} \exp(s'_i \|\mathbf{b}_i\|_2^2) \det(\mathbf{I}_3 - 2\mathbf{W}\mathbf{R}_{\mathbf{y}'_3\mathbf{y}'_3, i\ell})^{-1/2} \times \exp \left[ -\frac{1}{2} \mathbf{r}^T \mathbf{R}_{\mathbf{y}'_3\mathbf{y}'_3, i\ell} (\mathbf{I}_3 - 2\mathbf{W}\mathbf{R}_{\mathbf{y}'_3\mathbf{y}'_3, i\ell})^{-1} \mathbf{r} \right] \quad (33)$$

em que a matriz de correlação  $\mathbf{R}_{\mathbf{y}'_3\mathbf{y}'_3, i\ell}$  é calculada de maneira análoga à (19), que é:

$$\mathbf{R}_{\mathbf{y}'_3\mathbf{y}'_3, i\ell} = \begin{pmatrix} \mathbf{R}_{uu} & \mathbf{R}_{uu,i} & \mathbf{R}_{uu,m} \\ \mathbf{R}_{uu,i}^T & \mathbf{R}_{uu} & \mathbf{R}_{uu,h'} \\ \mathbf{R}_{uu,m}^T & \mathbf{R}_{uu,h'}^T & \mathbf{R}_{uu} \end{pmatrix} \quad (34)$$

com  $h' = |i - \ell|$ .

A partir de [17] é possível obter cada matriz  $\mathbf{R}_{uu,p}$ , com  $p = 1, \dots, M$ . Isso conclui os cálculos necessários para obtenção de  $\mathbf{p}_{\kappa d}$  usado em (8).

## REFERÊNCIAS

- [1] G. Kimeldorf and G. Wahba, "Some results on Tchebycheffian spline functions," *J. Math. Anal. Appl.*, vol. 33, pp. 82–95, 1971.
- [2] D. L. Duttweiler and T. Kailath, "An RKHS approach to detection and estimation theory: Some parameter estimation problems (Part V)," *IEEE Trans. Inf. Theory*, vol. 19, no. 1, pp. 29–37, 1973.
- [3] W. Liu, J. C. Principe, and S. Haykin, *Kernel Adaptive Filtering*. John Wiley & Sons, Inc., 2010.
- [4] W. Liu, P. P. Pokharel, and J. C. Principe, "The kernel least-mean-squares algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 543–554, February 2008.
- [5] Y. Engel, S. Mannor, and R. Meir, "Kernel recursive least squares," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2275–2285, 2004.
- [6] P. Honeine, C. Richard, and J. C. M. Bermudez, "On-line nonlinear sparse approximation of functions," in *Proc. IEEE ISIT'07*, Nice, France, June 2007, pp. 956–960.
- [7] C. Richard, J. C. M. Bermudez, and P. Honeine, "Online prediction of time series data with kernels," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1058–1067, March 2009.
- [8] W. D. Parreira, J. Bermudez, C. Richard, and J.-I. Tourneret, "Stochastic behavior analysis of the gaussian kernel least-mean-square algorithm," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2208–2222, March 2012.
- [9] J. Chen, W. Gao, C. Richard, and J. C. M. Bermudez, "Convergence analysis of kernel lms algorithm with pre-tuned dictionary," in *Proc. IEEE ICASSP'13*, Vancouver, Canada, May 2013.
- [10] W. Gao, J. Chen, C. Richard, J. C. M. Bermudez, and J. Huang, "Convergence analysis of the augmented complex klms algorithm with pre-tuned dictionary," in *Proc. IEEE ICASSP'15*, April 2015, pp. 2006–2010.
- [11] P. Honeine, "Analyzing sparse dictionaries for online learning with kernels," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6343–6353, December 2015.
- [12] J. Omura and T. Kailath, "Some useful probability distributions," Stanford Electronics Laboratories, Stanford University, Stanford, California, USA, Tech. Rep. 7050-6, 1965.
- [13] J. Minkoff, "Comment: On the unnecessary assumption of statistical independence between reference signal and filter weights in feedforward adaptive systems," *IEEE Trans. Signal Process.*, vol. 49, no. 5, p. 1109, May 2001.
- [14] A. H. Sayed, *Fundamentals of adaptive filtering*. Hoboken, NJ: John Wiley & Sons, 2003.
- [15] W. D. Parreira, J. C. M. Bermudez, C. Richard, and J.-Y. Tourneret, "Steady-state behavior and design of the gaussian klms algorithm," in *Proc. EUSIPCO'11*, Barcelona, Spain, September 2011.
- [16] —, "Stochastic behavior analysis of the gaussian kernel least mean square algorithm," in *Proc. IEEE ICASSP'11*, Prague, Czech Republic, May 2011.
- [17] V. B. Nicolau, "Cancelamento de realimentação acústica em aparelhos auditivos através do método de erro de predição com comprimento deficiente," Master's thesis, Federal University of Santa Catarina, 2010.