

MODELAGEM E TRATAMENTO ACÚSTICO DE AMBIENTES URBANOS

Lucas do Vale Santos

Orientadora: Mariane Rembold Petraglia



Laboratório de Processamento Analógico e Digital de Sinais (PADS)

Engenharia Eletrônica e de Computação

Escola Politécnica - UFRJ

INTRODUÇÃO

- Compreender como a acústica de um ambiente urbano funciona é de grande importância para gerenciar determinados problemas práticos.
- A modelagem para esse tipo de ambiente apresenta grande utilidade no estudo de algumas aplicações como as seguintes:
 - Localização espacial de veículos (DOA - *Direction of Arrival*);
 - Identificação e separação de fontes sonoras;
 - Formulação de assinaturas sonoras (para cada veículo);
 - Detecção e tratamento do nível de ruído.
- No momento, será abordado como utilizamos métodos para estimar o DOA.



DIREÇÃO DE CHEGADA (DOA)

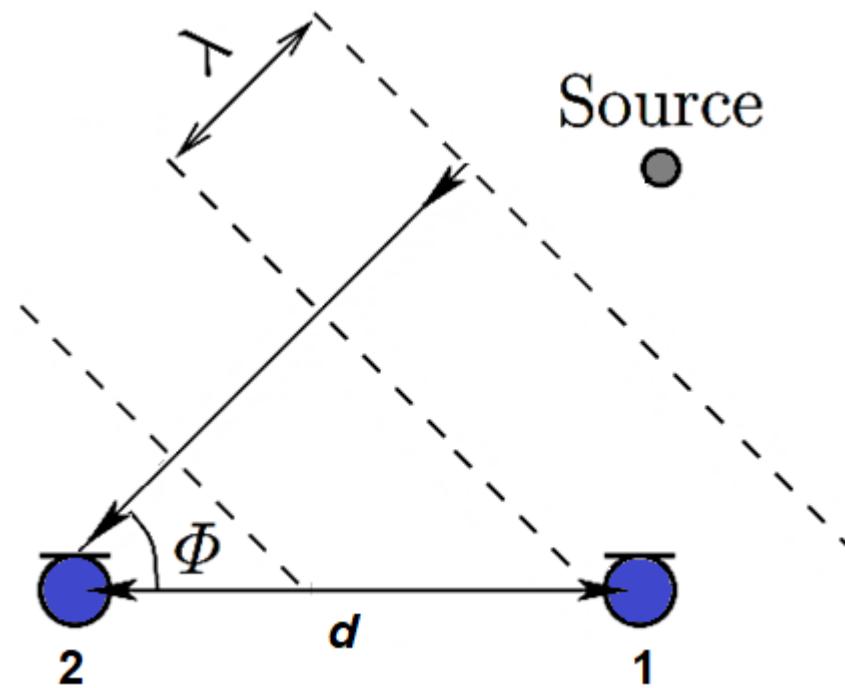
- Para estimativa do ângulo de chegada, foram aplicados alguns algoritmos de processamento de sinais:
 - Generalized cross-correlation with phase transform (GCC-PHAT);
 - Interaural time differences (ITD);
 - Adaptative eigenvalue decomposition (AEVD);
 - Fast block least mean square (FLMS).
- Em conjunto com tais métodos, dispomos do *beamforming* (formador de feixes) para redução de ruído e interferências, ou seja, maximização dos sinais das direções desejadas.



DIREÇÃO DE CHEGADA (DOA)

- A localização das fontes sonoras foi estudada avaliando-se as estimativas das DOAs das fontes principais em um conjunto de microfones.
- Sendo d a distância entre microfones, Δt a diferença de tempo de chegada e v_s a velocidade do som:

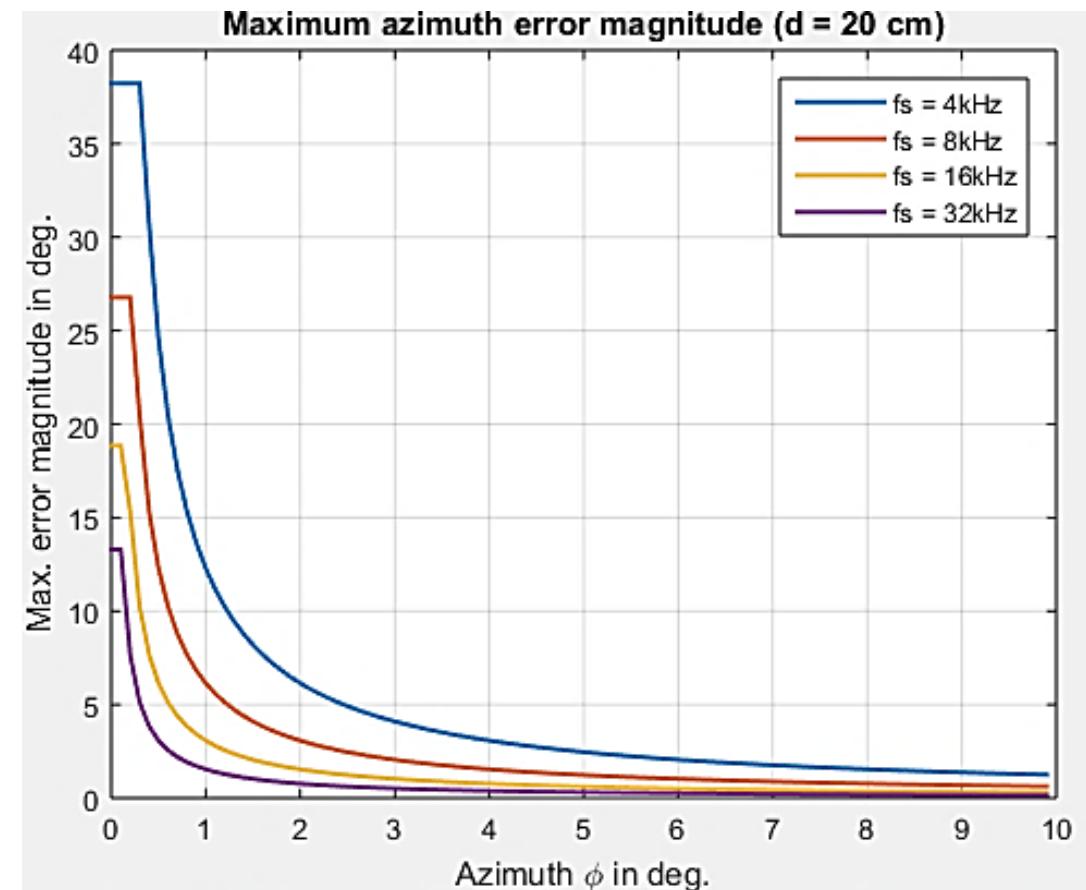
$$\Phi = \arccos \frac{v_s \Delta t}{d}$$



DIREÇÃO DE CHEGADA (DOA)

- Devido à relação não-linear, a estimativa apresenta pouca acurácia para ângulos próximos a 0 ou 180 graus;
- Assim, podemos estimar o erro máximo da estimativa do DOA da seguinte forma:

$$\delta\Phi_{\max}(\Phi) \approx \min \left(\arccos \left(1 - \frac{v_s}{2df_s} \right), \frac{\frac{v_s}{2df_s}}{|\sin \Phi|} \right)$$



MÉTODOS



GCC-PHAT

- Sejam os sinais analógicos (considerando ruído) que chegam aos microfones

$$\begin{array}{l} x_{a1}(t) = s_a(t) + v_{a1}(t) \\ x_{a2}(t) = s_a(t - \tau_0) + v_{a2}(t) \end{array} \xrightarrow{\substack{\text{Amostrando} \\ \text{a um período } T}} \begin{array}{l} x_1(n) = s(n) + v_1(n) \\ x_2(n) = \underbrace{s_a(nT - \tau_0)}_{s_{\tau_0}(n)} + v_2(n) \end{array}$$

- Pela transformada de Fourier

$$\begin{aligned} X_1(e^{j\Omega}) &= S(e^{j\Omega}) + V_1(e^{j\Omega}) \\ X_2(e^{j\Omega}) &= S(e^{j\Omega}) e^{-j\Omega\frac{\tau_0}{T}} + V_2(e^{j\Omega}) \end{aligned}$$

- A densidade espectral de potência cruzada é dada por

$$S_{x_1 x_2}(\Omega) = E\{X_1(e^{j\Omega}) X_2^*(e^{j\Omega})\} = S_{ss}(\Omega) e^{j\Omega\frac{\tau_0}{T}}$$



GCC-PHAT

- O *delay* pode ser estimado pelo índice referente ao máximo da função correlação cruzada:

$$\frac{\tau_0}{T} \approx n_0 = \arg \max_n R_{x_1 x_2}(n)$$

onde

$$R_{x_1 x_2}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \psi_{12}(e^{j\Omega}) S_{x_1 x_2}(\Omega) e^{j\Omega n} d\Omega$$

e a transformada de fase é obtida com

$$\psi_{12}(e^{j\Omega}) = \frac{1}{|S_{x_1 x_2}(\Omega)|}$$

- A densidade espectral é obtida por interpolação através das DFTs dos sinais janelados, a fim de acomodar a não-estacionariedade dos sons



ITD

- Utiliza a diferença *interaural* de tempo para determinar espacialmente a posição de uma fonte, dada por

$$\tau_i = \frac{d}{2v_s} \sin \left(\frac{i-1}{I-1} \pi - \frac{\pi}{2} \right), \quad i = 1, 2, \dots, I$$

- A ideia é encontrar a melhor coincidência entre os sinais de um par de microfones para estimar o *atraso ótimo*:

$$p_k(i) = e^{-j \frac{2\pi}{N} k f_s \tau_i}, \quad k = 0, 1, \dots, \frac{N}{2}, \quad i = 1, 2, \dots, I$$

$$P_k(\tau_i, m) = \alpha P_k(\tau_i, m-1) + \delta(i - i_{\text{opt}}(m, k))$$

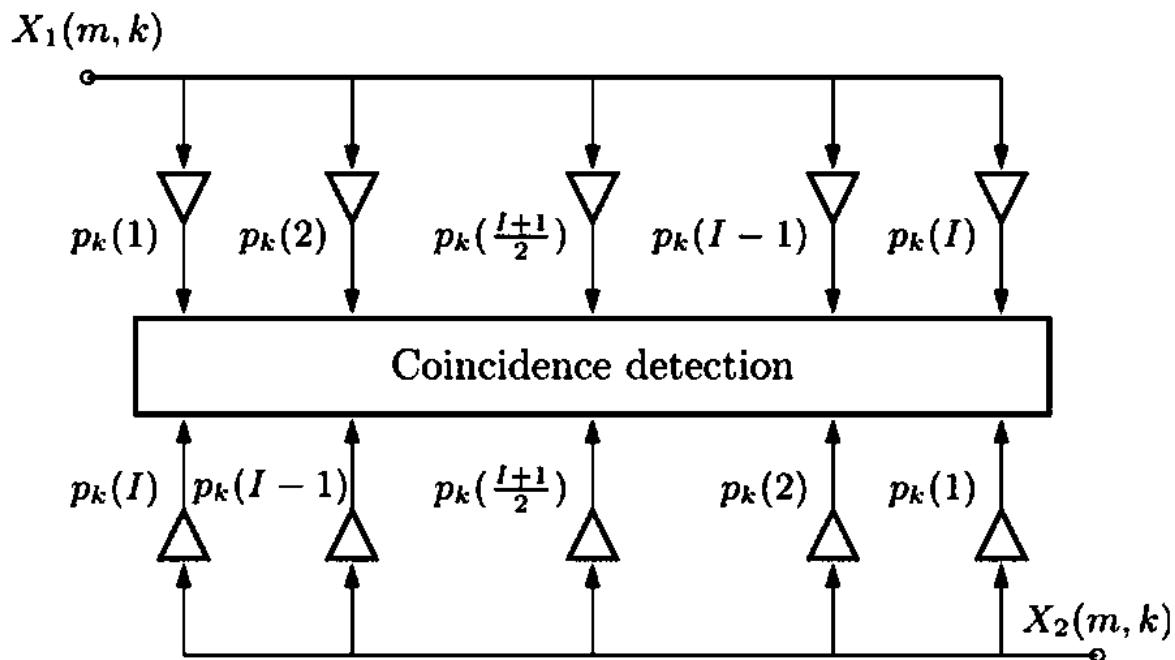


ITD

com índices ótimos:

$$i_{\text{opt}}(m, k) = \arg \min_i \Delta_i(m, k), \quad k = 0, 1, \dots, \frac{N}{2}$$

$$\Delta_i(m, k) = \left| p_k(i)X_1(m, k) - p_k(I - i + 1)X_2(m, k) \right|^2, \quad i = 1, 2, \dots, I$$

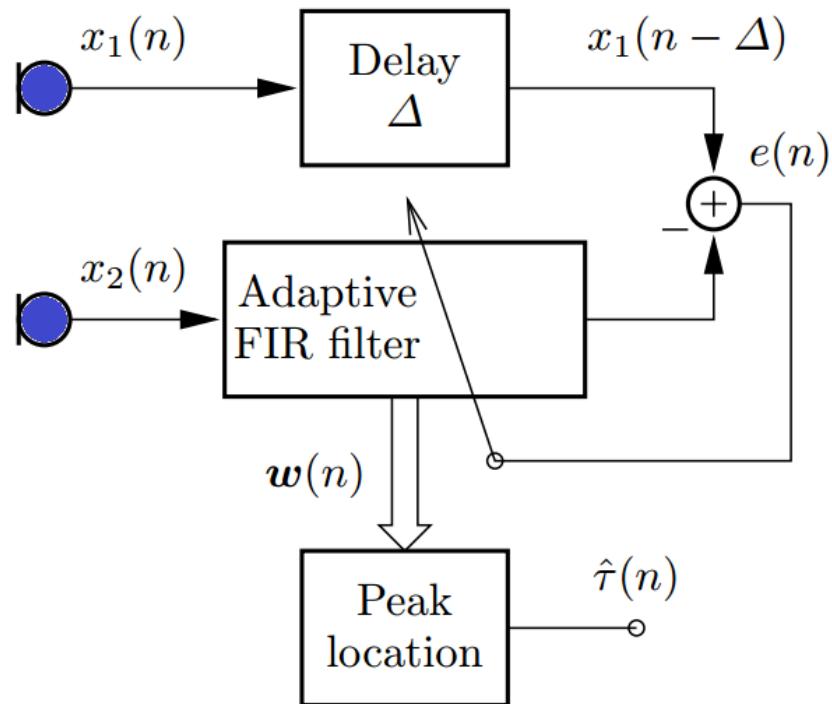


O atraso ótimo é obtido por:

$$\tau_{\text{opt}}(m) = \arg \max_{\tau_i} \sum_{k=0}^{\frac{N}{2}} P_k(\tau_i, m)$$



FLMS



- A ideia é reproduzir o filtro desejado FIR, encontrando seus coeficientes relacionados aos mínimos quadrados do sinal de erro (LMS).
- Sendo o sinal de erro e o vetor de pesos

$$e(n) = x_1(n - \Delta) - \mathbf{w}^T(n) \mathbf{x}_2(n),$$

$$\mathbf{w}(n + 1) = \mathbf{w}(n) + \mu_{\text{LMS}} e(n) \mathbf{x}_2(n)$$

- Implementa os filtros adaptativos pelo *fast block* LMS no domínio da frequência (FFT):

$$W(m + 1, k) = W(m, k) + \frac{\mu}{S_{x_2 x_2}(m, k) + \varepsilon} X_2^*(m, k) E(m, k), \quad k = 0, 1, \dots, N - 1$$



FLMS

- Esses coeficientes são atualizados por quadros conforme as relações:

$$X_2(m, k) = \sum_{n=0}^{N-1} x_2(mL + n) e^{-j\frac{2\pi}{N}nk}, \quad k = 0, 1, \dots, N-1$$

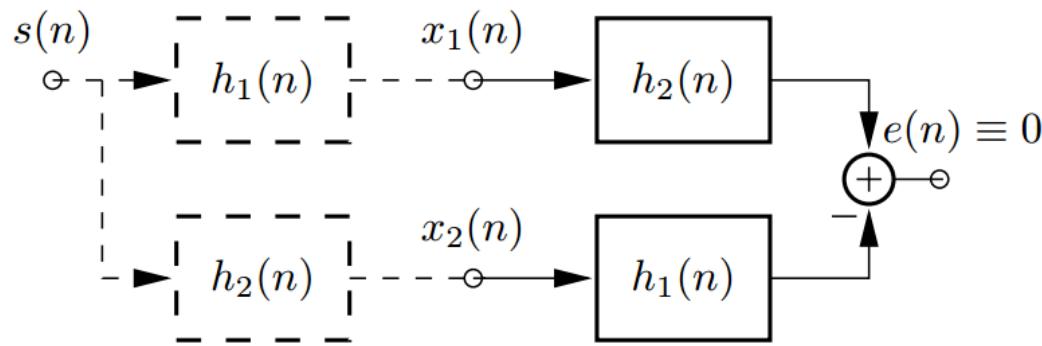
$$\tilde{e}(m, n) = \begin{cases} 0 & n = 0, 1, \dots, L-1 \\ x_1(mL + n - \Delta) - \frac{1}{N} \sum_{k=0}^{N-1} W(m, k) X_2(m, k) e^{j\frac{2\pi}{N}nk} & n = L, L+1, \dots, N-1 \end{cases}$$

- O delay é obtido pelos picos correspondentes a cada quadro

$$w(m', n) = \frac{1}{N} \sum_{k=0}^{N-1} W(m', k) e^{j\frac{2\pi}{N}nk}, \quad n = 0, 1, \dots, N-1.$$



AEVD



- Utilizando os filtros adaptativos, estima-se as repostas ao impulso por meio da decomposição em autovalores.

- Seguindo a relação

$$(x_1 * h_2)(n) = (s * h_1 * h_2)(n) = (x_2 * h_1)(n)$$

- Que pode ser reescrita na forma

$$\mathbf{x}^T(n)\mathbf{u} = \mathbf{x}_1^T(n)\mathbf{h}_2 - \mathbf{x}_2^T(n)\mathbf{h}_1 = 0.$$

$$\mathbf{R}_{\mathbf{x}\mathbf{x}}(n)\mathbf{u} = \mathbf{0}.$$

onde

$$\begin{aligned}\mathbf{x}(n) &= [\mathbf{x}_1^T(n) \ \mathbf{x}_2^T(n)]^T \\ \mathbf{u} &= [\mathbf{h}_2^T \ -\mathbf{h}_1^T]^T\end{aligned}$$

e $\mathbf{R}_{\mathbf{x}\mathbf{x}}(n) = E\{\mathbf{x}(n)\mathbf{x}^T(n)\} \approx \mathbf{x}(n)\mathbf{x}^T(n)$



AEVD

- Aplicando novamente os mínimos quadrados, temos que

o sinal de erro $e(n) = \mathbf{u}^T(n)\mathbf{x}(n)$

o vetor \mathbf{u} $\mathbf{u}(n+1) = \mathbf{u}(n) - \mu_{\text{LMS}} e(n)\mathbf{x}(n).$

- A atualização dos coeficientes é semelhante ao FLMS:

$$U_i(m+1, k) = U_i(m, k) - \frac{\mu}{S_{x_i x_i}(m, k) + \varepsilon} X_i^*(m, k) E(m, k) \quad k = 0, 1, \dots, N-1$$

$$e(m, n) = \frac{1}{N} \sum_{k=0}^{N-1} \left[U_1(m, k) X_1(m, k) + U_2(m, k) X_2(m, k) \right] e^{j \frac{2\pi}{N} nk}, \quad n = 0, 1, \dots, N-1$$

- O delay é obtido buscando-se o pico negativo mais expressivo em \mathbf{u}_2 :

$$u_2(m', n) = \frac{1}{N} \sum_{k=0}^{N-1} U_2(m', k) e^{j \frac{2\pi}{N} nk}, \quad n = 0, 1, \dots, N-1.$$

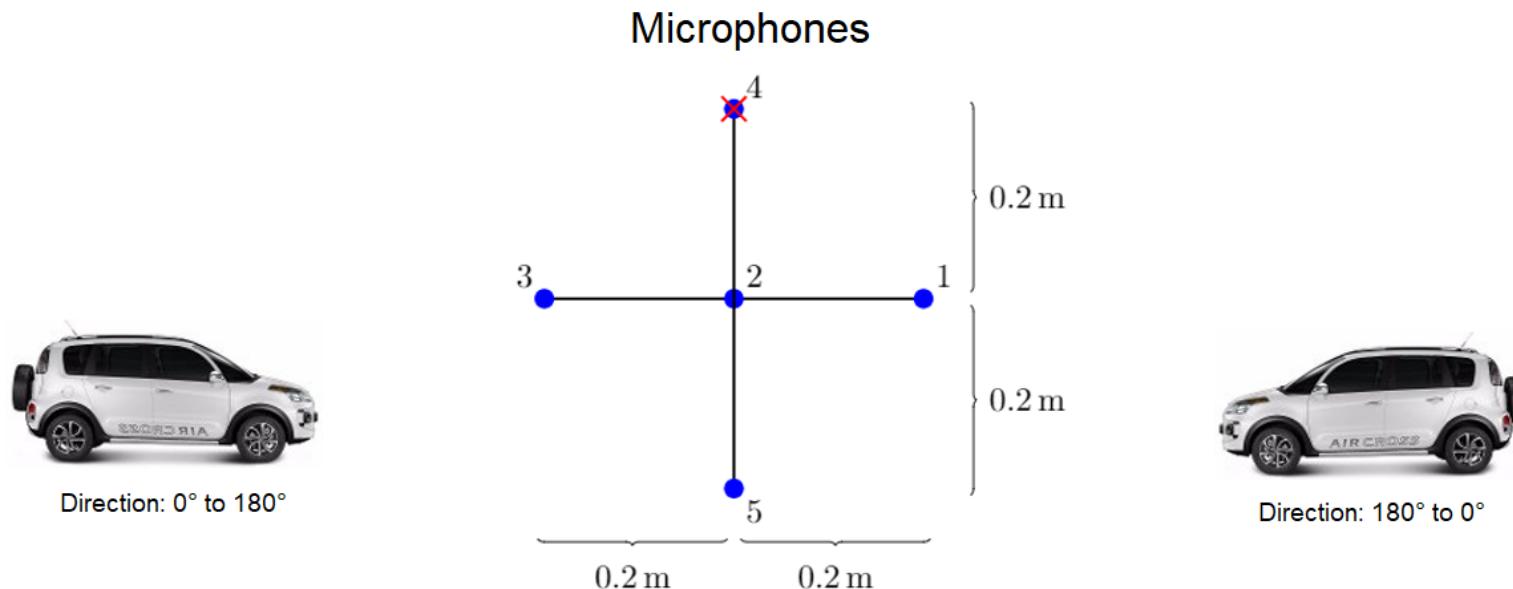


EXPERIMENTO

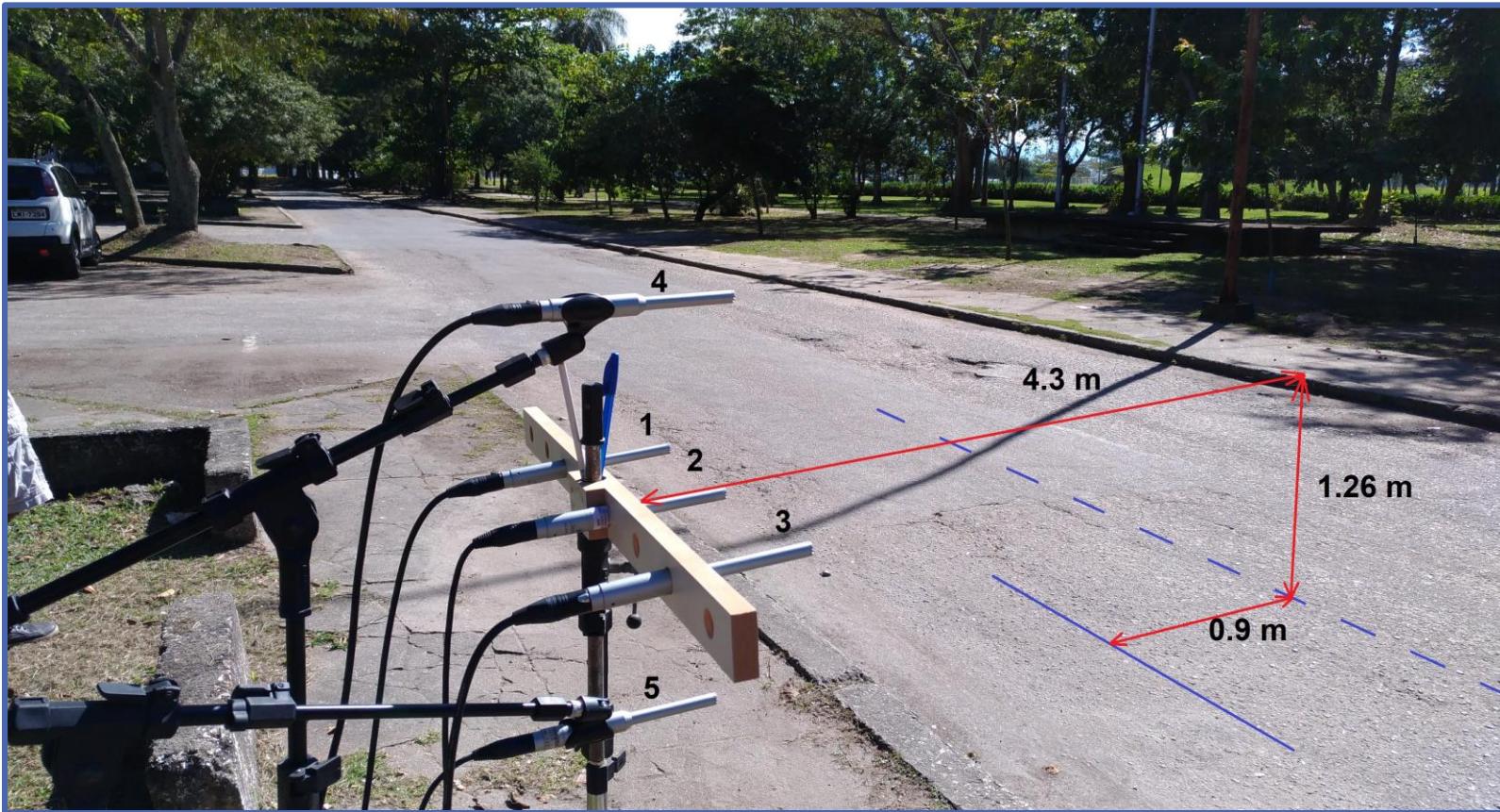


CONFIGURAÇÃO

- Um arranjo cruzado com 5 microfones foi utilizado, com 3 diferentes veículos passando (em configurações diversas):

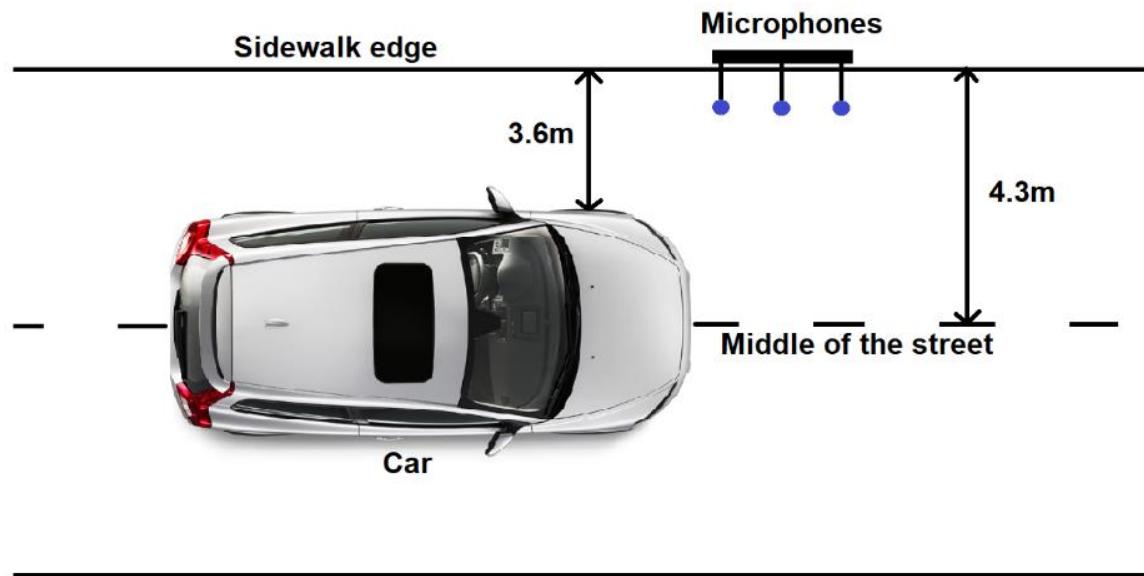


CONFIGURAÇÃO



CONFIGURAÇÃO

- Cada configuração foi realizada variando-se:
 - a quantidade de carros (1 a 3 modelos distintos);
 - a velocidade (30 a 70 Km/h);
 - a orientação inicial (0 ou 180 graus);
 - Espaçamento entre os carros (curto ou longo).



PARÂMETROS E CARACTERÍSTICAS

- Parâmetros utilizados em cada configuração do experimento:

| Test | Speed (Km/h) | Number of cars | ID | Spacing | Orientation |
|------|--------------|----------------|---------|---------|-------------|
| 1 | 30 | 2 | 1, 2 | Short | 180° to 0° |
| 2 | 30 | 2 | | Long | 0° to 180° |
| 3 | 40 | 2 | | Short | 180° to 0° |
| 4 | 40 | 2 | | Long | 0° to 180° |
| 5 | 50 | 2 | | Short | 180° to 0° |
| 6 | 50 | 2 | | Long | 0° to 180° |
| 7 | 60 | 2 | | Short | 180° to 0° |
| 8 | 60 | 3 | 1, 2, 3 | Short | 180° to 0° |
| 9 | 60 | 3 | | Long | 180° to 0° |
| 10 | 70 | 1 | 1 | ----- | 180° to 0° |



PARÂMETROS E CARACTERÍSTICAS

- Características dos veículos utilizados:

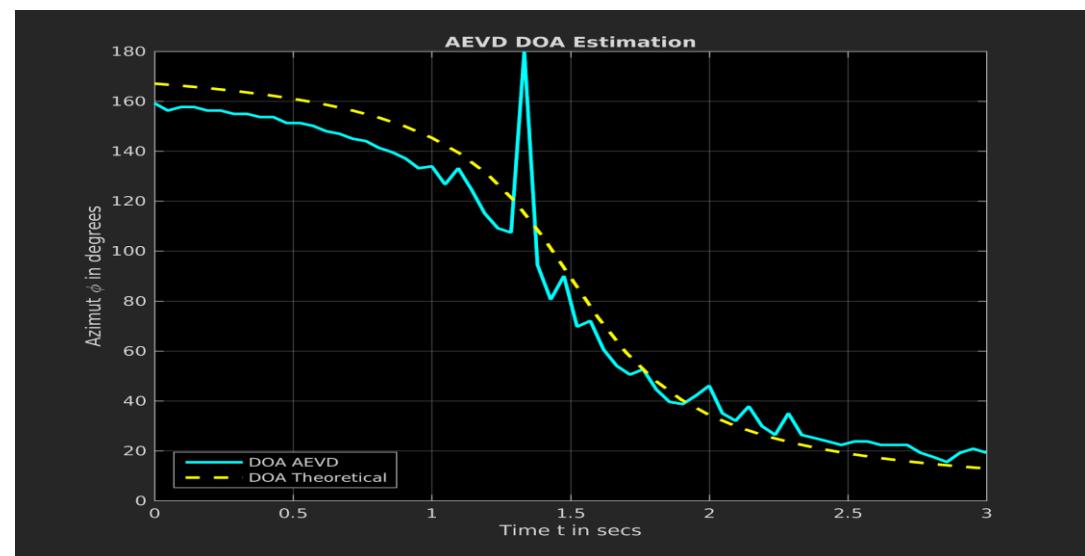
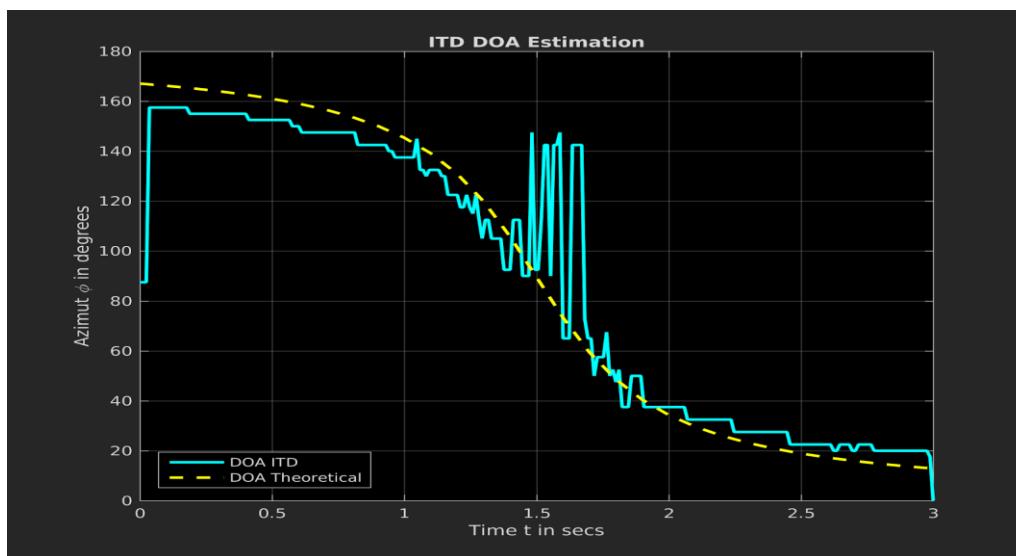
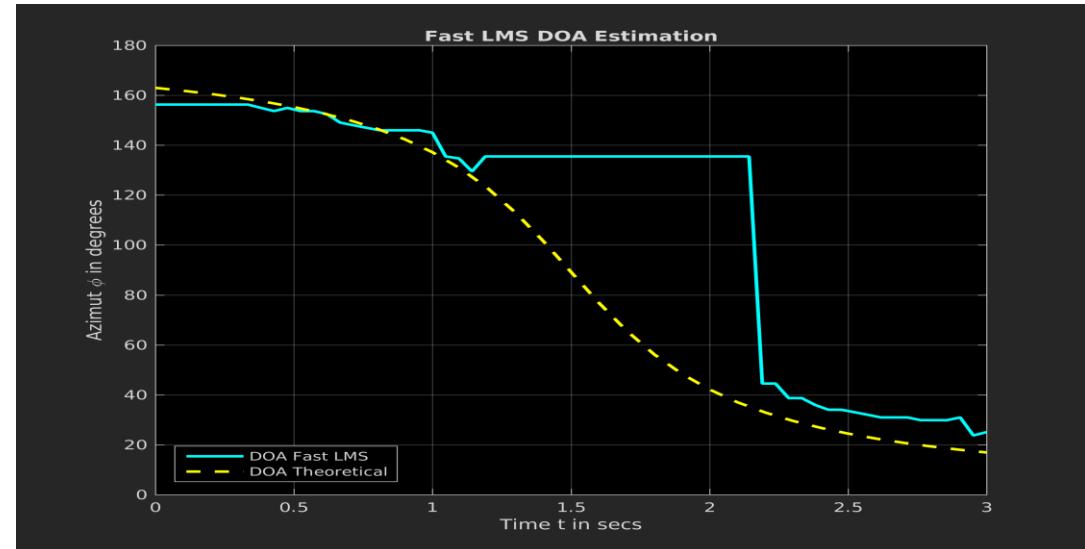
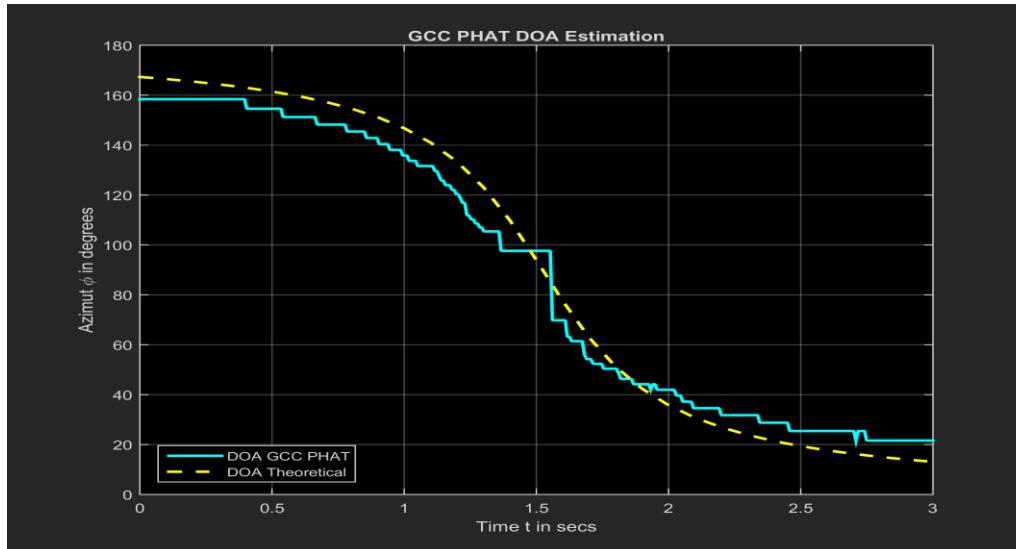
| Car | 1 | 2 | 3 |
|--------------------|-----------|-----------|--------|
| Wheels spacing (m) | 2.70 | 2.57 | 2.45 |
| Gearbox type | Automatic | Automatic | Manual |
| Engine Height | 0.75 | 0.90 | 0.60 |



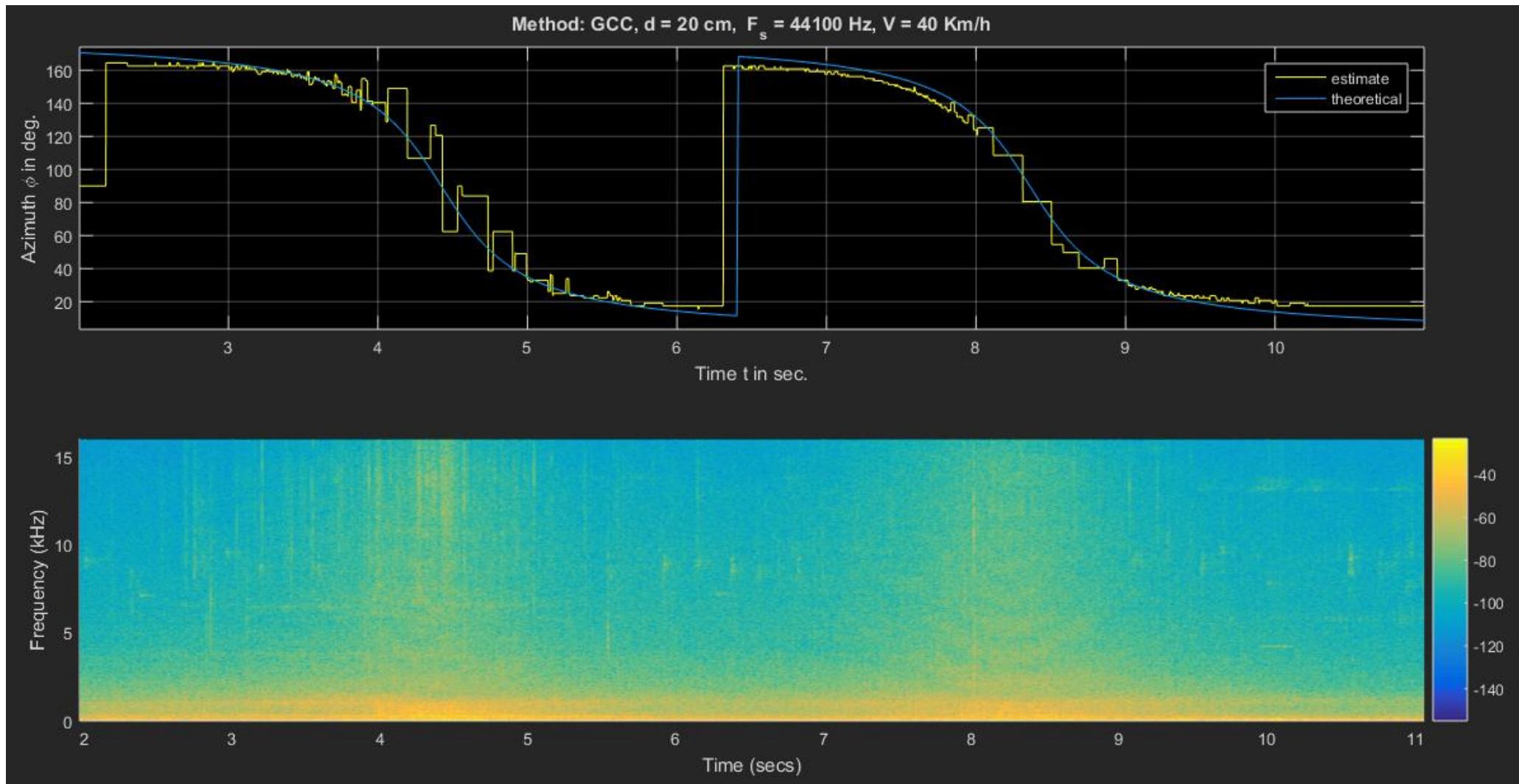


ANÁLISE

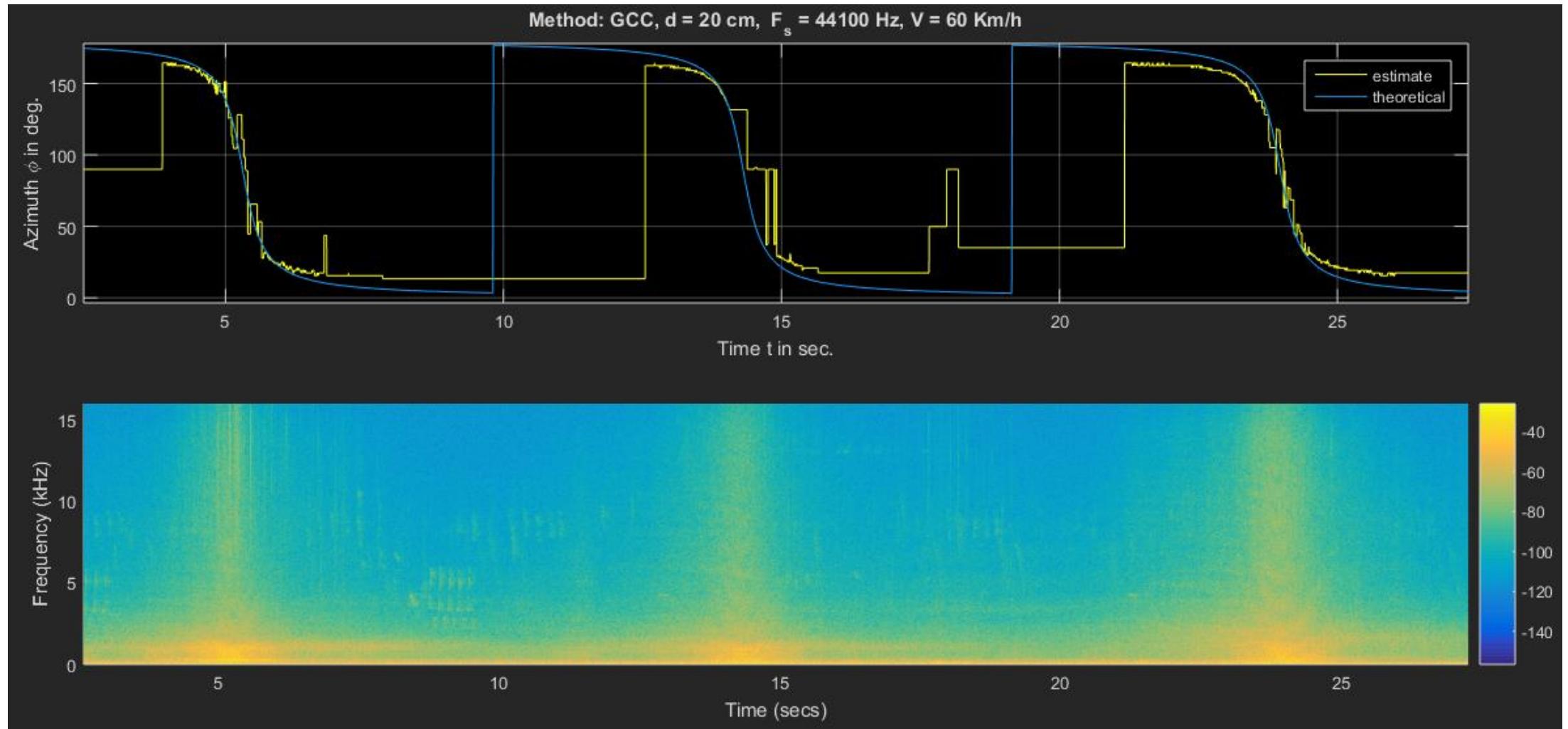
COMPARANDO OS ALGORITMOS



COMPARANDO CARROS DIFERENTES

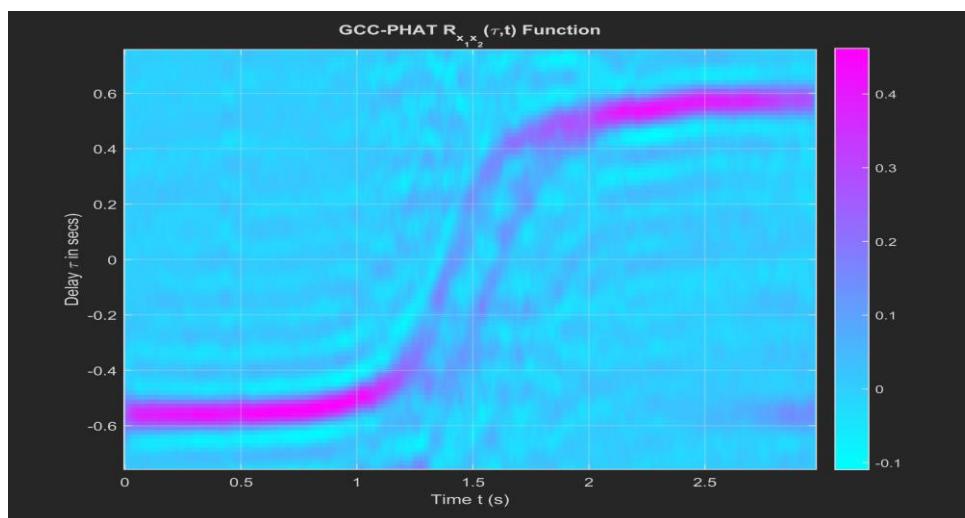
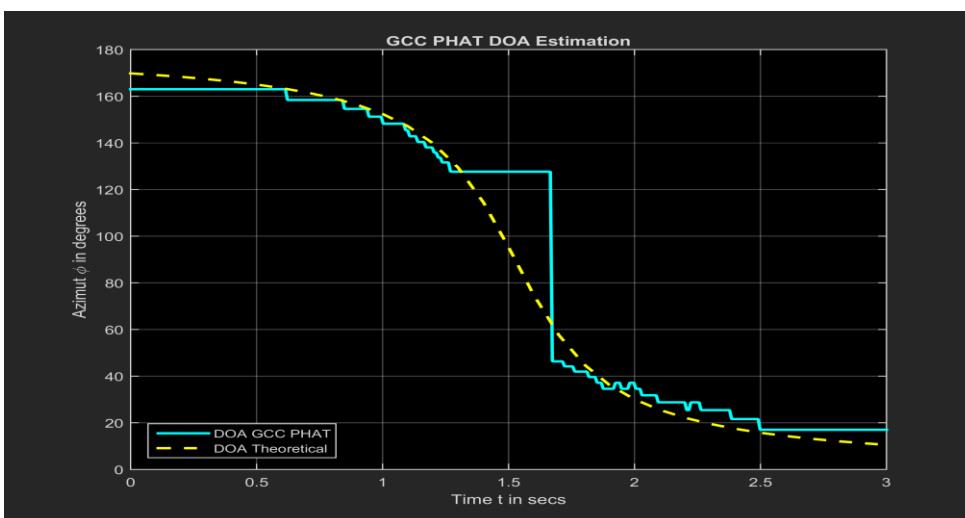
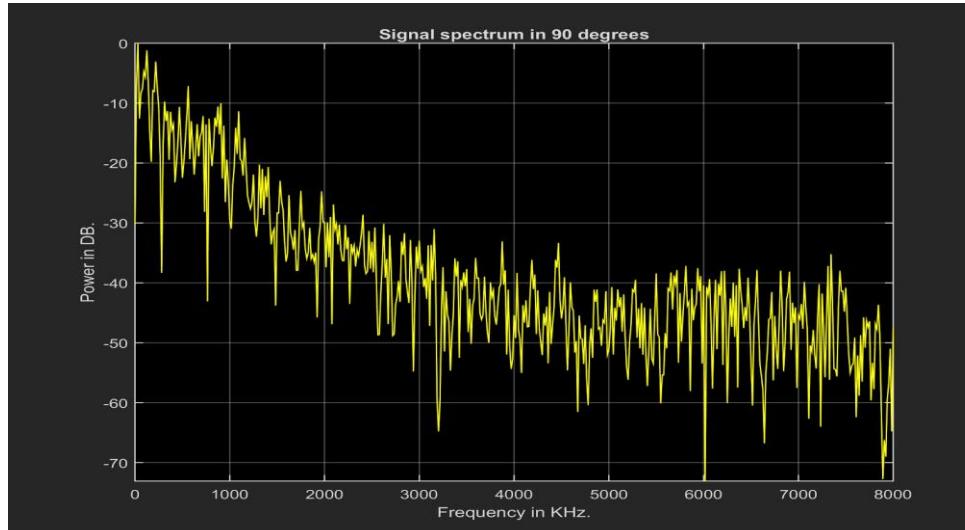
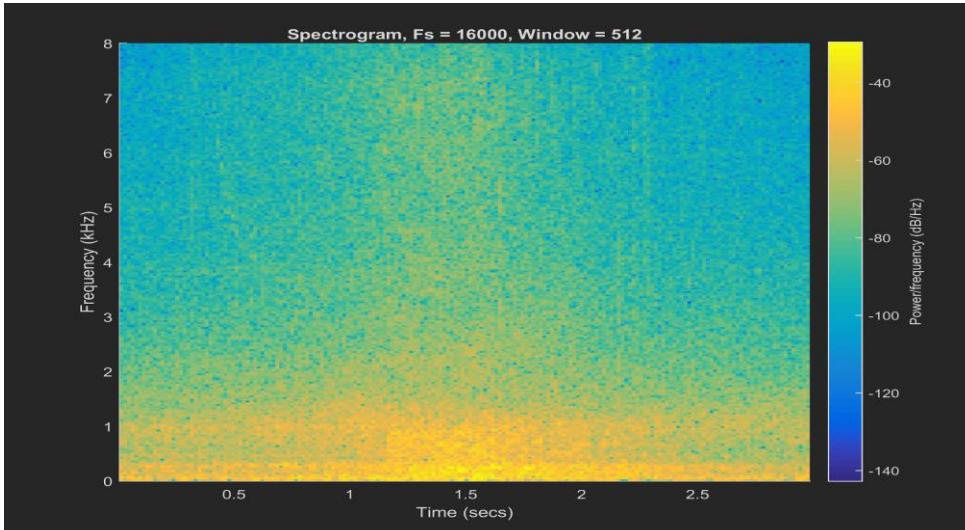


COMPARANDO CARROS DIFERENTES



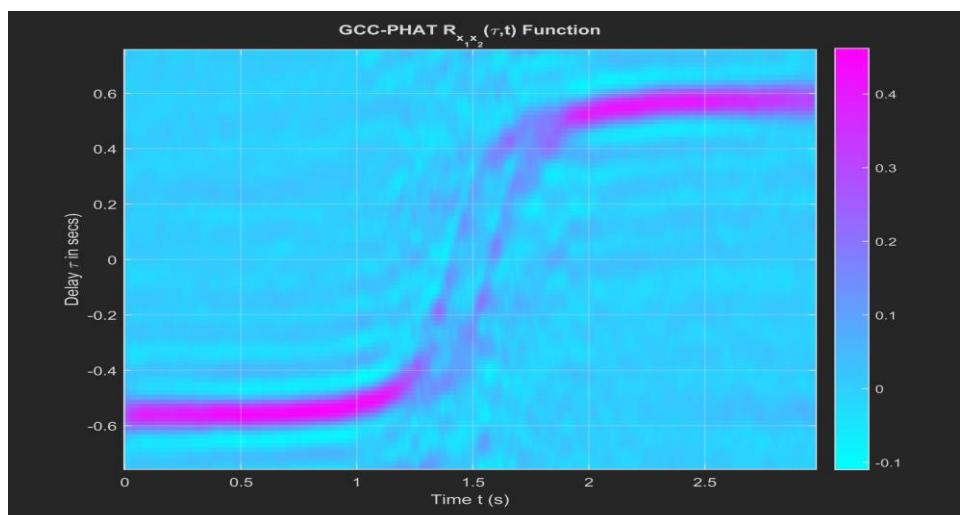
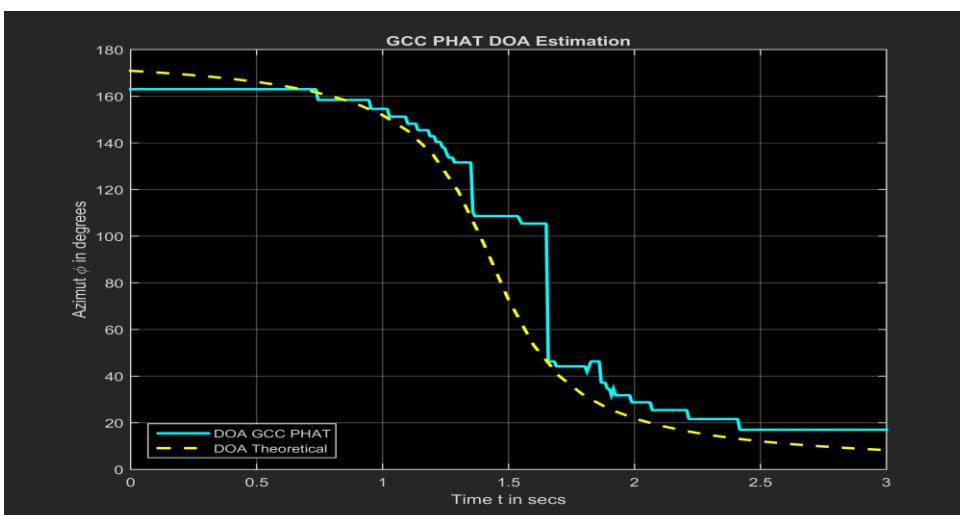
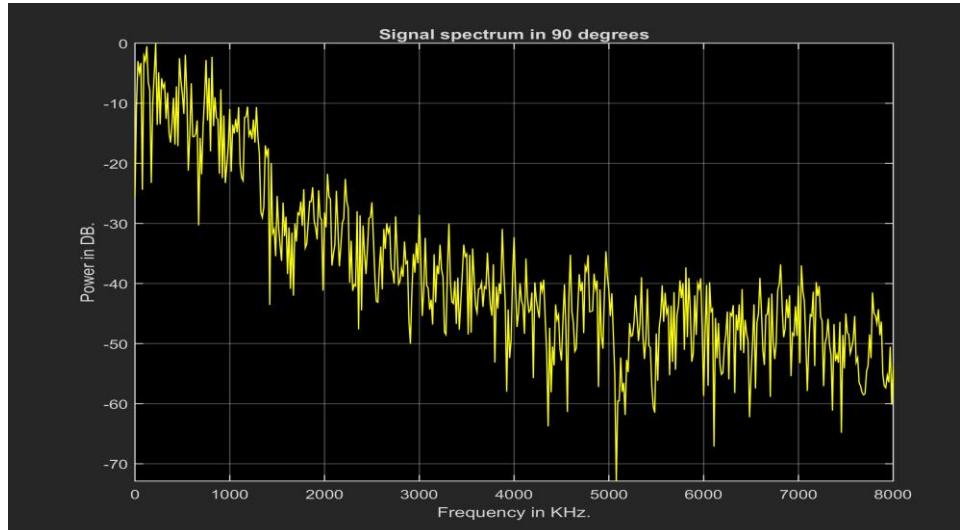
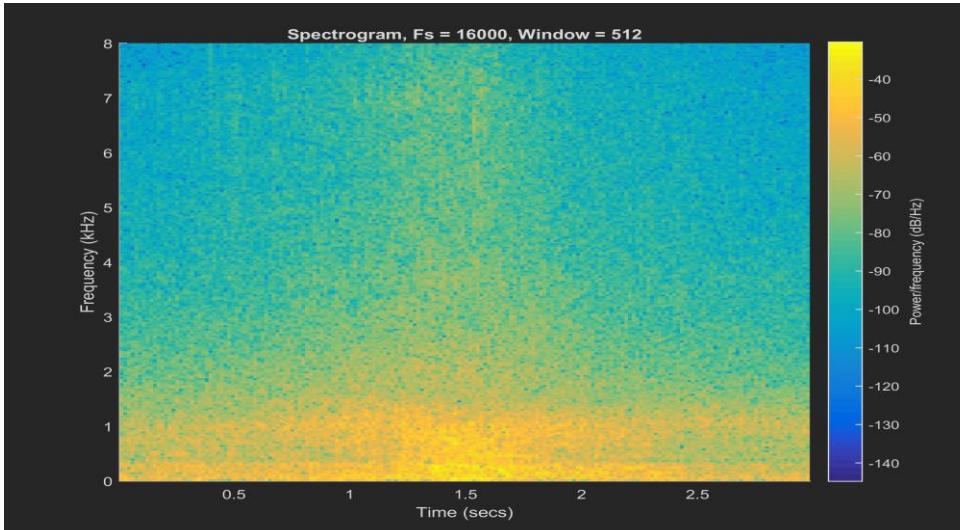
COMPARANDO VELOCIDADES DIFERENTES

- Carro 1: 50 Km/h – 180 a 0°



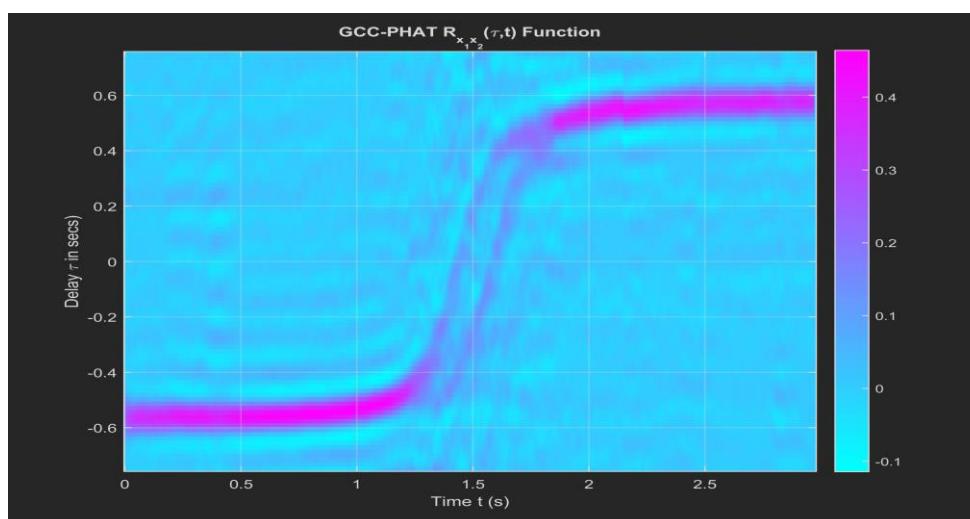
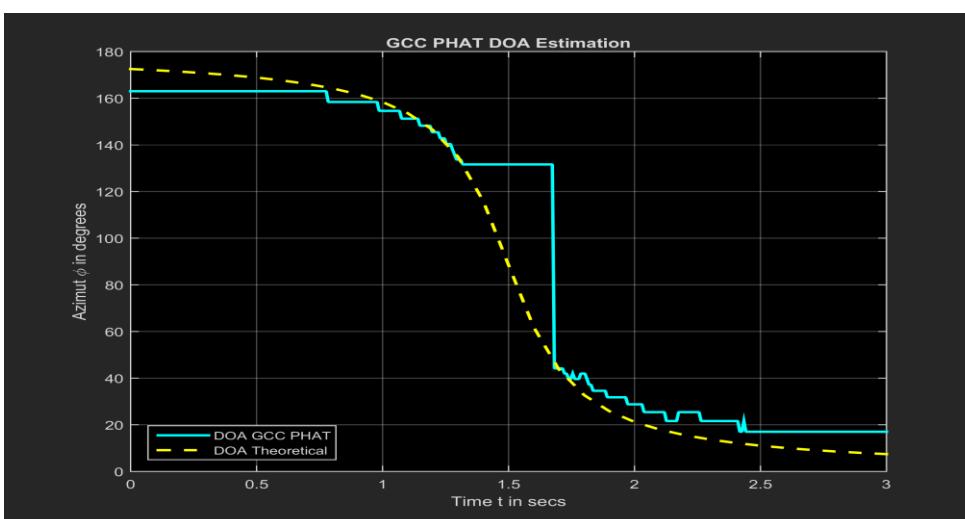
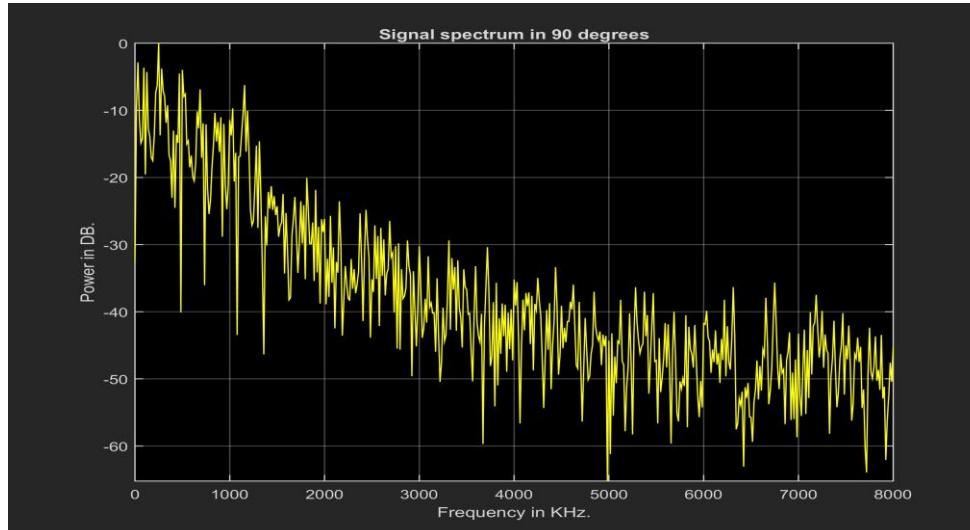
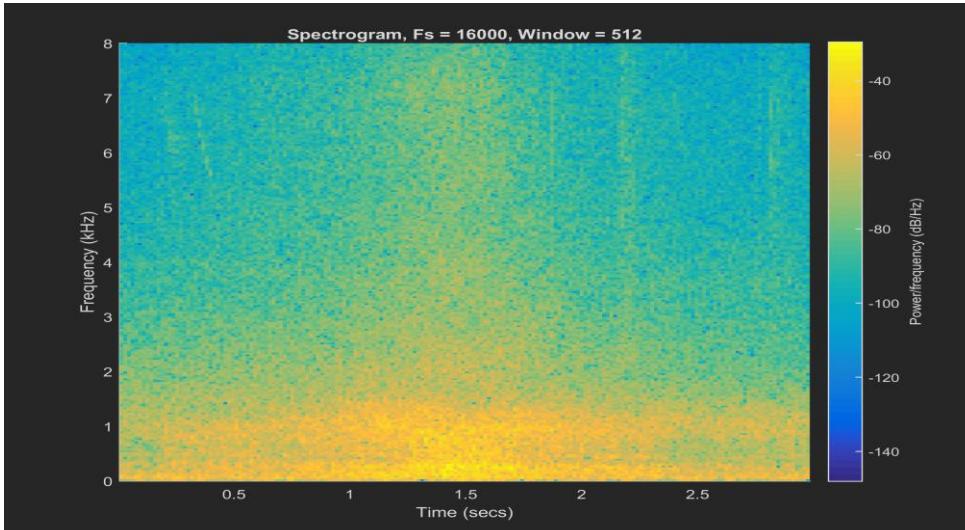
COMPARANDO VELOCIDADES DIFERENTES

- Carro 1: 60 Km/h – 180 a 0°



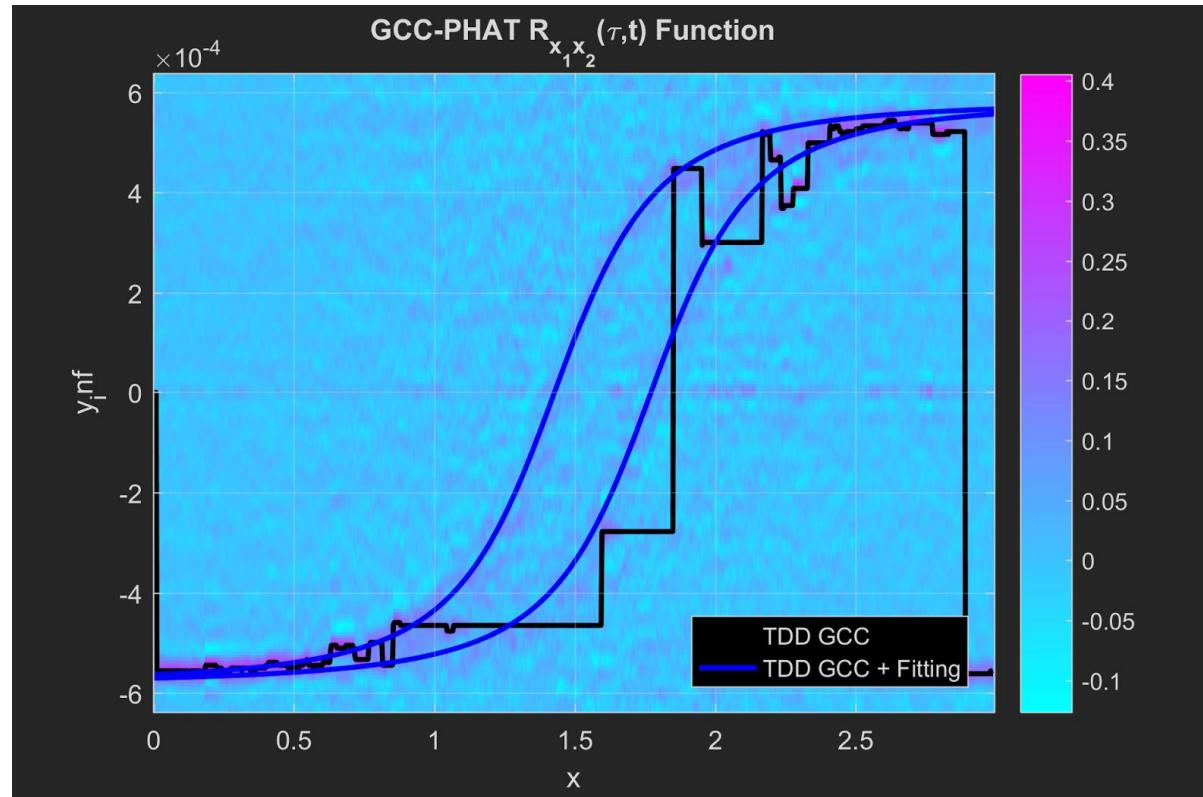
COMPARANDO VELOCIDADES DIFERENTES

- Carro 1: 70 Km/h – 180 a 0°



DESEMPENHO DO GCC-PHAT

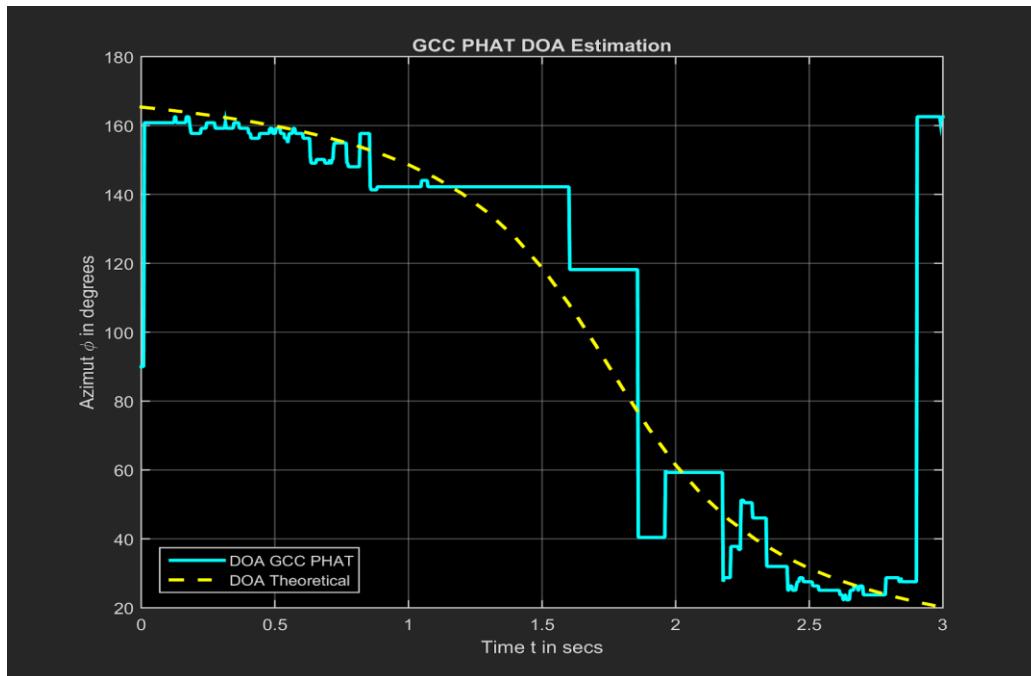
- De maneira geral, o GCC-PHAT apresentou o melhor desempenho, sendo entretanto, insuficiente para distinguir diferentes tipos de fontes de som inerentes a cada veículo:



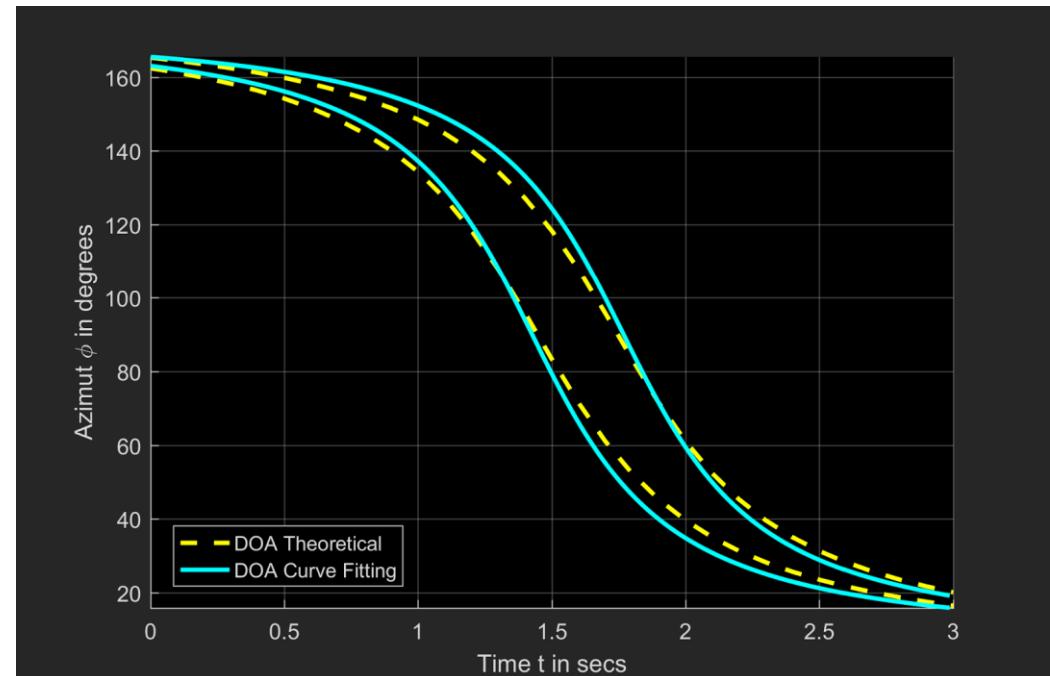
REALIZANDO O AJUSTE DAS CURVAS

- Velocidade: 30 Km/h

GCC-PHAT



GCC-PHAT + Curve Fitting



PRÓXIMOS OBJETIVOS

- A seguir utilizaremos as estimativas de DOA em conjunto com o *Beamforming*, dando preferência às direções de interesse e suprindo as fontes indesejadas;
- Também será importante extrair os dados dos microfones verticais;
- A partir disso será possível avaliar com mais detalhe as principais fontes sonoras presentes em cada veículo (motor, rolamento, etc).



REFERÊNCIAS

- Hänsler, E., Schmidt, G.. Topics in Acoustic Echo and Noise Control: Selected Methods for the Cancellation of Acoustical Echoes, the Reduction of Background Noise, and Speech Processing . Germany: Springer, 2006. 642 p. (Signals and Communication Technology)



OBRIGADO!

