

Análise estatística do algoritmo Shalvi-Weinstein

Renato Candido, Magno T. M. Silva, Vítor H. Nascimento e Maria D. Miranda

Resumo—Algoritmos autodidatas com boas propriedades de convergência, *tracking* e robustez numérica são desejáveis para garantir o bom desempenho de um sistema de comunicação. Recentemente, foi proposta uma versão do algoritmo Shalvi-Weinstein (SWA) que evita divergência e tem custo computacional reduzido. Neste trabalho é apresentada uma análise do regime e do transitório desse algoritmo. Além de algumas hipóteses clássicas usadas na análise de algoritmos adaptativos como a teoria da independência, são considerados modelos recentes propostos para analisar o conhecido algoritmo do módulo constante (CMA). Resultados de simulação confirmam a validade dos modelos obtidos.

Palavras-Chave—Filtros adaptativos, equalização autodidata, algoritmo Shalvi-Weinstein, análise estatística, análise de *tracking*.

Abstract—Blind equalization algorithms with good convergence and tracking properties, and numerical robustness are desirable to ensure the good performance of a communications system. Recently was proposed a version of the Shalvi-Weinstein algorithm (SWA) which avoids divergence and has a reduced computational cost. In this paper, we present steady-state and transient analyses for this algorithm. Besides some classical assumptions used in the analysis of adaptive filters, such as the independence theory, we consider some recent models proposed to analyze the constant modulus algorithm (CMA). Good agreement between analytical and simulation results is observed.

Keywords—Adaptive filters, blind equalization, Shalvi-Weinstein algorithm, statistical analysis, tracking analysis.

I. INTRODUÇÃO

Usualmente os receptores dos sistemas de comunicação empregam equalizadores adaptativos para remover a interferência intersimbólica introduzida por canais dispersivos. Dentre esses equalizadores, destacam-se os que operam de forma autodidata, considerando informações das estatísticas de ordem superior do sinal transmitido. De modo a garantir a eficiência da transmissão, é desejável que esses algoritmos apresentem boas propriedades de convergência, *tracking* e robustez numérica [1]–[3]. Os algoritmos de equalização autodidata mais conhecidos na literatura são os do tipo módulo constante como o *constant modulus algorithm* (CMA) [4], [5] e o algoritmo Shalvi-Weinstein (SWA) [6]. Em relação ao CMA, o SWA tem como vantagem uma maior velocidade de convergência, mas apresenta problemas de divergência, além de um maior custo computacional.

Em [7], verificou-se que a divergência do SWA pode ser causada pela inconsistência na estimativa não-linear do sinal transmitido, ou pela perda de positividade da estimativa da matriz de autocorrelação. A fim de evitar a primeira causa de divergência, foi proposto um SWA com dois modos de operação, denominado *dual-mode-SWA* (DM-SWA). No primeiro modo de operação, esse algoritmo funciona como o SWA convencional e no segundo modo, ele rejeita estimativas não consistentes do sinal transmitido. Além disso, para evitar a divergência devido à perda de positividade da estimativa da matriz de autocorrelação e reduzir o custo computacional, o algoritmo foi convenientemente descrito para ser implementado com as já conhecidas técnicas de algoritmos dos mínimos quadrados recursivos (RLS - *recursive least squares*) rápidos que evitam divergência. Cabe observar que o DM-SWA ainda apresenta os mesmos problemas de convergência para mínimos locais, comuns em algoritmos do tipo módulo constante. A vantagem é que, quando adequadamente

Os autores estão com a Escola Politécnica, Universidade de São Paulo, São Paulo, SP, Brasil. E-mails: {renatocan, magno, vitor}@lps.usp.br e maria@lcs.poli.usp.br. Este trabalho foi financiado pela FAPESP (2008/00773-1 e 2008/04828-5) e CNPq (302633/2008-1 e 303361/2004-2).

implementado, ele não diverge mesmo considerando aritmética de precisão finita.

A literatura contém alguns resultados sobre a análise estatística de algoritmos autodidatas. Expressões analíticas para o erro quadrático médio em excesso (EMSE - *excess mean-square error*) de algoritmos do tipo módulo constante têm sido obtidas, principalmente na última década (e.g., [8]–[14]). Devido à natureza não-linear desses algoritmos, a maior parte desses trabalhos apresenta análises em regime para o CMA com passo de adaptação pequeno. Analogamente, uma análise de *tracking* para o SWA foi apresentada em [11] assumindo um fator de esquecimento próximo de um. Recentemente, foi proposto um modelo para o erro de estimação de algoritmos do tipo módulo constante em [12]. Nesse modelo, o erro de estimação é função do erro *a priori* e de um termo que mede a variabilidade do módulo do sinal transmitido. Ele permitiu obter uma análise de estabilidade do CMA [13] e também uma expressão analítica mais precisa para o EMSE [14].

Neste artigo é apresentada uma análise estatística para o DM-SWA. São feitas análises em regime para os dois modos de operação e também uma análise de transitório no modo em que a equação de adaptação do SWA e DM-SWA coincidem. Parte da análise em regime é baseada na análise do algoritmo RLS de [15]. Pelo fato dos resultados de [15] serem mais precisos do que os de outras análises em regime existentes para o RLS, obtém-se resultados mais precisos para o SWA que os de [11] e [12]. Como no caso do RLS, obtém-se um modelo que prevê aproximadamente situações em que a probabilidade de divergência do SWA é elevada. Para facilitar a análise, assume-se que os sinais são reais. Na Seção II, apresenta-se a formulação do problema e o DM-SWA. O modelo de [12] é revisitado e estendido para a estimativa do sinal transmitido na Seção III. Nas Seções IV e V, são apresentadas as análises de *tracking* e do transitório para o DM-SWA. Os resultados de simulações e as conclusões são mostrados nas seções VI e VII, respectivamente.

II. A FORMULAÇÃO DO PROBLEMA E O DM-SWA

Um esquema simplificado de um sistema de comunicação em banda base com um equalizador fracionário, sobreamostrado por um fator L é mostrado na Figura 1. Sob certas condições bem conhecidas, esse modelo assegura equalização perfeita na ausência de ruído [16]. O sinal transmitido $a(n)$ é assumido iid (independente e identicamente distribuído) e não-gaussiano. É usual considerar o equalizador como um sistema com resposta impulsiva de duração finita (FIR - *finite impulse response*) contendo M coeficientes, tendo como entrada o vetor regressor $\mathbf{u}(n)$ e como saída $y(n) = \mathbf{u}^T(n)\mathbf{w}(n-1)$, sendo que $\mathbf{w}(n-1)$ representa o vetor de coeficientes do equalizador e o superescrito T indica transposição. Deseja-se que tal equalizador minimize o efeito do canal e recupere o sinal $a(n)$ para algum atraso τ_d , obtendo-se na saída do decisor a estimativa $\hat{a}(n - \tau_d)$.

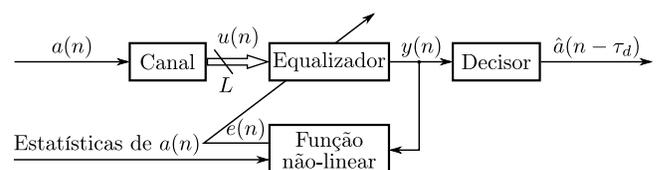


Fig. 1. Sistema de comunicação com um equalizador fracionário.

A equação de adaptação do SWA é dada por [6]

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \bar{e}(n)\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n) \quad (1)$$

sendo $\bar{e}(n) = e(n)/\bar{\gamma}$,

$$e(n) = [r - y^2(n)]y(n), \quad (2)$$

$\bar{\gamma} = 3\sigma_a^2 - r$, $r = E\{a^4(n)\}/\sigma_a^2$, $\sigma_a^2 = E\{a^2(n)\}$ e $E\{\cdot\}$ o operador esperança. A matriz $\hat{\mathbf{R}}(n)$ é uma estimativa da matriz de autocorrelação $\mathbf{R} = E\{\mathbf{u}(n)\mathbf{u}^T(n)\}$, ou seja,

$$\hat{\mathbf{R}}(n) = \lambda\hat{\mathbf{R}}(n-1) + \mathbf{u}(n)\mathbf{u}^T(n), \quad (3)$$

sendo $0 \ll \lambda < 1$ um fator de esquecimento, $\hat{\mathbf{R}}(0) = \delta\mathbf{I}$, \mathbf{I} a matriz identidade e δ uma constante positiva. A inversa $\hat{\mathbf{R}}^{-1}(n)$ é obtida aplicando-se o lema de inversão matricial a (3) [1].

É conhecido que o SWA quando implementado conforme apresentado em [6] apresenta problemas de divergência. Entretanto, esse problema foi recentemente superado na versão do DM-SWA proposta em [7]. A idéia desse algoritmo, que opera em dois modos distintos, surgiu devido a possibilidade de expressar (1) de forma semelhante à equação de adaptação do algoritmo RLS. Isso pode ser facilmente obtido se o “erro” $\bar{e}(n)$ for reescrito como

$$\bar{e}(n) = e(n)/\bar{\gamma} = d(n) - y(n), \quad (4)$$

sendo $d(n) = x(n)y(n)$ e $x(n) = [3\sigma_a^2 - y^2(n)]/\bar{\gamma}$. No RLS, $d(n)$ representa a resposta desejada. No SWA, reescrever $\bar{e}(n)$ como (4) é interessante porque a informação das estatísticas de ordem superior do sinal transmitido aparece na variável $d(n)$. Nesse caso, tanto $d(n)$ quanto $y(n)$ podem ser interpretados como estimativas do sinal transmitido. Tal fato sugeriu supor que $d(n)$ é consistente quando apresenta o mesmo sinal de $y(n)$ e que, nessa situação, $y(n)$ está numa *região de interesse*. Quando $d(n)$ apresenta um sinal diferente de $y(n)$, o que acontece quando $y^2(n) > 3\sigma_a^2$ ou $x(n) < 0$, a estimativa não é considerada consistente e, para evitar divergência, deve ser rejeitada. Dessa forma, fazendo $d(n) = 0$, a equação (1) se reduz a

$$\mathbf{w}(n) = \mathbf{w}(n-1) - y(n)\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n). \quad (5)$$

As equações (1) e (5) caracterizam os dois modos de operação do DM-SWA.

Também foi mostrado em [7] que se $y(n)$ deixa a região de interesse por alguma razão, como uma inicialização ruim ou uma relação sinal-ruído baixa, a equação (5) força $y(n)$ a retornar para a região de interesse depois de um intervalo de um tempo finito. Dessa forma, embora (5) não possibilite obter um resultado adequado para equalização, ela é importante para evitar a divergência devido à inconsistência na estimativa não-linear do sinal transmitido. Além disso, quando $y(n)$ está na *região de interesse*, mesmo $d(n)$ sendo limitado, o algoritmo pode divergir se não forem tomadas precauções na atualização da inversa da matriz de autocorrelação. Porém, técnicas eficientes de implementar a atualização de tal matriz são bem conhecidas, como, por exemplo, a solução com estrutura em treliça de [7].

III. UM MODELO ESTATÍSTICO PARA $d(n)$

Uma medida de desempenho do equalizador é dada pelo EMSE, definido como $\zeta(n) \triangleq E\{e_a^2(n)\}$, sendo $e_a(n)$ o erro *a priori* expresso como

$$e_a(n) = \mathbf{u}^T(n)\tilde{\mathbf{w}}(n-1), \quad (6)$$

em que $\tilde{\mathbf{w}}(n-1) = \mathbf{w}_o(n-1) - \mathbf{w}(n-1)$ é o vetor de erros dos coeficientes e \mathbf{w}_o é a solução ótima de *zero-forcing*. Num ambiente estacionário, o vetor \mathbf{w}_o é fixo, enquanto num ambiente não-estacionário, ele varia com o tempo. Neste caso, é usual modelar sua variação através do *random-walk model* [3, p. 359], dado por

$$\mathbf{w}_o(n) = \mathbf{w}_o(n-1) + \mathbf{q}(n), \quad (7)$$

sendo $\mathbf{q}(n)$ um vetor iid com matriz de autocorrelação positiva-definida $\mathbf{Q} = E\{\mathbf{q}(n)\mathbf{q}^T(n)\}$, independente das condições iniciais $\{\mathbf{w}_o(0), \mathbf{w}(0)\}$ e de $\{\mathbf{u}(l)\}$ para todo l .

Como no caso do CMA, uma análise estatística do SWA requer hipóteses simplificadoras. Algumas delas são revisitadas a seguir. Detalhes a respeito das mesmas podem ser encontrados em [9]–[14].

- A1. $\mathbf{u}(n)$ e $\tilde{\mathbf{w}}(n-1)$ são independentes. Essa hipótese faz parte da teoria da independência, amplamente usada na análise de filtros adaptativos [1], [8], [17].
- A2. $a(n-\tau_d)$ e $e_a(n)$ são independentes. Essa hipótese tem sido amplamente usada na análise em regime do CMA [9]–[12], [14].
- A3. $E\{a^k(n)\} = 0$, $k = 2m+1$, $m \in \mathbb{N}$, o que decorre da constelação ser simétrica. Além disso, $a(n)$ é sub-gaussiana, como é o caso da maioria das constelações usadas em comunicações digitais [6].
- A4. A relação sinal-ruído é alta o suficiente para que $a(n-\tau_d) \approx \mathbf{u}^T(n)\mathbf{w}_o(n-1)$. Entretanto, devido à variação do canal e ao ruído do gradiente, o vetor de coeficientes $\mathbf{w}(n-1)$ não é igual a $\mathbf{w}_o(n-1)$, mesmo em regime. Usando essa aproximação, $y(n)$ pode ser reescrito como $y(n) = \mathbf{u}^T(n)\mathbf{w}(n-1) = \mathbf{u}^T(n)[\mathbf{w}_o(n-1) - \tilde{\mathbf{w}}(n-1)]$, isto é, $y(n) \approx a(n-\tau_d) - e_a(n)$.

Usando A4 e assumindo que $e_a(n)$ é suficientemente pequeno para que os termos que dependem de combinações de $e_a^k(n)$, $k \geq 2$ possam ser desprezados em relação a $e_a(n)$, (2) pode ser reescrita como [12], [13]

$$e(n) \approx \gamma(n)e_a(n) + \beta(n), \quad (8)$$

sendo

$$\gamma(n) = 3a^2(n-\tau_d) - r, \quad (9)$$

$$\beta(n) = r a(n-\tau_d) - a^3(n-\tau_d). \quad (10)$$

Cabe notar que $\beta(n)$ é uma variável aleatória iid. Usando A3, observa-se que $E\{\beta(n)\} = 0$ e sua variância vale

$$\sigma_\beta^2 \triangleq E\{\beta^2(n)\} = E\{a^6(n) - r^2 a^2(n)\}. \quad (11)$$

Analogamente, $\gamma(n)$ também é iid e seus momentos de primeira e segunda ordens são dados respectivamente por $E\{\gamma(n)\} = \bar{\gamma}$ e

$$\xi \triangleq E\{\gamma^2(n)\} = 3r\sigma_a^2 + r^2. \quad (12)$$

Cabe observar que $\bar{\gamma}$ representa a curtose do sinal transmitido com o sinal trocado e para constelações sub-gaussianas é sempre positivo (curtose negativa). Uma consequência imediata da Hipótese A2 é que tanto $\gamma(n)$ quanto $\beta(n)$ são independentes do erro *a priori* $e_a(n)$.

Inspirando-se na análise em regime do CMA de [9] e [10], o modelo (8) foi proposto em [12]. Posteriormente, ele foi usado na análise de estabilidade do CMA em [13] e estendido para sinais complexos em [14]. Usando esse modelo, o “erro” $\bar{e}(n)$ de (4) pode ser reescrito aproximadamente como

$$d(n) - y(n) \approx \frac{\gamma(n)}{\bar{\gamma}} e_a(n) + \frac{\beta(n)}{\bar{\gamma}}. \quad (13)$$

Substituindo (6) em (13), o sinal $d(n)$ pode ser modelado por

$$d(n) \approx \left[1 - \frac{\gamma(n)}{\bar{\gamma}}\right] \mathbf{u}^T(n)\mathbf{w}(n-1) + \frac{\gamma(n)}{\bar{\gamma}} \mathbf{u}^T(n)\mathbf{w}_o(n-1) + \frac{\beta(n)}{\bar{\gamma}}. \quad (14)$$

Com $\gamma(n) \equiv 1$ e interpretando $\beta(n)$ como ruído de medida, (14) se reduz ao modelo de regressão linear usado em filtragem adaptativa supervisionada [3, p. 284]. Cabe observar que $\beta(n)$ é identicamente nulo para constelações de módulo constante. Dessa forma, a variação no módulo de $a(n)$, medida por $\beta(n)$, faz o papel de ruído de medida para algoritmos do tipo módulo constante.

IV. ANÁLISE DE *tracking*

A seguir, a análise de *tracking* do DM-SWA é feita separando os casos em que $y(n)$ está dentro ou fora da região de interesse, o que corresponde respectivamente às equações de atualização (1) ou (5). Embora muitas vezes o DM-SWA fique chaveando entre esses dois modos de operação, a fim de facilitar a análise, assume-se cada modo de operação de forma separada.

A. Análise dentro da região de interesse

Usando A1, o EMSE em regime do SWA pode ser calculado como

$$\zeta(\infty) \approx \lim_{n \rightarrow \infty} E\{\tilde{\mathbf{w}}^T(n-1)\mathbf{R}\tilde{\mathbf{w}}(n-1)\} = \lim_{n \rightarrow \infty} \text{Tr}(\mathbf{R}\mathbf{S}(n-1)),$$

sendo $\mathbf{S}(n) \triangleq E\{\tilde{\mathbf{w}}(n)\tilde{\mathbf{w}}^T(n)\}$ a matriz de covariância do vetor de erro dos coeficientes e $\text{Tr}(\cdot)$ o traço de uma matriz. Na análise em regime apresentada a seguir, $\lim_{n \rightarrow \infty}$ será suprimido para simplificar a notação. Na região de interesse, $y^2(n) < 3\sigma_a^2$ e as equações de atualização do DM-SWA e do SWA coincidem e são dadas por (1).

A Equação (1) pode ser reescrita em função do vetor de erro de coeficientes $\tilde{\mathbf{w}}(n)$. Subtraindo ambos os lados de (1) da solução ótima $\mathbf{w}_o(n)$ e usando (7), obtém-se

$$\tilde{\mathbf{w}}(n) = \tilde{\mathbf{w}}(n-1) - [d(n) - y(n)]\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n) + \mathbf{q}(n). \quad (15)$$

Substituindo $d(n)$ pelo modelo (14), após algumas manipulações algébricas chega-se a

$$\begin{aligned} \tilde{\mathbf{w}}(n) = & \left[\mathbf{I} - \frac{\gamma(n)}{\bar{\gamma}}\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)\mathbf{u}^T(n) \right] \tilde{\mathbf{w}}(n-1) \\ & - \frac{\beta(n)}{\bar{\gamma}}\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n) + \mathbf{q}(n). \end{aligned} \quad (16)$$

A matriz identidade que aparece em (16) pode ser reescrita usando (3). Assim, multiplicando (3) à esquerda por $\hat{\mathbf{R}}^{-1}(n)$ obtém-se

$$\mathbf{I} = \lambda\hat{\mathbf{R}}^{-1}(n)\hat{\mathbf{R}}(n-1) + \hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)\mathbf{u}^T(n). \quad (17)$$

Substituindo (17) em (16), chega-se a

$$\begin{aligned} \tilde{\mathbf{w}}(n) = & \left[\lambda\hat{\mathbf{R}}^{-1}(n)\hat{\mathbf{R}}(n-1) + \left(1 - \frac{\gamma(n)}{\bar{\gamma}}\right)\hat{\mathbf{R}}^{-1}(n) \right] \\ & \times \mathbf{u}(n)\mathbf{u}^T(n)\tilde{\mathbf{w}}(n-1) - \frac{\beta(n)}{\bar{\gamma}}\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n) + \mathbf{q}(n). \end{aligned} \quad (18)$$

As Equações (16) e (18) fornecem o mesmo $\tilde{\mathbf{w}}(n)$. Entretanto, termos que aparecem em (18) permitem utilizar aproximações na análise em regime que levam a resultados mais precisos, como os da análise do algoritmo RLS de [15].

Usando (18) e o fato de que $\mathbf{q}(n)$ é independente das condições iniciais e de $\mathbf{u}(n)$, obtém-se

$$\begin{aligned} E\{\tilde{\mathbf{w}}^T(n)\mathbf{R}\tilde{\mathbf{w}}(n)\} = & \\ & + \lambda^2 \overbrace{E\{\tilde{\mathbf{w}}^T(n-1)\hat{\mathbf{R}}(n-1)\hat{\mathbf{R}}^{-1}(n)\mathbf{R}\hat{\mathbf{R}}^{-1}(n)\hat{\mathbf{R}}(n-1)\tilde{\mathbf{w}}(n-1)\}}^{\mathcal{A}} \\ & + E\left\{ \overbrace{\left(1 - \frac{\gamma(n)}{\bar{\gamma}}\right)^2 \tilde{\mathbf{w}}^T(n-1)\mathbf{u}(n)\mathbf{u}^T(n)\hat{\mathbf{R}}^{-1}(n)\mathbf{R}\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)\mathbf{u}^T(n)\tilde{\mathbf{w}}(n-1)}^{\mathcal{B}} \right\} \\ & + E\left\{ \overbrace{\frac{\beta^2(n)}{\bar{\gamma}^2} \mathbf{u}^T(n)\hat{\mathbf{R}}^{-1}(n)\mathbf{R}\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)}^{\mathcal{C}} \right\} + E\{\mathbf{q}^T(n)\mathbf{R}\mathbf{q}(n)\} \\ & + 2\lambda \overbrace{E\left\{ \left(1 - \frac{\gamma(n)}{\bar{\gamma}}\right) \tilde{\mathbf{w}}^T(n-1)\hat{\mathbf{R}}(n-1)\hat{\mathbf{R}}^{-1}(n)\mathbf{R}\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)\mathbf{u}^T(n)\tilde{\mathbf{w}}(n-1) \right\}}^{\mathcal{E}} \\ & - \frac{2}{\bar{\gamma}} \lambda \overbrace{E\{\beta(n)\mathbf{u}^T(n)\hat{\mathbf{R}}^{-1}(n)\mathbf{R}\hat{\mathbf{R}}^{-1}(n)\hat{\mathbf{R}}(n-1)\tilde{\mathbf{w}}(n-1)\}}^{\mathcal{F}} \\ & - \frac{2}{\bar{\gamma}} \overbrace{E\left\{ \left(1 - \frac{\gamma(n)}{\bar{\gamma}}\right) \beta(n)\mathbf{u}^T(n)\hat{\mathbf{R}}^{-1}(n)\mathbf{R}\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)\mathbf{u}^T(n)\tilde{\mathbf{w}}(n-1) \right\}}^{\mathcal{G}}. \end{aligned} \quad (19)$$

Como em regime vale

$$E\{\tilde{\mathbf{w}}^T(n)\mathbf{R}\tilde{\mathbf{w}}(n)\} = E\{\tilde{\mathbf{w}}^T(n-1)\mathbf{R}\tilde{\mathbf{w}}(n-1)\}, \quad (20)$$

o termo do lado esquerdo de (19) também pode ser aproximado por $\zeta(\infty)$.

Na seqüência, são obtidas aproximações para os termos \mathcal{A} a \mathcal{G} em regime ($n \rightarrow \infty$). Para isso, é necessário assumir a seguinte hipótese

R1. $\hat{\mathbf{R}}^{-1}(n)\hat{\mathbf{R}}(n-1) \approx \mathbf{I}$. Essa aproximação é razoável em regime, principalmente quando $\lambda \lesssim 1^1$.

\mathcal{A} - Usando R1, o termo \mathcal{A} pode ser aproximado por

$$\mathcal{A} \approx E\{\tilde{\mathbf{w}}^T(n-1)\mathbf{R}\tilde{\mathbf{w}}(n-1)\} = \zeta(\infty). \quad (21)$$

\mathcal{B} - O termo \mathcal{B} pode ser reescrito como

$$\mathcal{B} = E\left\{ \left(1 - \gamma(n)/\bar{\gamma}\right)^2 e_a^2(n) \|\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)\|_{\mathbf{R}}^2 \right\} \quad (22)$$

sendo $\|\mathbf{x}\|_{\mathbf{A}}^2 \triangleq \mathbf{x}^T \mathbf{A} \mathbf{x}$. Para continuar, é necessário assumir que

R2. $e_a(n)$ é independente de $\|\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)\|_{\mathbf{R}}^2$. Uma hipótese semelhante foi usada na análise em regime do RLS [3, Cap. 7] e do SWA [11] com a conservação de energia. Ela é razoável em regime, uma vez que a norma do vetor $\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)$ varia lentamente face a $e_a(n)$.

Usando uma hipótese semelhante a R2, a variável $\gamma(n)$ também pode ser assumida independente de $\|\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)\|_{\mathbf{R}}^2$. Assim, usando A2 e R2, obtém-se

$$\mathcal{B} \approx \alpha E\left\{ \|\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)\|_{\mathbf{R}}^2 \right\} \zeta(\infty), \quad (23)$$

sendo $\alpha \triangleq E\left\{ \left(1 - \gamma(n)/\bar{\gamma}\right)^2 \right\} = \xi/\bar{\gamma}^2 - 1$.

Para obter uma aproximação para $E\left\{ \|\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)\|_{\mathbf{R}}^2 \right\}$, é necessário assumir que

R3. $\| (1 - \lambda)\mathbf{R}^{-1}\tilde{\mathbf{R}}(n) \| < 1$, sendo que $\| \cdot \|$ denota a norma espectral e $\tilde{\mathbf{R}}(n) \triangleq (1 - \lambda)\hat{\mathbf{R}}(n) - \mathbf{R}$ uma matriz de perturbação hermitiana de média nula [15].

R4. A matriz $\mathbf{F}(n) \triangleq (1 - \lambda)\mathbf{R}^{-1}\tilde{\mathbf{R}}^{-1}(n)$ é independente de $\mathbf{u}(n)$ e tem média nula, i.e., $E\{\mathbf{F}(n)\} = \mathbf{0}$ [15]. Essa matriz representa uma flutuação do produto $\hat{\mathbf{R}}^{-1}(n)\mathbf{R}$ em torno da matriz identidade já que é possível mostrar (ver [15, Eq. (A3.7)]) que

$$\hat{\mathbf{R}}^{-1}(n)\mathbf{R} \approx (1 - \lambda) [\mathbf{I} - \mathbf{F}(n)]. \quad (24)$$

Usando R3 e R4 e seguindo os cálculos de [15], [2, p. 214], obtém-se

$$E\left\{ \|\hat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)\|_{\mathbf{R}}^2 \right\} \approx (1 - \lambda)^2 \rho M \quad (25)$$

sendo

$$\rho \triangleq \left(1 + \frac{1 - \lambda}{1 + \lambda} \kappa_u \right) \quad (26)$$

e $\kappa_u \triangleq \sigma_{u^2}^2 / [\sigma_u^2]^2$, $\sigma_u^2 = E\{u^2(n)\}$ e $\sigma_{u^2}^2 = E\{u^4(n)\} - \sigma_u^4$. Se $u(n)$ for gaussiano, $\kappa_u = 2$ e para $\lambda \rightarrow 1$ observa-se que $\rho \rightarrow 1$.

Substituindo (25) em (23), chega-se a

$$\mathcal{B} \approx (1 - \lambda)^2 \rho M \alpha \zeta(\infty). \quad (27)$$

\mathcal{C} - Para canais com resposta impulsiva longa, foi mostrado em [12, Ap. I] que $E\{\beta^2(n)u^2(n)\} \approx \sigma_\beta^2 E\{u^2(n)\}$. Embora essa aproximação tenha sido justificada para canais longos, resultados de simulação sugerem que ela é razoável mesmo para canais com resposta impulsiva curta [12], [13]. Usando essa aproximação e (25), o termo \mathcal{C} pode ser aproximado por

$$\mathcal{C} \approx \frac{\sigma_\beta^2}{\bar{\gamma}^2} (1 - \lambda)^2 \rho M. \quad (28)$$

¹O símbolo \lesssim representa aproximadamente igual mas menor que.

\mathcal{D} - O termo \mathcal{D} é igual a $\text{Tr}(\mathbf{QR})$.

\mathcal{E} - Usando (3) e R1, vale $\mathbf{u}(n)\mathbf{u}^T(n) \approx (1-\lambda)\widehat{\mathbf{R}}(n)$. Com essa aproximação e A2, obtém-se $\mathcal{E} \approx 0$.

Para obter aproximações para os termos \mathcal{F} e \mathcal{G} é necessário assumir ainda que

R5. *Em regime, $E\{\widehat{\mathbf{R}}^{-1}(n)\} \approx (1-\lambda)\mathbf{R}^{-1}$. Para $\lambda \lesssim 1$ e regressores gaussianos, essa hipótese é razoável como mostrado em [3, Sec. 6.9.2].*

\mathcal{F} - Usando A2, R1 e R5, $\mathcal{F} \approx 0$.

\mathcal{G} - Usando A2, A3 e R5, $\mathcal{G} \approx 0$.

Substituindo essas aproximações em (19), após algumas manipulações algébricas chega-se a

$$\zeta(\infty) \approx \frac{\sigma_\beta^2(1-\lambda)M\rho(\bar{\gamma})^{-2} + \text{Tr}(\mathbf{QR})/(1-\lambda)}{1 + \lambda - (1-\lambda)\rho\alpha M}. \quad (29)$$

O termo $(1-\lambda)\rho\alpha M$ que aparece no denominador de (29) permite prever situações em que a probabilidade de divergência do SWA é elevada, como será mostrado nas simulações da Seção VI.

Para $\lambda \lesssim 1$, (29) se reduz a

$$\zeta(\infty) \approx \frac{\sigma_\beta^2(1-\lambda)M(\bar{\gamma})^{-2} + \text{Tr}(\mathbf{QR})/(1-\lambda)}{2}. \quad (30)$$

que foi obtida em [11] usando o método de análise baseado na conservação de energia e em [12] usando o método tradicional baseado no cálculo de uma recursão da matriz $\mathbf{S}(n)$.

B. Análise fora da região de interesse

Assumindo agora que o DM-SWA opera apenas fora da região de interesse, pode-se obter uma outra aproximação para $\zeta(\infty)$. Como mostrado em [7], o DM-SWA volta para a região de interesse após um intervalo de tempo finito. Por isso, a hipótese de que os coeficientes do equalizador são adaptados apenas com (5) não é realista, já que o DM-SWA fica chaveando entre os dois modos de operação quando a probabilidade de divergência do SWA é elevada. A análise a seguir pode ser interpretada como a de pior caso e não modela o funcionamento real do DM-SWA. Pretende-se apenas obter uma expressão analítica do EMSE de pior caso.

Reescrevendo (5) em função de $\tilde{\mathbf{w}}(n)$, obtém-se

$$\tilde{\mathbf{w}}(n) = \left[\mathbf{I} - \widehat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)\mathbf{u}^T(n) \right] \tilde{\mathbf{w}}(n-1) + \widehat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{w}_o(n-1) + \mathbf{q}(n). \quad (31)$$

Para obter uma recorrência para a matriz $\mathbf{S}(n) = E\{\tilde{\mathbf{w}}(n)\tilde{\mathbf{w}}^T(n)\}$, multiplica-se (31) por sua transposta e toma-se a esperança. Após algumas manipulações, chega-se a

$$\begin{aligned} \mathbf{S}(n) &= \mathbf{S}(n-1) - 2(1-\lambda)\mathbf{S}(n-1) \\ &+ (1-\lambda)^2\mathbf{R}^{-1}E\{\mathbf{u}(n)\mathbf{u}^T(n)\mathbf{S}(n-1)\mathbf{u}(n)\mathbf{u}^T(n)\}\mathbf{R}^{-1} \\ &+ (1-\lambda)^2\mathbf{R}^{-1}\sigma_a^2 + \mathbf{Q}. \end{aligned} \quad (32)$$

Para obter (32), utilizaram-se as Hipóteses A1-A4 e o fato de que $\mathbf{q}(n)$ é independente das condições iniciais e de $\mathbf{u}(n)$. Além disso, a matriz $\widehat{\mathbf{R}}^{-1}(n)$ foi simplesmente substituída por sua média usando R5. Essa aproximação pode não ser válida em muitos casos. No entanto, nessa análise de pior caso com $\lambda \lesssim 1$, pode ser considerada razoável.

Os três primeiros termos do lado direito de (32) são lineares em $\mathbf{S}(n-1)$. Dessa forma, para λ suficientemente próximo de 1, o terceiro termo do lado direito de (32) pode ser desprezado em relação aos dois primeiros. Assumindo regressor gaussiano, esse termo poderia ser aproximado usando, por exemplo, os resultados de

[18], [2, p. 83]. No entanto, optou-se por simplesmente desprezá-lo. Multiplicando ambos os lados de (32) por \mathbf{R} , calculando o traço e o limite para $n \rightarrow \infty$, chega-se a

$$\zeta(\infty) \approx \frac{\sigma_a^2(1-\lambda)M + \text{Tr}(\mathbf{QR})/(1-\lambda)}{2}. \quad (33)$$

V. ANÁLISE DE TRANSITÓRIO

A seguir, obtém-se uma aproximação para a matriz $\mathbf{S}(n)$ assumindo que o DM-SWA opere apenas dentro da região de interesse no caso estacionário.

Por conveniência, reescreve-se (16) assumindo-se $\mathbf{q}(n) = \mathbf{0}$, i.e.,

$$\begin{aligned} \tilde{\mathbf{w}}(n) &= \overbrace{\tilde{\mathbf{w}}(n-1)}^{\mathbf{x}_1(n)} - \overbrace{\frac{\gamma(n)}{\bar{\gamma}}\widehat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)\mathbf{u}^T(n)\tilde{\mathbf{w}}(n-1)}^{\mathbf{x}_2(n)} \\ &\quad - \underbrace{\frac{\beta(n)}{\bar{\gamma}}\widehat{\mathbf{R}}^{-1}(n)\mathbf{u}(n)}_{\mathbf{x}_3(n)}. \end{aligned} \quad (34)$$

Considerando válidas as Hipóteses A1-A4 e assumindo que o canal tenha resposta impulsiva longa, é possível mostrar que $E\{\tilde{\mathbf{w}}(n)\} \approx \mathbf{0}$ [12], [13]. Sob essas suposições, os termos $E\{\mathbf{x}_k(n)\mathbf{x}_k^T(n)\}$ e $E\{\mathbf{x}_3(n)\mathbf{x}_3^T(n)\}$, $k = 1, 2$ são aproximadamente nulos em todo instante de tempo. Assim, a matriz $\mathbf{S}(n)$ pode ser aproximada como

$$\begin{aligned} \mathbf{S}(n) &\approx E\{\overbrace{\mathbf{x}_1(n)\mathbf{x}_1^T(n)}^{\mathcal{H}}\} + E\{\overbrace{\mathbf{x}_2(n)\mathbf{x}_2^T(n)}^{\mathcal{I}}\} \\ &\quad + E\{\overbrace{\mathbf{x}_3(n)\mathbf{x}_3^T(n)}^{\mathcal{J}}\} + E\{\overbrace{\mathbf{x}_1(n)\mathbf{x}_2^T(n)}^{\mathcal{K}}\} \\ &\quad + E\{\overbrace{\mathbf{x}_2(n)\mathbf{x}_1^T(n)}^{\mathcal{L}}\}. \end{aligned} \quad (35)$$

A seguir são calculadas aproximações para os termos \mathcal{H} a \mathcal{L} , assumindo a seguinte hipótese

T1. *Apesar dos valores esperados de $\widehat{\mathbf{R}}(n)$ e $\widehat{\mathbf{R}}^{-1}(n)$ não manterem a mesma relação inversa, é possível aproximar $E\{\widehat{\mathbf{R}}^{-1}(n)\} \approx [E\{\widehat{\mathbf{R}}(n)\}]^{-1} \triangleq \overline{\mathbf{P}}(n)$. Resultados de simulação sugerem que, dependendo da inicialização da matriz de autocorrelação inversa, essa aproximação pode ser razoável, principalmente para $\lambda \lesssim 1$. Essa hipótese é usual na análise do RLS (e.g., [19], [20] e referências contidas).*

Lembrando que $\widehat{\mathbf{R}}(0) = \delta\mathbf{I}$, (3) pode ser reescrita como

$$\widehat{\mathbf{R}}(n) = \lambda^n\delta\mathbf{I} + \sum_{l=1}^n \lambda^{n-l}\mathbf{u}(l)\mathbf{u}^T(l). \quad (36)$$

Aplicando a esperança em ambos os lados de (36) e usando T1, obtém-se

$$\overline{\mathbf{P}}(n) \approx \left[\lambda^n\delta\mathbf{I} + \mathbf{R}\frac{1-\lambda^n}{1-\lambda} \right]^{-1}. \quad (37)$$

Como feito na Seção IV-B, na análise seguinte, a matriz $\widehat{\mathbf{R}}^{-1}(n)$ é simplesmente substituída por sua média aproximada $\overline{\mathbf{P}}(n)$. Essa aproximação pode não ser válida durante as primeiras iterações ($n < M$). No entanto, assumindo que T1 seja válida, essa substituição é uma aproximação razoável se $\widehat{\mathbf{R}}^{-1}(n)$ variar lentamente em relação a $\mathbf{u}(n)\mathbf{u}^T(n)$.

\mathcal{H} - O termo \mathcal{H} vale

$$\mathcal{H} = E\{\tilde{\mathbf{w}}(n-1)\tilde{\mathbf{w}}^T(n-1)\} = \mathbf{S}(n-1). \quad (38)$$

\mathcal{I} - O termo \mathcal{I} é dado por

$$\mathcal{I} = \frac{1}{\bar{\gamma}^2} \mathbb{E} \left\{ \gamma^2(n) \hat{\mathbf{R}}^{-1}(n) \mathbf{u}(n) \mathbf{u}^T(n) \tilde{\mathbf{w}}(n-1) \right. \\ \left. \times \tilde{\mathbf{w}}^T(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \hat{\mathbf{R}}^{-1}(n) \right\}. \quad (39)$$

Para continuar, assume-se que

T2. O canal tem resposta impulsiva longa o suficiente para que o regressor $\mathbf{u}(n)$ seja considerado gaussiano, tal que [18], [2, p. 83]

$$\mathbb{E} \left\{ \mathbf{u}(n) \mathbf{u}^T(n) \tilde{\mathbf{w}}(n-1) \tilde{\mathbf{w}}^T(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \right\} \\ \approx 2\mathbf{R}\mathbf{S}(n-1)\mathbf{R} + \mathbf{R}\text{Tr}(\mathbf{R}\mathbf{S}(n-1)). \quad (40)$$

Uma consequência da hipótese de canal longo é que $\mathbb{E} \{ \gamma^2(n) u^4(n) \} \approx \mathbb{E} \{ \gamma^2(n) \} \mathbb{E} \{ u^4(n) \}$ como foi mostrado em [12], [13]. Assim, usando A1-A4, T1 e T2, o termo \mathcal{I} pode ser aproximado por

$$\mathcal{I} \approx \frac{\xi}{\bar{\gamma}^2} \bar{\mathbf{P}}(n) [2\mathbf{R}\mathbf{S}(n-1) + \mathbf{R}\text{Tr}(\mathbf{R}\mathbf{S}(n-1))] \bar{\mathbf{P}}(n). \quad (41)$$

\mathcal{J} - Usando novamente R1 e R2, o termo \mathcal{J} pode ser aproximado por

$$\mathcal{J} = \frac{1}{\bar{\gamma}^2} \mathbb{E} \left\{ \beta^2(n) \hat{\mathbf{R}}^{-1}(n) \mathbf{u}(n) \mathbf{u}^T(n) \hat{\mathbf{R}}^{-1}(n) \right\} \\ \approx \frac{\sigma_\beta^2}{\bar{\gamma}^2} \bar{\mathbf{P}}(n) \mathbf{R} \bar{\mathbf{P}}(n). \quad (42)$$

\mathcal{K} - Usando os mesmos argumentos, o termo \mathcal{K} pode ser aproximado por

$$\mathcal{K} = \frac{1}{\bar{\gamma}} \mathbb{E} \left\{ -\gamma(n) \tilde{\mathbf{w}}(n-1) \tilde{\mathbf{w}}^T(n-1) \mathbf{u}(n) \mathbf{u}^T(n) \hat{\mathbf{R}}^{-1}(n) \right\} \\ \approx -\mathbf{S}(n-1) \mathbf{R} \bar{\mathbf{P}}(n). \quad (43)$$

\mathcal{L} - De forma análoga ao termo \mathcal{K} , o termo \mathcal{L} pode ser aproximado por

$$\mathcal{L} = \frac{1}{\bar{\gamma}} \mathbb{E} \left\{ -\gamma(n) \hat{\mathbf{R}}^{-1}(n) \mathbf{u}(n) \mathbf{u}^T(n) \tilde{\mathbf{w}}(n-1) \tilde{\mathbf{w}}^T(n-1) \right\} \\ \approx -\bar{\mathbf{P}}(n) \mathbf{R} \mathbf{S}(n-1). \quad (44)$$

Assim, usando esses resultados, chega-se a

$$\mathbf{S}(n) \approx \mathbf{S}(n-1) - \bar{\mathbf{P}}(n) \mathbf{R} \mathbf{S}(n-1) - \mathbf{S}(n-1) \mathbf{R} \bar{\mathbf{P}}(n) \\ + \frac{1}{\bar{\gamma}^2} \bar{\mathbf{P}}(n) \left\{ \xi [2\mathbf{R}\mathbf{S}(n-1) + \mathbf{R}\text{Tr}(\mathbf{R}\mathbf{S}(n-1))] \right. \\ \left. + \sigma_\beta^2 \mathbf{R} \right\} \bar{\mathbf{P}}(n). \quad (45)$$

O EMSE em cada instante de tempo pode ser estimado aproximadamente por $\zeta(n) \approx \text{Tr}(\mathbf{R}\mathbf{S}(n-1))$.

VI. RESULTADOS DE SIMULAÇÃO

Para verificar a validade das análises, foi considerado o canal de [10], i.e., $[0,1 \ 0,3 \ 1 \ -0,1 \ 0,5 \ 0,2]$ e um equalizador fracionário sobreamostrado com fator $L = 2$, contendo $M = 4$ coeficientes, inicializados com um único elemento diferente de zero e igual a um na segunda posição.

Na Figura 2-(a), são mostrados os resultados do EMSE em regime obtidos através de 50 realizações do DM-SWA em função do fator de esquecimento λ . Considerou-se um ambiente não-estacionário tal que $\mathbf{Q} = 10^{-4} \mathbf{R}^{-1}$ e a transmissão de um sinal com modulação 4-PAM (*pulse amplitude modulation*) com estatísticas $\mathbb{E}\{a^6\} = 365$, $\mathbb{E}\{a^2\} = 5$ e $r = 8,2$. Na Figura 2-(b), é mostrada a probabilidade de divergência do SWA (P_d) em função de λ . Para estimar essa probabilidade, foram feitas 100 realizações do SWA e considerou-se

a divergência do algoritmo sempre que observada uma saída do tipo NaN no Matlab. Para valores de $1 - \lambda$ menores que 5×10^{-2} , o algoritmo SWA tem probabilidade de divergência praticamente nula, o que faz com que o DM-SWA opere dentro da região de interesse. Nesse intervalo, o EMSE em regime do DM-SWA tem uma boa concordância com o valor teórico de (29). Vale notar que o resultado obtido com a conservação de energia (Eq. (30)) coincide com (29) para valores de λ próximos de 1, mas apresenta um erro mais elevado para valores de λ menores. A probabilidade de divergência do SWA aumenta para valores de $1 - \lambda$ maiores que 5×10^{-2} e o DM-SWA passa a chavar entre seus dois modos de operação. Nesse caso, não há concordância entre os resultados experimentais e resultados teóricos de (29) e (30). No entanto, uma característica que pode ser observada no resultado de (29), é que este consegue prever um aumento no EMSE à medida que o SWA tem maior probabilidade de divergência, apresentando uma descontinuidade para um valor de λ próximo do valor “crítico”, referente à região de transição da curva de P_b . Para valores de λ menores que o valor “crítico”, o resultado previsto por (29) deve ser desconsiderado já que a probabilidade de divergência do SWA é praticamente igual a 1. Em relação ao resultado da análise fora da região de interesse, como (33) foi obtida através de uma análise de pior caso, pode-se notar que o resultado estabelece um limite máximo para o EMSE do DM-SWA em situações onde a probabilidade de divergência do SWA é alta.

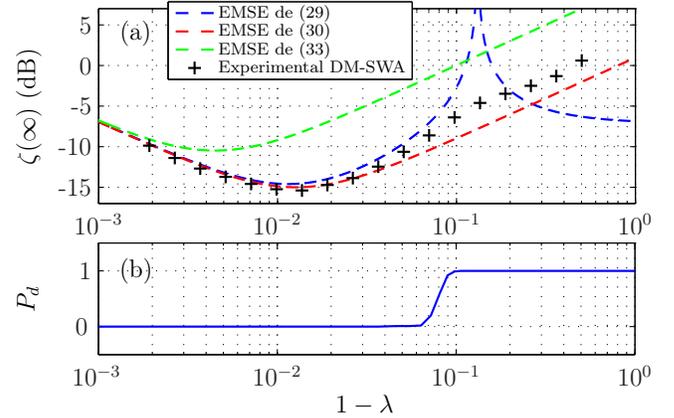


Fig. 2. (a) EMSE do DM-SWA e (b) Probabilidade de divergência do SWA em função de λ ; 4-PAM, $M = 4$, $L = 2$, $\delta = 10^2$, $\mathbf{Q} = 10^{-4} \mathbf{R}^{-1}$.

Com relação à análise de transitório no caso estacionário, foram feitas simulações para os algoritmos SWA e DM-SWA considerando o mesmo canal da simulação em regime e a transmissão de um sinal 6-PAM com estatísticas $\mathbb{E}\{a^6\} = 5,45 \times 10^3$, $\mathbb{E}\{a^2\} = 11,67$ e $r = 20,2$. As curvas experimentais foram obtidas com uma média de 500 realizações e para facilitar a visualização, utilizou-se um filtro de média móvel com 8 coeficientes. Na Figura 3, são mostrados os resultados obtidos com $\lambda = 0,999$ e a estimativa da matriz de autocorrelação inicializada com $\delta \mathbf{I} = 10^2 \mathbf{I}$. Pode-se notar que o valor teórico do EMSE instantâneo tem uma boa concordância com o valor experimental obtido com o SWA, considerando que uma pequena diferença em dB é comum em modelos de algoritmos autodidatas devido às hipóteses que são assumidas para possibilitar a análise. Nesse caso, o resultado obtido com o DM-SWA é semelhante ao obtido com o SWA pois a probabilidade de divergência do SWA é baixa nessas condições e o DM-SWA opera dentro da região de interesse durante a maior parte das iterações. No entanto, pelo fato do DM-SWA operar fora da região de interesse em algumas iterações, pode ser observada uma pequena diferença em sua taxa de convergência em relação à do SWA. Na Figura 4, são mostrados os resultados obtidos nas mesmas condições, mas com a estimativa

da matriz de autocorrelação inicializada com $\delta \mathbf{I} = 10^{-5} \mathbf{I}$. Nesse caso, a inversa da estimativa da matriz de autocorrelação tem norma mais elevada nas iterações iniciais, fazendo com que as hipóteses assumidas não sejam tão realistas, o que causa uma diferença maior entre o EMSE teórico e o experimental durante a convergência. Esse erro que surge em decorrência das aproximações utilizadas, é propagado ao longo das iterações e prejudica o modelo. Nessa figura, é mostrado também o resultado teórico obtido quando a curva teórica é calculada sem considerar a primeira iteração, a fim de diminuir o erro propagado ao longo das iterações. É possível verificar que o resultado teórico é mais coerente com o resultado experimental mas ainda apresenta uma diferença considerável. O estudo de outras aproximações para evitar o erro obtido quando δ é um valor pequeno será tratado num trabalho futuro.

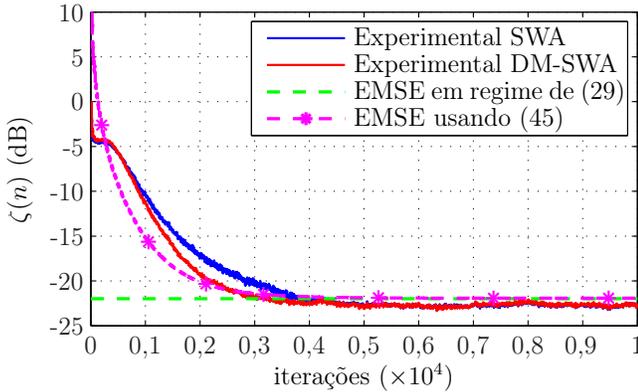


Fig. 3. EMSE teórico e experimental do SWA e DM-SWA; 6-PAM, $M = 4$, $L = 2$, $\lambda = 0,999$, $\delta = 10^2$, média de 500 realizações.

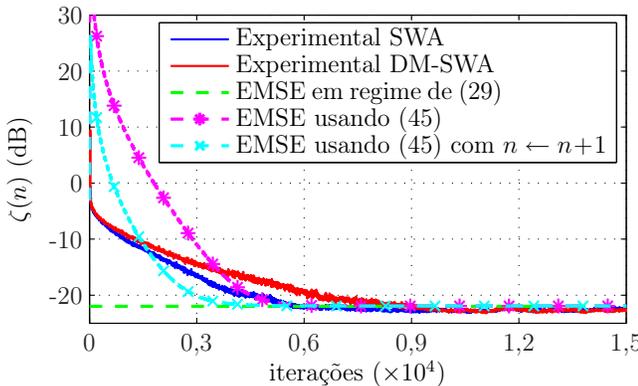


Fig. 4. EMSE teórico e experimental do SWA e DM-SWA; 6-PAM, $M = 4$, $L = 2$, $\lambda = 0,999$, $\delta = 10^{-5}$, média de 500 realizações.

VII. CONCLUSÕES

Neste trabalho foram feitas análises do regime e do transitório do algoritmo DM-SWA, sendo apresentadas expressões analíticas para o seu EMSE. Inicialmente, considerando que o DM-SWA opere dentro da região de interesse, tendo um desempenho idêntico ao do SWA, apresentou-se uma expressão teórica para o EMSE em regime. Neste caso, usando aproximações semelhantes às usadas na análise em regime do RLS de [15], chegou-se a um modelo para o SWA mais preciso que os existentes na literatura, principalmente quando se considera fatores de esquecimento com valores menores. Através de simulações, mostra-se que o modelo proposto possibilita uma estimativa razoável do intervalo de λ para o qual a probabilidade de divergência do SWA é aproximadamente nula. A fim de obter

uma expressão para o valor máximo do EMSE num caso extremo, foi feita também a análise do DM-SWA, considerando a operação apenas fora da região de interesse. O interessante é que esses resultados em conjunto permitem o cálculo teórico de uma faixa de valores para o EMSE em regime obtido com o DM-SWA numa situação qualquer, sendo o EMSE mínimo obtido quando este opera somente dentro da região de interesse e o EMSE máximo dado pela análise fora da região de interesse. Finalmente, foi feita uma análise de transitório no caso estacionário, considerando a operação apenas dentro da região de interesse. Foi obtida uma expressão para o EMSE instantâneo validada através de simulações. Verificou-se que quando a norma da estimativa da matriz de autocorrelação no instante inicial é pequena, o modelo apresenta um erro nas iterações iniciais, o que acaba prejudicando o resultado da análise de transitório. Quando essa norma é elevada, ocorre uma boa concordância entre os resultados teóricos e experimentais.

REFERÊNCIAS

- [1] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Upper Saddle River, 4th edition, 2002.
- [2] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, Springer, 3rd edition, 2008.
- [3] A. H. Sayed, *Fundamentals of Adaptive Filtering*, John Wiley & Sons, NJ, 2003.
- [4] D. N. Godard, "Self-recovering equalization and carrier tracking in two dimensional data communication system," *IEEE Trans. Commun.*, vol. 28, pp. 1867–1875, Nov. 1980.
- [5] J. R. Treichler and B. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. ASSP-28, pp. 334–358, Apr. 1983.
- [6] O. Shalvi and E. Weinstein, "Super-exponential methods for blind deconvolution," *IEEE Transactions on Information Theory*, vol. 39, pp. 504–519, Mar. 1993.
- [7] M. D. Miranda, M. T. M. Silva, and V. H. Nascimento, "Avoiding divergence in the Shalvi Weinstein Algorithm," *IEEE Trans. Signal Process.*, vol. 56, pp. 5403–5413, Nov. 2008.
- [8] I. Fijalkow, C. E. Manlove, and C. R. Johnson, "Adaptive fractionally spaced blind CMA equalization: Excess MSE," *IEEE Trans. Signal Process.*, vol. 46, pp. 227–231, Jan. 1998.
- [9] J. Mai and A. H. Sayed, "A feedback approach to the steady-state performance of fractionally spaced blind adaptive equalizers," *IEEE Trans. Signal Process.*, vol. 48, pp. 80–91, Jan. 2000.
- [10] N. R. Yousef and A. H. Sayed, "A feedback analysis of the tracking performance of blind adaptive equalization algorithms," in *Proc. of IEEE Conf. on Dec. and Control*, Dec. 1999, vol. 1, pp. 174–179.
- [11] M. T. M. Silva and M. D. Miranda, "Tracking issues of some blind equalization algorithms," *IEEE Signal Process. Lett.*, vol. 11, pp. 760–763, Sept. 2004.
- [12] M. T. M. Silva and V. H. Nascimento, "Improving the tracking capability of adaptive filters via convex combination," *IEEE Trans. Signal Process.*, vol. 56, pp. 3137–3149, Jul. 2008.
- [13] V. H. Nascimento and M. T. M. Silva, "Stochastic stability for the Constant Modulus Algorithm," *IEEE Trans. Signal Process.*, vol. 56, pp. 4984–4989, Oct. 2008.
- [14] M. T. M. Silva and V. H. Nascimento, "Tracking analysis of the Constant Modulus Algorithm," in *Proc. of ICASSP'08. IEEE*, 2008, pp. 3561–3564.
- [15] E. Eleftheriou and D. D. Falconer, "Tracking properties and steady-state performance of RLS adaptive filter algorithms," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, pp. 1097–1110, Oct. 1986.
- [16] J. R. Treichler, I. Fijalkow, and C. R. Johnson Jr., "Fractionally spaced equalizers," *IEEE Signal Process. Mag.*, vol. 13, pp. 65–81, May 1996.
- [17] J. E. Mazo, "On the independent theory of equalizer convergence," *Bell Syst. Tech. J.*, vol. 58, pp. 963–993, May/June 1979.
- [18] A. Feuer and E. Weinstein, "Convergence analysis of LMS filters with uncorrelated Gaussian data," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 1, pp. 222–229, Feb. 1985.
- [19] T. Adali and S. H. Ardalan, "On the effect of input signal correlation on weight misadjustment in the RLS algorithm," *IEEE Trans. Signal Process.*, vol. 43, pp. 988–991, Apr. 1995.
- [20] C. F. So, S. C. Ng, and S. H. Leung, "Gradient based variable forgetting factor RLS algorithm," *Signal Processing*, vol. 83, pp. 1163–1175, 2003.