

# Comparação entre os Algoritmos MUSIC de Análise por Banda Estreita e de Análise por Banda Larga na Localização de Fontes Sonoras

Fabio Zaccaro Scelza, Markus Vinicius Santos Lima e Luiz Wagner Pereira Biscainho

**Resumo**—Comparamos duas versões do algoritmo MUSIC já consolidadas na literatura, formuladas para localização de uma única fonte estática, quando aplicadas ao caso de múltiplas fontes sonoras móveis em ambiente aberto no campo distante. A literatura informa que elas apresentam desvantagens na análise da localização de mais de uma fonte emitindo sinais de natureza não-estacionária. Decidimos verificar numa série de experimentos as limitações dos algoritmos na prática. Expõem-se considerações e particularidades sobre o desempenho desses algoritmos em relação à definição e classificação de picos máximos em funções de saída e às características dos sinais de entrada empregados.

**Palavras-Chave**—MUSIC, movimento, fonte sonora, localização.

**Abstract**—We compare two well-known versions of the MUSIC algorithm, previously designed for localization of one static sound source, when used to localize multiple moving sound sources. We consider free-field environment and sources at far-field distances from the microphone array. The literature itself informs disadvantages when these algorithms aim to localize more than one sound source emitting non-stationary signals simultaneously. We decided, through several experiments, to verify their limitations in practice. Particularities of the performances of these strategies are reported regarding the readability and classification of maximum peaks in the output functions, as well as the different characteristics of the input signals used.

**Keywords**—MUSIC, movement, sound source, localization.

## I. INTRODUÇÃO

A demanda comercial pelo desenvolvimento de algoritmos localizadores de eventos por meios autônomos vem crescendo devido ao recente conceito de internet das coisas (*Internet of Things - IoT*) aplicado a serviços de controle e vigilância [1]. Além disso, o interesse público por cidades mais sustentáveis, seguras e organizadas despertou a iniciativa das cidades inteligentes (*Smart Cities*), as quais utilizam múltiplos sensores e análises de dados para identificação e controle de eventos urbanos [2]. Neste contexto, insere-se o estágio atual dos estudos sobre localização de fontes sonoras. Esforços de pesquisa em áudio focados na localização e controle em tempo real de fontes emissoras de ruído ou sinais específicos ganharam grande destaque nos últimos anos, oferecendo oportunidades para o desenvolvimento de novos trabalhos e soluções [3].

O estudo de localização de fontes de áudio se aplica a dois tipos de ambiente, fechado e aberto, cada qual com desafios

particulares quando se trata de encontrar a posição espacial de emissores. O ambiente fechado possui, em geral, grande incidência de reflexão, levando interferências construtivas do sinal refletido a serem confundidas com o sinal emitido. Entretanto, em ambientes fechados há maior controle de ruído e previsibilidade de eventos. Em ambiente aberto, a taxa de reflexão costuma ser muito baixa; com isso, a atenuação do sinal emitido pode ser um problema para sua captação e processamento. Além disso, como se trata de ambiente só parcialmente influenciado pelo homem, não há possibilidade de total controle sobre as interferências. Ambientes abertos se aplicam à maioria das situações de interesse quando falamos de eventos urbanos, pois há interesse público em identificar fontes de áudio em ruas, estradas e praças, por exemplo.

Nesse contexto, as fontes emissoras de som podem ser fixas ou móveis. Fontes fixas, na prática, podem ser estruturas urbanas que emitem ruído, como fábricas, festas ao ar livre, sinalizadores, incêndios etc. As fontes móveis podem se relacionar com automóveis, pessoas gritando em distúrbios sociais, disparos em movimento, drones etc. A atualização contínua da localização no segundo caso, ou rastreamento, baseia-se em características únicas (inferidas por meios probabilísticos ou algébricos) do sinal emitido por cada tipo de fonte. Desse modo, mais de uma fonte sonora móvel pode ser localizada simultaneamente ao longo de suas trajetórias.

Os principais métodos de localização de fonte sonora a partir do áudio captado por um arranjo de microfones se dividem em 3 tipos: i) Potência de Resposta Direcional (*Steered Response Power - SRP*) de *Beamformer*: utiliza um critério de filtragem espacial tal que a direção de captura com maior potência de chegada indicará a direção da fonte emissora; ii) Diferença de Tempo de Chegada (*Time Difference of Arrival - TDoA*) entre Pares de Microfones: determina a localização a partir de estimativas do conjunto de atrasos relativos de percurso até os microfones do arranjo; iii) Estimção Espectral de Alta Resolução [4]: faz a análise espectral das matrizes de correlação entre os sinais recebidos pelos microfones para regiões delimitadas no espaço, associando a posição da fonte à região com componentes mais acentuadas [4].

Pensando em aplicações típicas de ambiente urbano, como rastrear e localizar a ocorrência de fenômenos inesperados (acidentes, manifestações sociais, sinais de alerta etc.) capazes de interferir na normalidade da região, só consideramos emissores sonoros móveis em campo aberto, realizando sua trajetória em campo distante e a uma velocidade constante.

Após esta seção contextualizando seu tema, o trabalho está

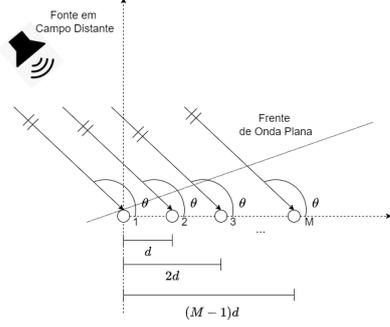


Fig. 1: Problema de localização de fonte sonora em campo distante usando um ULA com  $M$  microfones.

organizado como se segue: a Seção II apresenta o modelo descritivo dos sinais capturados; a Seção III resume os algoritmos de localização a serem comparados; a Seção IV relata e discute os resultados de experimentos simulados; e a Seção V apresenta conclusões do trabalho.

## II. MODELO DO SINAL

Considere o problema de localização de fonte sonora ilustrado na Fig. 1. Nessa figura, temos um arranjo composto de  $M$  microfones dispostos de maneira linear formando um arranjo linear uniforme (ULA, do inglês *Uniform Linear Array*). A distância entre dois microfones adjacentes é denotada por  $d$ . A tarefa de localização de fonte sonora consiste em determinar a posição angular  $\theta$  da fonte sonora em relação ao ULA. Nesse problema, assumimos que: (i) a fonte sonora está em campo distante (*far field*), i.e., a distância entre a fonte e o ULA é suficientemente grande para nos permitir assumir que a frente de onda que atinge os microfones é plana; e (ii) a onda se propaga em ambiente aberto, i.e., livre de reflexões.

A modelagem a seguir é fortemente baseada em [5]. Seja  $s[n]$  o sinal sonoro emitido pela fonte, com  $n$  inteiro. Ele se propaga pelo meio até os microfones do arranjo. Definindo o primeiro microfone do arranjo como referencial, o TDoA do sinal a cada elemento do arranjo é

$$\tau_{1i} = (i - 1)\tau, \quad i = 2, \dots, M, \quad (1)$$

com

$$\tau = \tau_{12} = \frac{d \cos(\theta)}{c}, \quad (2)$$

onde  $\tau$  denota o TDoA entre o primeiro e segundo microfones e  $c$  corresponde à velocidade de propagação do som no meio.

O sinal recebido pelo arranjo de microfones pode ser representado no domínio da frequência por

$$\mathbf{Y}(e^{j\Omega}) = \alpha e^{-j2\pi t_0 \Omega} S(e^{j\Omega}) \boldsymbol{\zeta}(\tau, e^{j\Omega}) + \mathbf{V}(e^{j\Omega}). \quad (3)$$

Se  $\mathbf{Y}(e^{j\Omega}) = [Y_1(e^{j\Omega}) \ Y_2(e^{j\Omega}) \ \dots \ Y_M(e^{j\Omega})]^T$ ,  $Y_i(e^{j\Omega})$  é a transformada de Fourier de tempo discreto (DTFT, do inglês *Discrete-Time Fourier Transform*) do sinal  $y_i[n]$  recebido pelo  $i$ -ésimo microfone;  $S(e^{j\Omega})$  é a DTFT de  $s[n]$ ;  $\mathbf{V}(e^{j\Omega})$  é um vetor contendo as DTFTs dos ruídos aditivos (assumidos brancos) referentes a cada microfone  $v_i[n]$ ; o vetor de direção

$$\boldsymbol{\zeta}(\tau, e^{j\Omega}) = [1 \ e^{-j2\pi\tau\Omega} \ \dots \ e^{-j2\pi(M-1)\tau\Omega}]^T, \quad (4)$$

sendo  $[\cdot]^T$  o operador de transposição, é formado pelos deslocamentos de fase do sinal recebido em cada microfone em relação ao sinal recebido pelo microfone de referência, calculado usando os TDoAs dados na equação (1);  $t_0$  é o tempo de chegada (ToF, do inglês *Time-of-Flight*, que é o tempo que leva para o sinal emitido chegar ao microfone de referência); e  $\alpha$  é um número real que representa a atenuação que o sinal da fonte sofre ao longo do percurso (assumido igual para todos os microfones, pois a distância entre fonte e arranjo é muito maior que  $(M-1)d$ ). Assim, o sinal recebido pelo  $i$ -ésimo microfone pode ser representado por

$$Y_i(e^{j\Omega}) = \alpha e^{-j2\pi[t_0 + (i-1)\tau]\Omega} S(e^{j\Omega}) + V_i(e^{j\Omega}), \quad (5)$$

que no domínio do tempo equivale a

$$y_i[n] = \alpha s[n - \lfloor (t_0 + (i-1)\tau)F_s \rfloor] + v_i[n], \quad (6)$$

para  $i = 1, \dots, M$ , onde  $F_s$  representa a frequência de amostragem do sinal e  $\lfloor \cdot \rfloor$  é a operação de arredondamento para o inteiro mais próximo.

## III. MUSIC

Neste trabalho, duas versões do algoritmo de Classificação Múltipla de Sinais (*Multiple Signal Classification - MUSIC*) são estudadas [5]: o algoritmo de varredura por banda estreita *Narrowband MUSIC* (NbMUSIC) e o de varredura por banda larga *Broadband MUSIC* (BbMUSIC). Essas duas técnicas são brevemente descritas nas subseções a seguir sob a hipótese de existência de uma única fonte. Na Subseção III-D, ambas as técnicas são generalizadas para múltiplas fontes.

### A. Narrowband MUSIC

Como o próprio nome já diz, a utilização deste algoritmo é restrita a sinais de banda estreita. Usando a equação (3), a matriz de covariância  $\mathbf{R}_Y(e^{j\Omega}) = E[\mathbf{Y}(e^{j\Omega}) \mathbf{Y}^H(e^{j\Omega})]$ , sendo  $E[\cdot]$  o valor esperado e  $[\cdot]^H$  o operador hermitiano, é dada por

$$\mathbf{R}_Y(e^{j\Omega}) = \mathbf{R}_S(e^{j\Omega}) + \sigma_V^2(e^{j\Omega})\mathbf{I}, \quad (7)$$

onde

$$\mathbf{R}_S(e^{j\Omega}) = \alpha^2 \sigma_S^2(e^{j\Omega}) \boldsymbol{\zeta}(\tau, e^{j\Omega}) \boldsymbol{\zeta}^H(\tau, e^{j\Omega}), \quad (8)$$

e  $\sigma_S^2(e^{j\Omega})$  e  $\sigma_V^2(e^{j\Omega})$  são as variâncias do sinal emitido pela fonte e do ruído, respectivamente.

Ao realizar a decomposição espectral da matriz  $\mathbf{R}_Y$ , temos

$$\mathbf{R}_Y(e^{j\Omega}) = \mathbf{B} \boldsymbol{\Lambda} \mathbf{B}^H, \quad (9)$$

onde  $\boldsymbol{\Lambda}$  é uma matriz diagonal cujos elementos não-nulos são os autovalores de  $\mathbf{R}_Y$  em ordem decrescente, i.e., tais que  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M$ , e  $\mathbf{B}$  é uma matriz cujas colunas são os respectivos autovetores, denotados por  $\mathbf{b}_i$ , para  $i = 1, \dots, M$ . Analisando as equações (7) e (8), observamos que  $\mathbf{R}_S(e^{j\Omega})$  possui posto 1 e, portanto, só contribui com o primeiro autovalor de  $\mathbf{R}_Y$ , enquanto a componente de ruído  $\sigma_V^2(e^{j\Omega})\mathbf{I}$  influencia todos os autovalores, pois tem posto completo. Dessa forma, temos que  $\lambda_i = \sigma_V^2(e^{j\Omega})$  para todo  $i \geq 2$ , i.e.,

$$\mathbf{R}_Y(e^{j\Omega}) \mathbf{b}_i = \lambda_i(e^{j\Omega}) \mathbf{b}_i = \sigma_V^2(e^{j\Omega}) \mathbf{b}_i, \quad \forall i \neq 1. \quad (10)$$

A partir das equações (8) e (10), podemos concluir que

$$\alpha^2 \sigma_s^2 (e^{j\Omega}) \zeta(\tau) \zeta^H(\tau) \mathbf{b}_i = \mathbf{0} \Rightarrow \mathbf{b}_i^H \zeta(\tau) = 0, \quad \forall i \neq 1, \quad (11)$$

ou seja, que os autovetores pertencentes ao subespaço onde existe apenas ruído são ortogonais ao vetor de direção  $\zeta(\tau)$ .

Portanto, na implementação do NbMUSIC, buscamos  $\tau'$  que maximiza a função pseudoespectral

$$J_{\text{music}}(\tau') = \frac{1}{\sum_{i=2}^M |\mathbf{b}_i^H \zeta(\tau')|^2}, \quad (12)$$

após o que o respectivo  $\theta$  é determinado via equação (2).

### B. Narrowband MUSIC para sinais de banda larga

O NbMUSIC pode ser facilmente estendido para aplicações envolvendo sinais de banda larga. Para isso, basta calcular  $J_{\text{music}}(\tau')$  para diversas frequências  $\Omega$ , varrendo principalmente as frequências onde o sinal tem mais energia. Em seguida, buscamos o  $\tau'$  correspondente ao maior pico em  $J_{\text{music}}(\tau')$  considerando todas as frequências analisadas.

### C. Broadband MUSIC

Para lidar com sinais de banda larga, o BbMUSIC utiliza os sinais representados no domínio do tempo e calcula a matriz de correlação espacial do sinal  $\mathbf{R}_y[\rho] = E[\mathbf{y}_\rho \mathbf{y}_\rho^T]$ , onde as  $M$  entradas do vetor  $\mathbf{y}_\rho$  possuem a seguinte lei de formação:  $y_{\rho,i}[n] = y_i[n + (i-1)\rho]$ . A ideia aqui é que através deste  $\rho$  possamos estimar  $n_\tau = \lfloor \tau F_s \rfloor$ , que é o TDoA dado em amostras. Note que para  $\rho = n_\tau$ , o vetor  $\mathbf{y}_\rho$  compensa os atrasos intermicrofones presentes na equação (6) de forma a alinhar temporalmente todos os  $M$  sinais, o que faz com que a componente de sinal em  $\mathbf{R}_y[\rho]$  tenha posto 1, de forma semelhante ao que foi explicado na Subseção III-A. Consequentemente, uma análise por meio dos autovalores e autovetores de  $\mathbf{R}_y[\rho]$  nos leva à seguinte conclusão [5]:

$$\mathbf{b}_i^T[\rho] \mathbf{R}_y[\rho] \mathbf{b}_i[\rho] \begin{cases} = \sigma_v^2, & \text{se } \rho = n_\tau \\ \geq \sigma_v^2, & \text{se } \rho \neq n_\tau, \end{cases} \quad (13)$$

para todo  $i \neq 1$ . Novamente,  $\mathbf{b}_i[\rho]$  é o  $i$ -ésimo autovetor de  $\mathbf{R}_y[\rho]$  e os autovalores são organizados em ordem decrescente.

Assim, no algoritmo BbMUSIC buscamos encontrar o argumento  $\rho$  que maximiza a função-objetivo

$$J_{\text{Bbmusic}}[\rho] = \frac{1}{\sum_{i=2}^M \mathbf{b}_i^T[\rho] \mathbf{R}_y[\rho] \mathbf{b}_i[\rho]}. \quad (14)$$

### D. Múltiplas Fontes

Seja  $Q$  o número de fontes, com  $1 \leq Q < M$ . As funções-objetivo descritas anteriormente podem ser adaptadas para lidar com múltiplas fontes da seguinte forma [5]:

$$J_{\text{music}}(\tau') = \frac{1}{\sum_{i=Q+1}^M |\mathbf{b}_i^H \zeta(\tau')|^2}, \quad (15)$$

$$J_{\text{Bbmusic}}[\rho] = \frac{1}{\sum_{i=Q+1}^M \mathbf{b}_i^T[\rho] \mathbf{R}_y[\rho] \mathbf{b}_i[\rho]}. \quad (16)$$

Além disso, em vez de buscarmos um único pico (máximo), agora devemos buscar os argumentos correspondentes aos  $Q$  maiores picos nessas funções-objetivo.

## IV. SIMULAÇÕES E RESULTADOS

O objetivo desta seção é avaliar o desempenho dos algoritmos NbMUSIC e BbMUSIC na localização de múltiplas fontes sonoras. Nas simulações a seguir, foi utilizado um ULA com  $d = 1$  cm,  $c = 340$  m s<sup>-1</sup>, assumiram-se fontes no *far-field* e razão-sinal-ruído (SNR de cada um dos sinais sobre o ruído aditivo de fundo, branco e gaussiano) de 50 dB. Na Subseção IV-A, as fontes são artificiais e favorecem as técnicas; na Subseção IV-B utilizamos fontes realísticas com características mais desafiadoras e em movimento.

### A. Definição dos picos da função objetivo

Os experimentos realizados aqui visam a ilustrar o formato das funções objetivo do NbMUSIC e do BbMUSIC. Dado que estamos resolvendo um problema de localização de fonte, é mais útil reparametrizar essas funções-objetivo em função da posição angular  $\theta$ , através da equação (2), formando então a função  $P(\theta)$ .

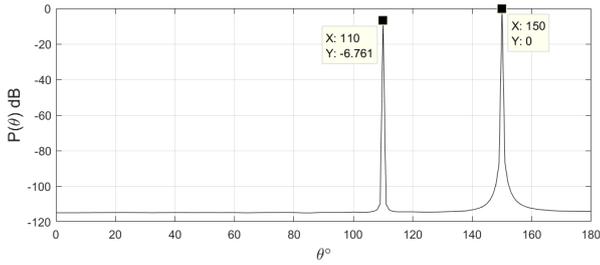
A Fig. 2 mostra as funções-objetivo  $P(\theta)$  para ambas as técnicas. Nessas simulações, consideramos a existência de duas fontes estáticas, uma a 110° e outra a 150° de posição angular, emitindo som simultaneamente. Na escolha dos sinais sonoros, buscamos favorecer as técnicas: para o NbMUSIC escolhemos senoides de 1 kHz e 3 kHz, respectivamente, e para o BbMUSIC, escolhemos ruídos brancos gaussianos independentes, sempre a uma taxa de amostragem de  $F_s = 15000$  amostras por segundo.

Os resultados exibidos na Fig. 2 mostram que ambas as técnicas conseguem localizar as duas fontes de forma satisfatória. Contudo, o  $P(\theta)$  na Fig. 2a exibe picos mais íngremes (mais definidos) que os picos observados na Fig. 2b. De fato, isso era esperado uma vez que a relação entre  $\rho$  e  $\theta$  é

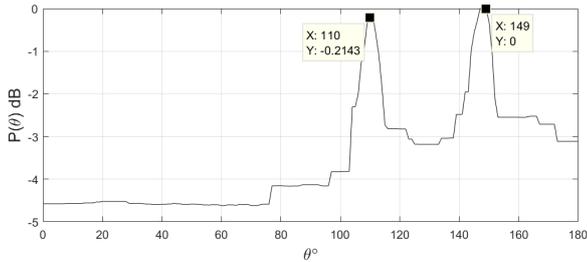
$$\rho = \lfloor \frac{d \cos(\theta)}{c} F_s \rfloor. \quad (17)$$

Devido ao arredondamento existente nessa relação, um mesmo  $\rho$  pode corresponder a diferentes valores de  $\theta$ , o que gera o padrão mais “quadriculado” de  $P(\theta)$  exibido na Fig. 2b.

Esse problema de resolução do BbMUSIC pode ser mitigado pelo aumento da frequência de amostragem. Aplicamos a  $F_s$  fatores de interpolação (FI) de 10, 100 e 1000 vezes. A Fig. 3, que por clareza mostra apenas FI = 100 e FI = 1000, confirma tal efeito. Naturalmente, ao se aumentar  $F_s$  para reduzir os efeitos da operação de arredondamento em (17), o algoritmo precisará lidar com uma quantidade bem maior de amostras, aumentando consideravelmente o seu tempo de processamento. Como ilustração, nas simulações realizadas o tempo de processamento para o caso de FI = 1000 foi aproximadamente 600 vezes maior que para o caso de FI = 10.



(a) *Narrowband MUSIC* para fontes senoidais.



(b) *Broadband MUSIC* para fontes de ruído branco.

Fig. 2: Localização de duas fontes estáticas em 110° e 150°.

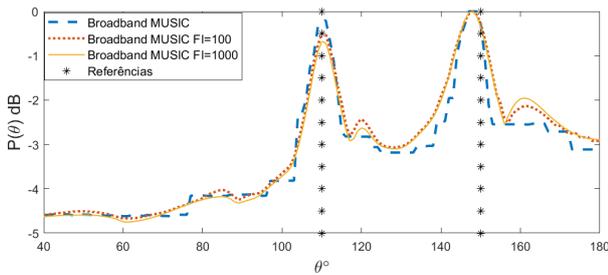
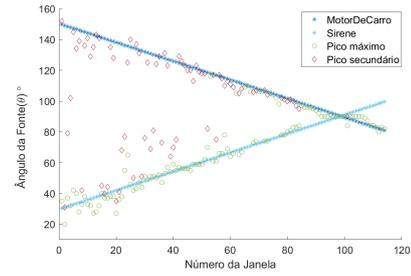


Fig. 3: Algoritmos *Broadband MUSIC* em comparação de desempenho para fontes estáticas em 110° e 150°.

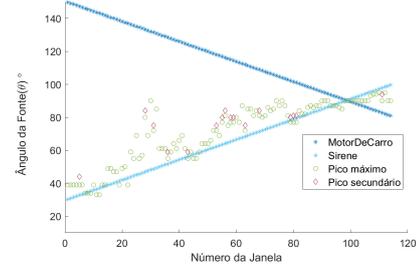
### B. Fontes Móveis

Nesta subseção, consideramos um cenário mais desafiador: duas fontes reais em movimento emitindo sinais de banda larga. Seu deslocamento é modelado por um movimento circular uniforme com raio de 50 m ao redor do sensor de referência. Uma das fontes tem posição angular inicial em 30°, a outra iniciando em 150°. Essas fontes se deslocam em sentidos contrários a uma velocidade de 20 km h<sup>-1</sup> e, em algum ponto, elas se cruzam, o que aumenta a dificuldade de localizá-las. A taxa de amostragem é  $F_s = 41000$  amostras por segundo. O áudio processado pelo ULA é subdividido em pequenos blocos sequenciais e sem sobreposição, chamados janelas, cada qual com 4250 amostras do sinal. Como o número de fontes (2) é conhecido, para cada janela geramos 2 resultados de localização correspondentes aos 2 maiores picos da função objetivo (sempre que eles existirem): o pico principal ou máximo e o pico secundário. Quanto aos sinais emitidos, uma das fontes emite o som de um motor de carro em aceleração, enquanto a outra fonte emite uma sirene de ambulância. Ambos os sinais de áudio são gravados em condições reais, e foram obtidos de [6] e [7].

Os resultados da Fig. 4 mostram que o NbMUSIC consegue localizar razoavelmente ambas as fontes ao longo da trajetória,



(a) *Narrowband MUSIC*.



(b) *Broadband MUSIC*.

Fig. 4: Localização das fontes de Motor de Carro e Sirene em movimento.

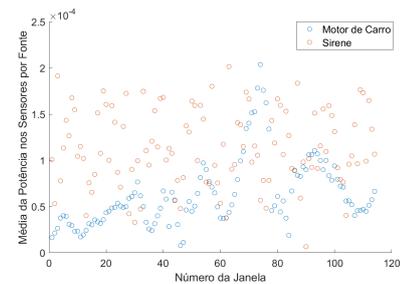
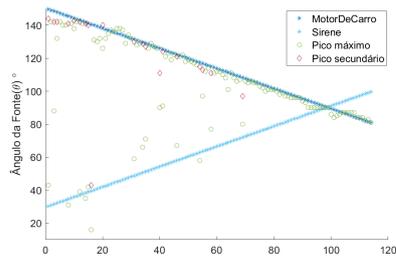


Fig. 5: Potência isolada dos sinais que chegam nos sensores.

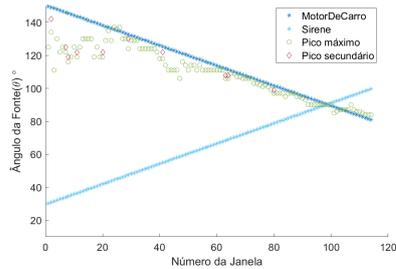
exceto nas janelas próximas ao cruzamento das fontes, quando apenas uma das fontes é localizada. Nota-se também que os picos principais (em verde) estão majoritariamente relacionados ao sinal da sirene, e os picos secundários ao motor de carro. Isso se justifica pelo fato de que a potência do sinal da sirene é ligeiramente superior à do motor de carro durante quase todo o tempo, como pode ser visto na Fig. 5. Já o BbMUSIC, por outro lado, só conseguiu localizar uma das fontes (a sirene) e, ainda assim, suas estimativas de localização possuem um grande viés.

Com a finalidade de averiguar melhor a influência da potência dos sinais na acurácia da localização, mais dois experimentos foram realizados. Em cada um desses experimentos, um dos sinais é mantido exatamente igual ao que já foi descrito (com potência por janela mostrada na Fig. 5), enquanto o outro sinal tem sua amplitude aumentada em 10 vezes.

Na Fig. 6, temos os resultados de localização para o caso em que o sinal do motor foi amplificado de 10 vezes. Em comparação com os resultados da Fig. 4, observa-se que os picos principais do NbMUSIC localizaram o motor de carro, que agora é a fonte dominante, enquanto a sirene passou a ser muito mal localizada pela técnica. O BbMUSIC, por sua vez, novamente só localizou uma das fontes, desta vez o motor de carro (dominante). Mais uma vez, é possível



(a) Narrowband MUSIC.



(b) Broadband MUSIC.

Fig. 6: Localização das fontes em movimento - Sinal do Motor com amplitude multiplicada por 10.

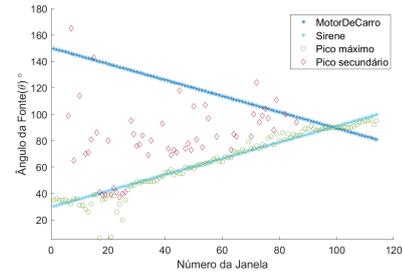
notar que as estimativas são enviesadas na direção da outra fonte (especialmente nas primeiras janelas), como se existisse uma disputa entre as potências das fontes janela-a-janela. Conforme ilustrado na Fig. 5, a potência do motor de carro é muito pequena nas primeiras janelas; assim, mesmo com a amplificação do sinal em 10 vezes, sua potência resultante não resulta tão maior que a potência da sirene.

Para confirmar a nossa tese da fonte dominante, amplificou-se em outra simulação o sinal da sirene (que na configuração original já era mais potente, como se vê na Fig. 5) por um fator de 10. Os resultados podem ser vistos na Fig. 7. Confirmando nossa expectativa, ambas as técnicas conseguiram localizar a fonte dominante satisfatoriamente. O próprio viés observado nas estimativas de localização do BbMUSIC nas Figs. 4 e 6 é praticamente inexistente na Fig. 7.

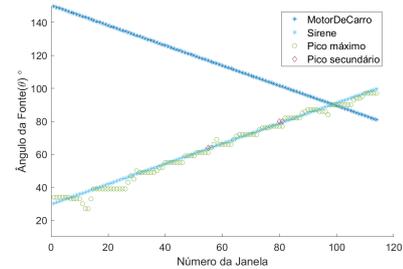
## V. CONCLUSÕES

Neste artigo, nós estudamos o desempenho dos algoritmos NbMUSIC e BbMUSIC na localização de fontes sonoras. Consideramos a existência de duas fontes sonoras emitindo sinais reais de banda larga e em deslocamento. Além desse cenário mais desafiador em que testamos os algoritmos, também analisamos a influência da potência dos sinais no desempenho das técnicas. Este aspecto, embora pouco explorado na literatura, mostrou-se determinante para o desempenho de uma das técnicas.

Para o NbMUSIC, notamos que ele consegue localizar a fonte dominante com boa acurácia, enquanto que a localização da outra fonte (de menor intensidade) depende de ela não ser muito menos potente que a dominante. Além disso, como esta técnica realiza uma varredura ao longo das frequências que compõem a banda do sinal, fontes com características espectrais diferentes tendem a favorecer a técnica. Este é



(a) Narrowband MUSIC.



(b) Broadband MUSIC.

Fig. 7: Localização das fontes em movimento - Sinal da Sirene com amplitude multiplicada por 10.

o caso, por exemplo, quando o pico de energia dos sinais emitidos pelas fontes estão em frequências bem distintas.

Já para o BbMUSIC, verificamos que, embora ele seja capaz de lidar com múltiplas fontes (como no caso simulado com emissões de ruído branco), nos cenários práticos aqui simulados ele só conseguiu estimar uma das fontes (a de maior potência). Notamos também que, em geral, a estimativa de posição é enviesada na direção da outra fonte. Portanto, deve-se ter algum cuidado ao utilizar este algoritmo em ambientes com múltiplas fontes. Neste caso, o BbMUSIC pode ser útil caso haja interesse apenas na fonte de maior potência. Caso o objetivo seja localizar todas as múltiplas fontes, o BbMUSIC pode funcionar caso as fontes tenham potências similares e banda larga/cheia (similar a um ruído branco). Para esta tarefa, entretanto, os autores optariam pelo NbMUSIC ao invés do BbMUSIC, pois a primeira técnica é menos sensível às características dos sinais emitidos e, portanto, tende a ser mais efetiva na prática.

## REFERÊNCIAS

- [1] L. Hu and Q. Ni, "IoT-driven automated object detection algorithm for urban surveillance systems in smart cities," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 747–754, 2018.
- [2] C. Yin, Z. Xiong, H. Chen, and J. Wang, "A literature survey on smart cities," *Sci. China Inf. Sci.*, vol. 58, pp. 1–18, 2015.
- [3] C. Evers, H. W. Löllmann, H. Mellmann, A. Schmidt, H. Barfuss, P. A. Naylor, and W. Kellermann, "The LOCATA challenge: Acoustic source localization and tracking," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1620–1643, 2020.
- [4] M. S. Brandstein and D. B. Ward, *Microphone Arrays: Signal Processing Techniques and Applications*. Springer, 2001.
- [5] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*. Springer, 2008.
- [6] modman34, "car engine.mp3." Disponível em: <https://freesound.org/people/modman34/sounds/512558/>. Acesso em: 29 de Junho de 2022, 2020.
- [7] Robinhood76, "09759 ambulance siren loop.wav." Disponível em: <https://freesound.org/people/Robinhood76/sounds/540153/>. Acesso em: 01 de fevereiro 2022, 2020.