ScuDo
Scuola di Dottorato ~ Doctoral School
WHAT YOU ARE, TAKES YOU FAR

XXXVII Cycle

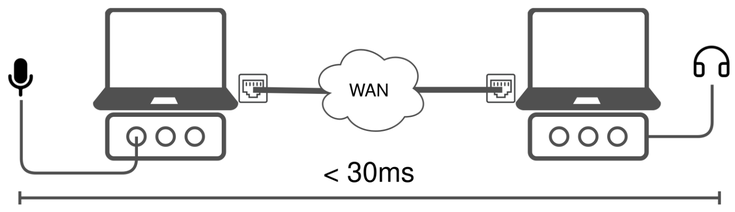
Low-latency & accessible technologies for networked music performance

Matteo Sacchetto

Supervisors: Prof. Cristina Rottondi, Prof. Andrea Bianco

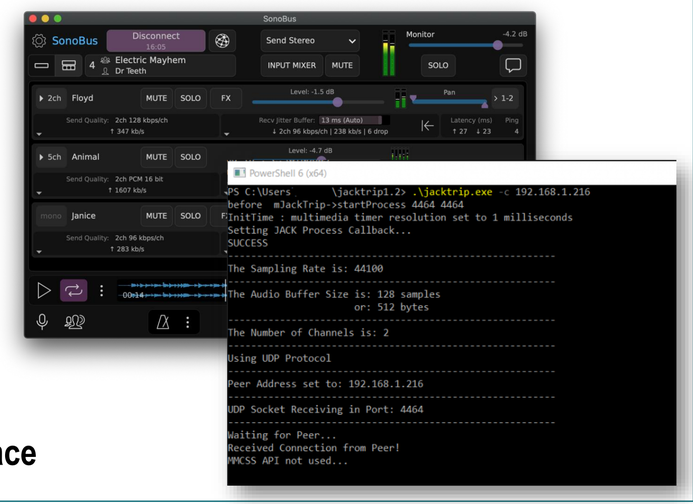
Research context and motivation

- The SARS-CoV-2 pandemic **encouraged the use of videoconferencing platforms** also for **music teaching and networked music performance** (NMP)
- NMP is "a **real-time** interaction **over the network** which enables musicians to play together **as if they are in the same room**"
- In NMP **latency** and **audio quality** are the two most important factors



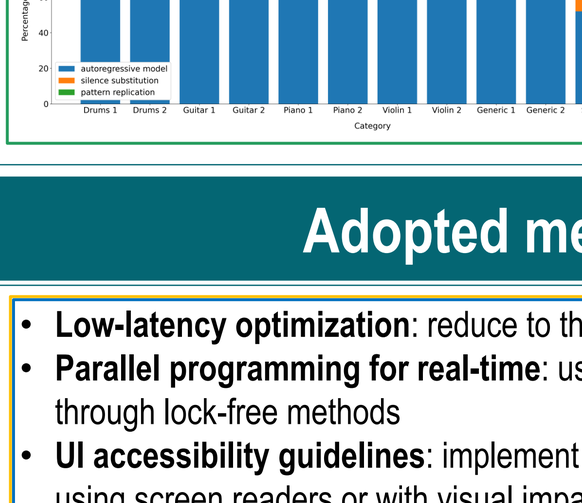
- Latency **< 30ms**
- **Uncompressed audio** (20Hz-20kHz).

- Standard videoconferencing platforms **lack in performance**
  - Latency **> 100ms**
  - Voice optimized codecs/settings
  - **Heavy processing** (noise suppression, …)
- Traditional NMP solutions **lack in usability and accessibility**
  - Difficult installation
  - **Complex/Not accessible interface**

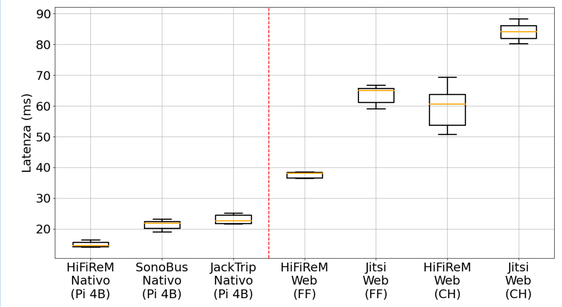


Novel contributions

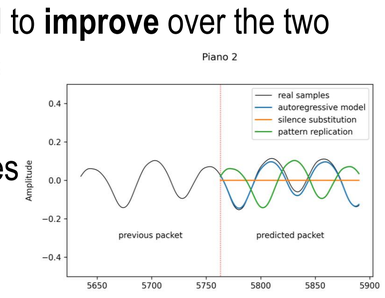
- Our web app (JackTrip-WebRTC) was able to **improve** over traditional WebRTC based solutions **by 20-60ms**
  - **< 40ms** with Firefox on macOS

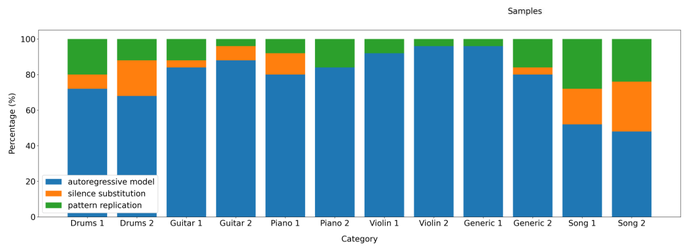


- Our native solution was able to **reduce latency by 5-12ms** w.r.t. state-of-the-art NMP solutions
  - **12-15ms** on a local network



- AR models showed to **improve** over the two traditional solutions
  - **Lower errors**
  - $\frac{1}{2}$ discontinuities
  - **Qualitatively better**

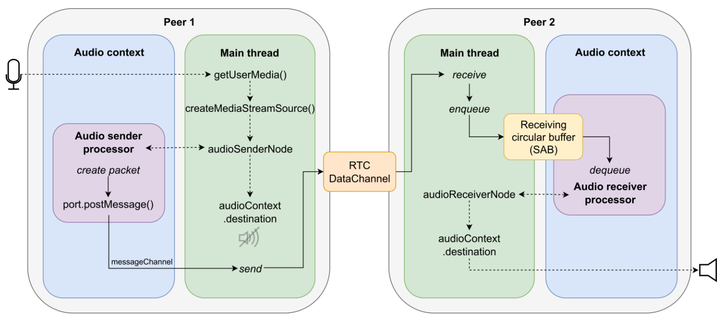




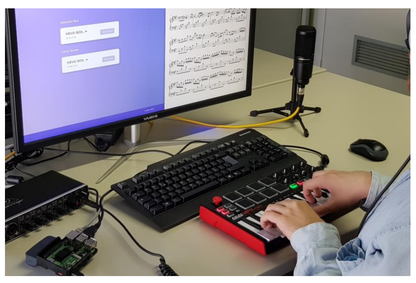
- Final interface

Addressed research questions/problems

- Web applications are easy to use and **designed to be accessible**
- Design a web app focused on NMP
  - Use **WebRTC's DataChannel (UDP)**
  - Use **Web Audio API's AudioWorklet**



- Web applications don't have access to low-latency audio drivers (ALSA, ASIO, ..)
- **Overcome limitations of web apps** by developing a **native solution controllable through a web app**
  - Based on a **Raspberry Pi 4B**
  - **Accessible** and **easy to use interface**



- Traditional NMP packet loss concealment (PLC) solutions lead to **suboptimal audio quality**
  - Silence substitution

0.9	0.8	...	0.8	0.5	0	0	...	0	0

  - Pattern replication

0.9	0.8	...	0.8	0.5	0.9	0.8	...	0.8	0.5

- **Propose a NMP specific PLC solution based on autoregressive (AR) models**

- Experiments on how to integrate **haptic feedback** and visual stimuli to enhance **audio accessibility**

Adopted methodologies

- **Low-latency optimization**: reduce to the bare minimum the processing time
- **Parallel programming for real-time**: use minimum number of threads synchronized through lock-free methods
- **UI accessibility guidelines**: implement a user interface to be accessible to people using screen readers or with visual impairments

- **Autoregressive Models**: type of process where the forecast variable is expressed as a linear combination of its past values

$$\mathrm{AR}(\rho):\ y_t = c + \psi_1 y_{t-1} + \psi_2 y_{t-2} + \ \ldots\ + \psi_\rho y_{t-\rho} + \epsilon_t$$

- **Augmented Dickey fuller test**: tests the null hypothesis that a unit root is present in a time series. Used to check if the time series is stationary

- **Psycho-perceptual tests**: for associating audio tracks to attributes belonging to **different human senses** other than earing (touch, color, smell, ...)

Future work

- Define a strategy to deal with **audio card drift**
- **Auto-estimate parameters** based on the network connection

- Add **NAT traversal**
- Add some **remote mixing capabilities**
- Finish development of **custom HW audio card/accelerator**

- Define a strategy and **deploy AR models** in a real case

- Perform a **qualitative analysis** of the resulting interface

Submitted and published works

- Sacchetto, M., Gastaldi, P., Chafe, C., Rottondi, C., Servetti, A., "*Web-Based Networked Music Performances via WebRTC: a Low-latency PCM Audio Solution*", Journal of the Audio Engineering Society, 2022. *(ongoing publication)*
- Sacchetto, M., Huang, Y., Bianco, A., Rottondi, C., "*Using Autoregressive Models for Real-Time Packet Loss Concealment in Networked Music Performance Applications*", AudioMostly 2022, St. Pölten, 6-9 September 2022. *(ongoing publication)*
- Sacchetto, M., Rottondi, C., Servetti, A., Shtrepi, L., Masoero, M., Valle, A., "*HiFiReM: un approccio unificato, web e nativo, per la didattica musicale remota*", XXIII CIM, Ancona, 25-28 October 2022. *(submitted)*

List of attended classes

- 01UJBRV – Adversarial training of neural networks (April 2022, 3 CFU)
- 01RGBRV – Optimization methods for engineering problems (May 2022, 6 CFU)
- 01MNFIU – Parallel and distributed computing (June 2022, 5 CFU)
- 01SCSIU – Machine learning for pattern recognition (July 2022, 4 CFU)

Politecnico di Torino

PhD program in
Electrical, Electronics and Communications Engineering