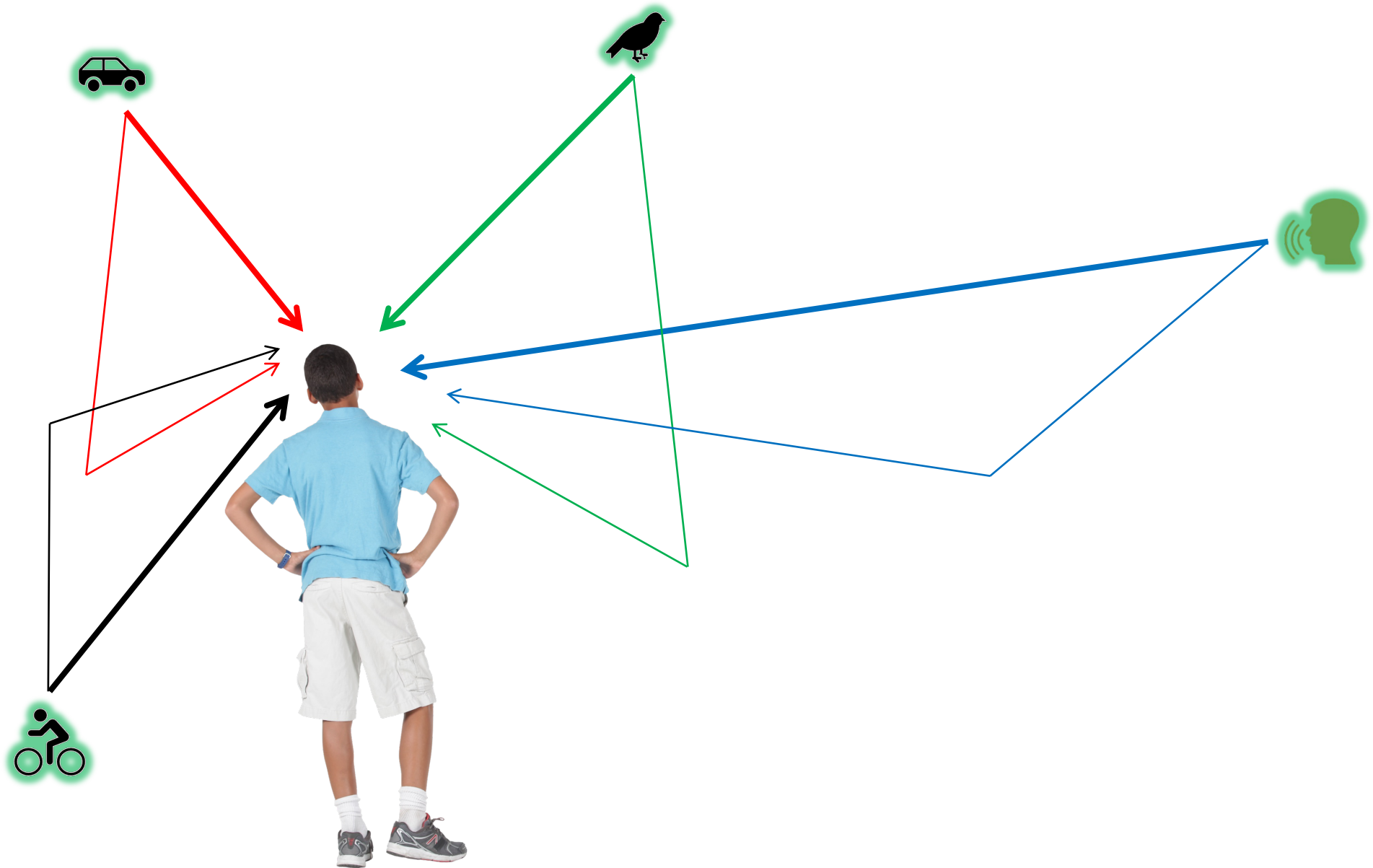
Source separation

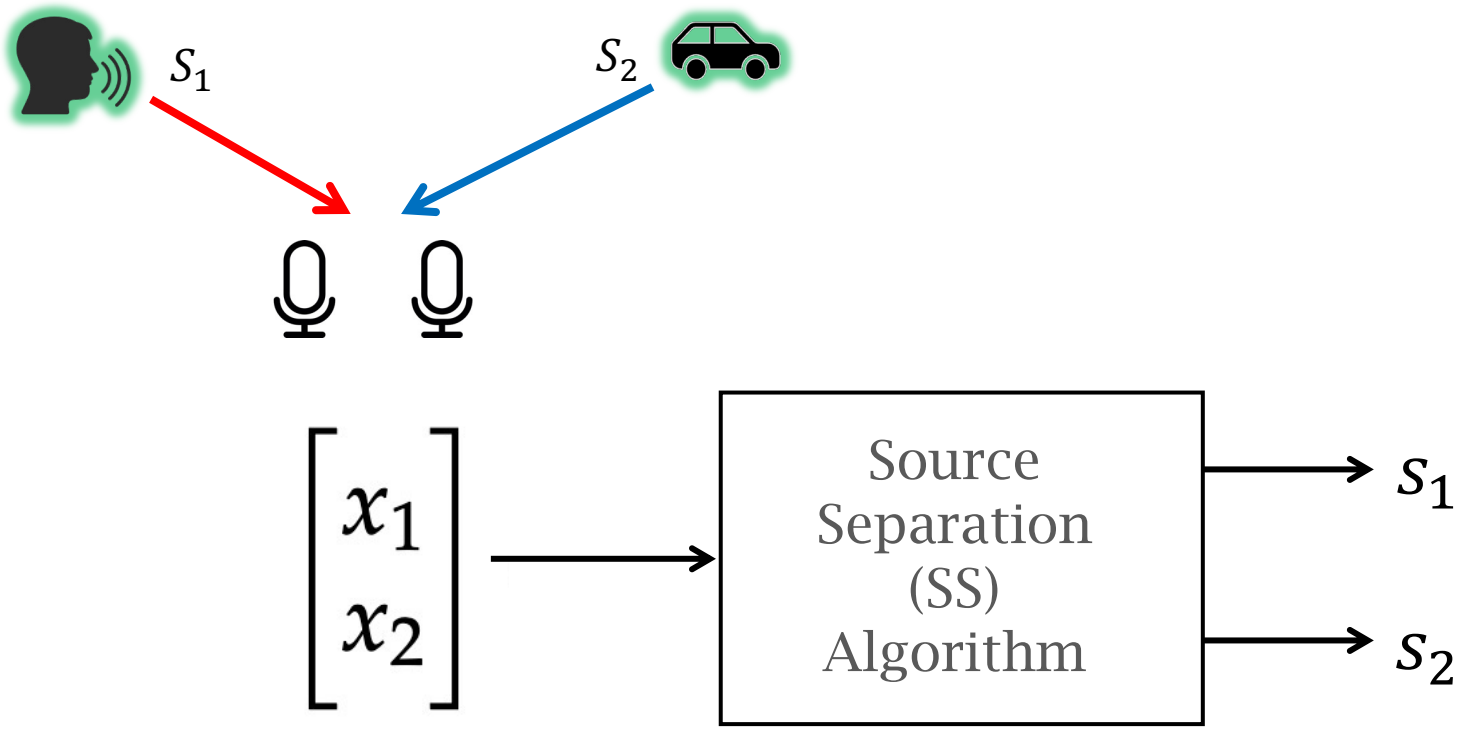
Source separation



Source separation

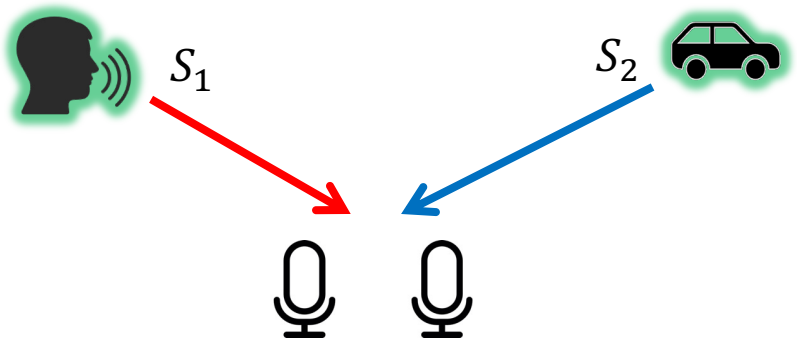
Source separation preliminaries

Source separation: The general problem statement



Source separation preliminaries

Source separation: The general problem statement

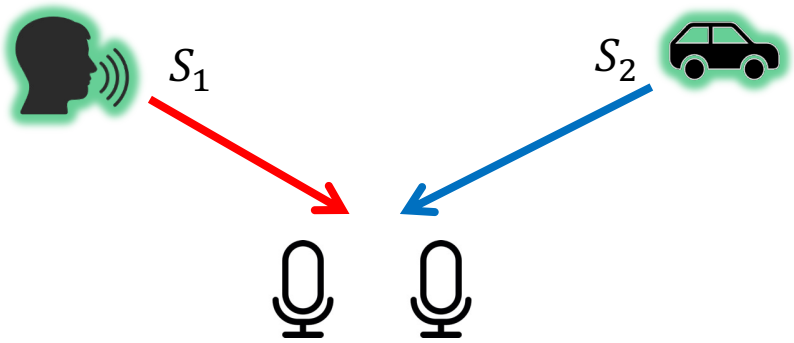


$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}}_{\text{Mixing matrix}} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

Unknown mixing matrix, unknown source signals → heavily under-determined

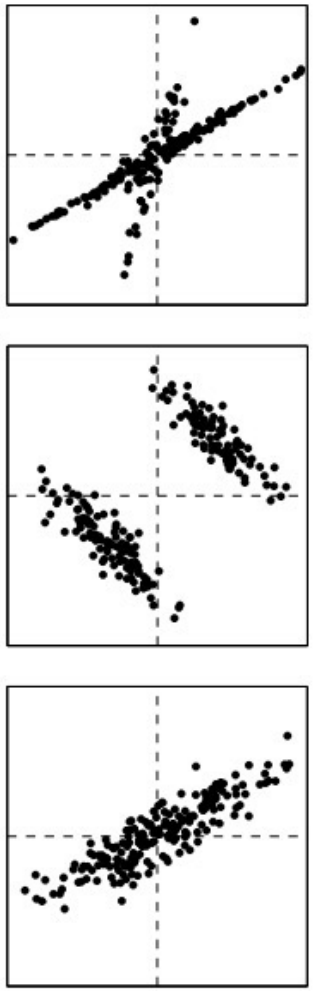
Source separation preliminaries

Source separation: The general problem statement



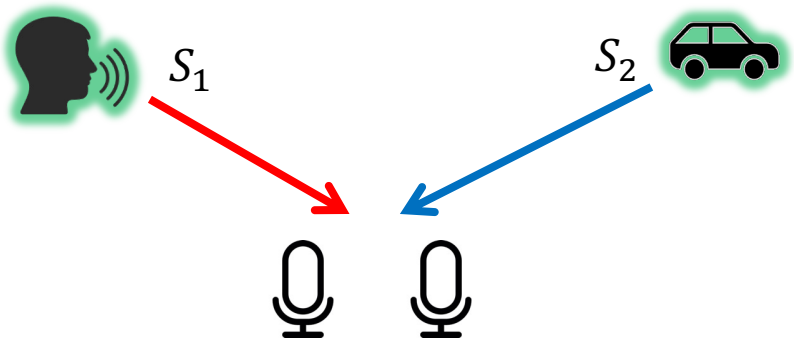
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}}_{\text{Mixing matrix}} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} =$$



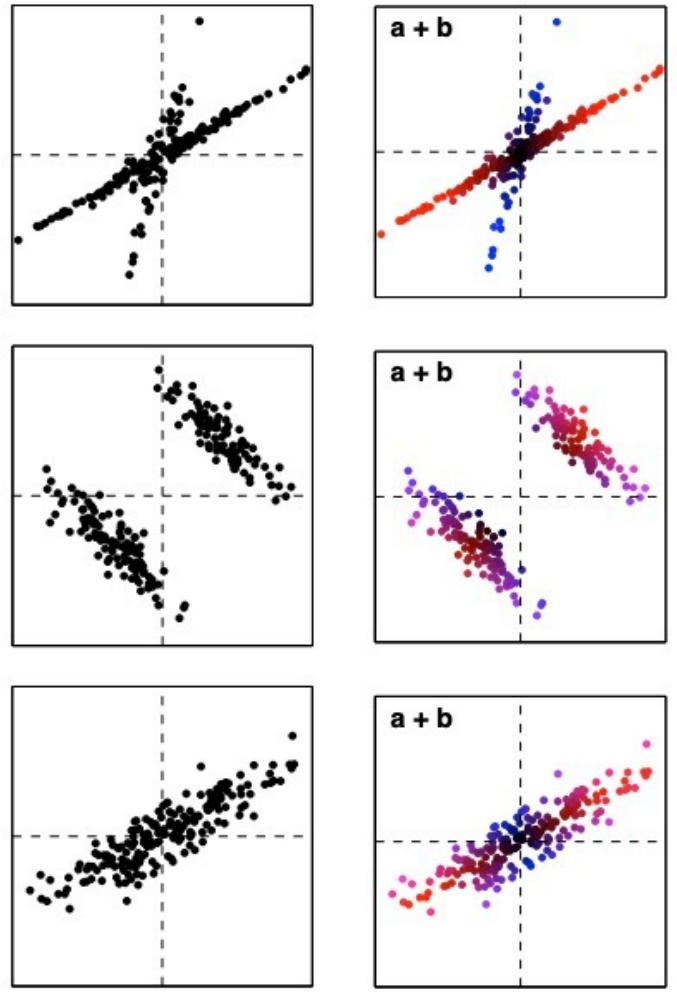
Source separation preliminaries

Source separation: The general problem statement



$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}}_{\text{Mixing matrix}} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} =$$

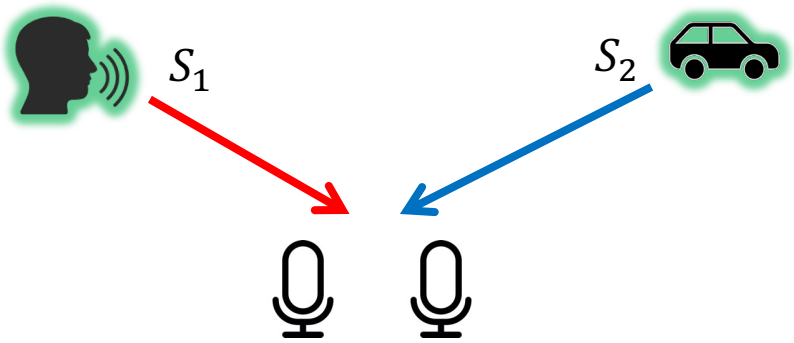


Hard to separate the sources even visually

When can we solve SS?

Let's make some simplifications: Mixing matrix is known based on **Angle of Arrival (AOA)**

- What is angle of arrival (AoA)? How do you quantify it?
- Relation between AoA and FFT
- How do you get AoA? From camera or from audio itself?
- How to solve X = A S + N even when A is known

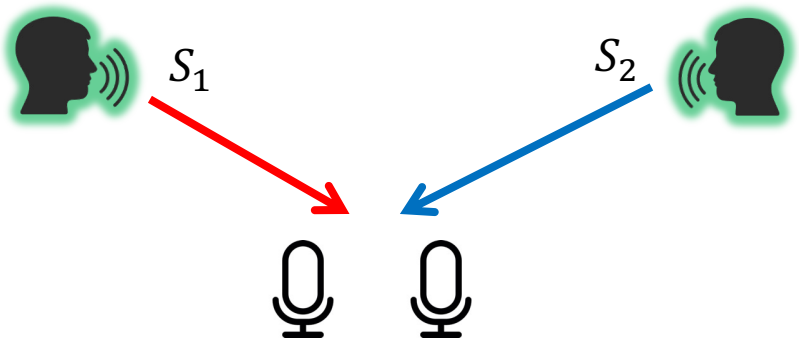


$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} | & | \\ \vec{a}_{\theta_1} & \vec{a}_{\theta_2} \\ | & | \end{bmatrix}}_{\text{AoA matrix A}} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

$$X = A.S + N$$

Source = Speech

But what if AoA unknown? **It's hard to solve for S … but what if S is speech signals?**



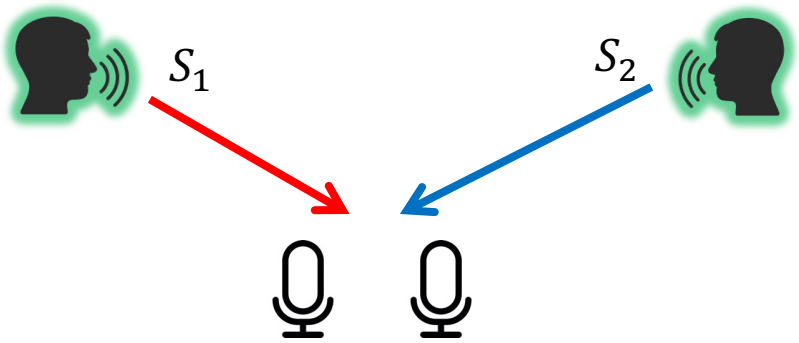
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} | & | \\ \vec{a}_{\theta_1} & \vec{a}_{\theta_2} \\ | & | \end{bmatrix}}_{\text{AoA matrix A}} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

$$X = A.S + N$$

Source = Speech

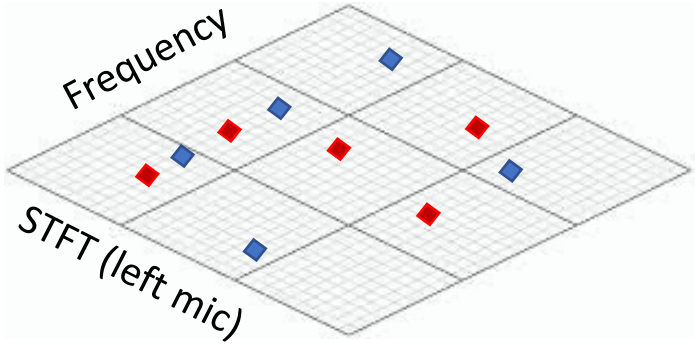
But what if AoA unknown? **It's hard to solve for S … but what if S is speech signals?**

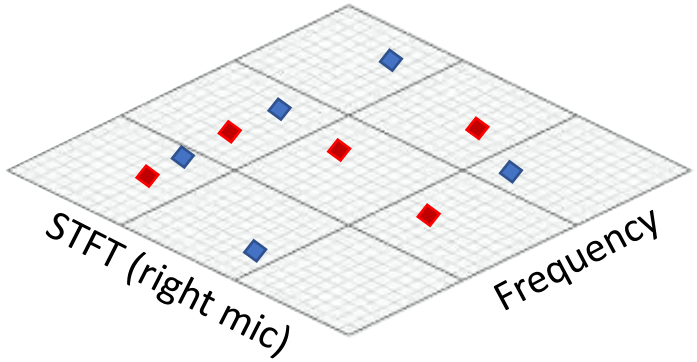
When the source signal is speech, exploit TF-disjointness



$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} | & | \\ \vec{a}_{\theta_1} & \vec{a}_{\theta_2} \\ | & | \end{bmatrix}}_{\text{AoA matrix A}} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

$$X = A.S + N$$

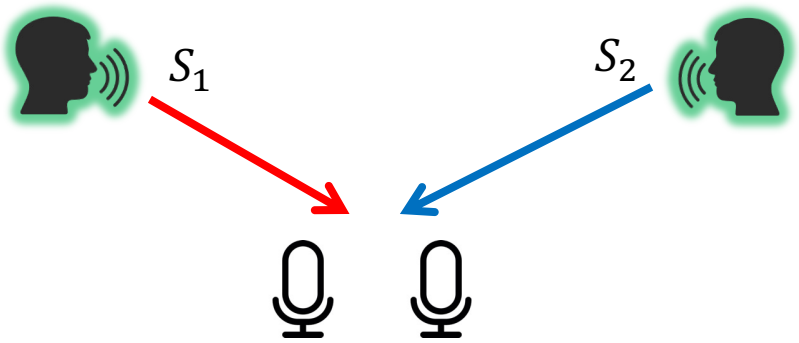




Source = Speech

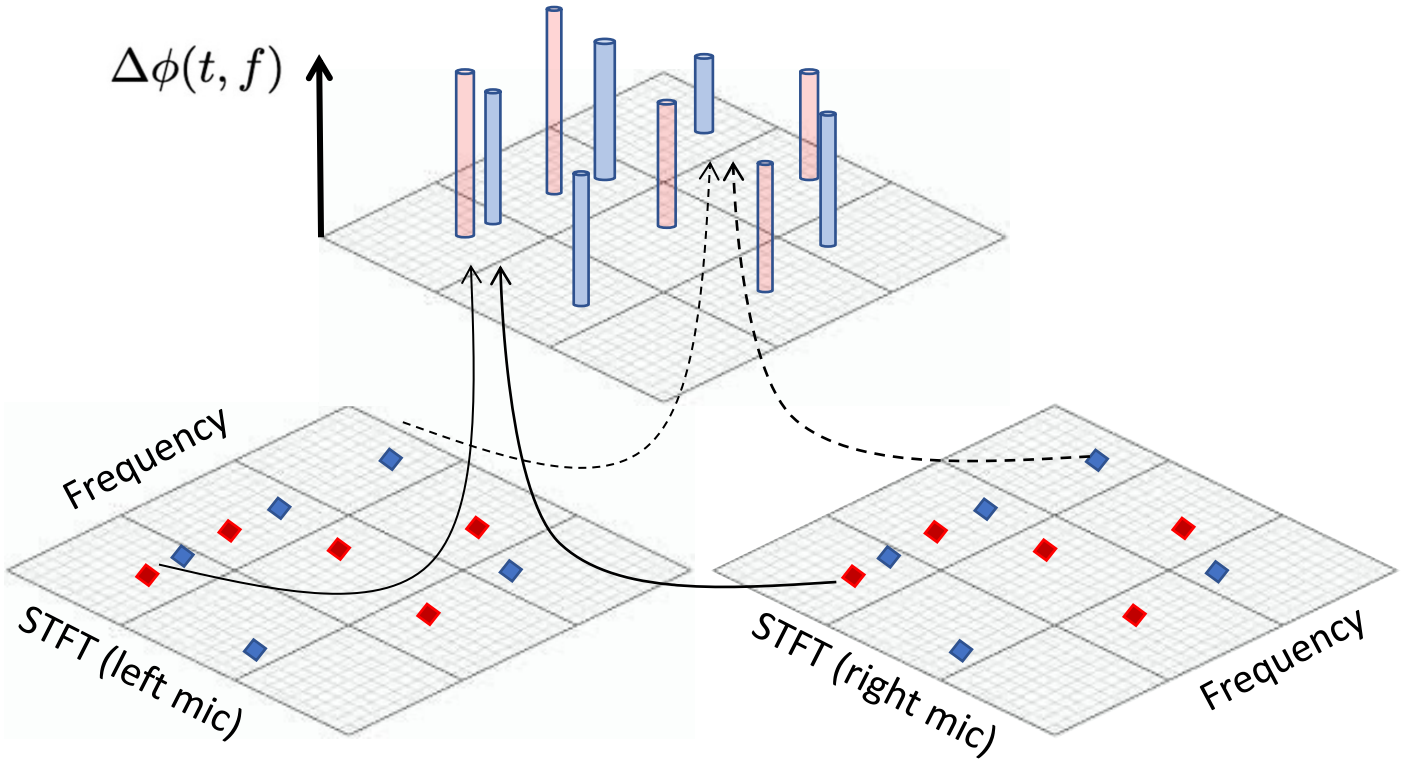
But what if AoA unknown? **It's hard to solve for S … but what if S is speech signals?**

When the source signal is speech, exploit TF-disjointness



$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} | & | \\ \vec{a}_{\theta_1} & \vec{a}_{\theta_2} \\ | & | \end{bmatrix}}_{\text{AoA matrix A}} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

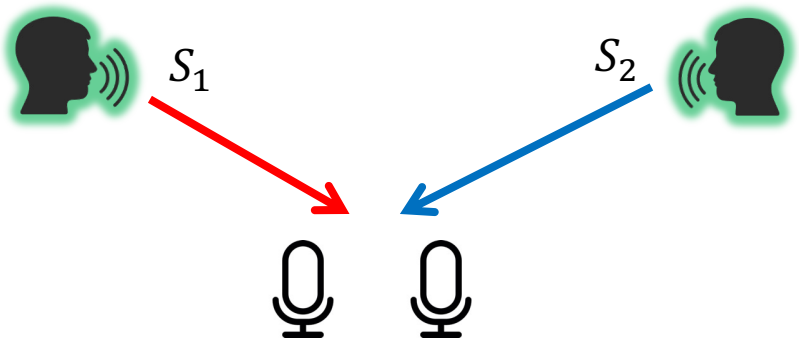
$$X = A.S + N$$



Source = Speech

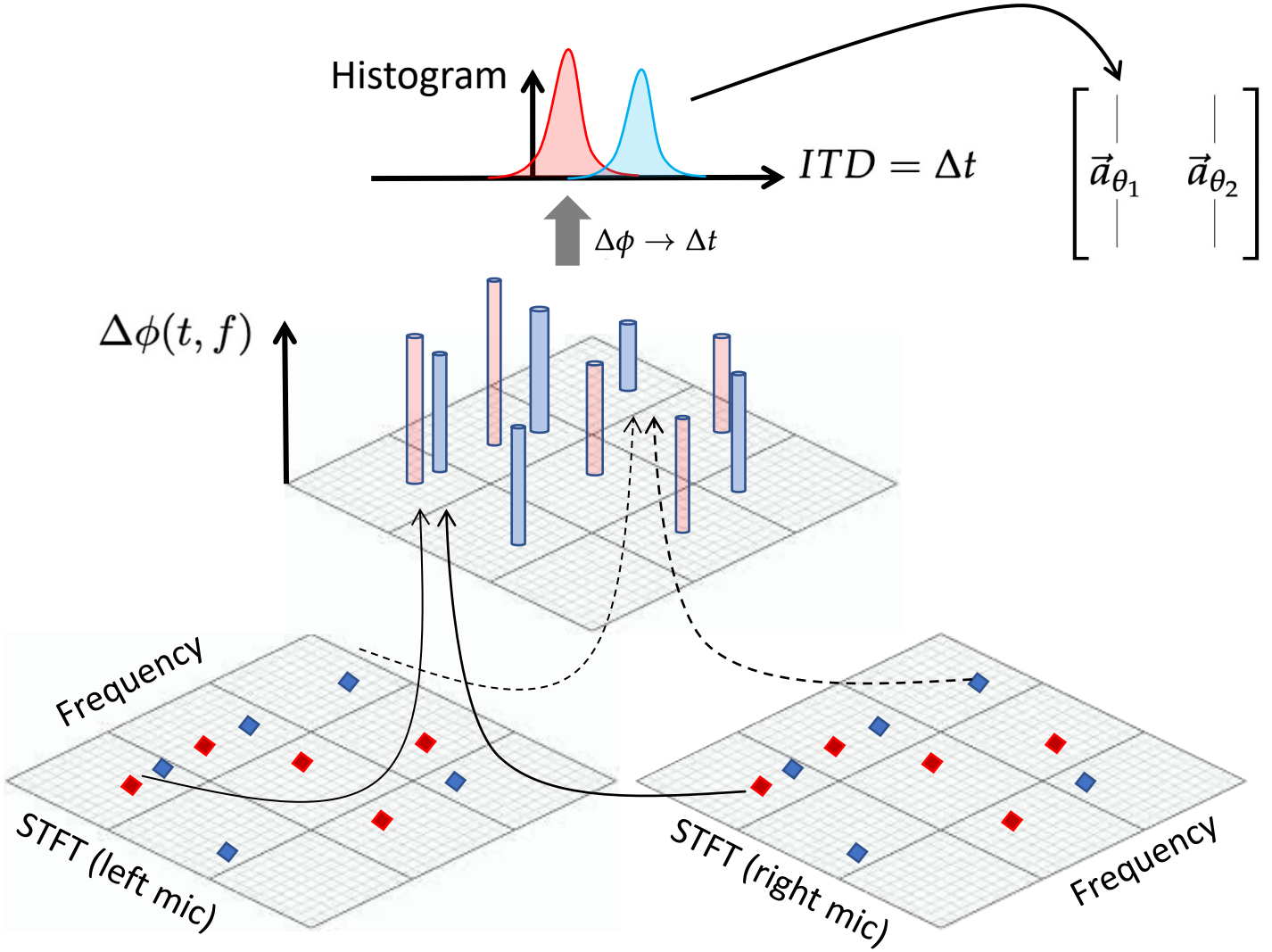
But what if AoA unknown? **It's hard to solve for S … but what if S is speech signals?**

When the source signal is speech, exploit TF-disjointness



$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} | & | \\ \vec{a}_{\theta_1} & \vec{a}_{\theta_2} \\ | & | \end{bmatrix}}_{\text{AoA matrix A}} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

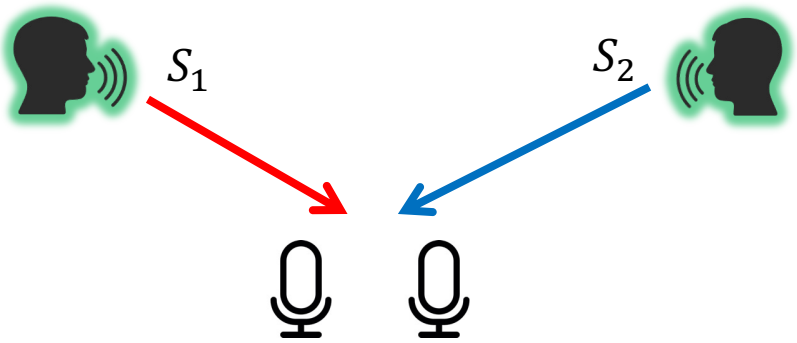
$$X = A.S + N$$

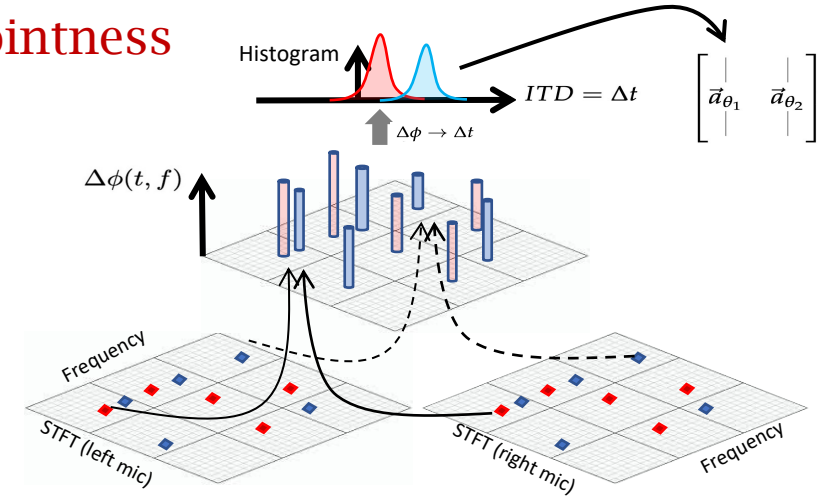


Source = Speech

But what if AoA unknown? **It's hard to solve for S … but what if S is speech signals?**

When the source signal is speech, exploit TF-disjointness





$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} | & | \\ \vec{a}_{\theta_1} & \vec{a}_{\theta_2} \\ | & | \end{bmatrix}}_{\text{AoA matrix A}} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

$$X = A.S + N$$

DUET Algorithm Steps:

1.

2.

3.

4.

5.