

Controle Ativo de Ruído com Baixo Custo Computacional em Headsets com Reprodução de Fala

Vinícius da P. Wessler e Bruno C. Bispo

Resumo— Este trabalho avalia o desempenho dos algoritmos adaptativos LMS, NLMS e duas implementações do APA no controle ativo de ruído (ANC) em headsets com reprodução de fala e contaminação por ruído branco e de balbuciação. Além do erro quadrático médio, os algoritmos são avaliados pela qualidade e inteligibilidade dos sinais de fala processados. Simulações são feitas em cenários variante no tempo, onde os caminhos acústicos são alterados durante a operação. Os resultados mostram que os algoritmos proporcionam equivalente atenuação de ruído branco, próxima a 10,7 dB quando SNR é 0 dB, com leve superioridade dos APAs na qualidade e inteligibilidade em baixas SNRs. Para ruído de balbuciação, os APAs se destacam por alcançarem atenuação superior a 16 dB quando SNR é 0 dB. Conclui-se que esses algoritmos com baixo custo computacional conseguem melhorar a inteligibilidade da fala em baixas SNRs, fator de suma importância em headsets projetados para serviços de comunicação como telemarketing e teleatendimento, e a qualidade da fala em altas SNRs.

Palavras-Chave— Controle ativo de ruído, sinais de fala, ruído branco, ruído de balbuciação, LMS, NLMS, APA.

Abstract— This work evaluates the performance of the adaptive algorithms LMS, NLMS and two implementations of APA for active noise control (ANC) in headsets with speech reproduction and contamination by white and babble noises. Besides the mean square error, the algorithms are evaluated by the quality and intelligibility of the processed speech signals. Simulations are performed in time-varying scenarios, where acoustic paths are changed during operation. Results show that the algorithms provide equivalent white noise attenuation, close to 10.7 dB when SNR is 0 dB, with slight superiority of APAs in quality and intelligibility at low SNRs. For babble noise, the APAs stand out for attenuating more than 16 dB when SNR is 0 dB. It is concluded that these low-cost algorithms can improve speech intelligibility at low SNRs, a crucial factor for headsets designed for communication services such as telemarketing and call centers, and quality at high SNRs.

Keywords— Active noise control, speech signal, white noise, babble noise, LMS, NLMS, APA.

I. INTRODUÇÃO

O controle ativo de ruído (ANC, do inglês *active noise control*) visa mitigar um sinal indesejado, dito ruído, por meio da superposição com um sinal antirruído, de igual magnitude e fase oposta [1]. O ANC pode ser aplicado em dutos, cabines de aeronaves, interiores de automóveis [2] e, de maior interesse para o presente estudo, em fones de ouvido [3,4,11].

As duas abordagens principais para o ANC empregam filtros adaptativos [2] ou aprendizado de máquina (AM) [15]. Apesar de mais recentes, os métodos baseados em AM têm custo computacional elevado, sendo inviáveis para aplicações em tempo real em dispositivos com baixo poder de processamento [15].

O algoritmo de filtragem adaptativa mais utilizado no ANC é o LMS (*least mean squares*), na estrutura FxLMS (*filtered-X LMS*) [5]. Nele, um passo de adaptação fixo dita a velocidade de convergência. Apesar de simplificar, o passo fixo gera limita-

ções: se grande, o filtro se torna sensível a sinais de entrada não-estacionários; se pequeno, a convergência se torna lenta. Assim, o desempenho pode ser comprometido em cenários dinâmicos.

O NLMS (*normalized LMS*) é o algoritmo com passo de adaptação variável mais empregado, proporcionando maior taxa de convergência e menor sensibilidade ao sinal de entrada [6]. Quando o ruído é autocorrelacionado, a taxa de convergência pode ser aumentada utilizando algoritmos com reutilização de dados, como o APA (*affine projection algorithm*) [7,8]. Porém, o APA não pode ser implementado de maneira direta no ANC, requerendo um processo de estimação de alguns sinais [8].

Independente do algoritmo empregado, o ANC pode ser implementado nas estruturas *feedforward* e *feedback*. A primeira utiliza dois microfones, para captar os sinais de referência e de erro, e a redução de ruído é limitada pelo passo de adaptação: quanto maior o passo, mais rápida a convergência, porém maior o erro em regime permanente [5]. A segunda utiliza apenas o microfone de erro, sendo de mais fácil implementação, porém consegue eliminar apenas a parcela determinística do ruído [2].

Os algoritmos LMS, NLMS e APA se encontram bem estudados na literatura de ANC quanto ao erro quadrático médio (MSE) [5,6,8]. No entanto, poucos estudos com reprodução de fala em headsets, a saber [15,16,17], avaliam também a qualidade e inteligibilidade da fala processada, fatores relevantes para comunicação pouco analisados na literatura de ANC [11]. Ainda mais, somente um desses estudos, a saber [15], avalia o desempenho desses algoritmos com reprodução de fala e contaminação por ruído de balbuciação (*babble*), cenário tipicamente encontrado em salas de telemarketing e teleatendimento.

Nesse sentido, esse trabalho tem o objetivo de avaliar o desempenho dos algoritmos LMS, NLMS e duas implementações do APA no contexto de ANC, com estrutura *feedforward*, em um headset com reprodução de fala e sujeito a ruídos branco e *babble*. A avaliação considera não somente o MSE, mas também a qualidade e inteligibilidade das falas processadas, em cenários variantes no tempo onde os caminhos acústicos são alterados durante a operação. Dessa forma, o trabalho contribui para literatura de ANC ao avaliar o impacto desses algoritmos na qualidade e inteligibilidade da fala ouvida em headsets com baixo poder de processamento e voltados ao teleatendimento.

O presente artigo está organizado da seguinte maneira: a seção II apresenta a estrutura de ANC em headsets e os algoritmos adaptativos empregados; a seção III descreve as duas estratégias para a implementação do APA; a seção IV traz a configuração das simulações realizadas; a seção V apresenta e discute os resultados; por fim, a seção VI conclui o trabalho.

II. CONTROLE ATIVO DE RUÍDO EM HEADSETS

Essa seção apresenta brevemente a estrutura de ANC em headsets e os algoritmos de filtragem adaptativa utilizados.

A. Estrutura de Redução de Ruído

Uma ilustração e o diagrama de blocos do problema são mostrados nas Fig. 1 e 2, respectivamente, onde [4,5]:

C. Estimação do Caminho Secundário

A estimativa $\hat{S}(z)$ pode ser calculada por *Adp2* de maneira *off-line* ou *on-line*, isto é, antes ou durante a operação do ANC [19]. Para uma melhor estimativa, deseja-se que $\mathbf{x}[n] \approx \mathbf{0}$ e a referência $\mathbf{a}[n]$ tenha banda larga. Essas condições, contudo, normalmente são apenas satisfeitas no caso *off-line*.

Substituindo (6) em (4), o sinal de erro pode ser escrito como

$$e[n] = d[n] - \mathbf{s}^T \mathbf{y}[n] - \mathbf{s}^T \mathbf{a}_M[n] + \hat{\mathbf{s}}^T \mathbf{a}_N[n]. \quad (20)$$

Se $\hat{S}(z)$ não modelar adequadamente $S(z)$, então um resquício do áudio $\tilde{\mathbf{a}}[n] = -\mathbf{s}^T \mathbf{a}_M[n] + \hat{\mathbf{s}}^T \mathbf{a}_N[n]$ será realimentado em *Adp*, comprometendo a efetividade do ANC. O impacto de $\tilde{\mathbf{a}}[n]$ em $e[n]$ é inversamente proporcional a potência do ruído $d[n]$.

Com o processo de atualização *on-line*, $\hat{\mathbf{s}}$ se torna $\hat{\mathbf{s}}[n]$.

III. IMPLEMENTAÇÕES DO APA

Para empregar o APA em sistemas ANC, é necessário obter $\mathbf{e}[n]$ sem o conhecimento de $d[n]$. Esse trabalho compara duas implementações amplamente utilizadas na literatura.

A. Recuperação Utilizando Amostras Passadas do Erro

A abordagem mais simples consiste em reconstruir o vetor de erro utilizando suas amostras passadas. Dessa maneira, o conhecimento de $d[n]$ não é necessário, bastando armazenar os valores passados de $e[n]$, obtidos por (4), resultando em [8]

$$\mathbf{e}[n] \approx [e[n] \ e[n-1] \ \dots \ e[n-L+1]]^T. \quad (21)$$

Essa aproximação tem melhor desempenho para valores pequenos de L e em ambientes estacionários [8]. Nesse trabalho, o APA com essa implementação será denominado APA1.

B. Vetor Desejado a Partir do Erro Acústico

Essa abordagem calcula o vetor $\mathbf{e}[n]$ após estimar o vetor $\mathbf{d}[n]$ a partir do erro acústico. Como o erro compensado a cada ciclo de adaptação é dado por (20), o valor de $d[n]$ pode ser estimado como [14]

$$d[n] = e[n] - \mathbf{s}^T \mathbf{y}[n] - \mathbf{s}^T \mathbf{a}_M[n] + \hat{\mathbf{s}}^T \mathbf{a}_N[n] \approx e[n] + \hat{\mathbf{s}}^T \mathbf{y}[n]. \quad (22)$$

Obtendo os valores de $d[n]$ em cada ciclo de adaptação, o sinal de erro pode ser reconstruído como [14]

$$\mathbf{e}[n] \approx \begin{bmatrix} d[n] - \mathbf{w}^T[n] \hat{\mathbf{x}}[n] \\ d[n-1] - \mathbf{w}^T[n] \hat{\mathbf{x}}[n-1] \\ \vdots \\ d[n-L+1] - \mathbf{w}^T[n] \hat{\mathbf{x}}[n-L+1] \end{bmatrix}. \quad (23)$$

Para que essa aproximação seja acurada, a estimativa do caminho secundário $\hat{S}(z)$ deve ser precisa [14]. Nesse trabalho, o APA com essa implementação será chamado de APA2.

IV. CONFIGURAÇÃO DAS SIMULAÇÕES

Essa seção descreve a configuração das simulações computacionais realizadas no *software* Matlab.

A. Sinais de Fala e Ruído

Foram utilizados 10 sinais de fala, 5 de locutores masculinos e 5 de femininos, da base de dados descrita em [10], com cerca de 8,5 s de duração e amostrados em 16 kHz. Os sinais possuem silêncio com duração de 0,5 a 1 s no início e de 0,5 s no final.

Foram utilizadas 10 realizações de ruído branco Gaussiano geradas por *software* e 10 realizações de ruído de balbuciação

do banco de dados MMS-SNSD [20], cujas amplitudes foram manipuladas para obter SNRs de 0, 10 e 20 dB no sinal $e_a[n]$.

B. Caminhos Acústicos

Os caminhos acústicos $P(z)$ e $S(z)$ foram modelados por respostas ao impulso, medidas num fone de ouvido intra-auricular, da base PANDAR [12]. As respostas foram reamostradas para 16 kHz e então truncadas em $M = 512$ amostras. Dois pares de caminhos acústicos foram usados: $P_1(z)$ e $S_1(z)$; $P_2(z)$ e $S_2(z)$. Para emular um cenário variante no tempo, após um terço de operação, o primeiro par foi alterado para o segundo.

O caminho secundário foi estimado usando o algoritmo LMS. Uma estimativa inicial $\hat{S}(z)$ foi obtida de forma *off-line* e, após o início da operação, foi atualizada de maneira *on-line*.

C. Parâmetros dos Algoritmos

Os parâmetros dos algoritmos foram fixados em $N = 256$, $\delta = 0,001$ e $L = 10$. De forma comparativa, foram usados $N = \{50, 300\}$, $\delta = 0,001$ e $L = \{5, 10\}$ em [14], enquanto $N = 50$ e $L = 5$ em [8]. O passo μ foi otimizado, através de uma pesquisa em grade, de forma a minimizar a média do MSE para cada algoritmo e ruído na SNR de 0 dB, cenário mais crítico entre os analisados, resultando nos valores exibidos na Tabela I.

TABELA I. PASSOS DE ADAPTAÇÃO UTILIZADOS.

Ruído	LMS	NLMS	APA1	APA2
Branco	0.01	0.125	0.025	0.025
Babble	0.00075	0.035	0.0125	0.0125

D. Métricas de Avaliação

O desempenho dos algoritmos foi mensurado pelo MSE entre a fala convoluída com o caminho secundário, $\mathbf{s}^T \mathbf{a}_M[n]$, e o erro acústico, $-e_a[n]$. O primeiro é o áudio que se deseja ouvir e o segundo é o áudio ouvido de fato. Como contribuição do presente trabalho, os algoritmos também foram avaliados pela qualidade e inteligibilidade dos sinais de fala processados.

1) *MSE*: O erro quadrático proporcionado pelo ANC é

$$E[n] = (e_a[n] + \mathbf{s}^T \mathbf{a}[n])^2. \quad (24)$$

O *MSE*[n] é a média, ao longo de todos os sinais, de $E[n]$. O *MSE*[n] subtraído da SNR indica a atenuação global.

2) *PESQ*: O WPESQ (*Wideband Perceptual Evaluation of Speech Quality*) é um algoritmo para avaliação objetiva da qualidade de sinais de fala amostrados a 16 kHz [13]. Ele compara representações psicoacústicas de um sinal de fala degradado e sua referência não corrompida. A pontuação do WPESQ pode ser mapeada para a escala 1-5 de opinião média (MOS, do inglês *Mean Opinion Score*), resultando na pontuação MOS-LQO (MOS-*Listening Quality Objective*). Neste trabalho, o WPESQ foi utilizado para mensurar a qualidade perceptual das falas processadas. Para tanto, os sinais $\mathbf{s}^T \mathbf{a}_M[n]$ e $-e_a[n]$ foram usados como referência e sinal degradado, respectivamente.

3) *NCM*: O NCM (*Normalized Covariance Metric*) é uma métrica para avaliação objetiva da inteligibilidade de sinais de fala [13]. Ela se baseia na covariância entre envelopes temporais de sub-bandas de um sinal de fala degradado e sua referência não corrompida [13]. Sua pontuação varia entre 0 e 1, onde valores mais altos indicam maior inteligibilidade. Neste trabalho, o NCM foi usado para mensurar a inteligibilidade das falas processadas. Para tanto, os sinais $\mathbf{s}^T \mathbf{a}_M[n]$ e $-e_a[n]$ foram usados como referência e sinal degradado, respectivamente.

V. RESULTADOS E DISCUSSÃO

Essa seção apresenta os resultados dos quatro algoritmos adaptativos no headset com reprodução de fala.

A. Ruído Branco

A Fig. 3 exibe as curvas de $MSE[n]$ obtidas pelos algoritmos para ruído branco e $SNR = 0$ dB. A linha tracejada vermelha na vertical indica o início da reprodução da fala. A média e o desvio padrão da atenuação nas três SNRs são mostrados na Tabela II. As curvas do APA1 e APA2 praticamente se sobrepõem.

Observa-se que, em regime permanente, foram obtidos valores de MSE próximos a -16 dB para o primeiro par de caminhos acústicos, mas apenas próximos a $-8,5$ dB para o segundo par. Isso ocorreu porque a solução ótima para o segundo par de caminhos acústicos é mais longa que a do primeiro, requerendo filtros mais longos ($N > 512$) para obter similar redução de ruído. Isso ilustra que, para um filtro com tamanho fixo, a redução alcançável de ruído é dependente dos caminhos acústicos.

Nota-se que, no início do $MSE[n]$, os APAs convergiram mais rapidamente e o NLMS foi ligeiramente superior ao LMS. Em regime permanente, todos atingiram valores semelhantes, como esperado. Esses resultados são consistentes com os apresentados em [8,14]. Após a mudança de caminhos acústicos, os algoritmos reconvergiram com velocidade similar.

Atenuações médias entre 10,3 e 10,7 dB foram obtidas para $SNR = 0$ dB. Os níveis de atenuação diminuíram com o aumento da SNR, conforme esperado, devido a menor potência de $\tilde{a}[n]$. Entre os algoritmos, a pouca diferença nas curvas de $MSE[n]$ se refletiu em níveis semelhantes de atenuação de ruído.

A média e o desvio-padrão do MOS-LQO e do NCM, nos três níveis de SNR, são apresentados na Tabela III. Nota-se que, em comparação proporcional ao caso sem ANC, a melhoria obtida na qualidade aumenta com o aumento da SNR, enquanto que na inteligibilidade diminui. Em $SNR = 0$ dB, os APAs elevaram o NCM de 0,54 para até 0,83, mas o MOS-LQO de 1,02 para apenas 1,096. Por outro lado, em $SNR = 20$ dB, eles subiram o MOS-LQO de 1,35 para até 2,46, mas o NCM de 0,94 para 0,974. A redução da melhoria na inteligibilidade, com o aumento da SNR, se deve à subida significativa da inteligibilidade da fala contaminada, a qual atinge um valor médio de 0,94 (próximo ao topo de escala) em $SNR = 20$ dB.

Assim, pode-se inferir que para ruído branco, em SNRs muito baixas, os algoritmos não são capazes de tornar a fala mais agradável, mas sim de torná-la mais compreensível, fator de suma importância em headsets projetados para comunicação. Por outro lado, em SNRs mais altas, os algoritmos não são capazes de promover substancial melhoria na inteligibilidade da fala, mas sim de tornar a fala significativamente mais agradável. Essa evidência sobre a relação entre a variação da SNR global e a melhoria fornecida pelos algoritmos adaptativos estudados na inteligibilidade e qualidade dos sinais de fala não foi encontrada na literatura de ANC e é uma contribuição do presente trabalho.

Portanto, para contaminação por ruído branco, os algoritmos proporcionaram semelhante redução de ruído, mas os APAs obtiveram um melhor desempenho em relação à qualidade e inteligibilidade da fala processada.

B. Ruído de Balbuciação

As curvas de $MSE[n]$ obtidas pelos algoritmos para ruído *babble* e $SNR = 0$ dB são apresentados na Fig. 4. A média e o desvio padrão da atenuação nas três SNRs são mostrados na Tabela IV. As curvas dos APAs se sobrepõem mais uma vez. Observa-se agora uma convergência mais lenta de todos os algo-

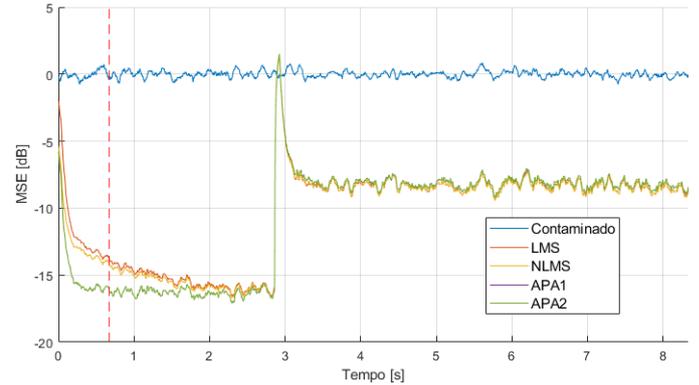


Fig. 3. Convergência dos algoritmos para ruído branco e $SNR = 0$ dB.

TABELA II. ATENUAÇÃO (dB) PARA RUÍDO BRANCO.

SNR (dB)	LMS	NLMS	APA1	APA2
0	10,31 ± 0,10	10,58 ± 0,09	10,68 ± 0,09	10,67 ± 0,09
10	9,75 ± 0,47	10,01 ± 0,43	10,01 ± 0,31	9,99 ± 0,31
20	6,59 ± 1,96	6,79 ± 1,82	6,25 ± 1,30	6,21 ± 1,32

TABELA III. PONTUAÇÕES DE MOS-LQO E NCM PARA RUÍDO BRANCO.

Adp.	Métrica	SNR = 0 dB	SNR = 10 dB	SNR = 20 dB
—	MOS-LQO	1,021 ± 0,004	1,058 ± 0,025	1,351 ± 0,139
	NCM	0,542 ± 0,111	0,829 ± 0,053	0,940 ± 0,013
LMS	MOS-LQO	1,069 ± 0,030	1,399 ± 0,170	2,316 ± 0,280
	NCM	0,696 ± 0,095	0,909 ± 0,032	0,970 ± 0,006
NLMS	MOS-LQO	1,070 ± 0,031	1,408 ± 0,173	2,332 ± 0,280
	NCM	0,726 ± 0,089	0,917 ± 0,028	0,971 ± 0,006
APA1	MOS-LQO	1,096 ± 0,046	1,509 ± 0,192	2,462 ± 0,247
	NCM	0,830 ± 0,055	0,940 ± 0,013	0,974 ± 0,008
APA2	MOS-LQO	1,096 ± 0,045	1,507 ± 0,191	2,458 ± 0,245
	NCM	0,830 ± 0,055	0,939 ± 0,013	0,974 ± 0,008

ritmos, uma evidente superioridade dos APAs, e uma substancial inferioridade do LMS. Nota-se também que até os APAs não atingiram a convergência com o primeiro par de caminhos acústicos, por causa do menor passo de adaptação. Simulações mais longas indicaram que os algoritmos convergem perto de -21 dB com o primeiro par e de -18 dB com o segundo. No entanto, o menor valor de μ contribuiu para os algoritmos terem um desempenho em regime permanente melhor do que com ruído branco.

Atenuações médias de 9,2, 13,7 e 16,4 dB foram obtidas com LMS, NLMS e APAs, respectivamente, para $SNR = 0$ dB. Mais uma vez, os níveis de atenuação diminuíram com o aumento da SNR devido a menor potência do ruído. Desvios-padrões mais elevados foram obtidos devido à não estacionariedade do ruído *babble*. Entre os algoritmos, a vantagem dos APAs no $MSE[n]$ se refletiu em níveis superiores de atenuação de ruído. A superioridade dos APAs pode ser explicada pelo ruído de balbuciação ser correlacionado, ao contrário do ruído branco, característica que favorece algoritmos com reutilização de dados.

A média e o desvio-padrão do MOS-LQO e do NCM, nos três níveis de SNR, são apresentados na Tabela V. Observa-se que a superioridade dos APAs no MSE se refletiu na qualidade e inteligibilidade da fala. Eles obtiveram o melhor resultado nas métricas específicas de fala nos três níveis de SNR. Percebe-se também que, em comparação com o ruído branco, níveis superiores de qualidade e inteligibilidade foram geralmente obtidos para o mesmo algoritmo e mesmo valor de SNR. Isso é consequência dos melhores níveis de atenuação de ruído fornecidos pelos algoritmos, com exceção do LMS.

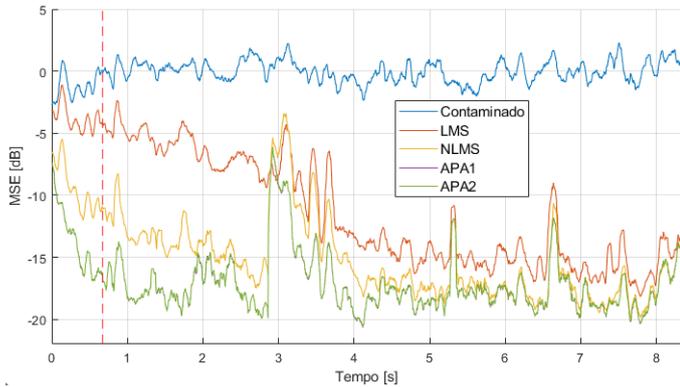


Fig. 4. Convergência dos algoritmos para ruído babble e SNR = 0 dB.

TABELA IV. ATENUAÇÃO (dB) PARA RUÍDO BABBLE.

SNR (dB)	LMS	NLMS	APA1	APA2
0	9,22 ± 1,90	13,68 ± 1,97	16,38 ± 1,91	16,36 ± 1,92
10	9,22 ± 1,85	12,74 ± 1,62	14,72 ± 1,52	14,69 ± 1,53
20	8,68 ± 1,63	8,42 ± 1,87	8,83 ± 1,79	8,78 ± 1,81

TABELA V. PONTUAÇÕES DE MOS-LQO E NCM PARA RUÍDO BABBLE.

Adp.	Métrica	SNR = 0 dB	SNR = 10 dB	SNR = 20 dB
—	MOS-LQO	1,088 ± 0,039	1,124 ± 0,291	1,460 ± 0,157
	NCM	0,398 ± 0,133	0,732 ± 0,110	0,939 ± 0,051
LMS	MOS-LQO	1,119 ± 0,065	1,417 ± 0,163	2,373 ± 0,340
	NCM	0,561 ± 0,126	0,841 ± 0,083	0,969 ± 0,028
NLMS	MOS-LQO	1,200 ± 0,089	1,834 ± 0,250	2,961 ± 0,267
	NCM	0,735 ± 0,097	0,925 ± 0,046	0,984 ± 0,012
APA1	MOS-LQO	1,447 ± 0,185	2,388 ± 0,328	3,322 ± 0,245
	NCM	0,845 ± 0,077	0,967 ± 0,026	0,994 ± 0,006
APA2	MOS-LQO	1,446 ± 0,185	2,386 ± 0,327	3,320 ± 0,245
	NCM	0,844 ± 0,077	0,967 ± 0,026	0,994 ± 0,006

É relevante notar que, em comparação proporcional ao caso sem ANC, o aumento da SNR levou novamente a um aumento da melhoria obtida na qualidade e uma redução da melhoria na inteligibilidade. Em SNR = 0 dB, os APAs elevaram o NCM de 0,40 para até 0,844, mas o MOS-LQO de 1,124 para apenas 1,447. Por outro lado, em SNR = 20 dB, eles subiram o MOS-LQO de 1,46 para até 3,32, mas o NCM de 0,94 para 0,994.

Assim, para ruído de balbuciação, também pode-se inferir que, em SNRs muito baixas, os algoritmos não conseguem tornar a fala mais agradável, mas sim mais compreensível. Em contrapartida, em SNRs mais altas, os algoritmos não conseguem promover substancial melhoria na inteligibilidade da fala, mas sim na qualidade. Essa relação entre a variação da SNR global e a melhoria fornecida pelos algoritmos adaptativos estudados na inteligibilidade e qualidade de sinais de fala contaminados por ruído de balbuciação não foi encontrada na literatura de ANC e é uma contribuição do presente trabalho.

Portanto, para contaminação por ruído de balbuciação, os APAs superaram os outros algoritmos, tanto no nível de redução de ruído quanto na inteligibilidade e qualidade da fala processada. O NLMS teve um desempenho intermediário que, dado sua menor carga computacional, justifica ser uma boa opção para sistemas com baixo poder de processamento.

VI. CONCLUSÕES

Este trabalho avaliou o desempenho dos algoritmos adaptativos LMS, NLMS e duas implementações do APA para controle ativo de ruído em headsets com reprodução de fala e contaminação por ruído branco e de balbuciação. A avaliação consi-

derou, além do tradicional erro médio quadrático, a qualidade e a inteligibilidade da fala processada, preenchendo uma lacuna na literatura. Simulações foram feitas em um cenário dinâmico, onde os caminhos acústicos foram alterados durante a operação.

Resultados mostraram que os algoritmos proporcionam níveis semelhantes de atenuação de ruído branco, superiores a 10 dB para SNR = 0 dB, com leve superioridade dos APAs na qualidade e inteligibilidade em baixas SNRs. Para ruído de balbuciação, os APAs se destacaram por alcançarem atenuações superiores a 16 dB em SNR = 0 dB, e por superarem os outros algoritmos na inteligibilidade e qualidade da fala processada.

Por fim, concluiu-se que os algoritmos adaptativos com baixo custo computacional estudados conseguem melhorar a inteligibilidade do sinal de fala em baixas SNRs, fator de suma importância em headsets projetados para serviços de comunicação como telemarketing e teletendimento, e a qualidade da fala em altas SNRs. Pesquisa sobre métodos para aumentar a qualidade da fala em baixas SNR deve ser realizada no futuro.

REFERÊNCIAS

- [1] S. J. Elliott, P. A. Nelson, “Active noise control”, *IEEE Signal Processing Magazine*, v. 10, no. 4, pp. 12-35, Oct. 1993.
- [2] S. M. Kuo, D. R. Morgan, “Active noise control: a tutorial review”, *Proceedings of the IEEE*, v. 87, no. 6, Jun. 1999.
- [3] W. S. Gan, S. Mitra, S. M. Kuo, “Adaptive feedback active noise control headset: implementation, evaluation and its extensions”, *IEEE Trans. Consum. Electron.*, v. 51, no. 3, Aug. 2005.
- [4] Y. K. Chong *et al.*, “Integrated headset using the adaptive hybrid active noise control system”, *International Conference on Information Communications & Signal Processing*, 2005, pp. 1324-1328.
- [5] I. T. Ardekani, W. H. Abdula., “FxLMS-based active noise control: a quick review”, *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, 2011.
- [6] D. T. M. Slock, “Modified LMS and NLMS algorithms with a new variable step size”, *International Conference on Control, Automation, Robotics and Vision*, 2006.
- [7] M. L. R. de Campos, P. S. R. Diniz, J. A. Apolinário Jr., “On normalized data-reusing and affine-projections algorithms”, *IEEE ICECS*, 1999.
- [8] A. González *et al.*, “Fast filtered-x affine projection algorithm for active noise control”, *IEEE WASPAA*, 2005.
- [9] B. Farhang-Boroujeny. *Adaptive Filters: Theory and Applications*. John Wiley & Sons, 2nd ed., 2013.
- [10] V. Alencar, A. Alcaim, “LSF and LPC - derived features for large vocabulary distributed continuous speech recognition in Brazilian Portuguese”, *Asilomar Conference on Signals, Systems and Computers*, 2018, pp.1237-1241.
- [11] S. Liebich *et al.*, “Signal processing challenges for active noise cancellation headphones”, *ITG-Symposium Speech Communication*, 2018.
- [12] S. Liebich *et al.*, “Acoustic path database for ANC in-ear headphone development”, *International Congress of Acoustics*, 2019, pp. 4326-4333.
- [13] P. C. Loizou, *Speech Enhancement: Theory and Practice*. CRC, 2013.
- [14] M. Ferrer *et al.*, “Fast affine projection algorithms for filtered-x multichannel active noise control”. *IEEE/ACM Trans. Audio Speech Lang. Process.*, v. 16, no. 8, Nov. 2008.
- [15] H. Zhang, D. L. Wang, “Deep ANC: a deep learning approach to active noise control”, *Neural Networks*, v. 141, Sep. 2021.
- [16] G. Vaziri *et al.*, “Evaluating noise suppression methods for recovering the Lombard speech from vocal output in an external noise field”, *Int. J. Speech Technol.*, v. 22, pp 31-46, Nov. 2018.
- [17] S. M. Potirakis *et al.*, “Assessment of military intercom headsets for maximum voice reproduction level in high noise conditions”, *Applied Acoustics*, v. 74, no. 6, pp 870-881, Jun. 2013.
- [18] M. Guldenschuh. “Secondary-path models in adaptive-noise-control headphones”, *IEEE ICSC*, 2013, pp. 653-658.
- [19] A. A. Milani, I. M. S. Panahi, R. Briggs, “Performance analysis of sub-band nLMS, APA and RLS in fMRI ANC with a non-minimum phase secondary path”, *IEEE ICASSP*, 2007, pp. 353- 356.
- [20] C. K. A. Reddy *et al.*, “A scalable noisy speech dataset and online subjective test framework”, *Interspeech*, 2019, pp. 1816-1820.