

# Distributed speech separation in spatially unconstrained microphone arrays

Nicolas Furnon<sup>1</sup>, Romain Serizel<sup>1</sup>, Irina Illina<sup>1</sup>, Slim Essid<sup>2</sup>

<sup>1</sup>Université de Lorraine, CNRS, Inria, Loria, F-54000 Nancy, France

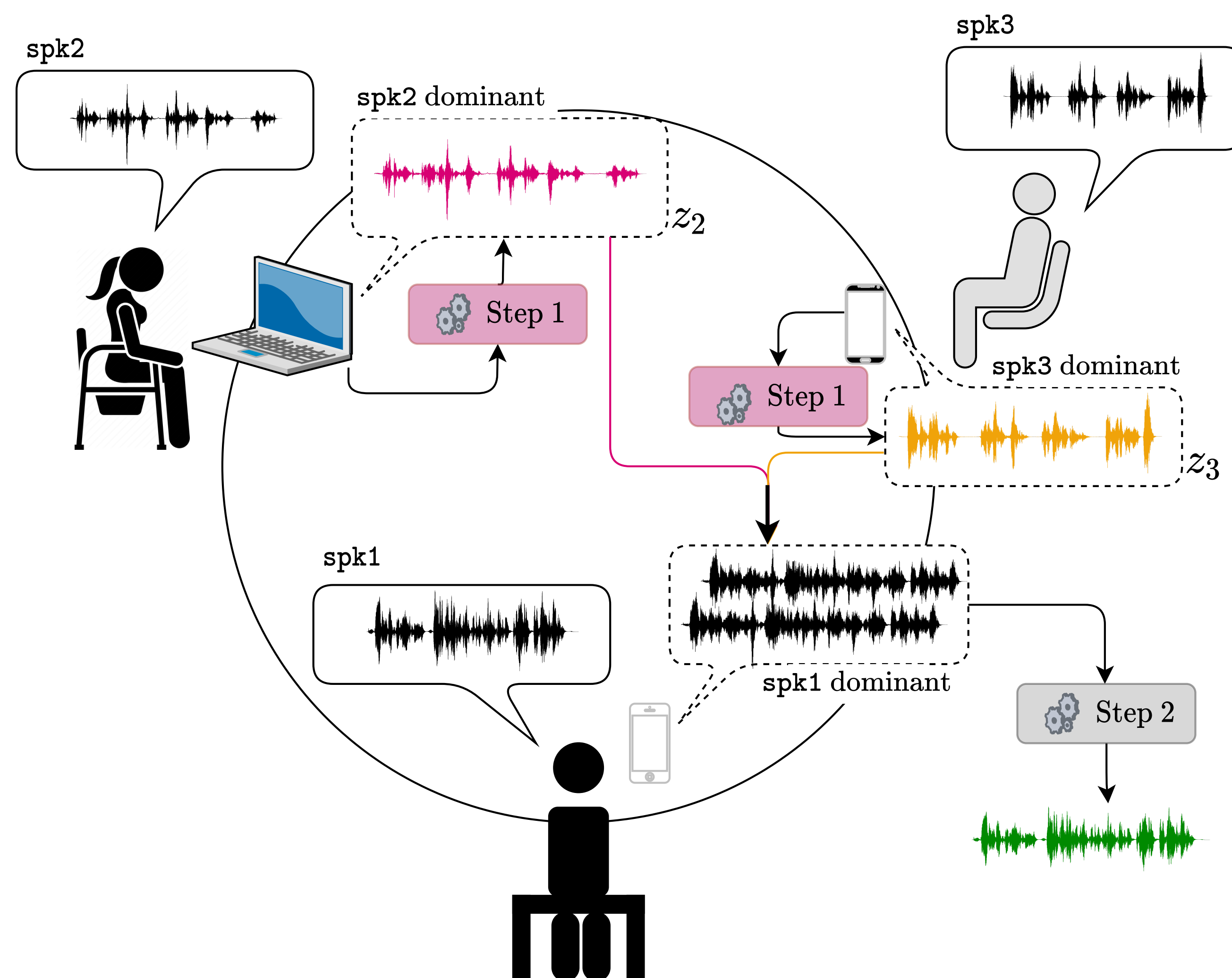
<sup>2</sup>LTCI, Télécom Paris, Institut Polytechnique de Paris, Palaiseau, France

## Motivation

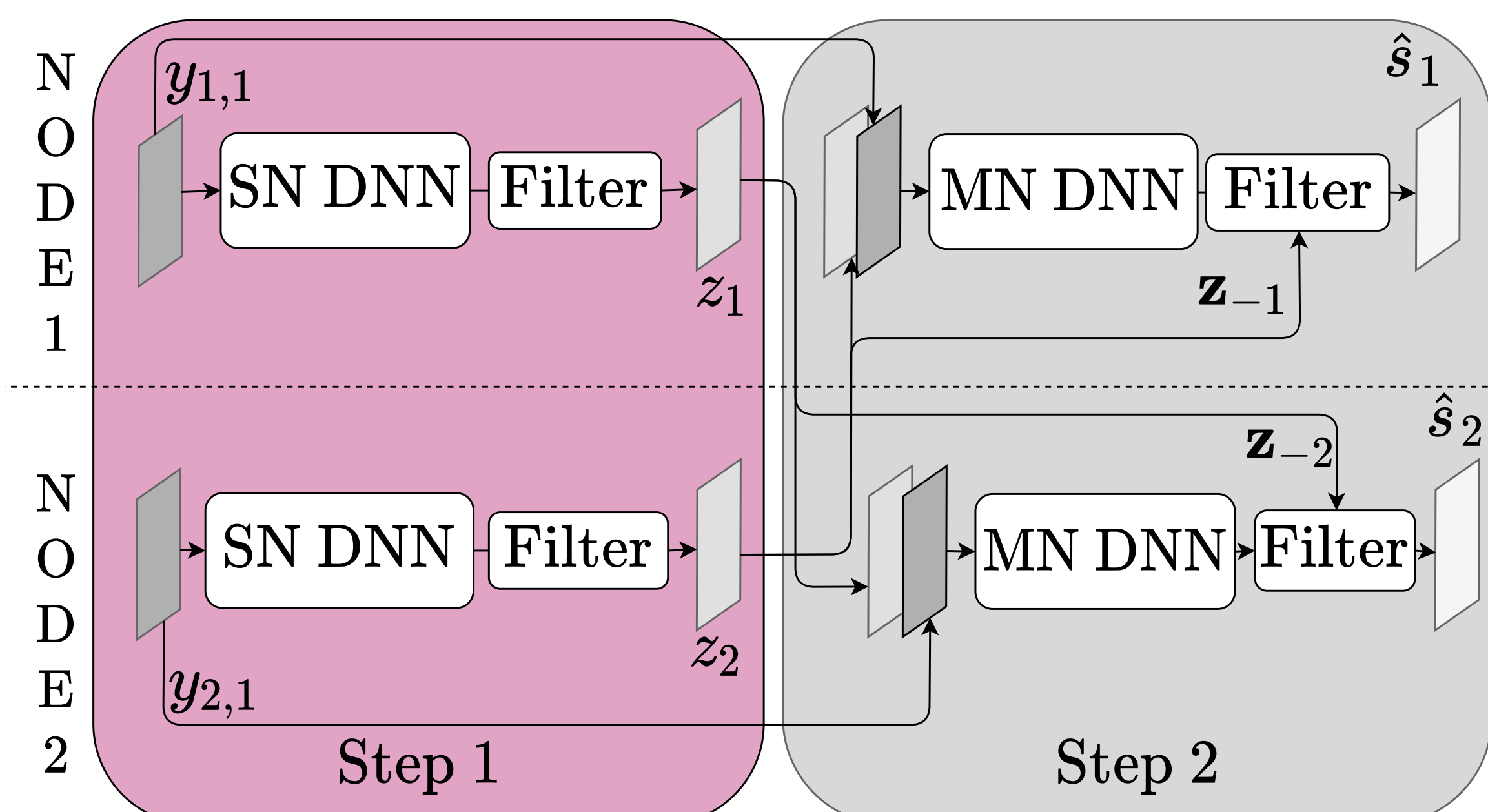
- **Context:** Source separation in meeting scenarios
- **Objective:** Use the already available microphones (smartphones, computers, ...)
- **Constraint:** Limited computational power → Distributed processing
- **Solution:** Tango, a distributed processing for source separation

## Proposed solution

Schematic representation focusing on node 1:



Block diagram of Tango with two nodes:



## Experimental setup

1 node = 4 microphones

$N = \{2, 3, 4\}$  sources

$K = \{2, 3, 4\}$  nodes

$0.3 \text{ m} \leq r \leq 2.5 \text{ m}$

$3 \text{ m} \leq w \leq 7 \text{ m}$

$3 \text{ m} \leq l \leq 9 \text{ m}$

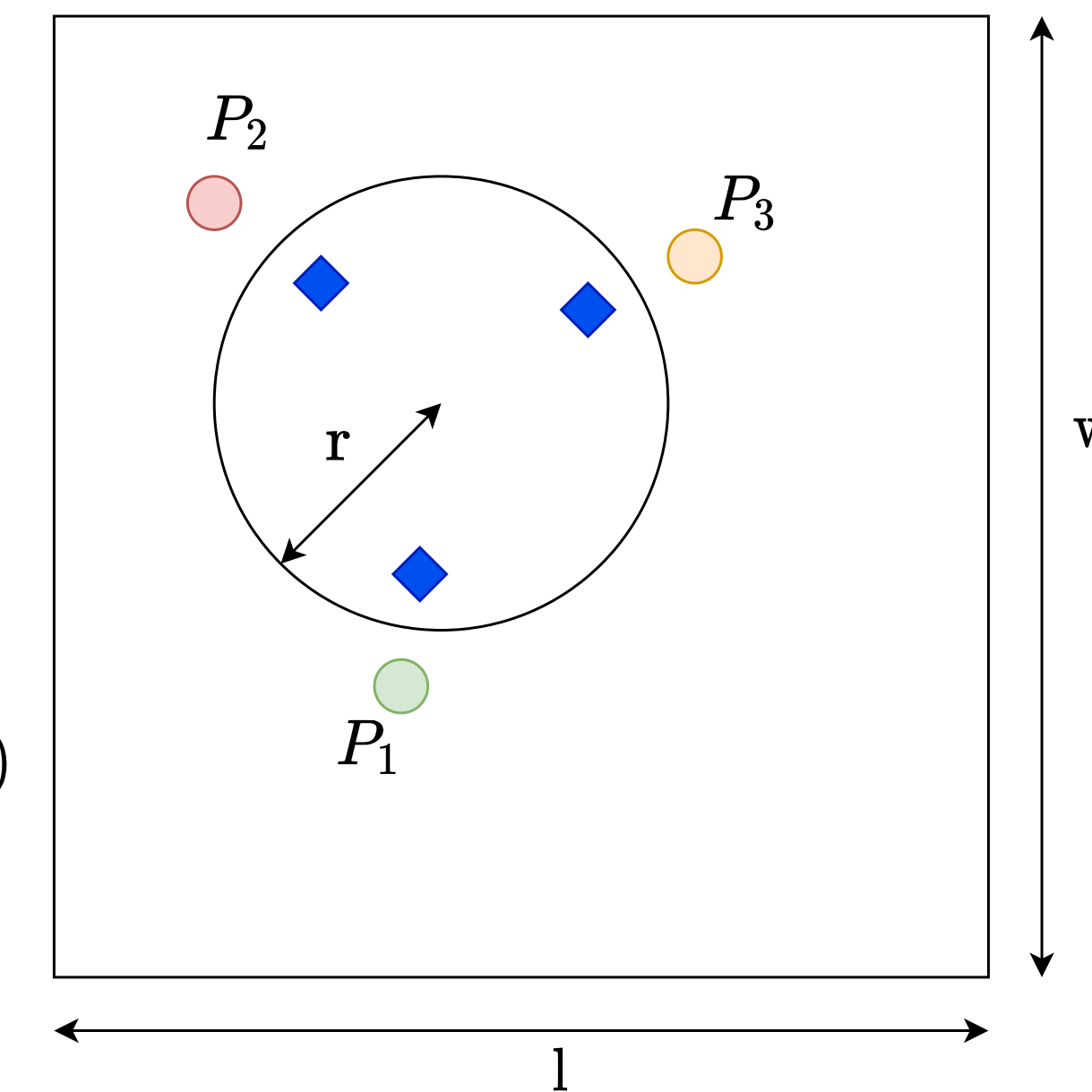
$150 \text{ ms} \leq T_{60} \leq 400 \text{ ms}$

$P_1 = P_2 = P_3$  (dry source levels)

~ 30 h train;

~ 3 h val;

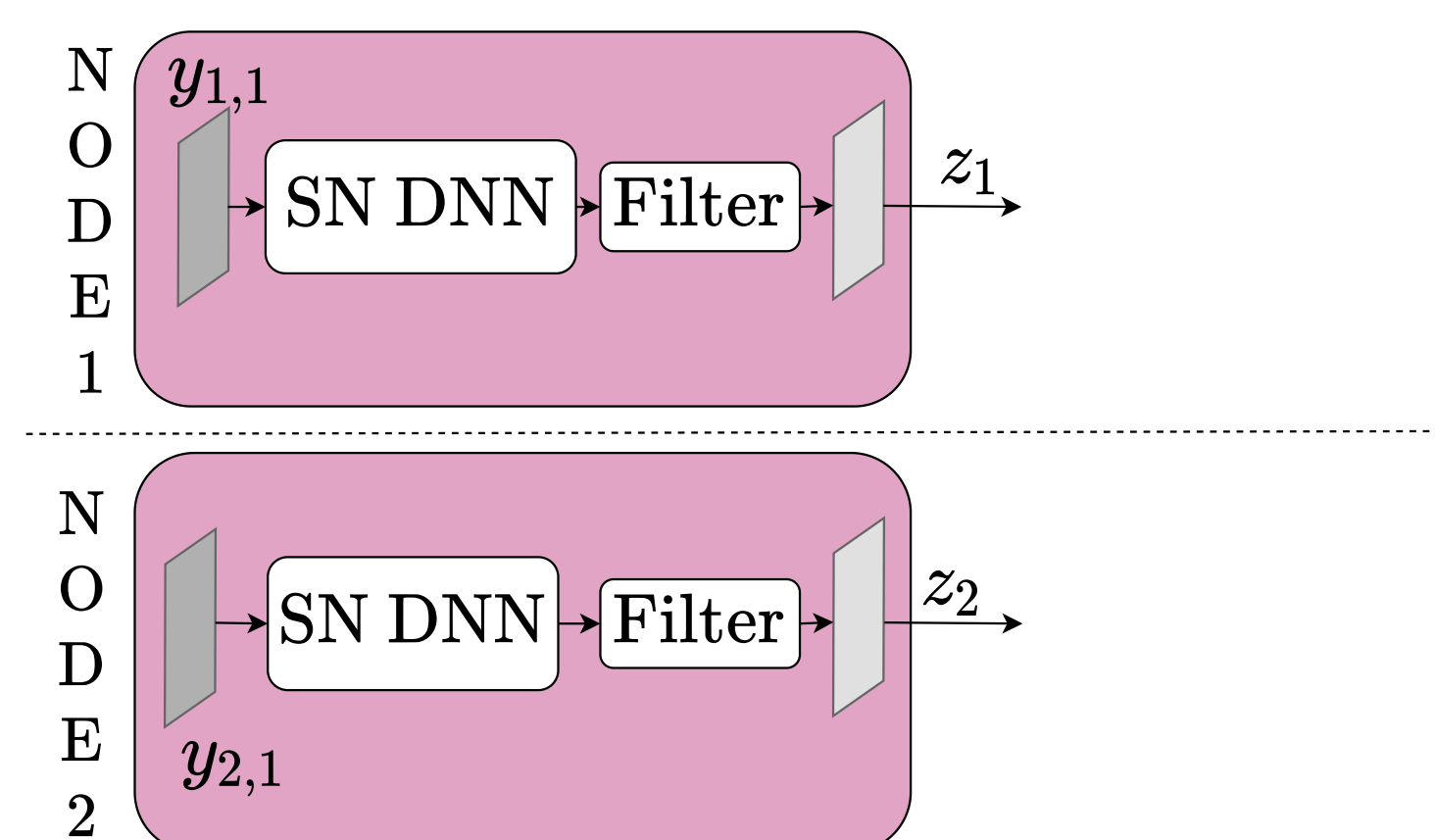
~ 3 h test



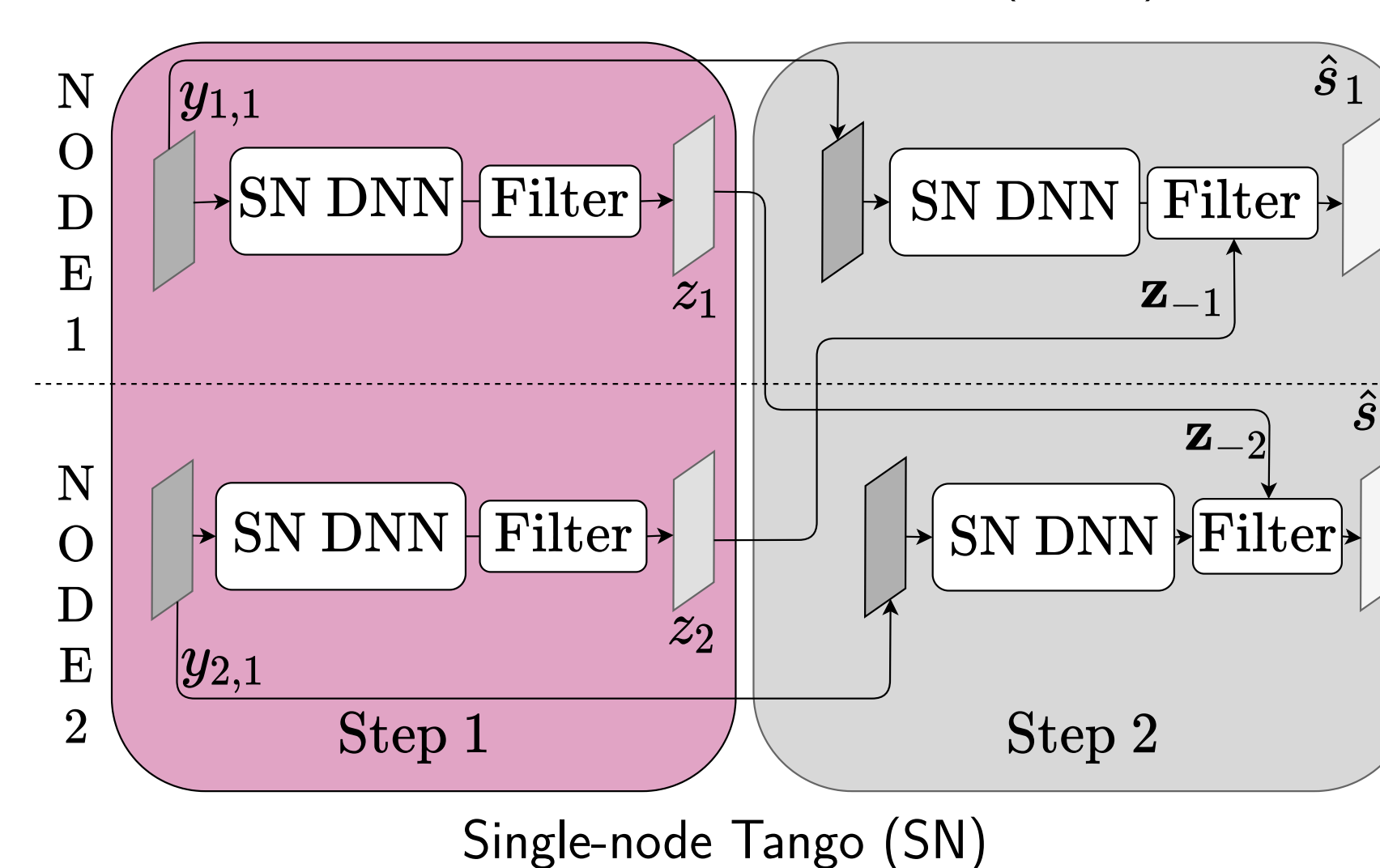
## Advantages

- Exploitation of spatial information
- Distributed processing
- Spatial information sent as pre-filtered estimates
- Usage of a priori knowledge (local SNR): each node estimates and sends a different source

## Compared methods



Local multichannel Wiener filter (MWF)



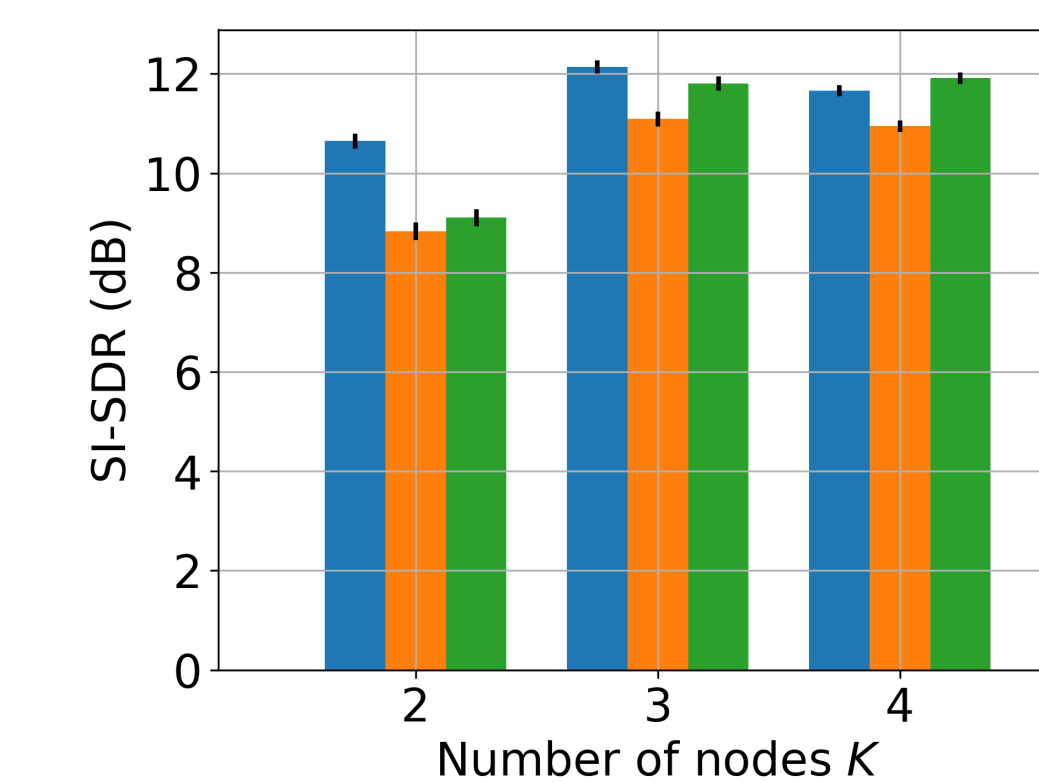
Single-node Tango (SN)

## Results

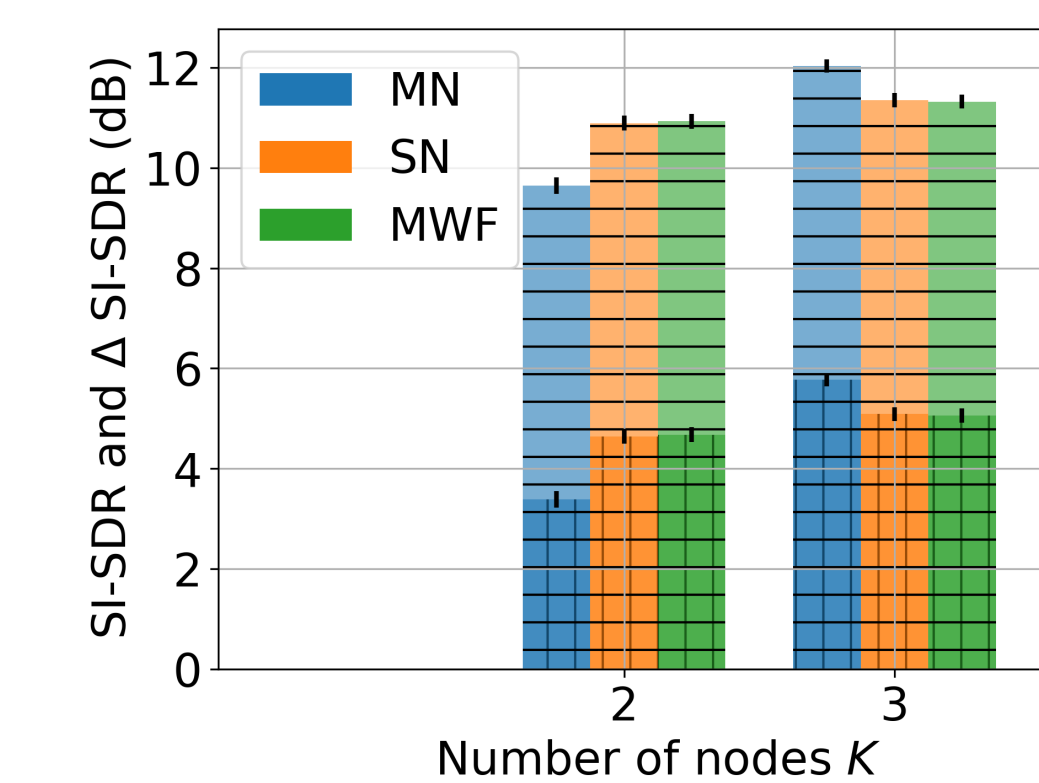
With  $K$  nodes and  $N$  sources, compare in three cases

- Balanced cases ( $K = N$ )
- Over-determined cases ( $K > N$ )
- Under-determined cases ( $K < N$ )

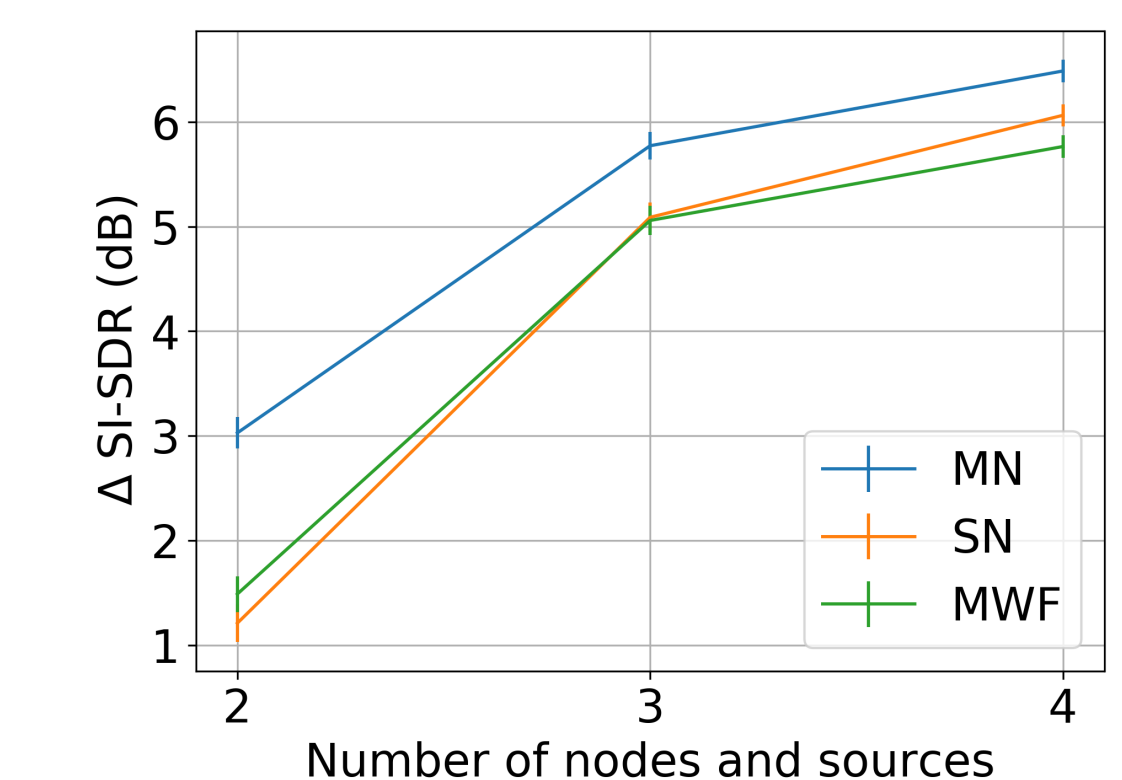
“MN” refers to our solution; “SN” to the single-node solution; “MWF” to the local MWF.



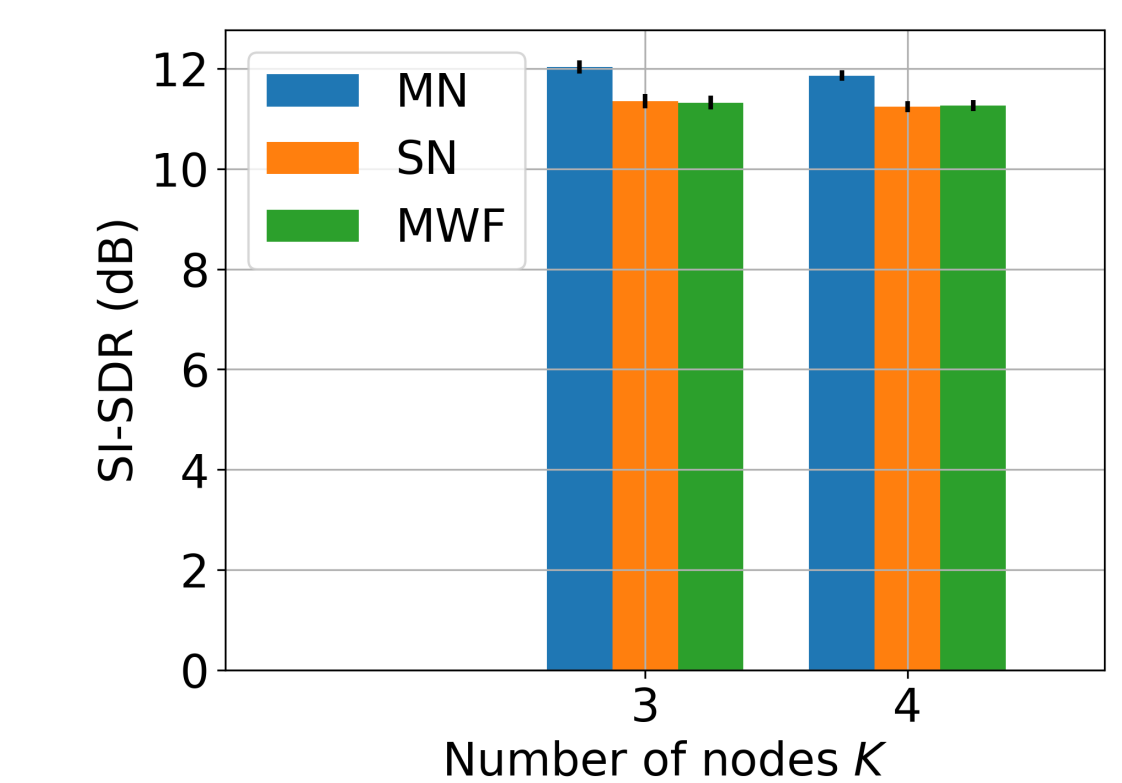
$N = 2$  sources



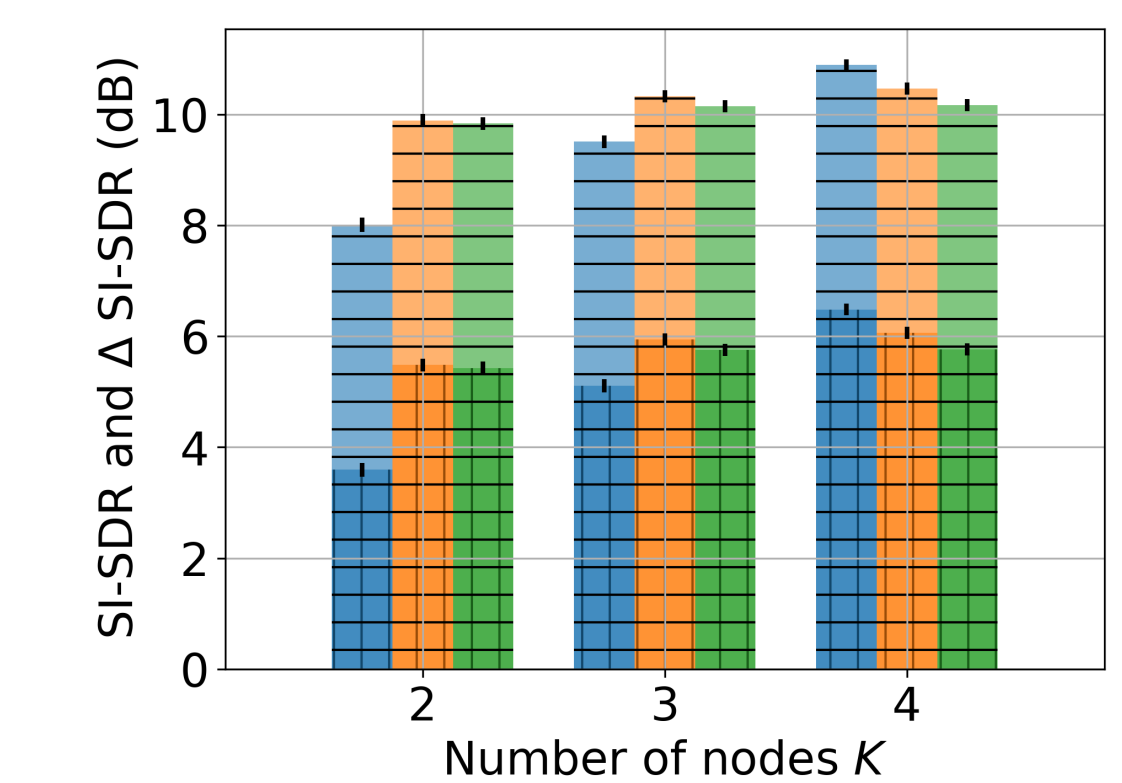
$N = 3$  sources



$N = 2$  sources



$N = 3$  sources



$N = 4$  sources

## Conclusion

**Tango:** a distributed processing for source separation

- Can process spatial information
- Evaluated on realistic meeting scenarios
- Improves performance when the number of nodes (and sources) increases
- Restricted to equally-determined or over-determined cases

## Links

- Paper: <https://hal.archives-ouvertes.fr/hal-02985794v3>
- Code to generate the dataset: <https://git.io/J014g>
- Code to Tango <https://git.io/J0144>