# Real-time Audio Separation of Human Voices

ELECTRICAL & COMPUTER ENGINEERING
UNIVERSITY OF MICHIGAN

EECS 452: Digital Signal Processing Design Lab – Winter 2021
Tainon Chen, Zhirui He, Nathan Hellstedt, Haotian Qiao, Erik Sangeorzan, Yuxin Zhong

## Introduction

Consider the cocktail party problem: humans can easily isolate a certain conversation in a noisy room, while computers cannot — we attempt to solve this problem, which is demonstrated in Figure 1. Our project aims to separate three individual human voices from mixed audio inputs. The hardware system we used in our project is shown in Figure 2.

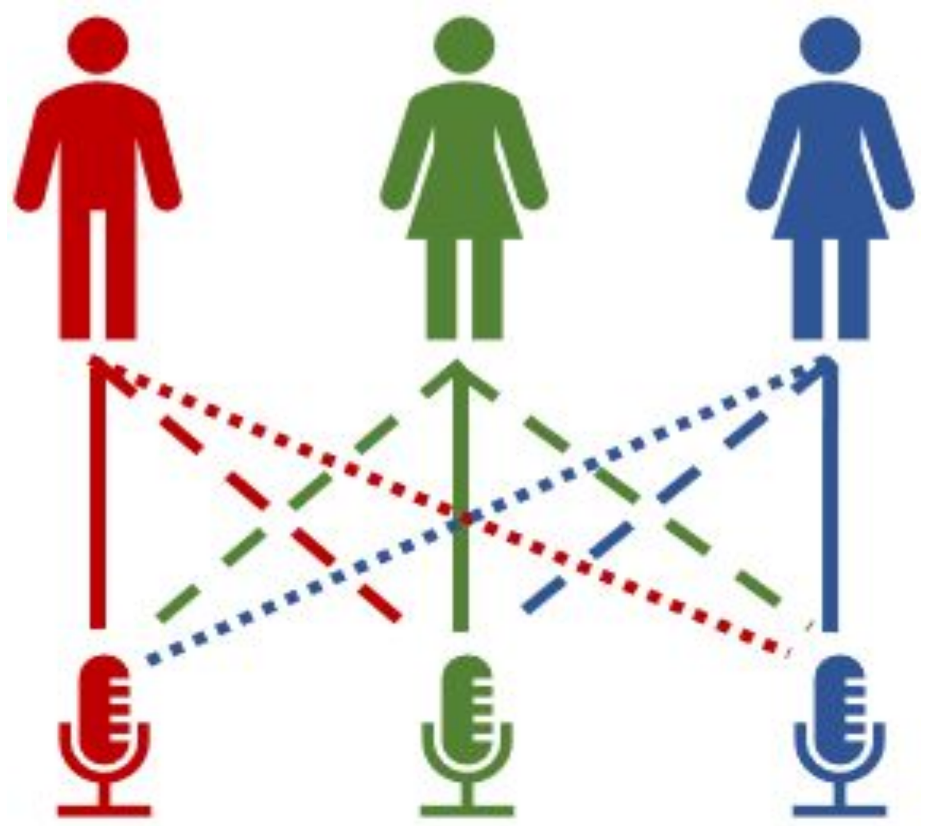

**Figure 1: Cocktail party problem**

## Independent Component Analysis (ICA)

Our project uses ICA to separate mixed audio into its individual sources. ICA uses the knowledge that the individual sources are non-gaussian and independent to isolate individual sources through the following algorithm:

1) Take N concurrent samples from each of the M microphones to form a MxN data array
2) Make the mean of each row in the array zero
3) Whiten the data to make the covariance between each array row zero
4) Iterate until a transform matrix that maximizes the non-gaussianity of each signal is found
5) Apply transform matrix to signal

## System Architecture

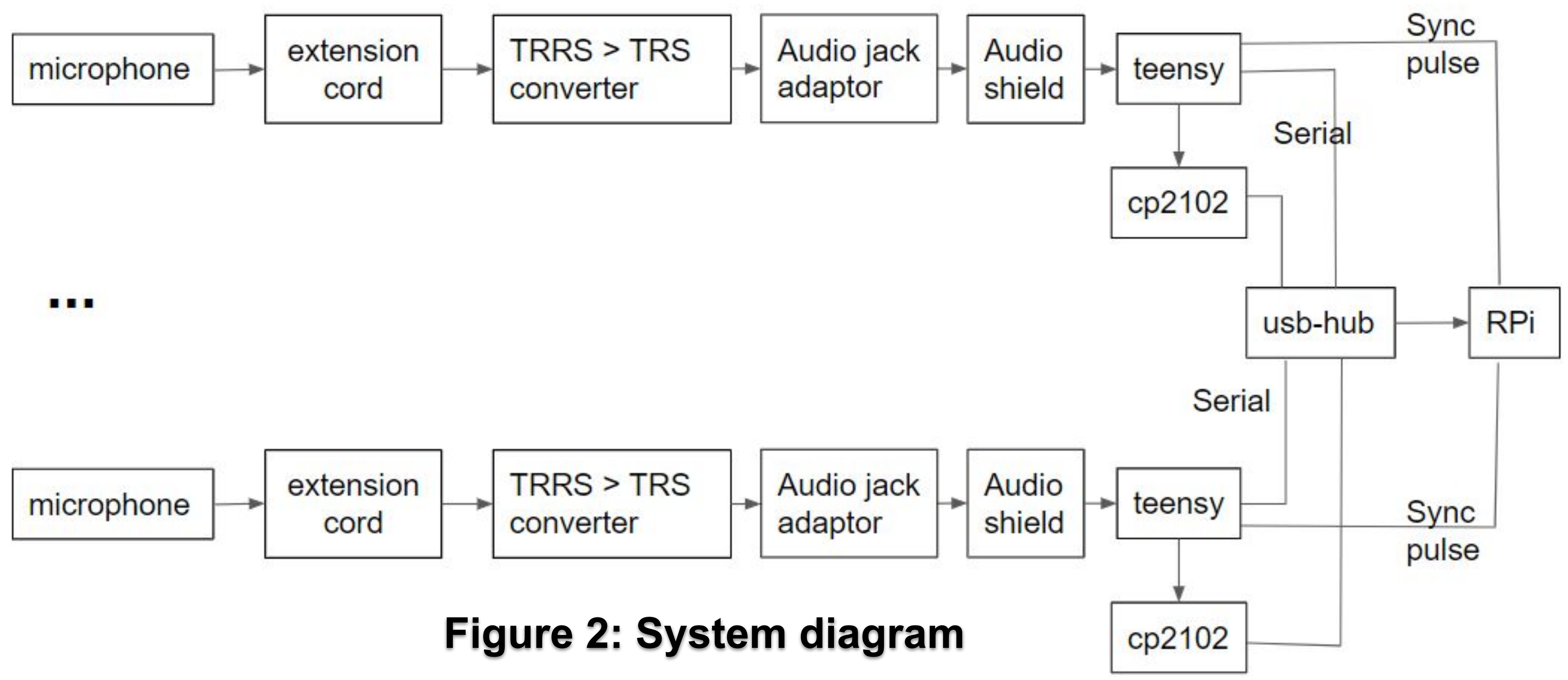


**Figure 2: System diagram**

## ICA Simulation Results

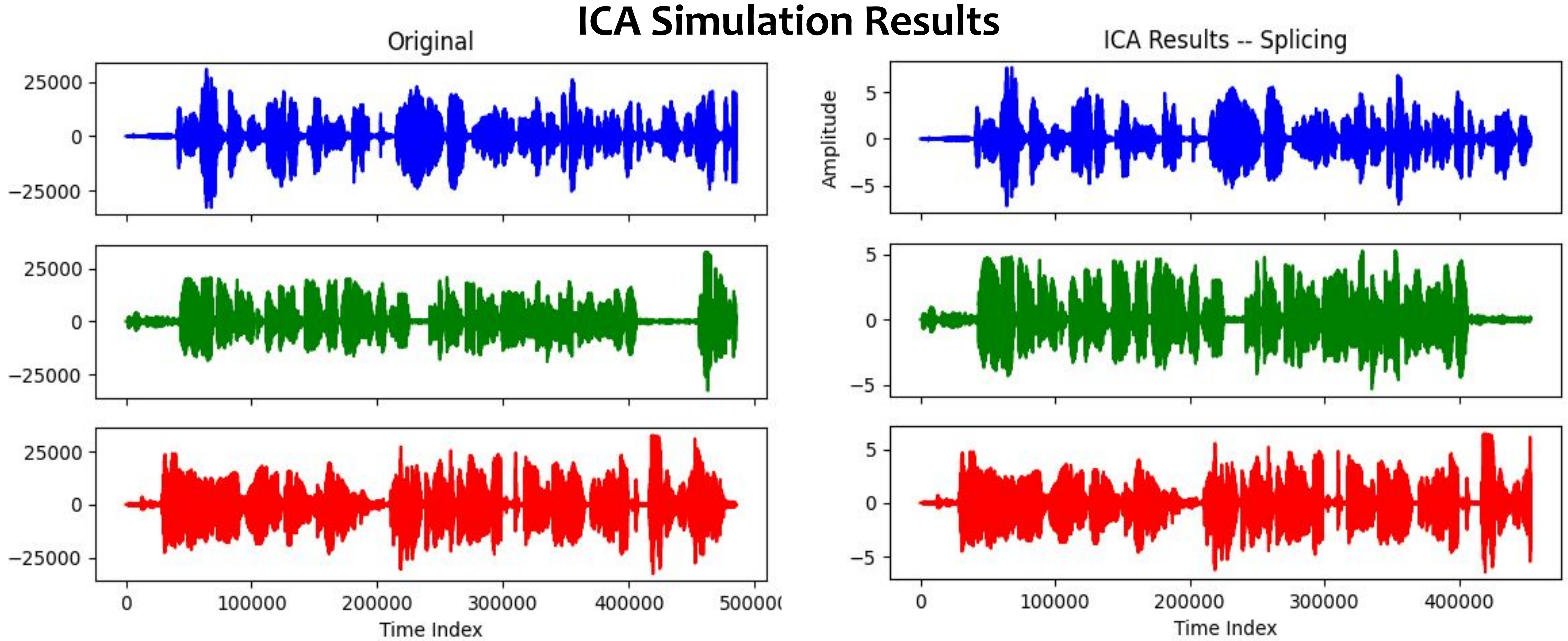


**Figure 3: ICA simulation results — original vs. recovered source signals**

## Results / Evaluation

Our system was able to successfully transmit and separate linearly mixed audio, as seen in Figure 3. Audio recorded from the microphone was transmitted to the Raspberry Pi at 12 kHz. ICA performed well under ideal conditions, but when the mixed audio had nonlinearities, it didn't perform well. Our next step is to incorporate the handling of nonlinearity into our source separation algorithm.

## Innovation

The main innovation of our project is performing source separation in real time. This was accomplished using our method of windowing and splicing to separate the signals. We also developed a user interface to playback and plot separated audio with minimal real-time delay.

## Acknowledgements

We would like to thank Professor Shai Revzen, David Kucher, and Rishank Nair for their guidance and help on this project. Additionally, we recognize the ECE undergraduate program for providing funding for this project.