

A Runtime Reconfigurable FPGA-Based Microphone Array for Sound Source Localization

Bruno da Silva*, Laurent Segers, An Braeken and Abdellah Touhafi
Dept. of Industrial Sciences (INDI), Vrije Universiteit Brussel (VUB), Brussels, Belgium



Abstract

Microphone arrays are able to recognize, profile and locate sound-sources in noisy environments, but their quality is determined by the quantity of microphones [1],[2].

A higher number of microphones increases the computational demand, making real-time response challenging.

In this demo, we present a scalable and runtime reconfigurable architecture able to support a variable number of microphones and orientations in order to provide accurate sound-source localization in real-time.

Microphone Sensor Array

The SoundCompass [3] is composed by 52 ADMP521 MEMS microphones [4] in a 20-cm circular printed board (PCB). The planar microphone array geometry is composed of four concentric sub-arrays of 4, 8, 16 and 24 MEMS microphones.

A continuous steering of the focus direction in a 360° sweep allows to measure the variations of the surrounding sound field in all directions.

The distributed geometry of the MEMS microphones allows to adapt the sensor to different sound-source profiles. For instance, not all the 52 MEMS microphones need to be active to detect a particular sound-source profile. Therefore, only a few number of sub-arrays may be active, decreasing the computational requirements and becoming more power efficient.

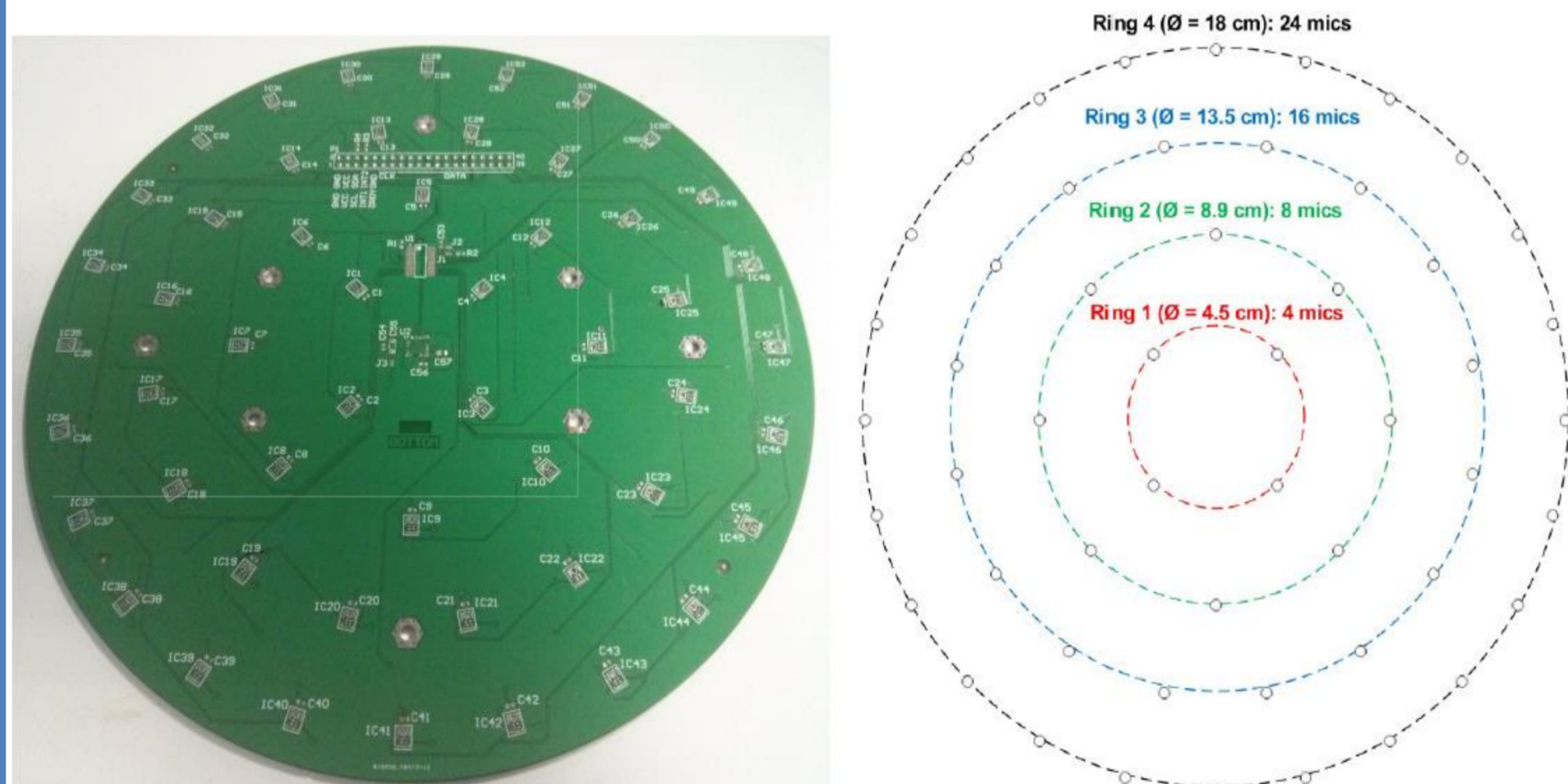


Figure 1: Our Sound-source localization device is composed of 52 MEMS microphones distributed in 4 sub-arrays.

Delay-and-Sum Beamforming

The beamforming technique applied in our sensor array is a widely adopted Delay-and-Sum beamforming [5].

The Delay-and-Sum beamforming delays the output signal of each microphone by a specific amount of time before adding all the output signals together.

Signals coming in the same direction as the focus direction will be amplified after the addition of all shifted outputs.

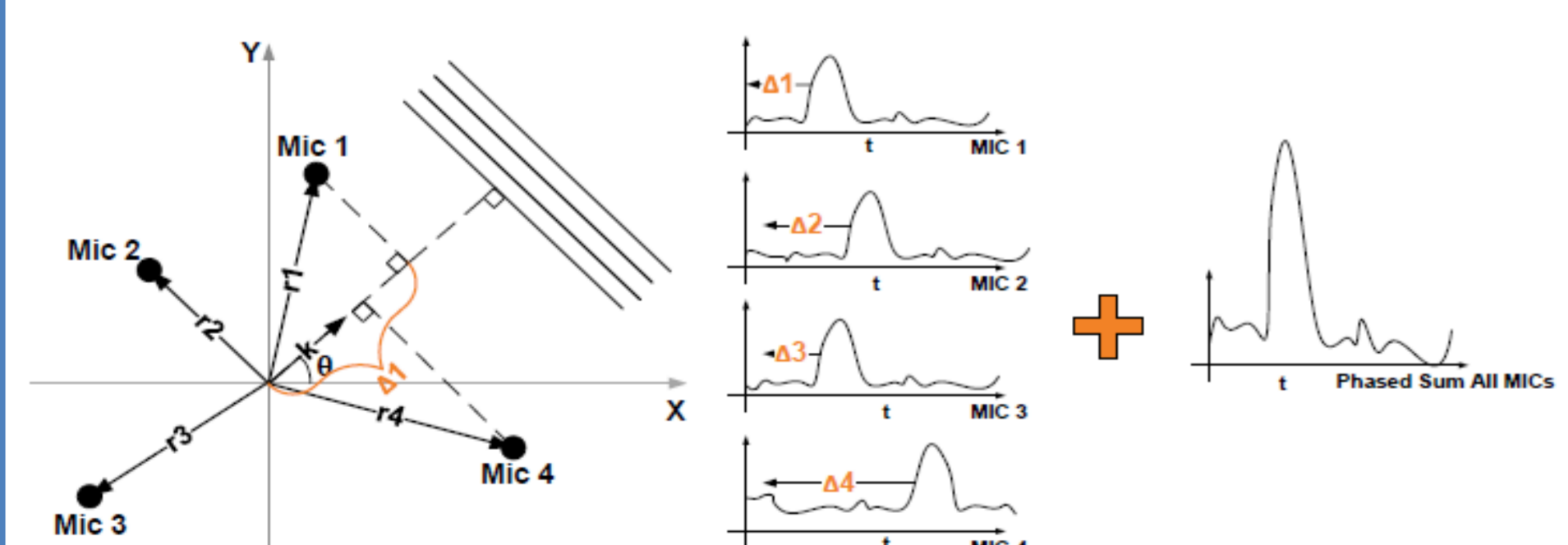


Figure 2: The output of each microphone in an array is shifted and then added to amplify sound signals arriving from a specific direction.

Acknowledgements

This work is a result of the CORNET project "DynamIA: Dynamic Hardware Reconfiguration in Industrial Applications". It is funded by IWT Flanders with reference number 140389.

A Scalable Architecture

The filter chain consists of a cascade of filters to down-sample and filter each microphone signal. This stage consists of a 2nd order decimator CIC filter, a low-pass compensation FIR filter and a running average block.

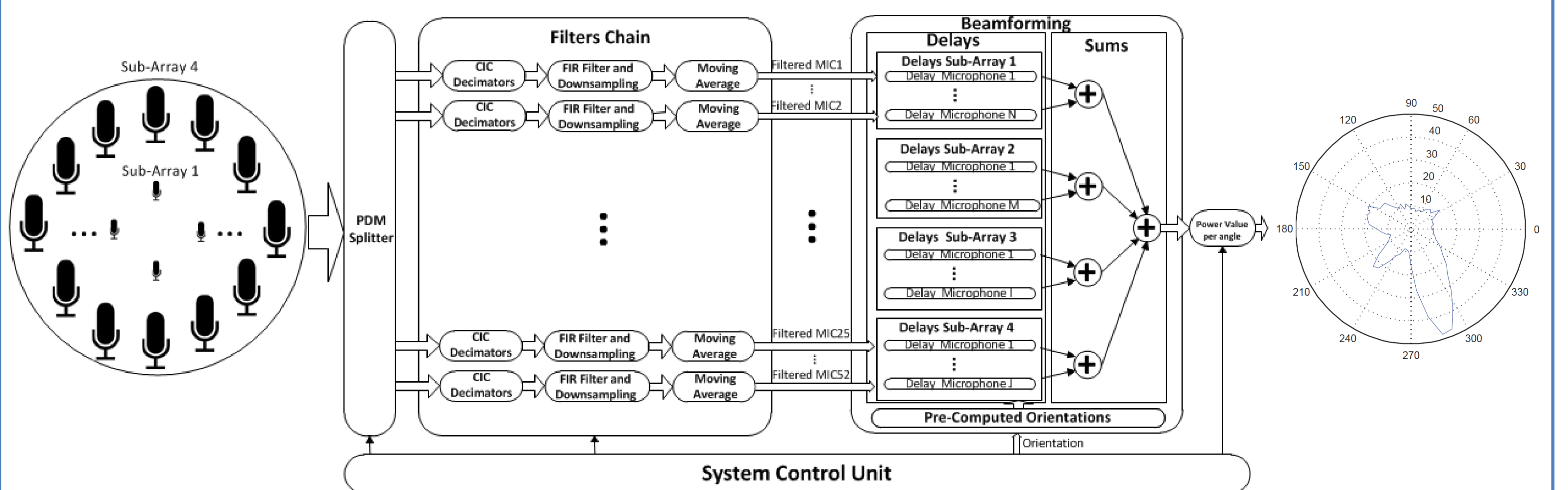
Our design supports up to 52 filter chain blocks working in parallel, but their number is defined by the number of active microphones. The unnecessary filter chain blocks are just disabled.

The beamforming stage depends on the number of microphones and sub-arrays.

The filtered microphone signals are grouped following their sub-array structure. Thus, instead of implementing one simple Delay-and-Sum of 52 microphones, there are four Delay-and-Sum operations in parallel for the 4, 8, 16 and 24 microphones.

Only the Delay-and-Sum beamforming block linked to an active sub-array is enabled. The remaining not active beamformers are set to zero in order to avoid any negative impact on the beamforming operation.

Consequently, any sub-array can be disabled, which directly deactivate the related filter and beamforming components, which is expected to have a direct impact over the power consumption.



Frequency Response when combining Sub-Arrays

The frequency response of the sub-arrays improves when they are combined since their frequency response are superposed. Thanks to this superposition, the side lobe and the major peaks of the smaller inner sub-arrays (sub-arrays 1 and 2) tend to fade out. Once all sub-arrays are combined (right figure in 3), the major side lobes are faded out when compared to sub-array 4 standalone. Only the common peaks of the outer sub-arrays remain visible.

The study of the frequency response of the microphone array confirms that, to achieve the best frequency response, the combination of sub-arrays to use as many microphones as possible is preferable in most of cases if there is not a resource limitation.

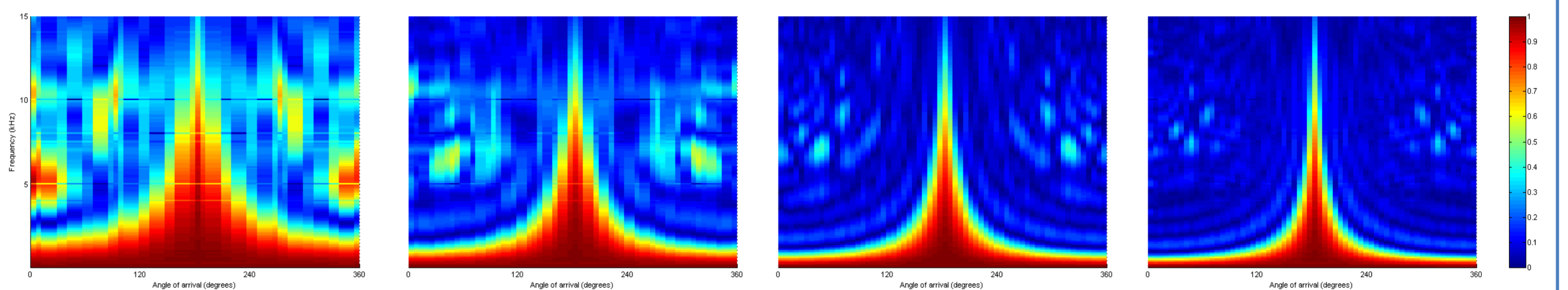


Figure 3: The waterfall diagrams show the power output of the most combined sub-arrays in all directions for all the responses. The left figure depicts the frequency response when combining the two most inner sub-arrays. The middle figure shows the response by also combining the sub-array 3, which is composed by 16 microphones. The right figure shows the frequency response when combining all sub-arrays. The values are normalized to 1.

Resource Consumption

The resource consumption is determined by the number of microphones further than the combined sub-arrays.

The flexibility of our architecture allows the creation of heterogeneous source-sound locators.

The architecture can be scaled for small FPGAs based on the target sound-source profile or a particular desirable power consumption.

For instance, the combination of the two inner sub-arrays would use 12 microphones while consuming less than 10% of the available resources.

Resources	Available Resources	Sub-Array 1 Standalone	Sub-Arrays 1 and 2	Sub-Arrays 1, 2 and 3	Complete
Slice Registers	106400	4699	14074	36323	66989
Slice LUTs	53200	2952	8573	25421	46655
LUT-FF pairs	86689	2526	7316	15568	26955
BRAM	140	4	12	28	52
DSPs48	220	8	28	60	108
Frequency [MHz]	-	130.365	118.061	118.061	118.061

Table 1: Resource consumption and maximum frequency when combining sub-arrays. The target FPGA is a Zynq 7020.

Conclusion and Future Work

The scalable approach in our design allows to vary the microphone array response based on performance or power efficiency.

Nevertheless, this is only our first scalable and reconfigurable approach and several tasks such as power analysis or acoustic measurements under realistic conditions must still be explored.

References

- [1] Zwysig, E., Lincoln, M. and Renals, S. "A digital microphone array for distant speech recognition." Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on. IEEE, 2010.
- [2] Zhang, X., Huang, J., Song, E., Liu, H., Li, B. and Yuan, X. "Design of small MEMS microphone array systems for direction finding of outdoors moving vehicles." Sensors 14.3 (2014): 4384-4398.
- [3] Tiete, J., Dominguez, F., Silva, B. D., Segers, L., Steenhaut, K., & Touhafi, A. (2014). "SoundCompass: a distributed MEMS microphone array-based sensor for sound localization". Sensors, 14(2), 1918-1949.
- [4] AnalogDevices. ADMP521 datasheet Ultralow Noise Microphone with Bottom Port and PDM Digital Output, Technical Report, Analog Devices: Norwood, MA, USA, 2012. 2526
- [5] Johnson, Don H., and Dan E. Dudgeon. "Array signal processing: concepts and techniques". Simon & Schuster, 1992.