# Fast Black-Box Optimizers for Low Delay Audio Source Separation

Gerald Schuller
*Ilmenau University of Technology, Institute for Media Technology, Ilmenau, Germany*

NYU 2024

# Introduction

- Black-box optimizers have made significant progress in recent years.
- But are these optimizers applicable for real-time signal processing tasks, with a fixed time budget?
- This presentation explores the use of black-box optimizers in audio source separation.

# Digital Signal Processing and Optimization

- Digital signal processing has long been linked to optimization.
- Early methods focused on filter design and adaptive filters (e.g., LMS, RLS).
- These methods are based on gradient descent, which works best with convex loss functions.

# Filter Banks and Non-Convex Functions

- Later developments: Filter banks for perfect reconstruction.
- Filter banks often lead to non-convex objective functions.
- Non-convexity makes optimization more challenging.

# Multichannel Blind Source Separation

- Multichannel blind source separation is another example of non-convex optimization.
- This problem is often solved in the STFT domain.
- Time-domain processing avoids delay but has higher rates of local minima.

# Recurrent Neural Networks and Vanishing Gradients

- Recurrent Neural Networks (RNNs) suffer from the vanishing gradient problem.
- Black-box optimizers can help with optimizing RNNs.
- LSTMs were introduced to mitigate this problem but have higher complexity.

# Applications of Black-Box Optimizers

- Speech recognition on embedded devices, adapted to the user.
- Non-differentiable loss functions in adversarial machine learning.
- Embedded reinforcement learning for small robots.

# Black-Box Optimization Methods

- Black-box optimization treats the objective function as a "black box."
- Gradient information is not required.
- These optimizers search over a probability distribution of potential solutions.

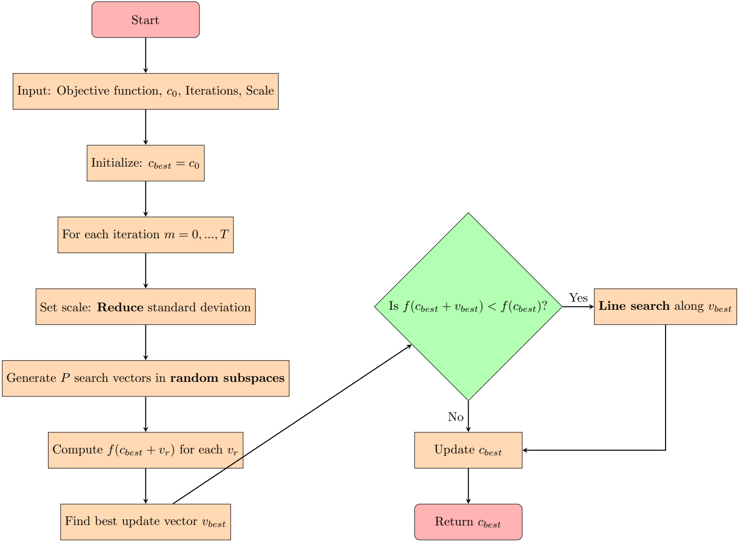
# The Method of Random Directions

- The proposed method was developed for optimizing low delay filter banks and blind audio source separation.
- Uses normally distributed search vectors and shrinking standard deviation.
- Introduces random subspaces for high-dimensional problems.

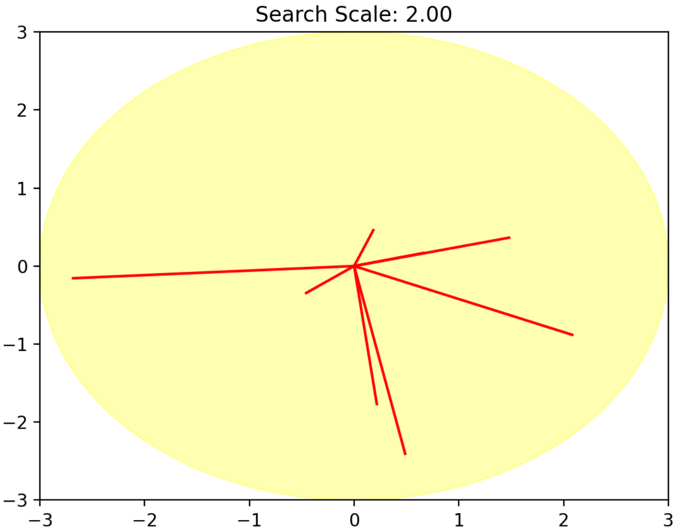
# Algorithm of Random Directions

- Pseudo-code for the Random Directions method:
- The algorithm guarantees that optimization does not lead to worse results.
- Random sub-space updates speed up high dimensional problems.
- Line search increases speed in smoother areas.

# Algorithm of Random Directions



# Random Directions Algorithm - Video Demonstration



# Compared Black-Box Optimizers

- Mixture Model-based Evolution Strategy (MMES).
- Rank-One Evolution Strategy (R1ES).
- Limited-Memory Matrix Adaptation Evolution Strategy (LMMAES).
- Gradientless Descent (GLD).
- Separable Natural Evolution Strategy (SNES)
- BErnoulli Smoothing (BES)

# Application 1: Blind Audio Source Separation

- Time-domain source separation using an unmixing matrix.
- Optimized using the Kullback-Leibler Divergence as the objective function.
- Parallel optimization during audio processing.
- More Information and examples in our repository `https://github.com/TUIlmenauAMS/LowDelayMultichannelSourceSeparation_Random-Directions_Demo`

# Results: Source Separation

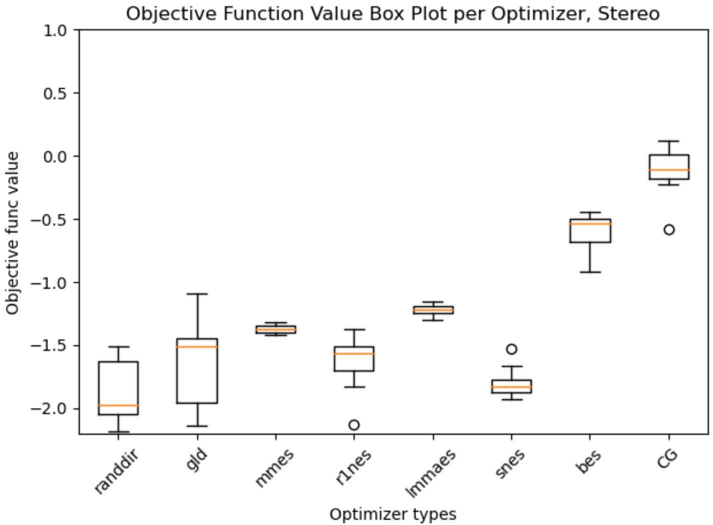
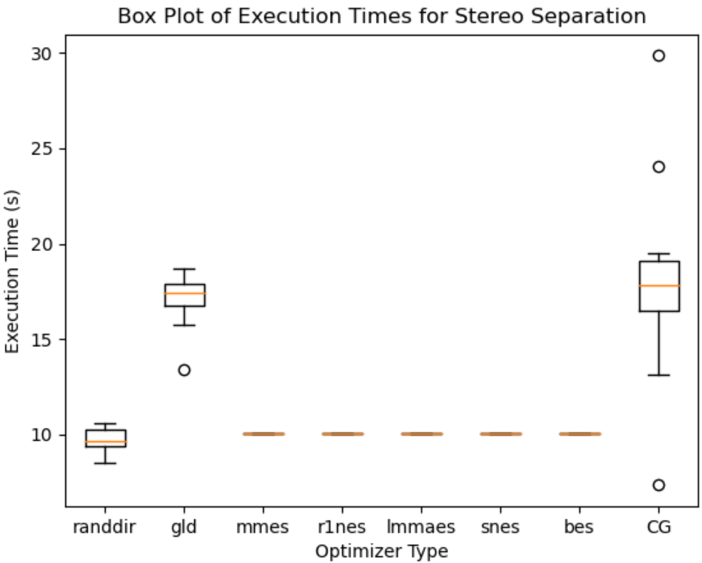


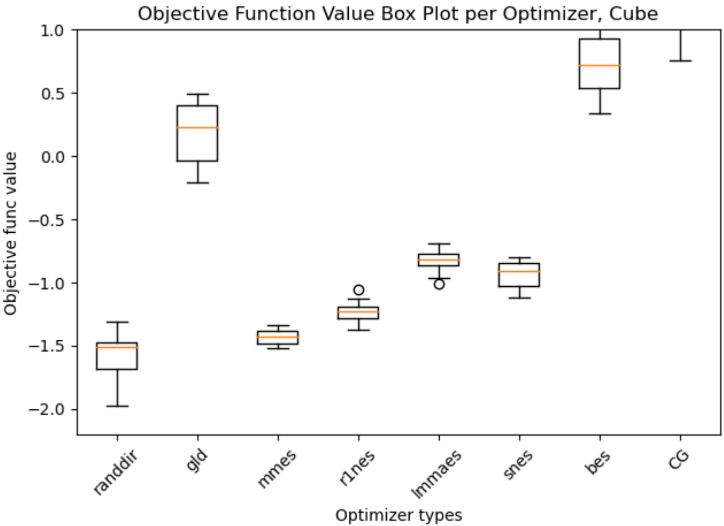

Figure: Stereo setup: Achieved minimum objective value for different optimizers.

# Results: Processing Time



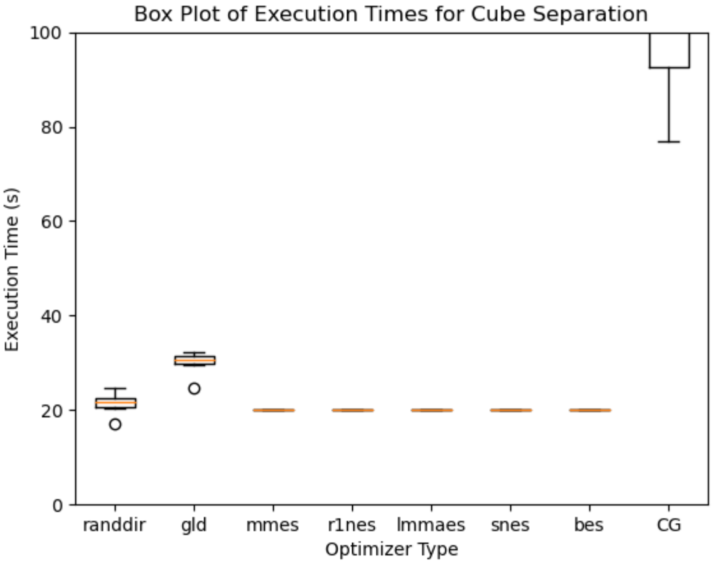

Processing times for different optimizers in stereo setup.

# Results: Cube Microphone Setup




Achieved minimum objective value for different optimizers.

# Results: Processing Time

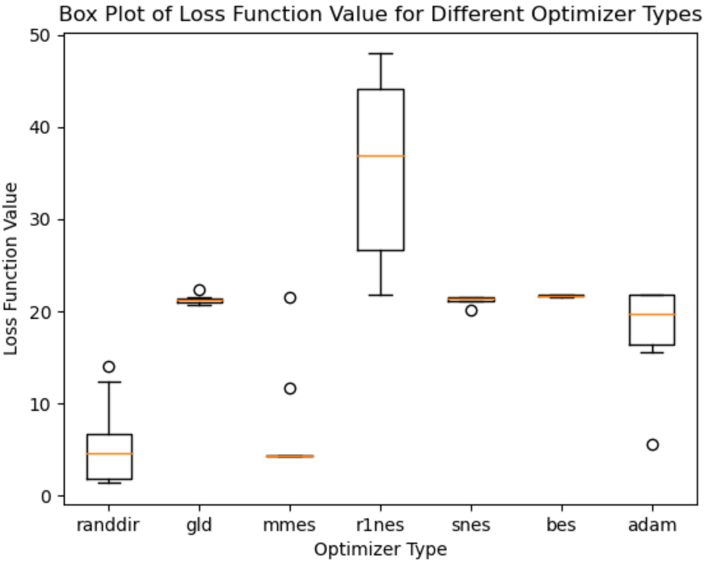



Processing times for different optimizers in cube setup.

# Application 2: Recurrent Neural Network

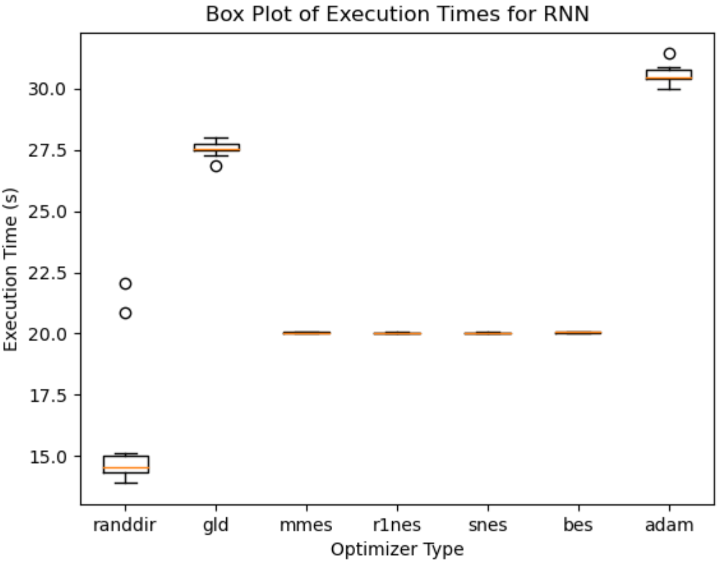
- Target signal: decaying sinusoid modeled by a 2nd order IIR filter.
- Optimized using mean squared error as the loss function.
- The optimizer must handle long-term dependencies.

# Results: Recurrent Neural Network




Achieved minimum objective value for RNN optimization.

# Results: Processing Time




Processing times for different optimizers in cube setup.

# Conclusion

- Black-box optimization can be applied to real-time audio signal processing.
- The Random Directions optimizer consistently performed among the best.
- Future work: Apply these methods to more complex neural networks and signal processing tasks.

# Thank You

Software and description available at:
https://github.com/TUIlmenauAMS/BlackBoxOptimizerSPcomparison

Thank you for your attention!