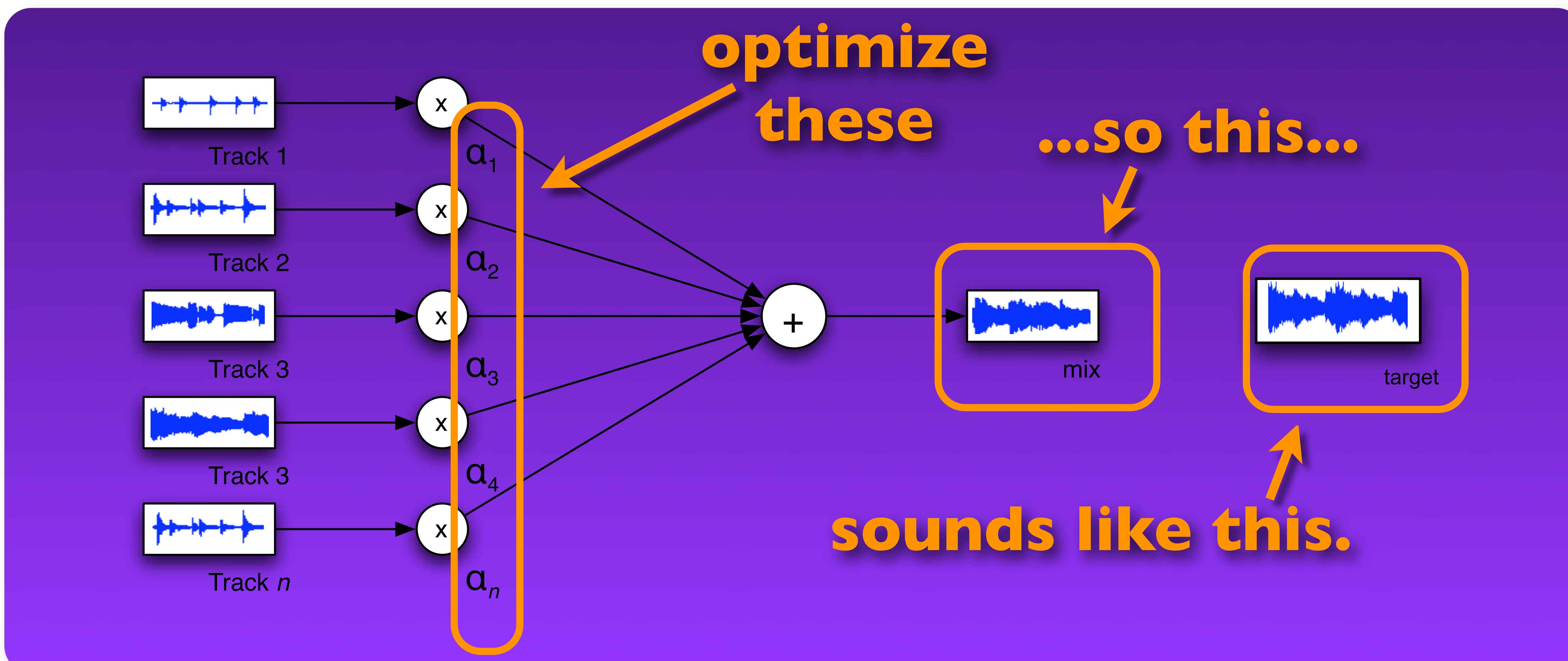
Music Technology Research Group
NEW YORK UNIVERSITY

AES 124th Convention Paper 7496

A Framework for Automatic Mixing Using Timbral Similarity Measures and Genetic Optimization

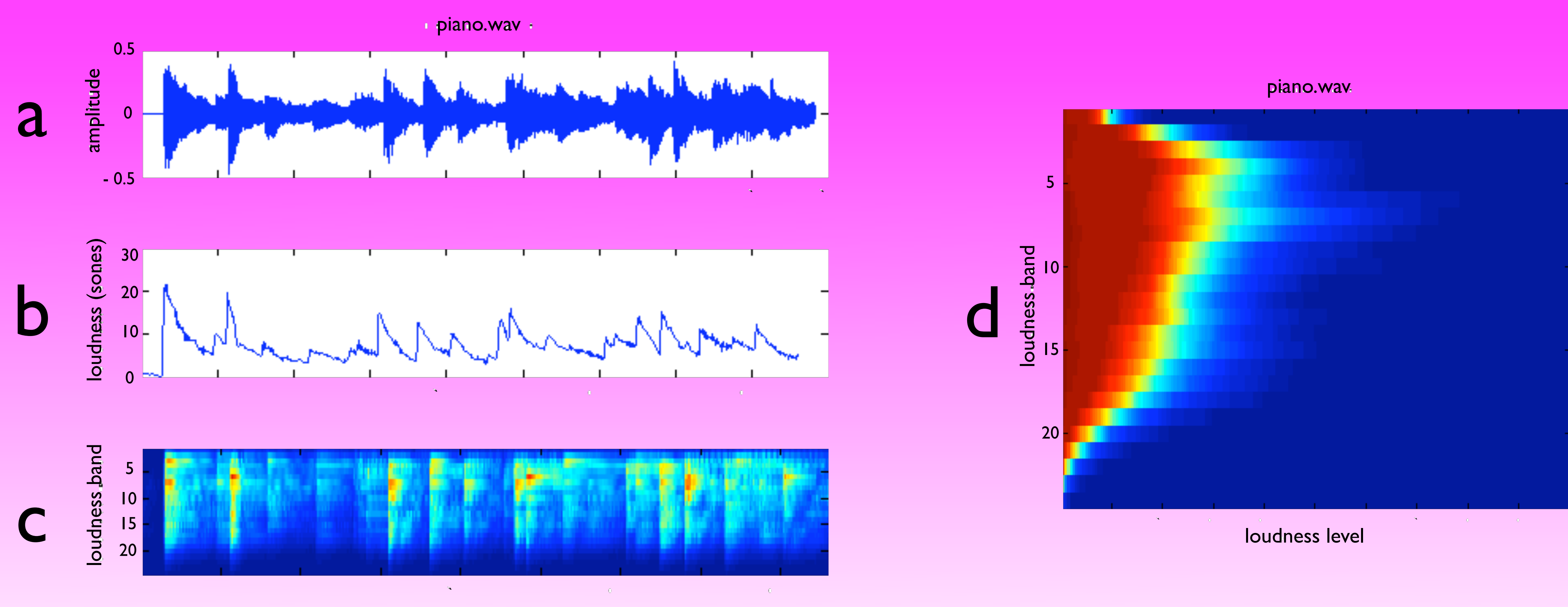
Bennett Kolasinski
New York University, New York, USA
bennett@bennettk.com

Project Description: A novel method is introduced for automatic mix recreation using timbral classification techniques and an optimization algorithm. This approach uses the Euclidean distance between modified Spectral Histograms to calculate the distance between a mix and a target sound and uses a genetic optimization algorithm to figure out the best coefficients for that mix. The implementation has been shown to successfully recreate multitrack mixes accurately and may pave the way toward the automatic mixing of novel multitrack sessions based on a desired target sound.



Spectral Histogram (SH)[1]

- a histogram of the number of times loudness levels in sones have been exceeded across the frequency bands of the Bark scale
- gives the ‘distance’ between two mixes

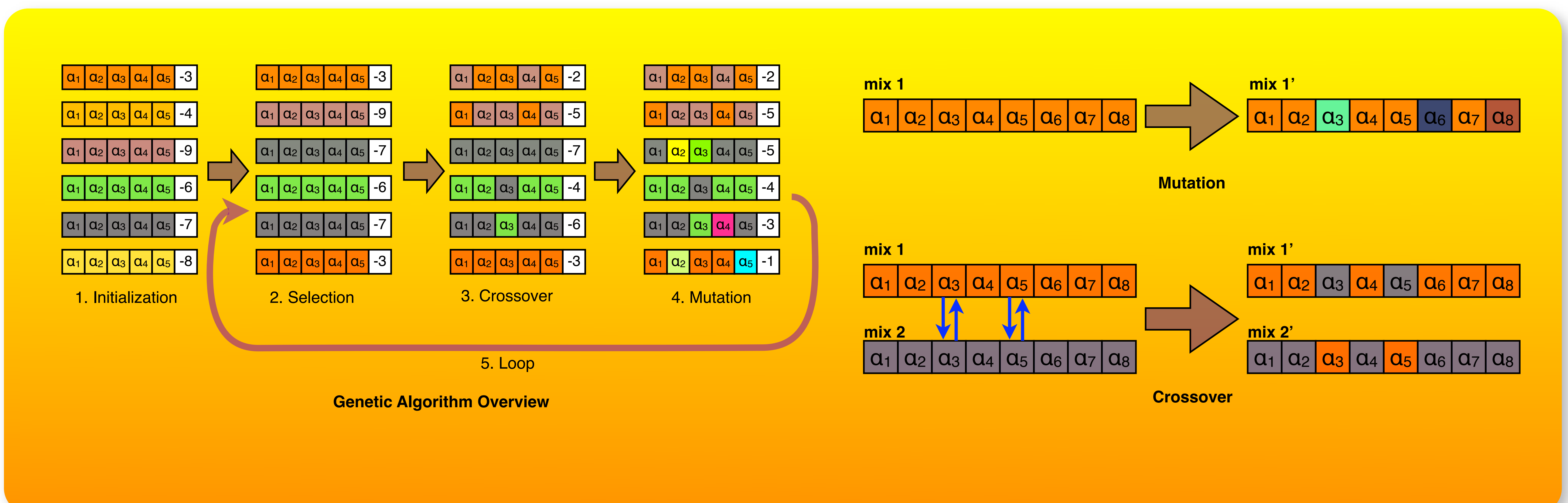


Four representations of the same piano signal: (a) waveform; (b) loudness in sones; (c) loudness in sones across the critical bands of the Bark scale; (d) Spectral Histogram

Genetic Optimization[2]

- based on biological field of genetics, generates a number of random mixes that evolve over generations to get closer to a target mix
- similar work done on estimating the parameters of FM Synthesis [3], [4]
- each mix is represented as a series of 'genes' representing the gains to be applied to each track
- Uses Euclidean Distance between SHs as the cost (distance) function:

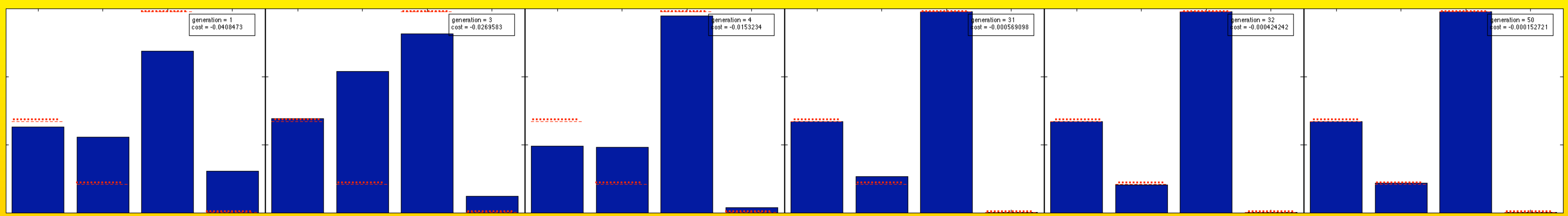
$$\sqrt{SH(mix)^2 - SH(target)^2}$$



Genetic Optimization Behavior

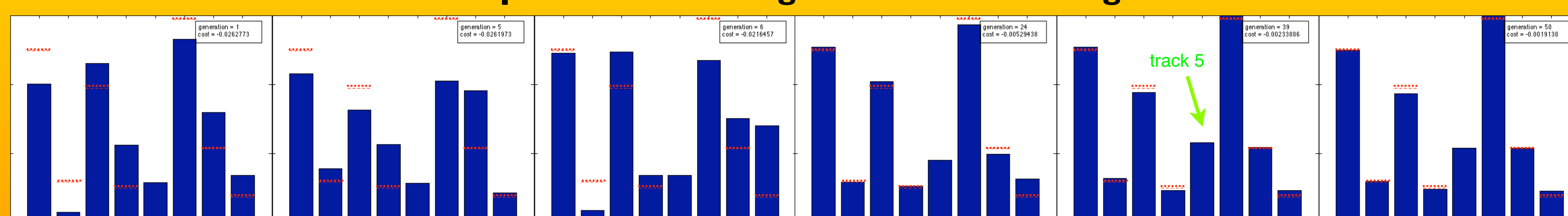
- Designed to start off with a broad search to find regions of global minima then taper off with local searches to find optimal values
- Blue bars represent the gain values applied to each track; red lines indicate target value

Optimization of Four-Track Recording



Generation: 1 3 4 31 32 50

Optimization of Eight-Track Recording

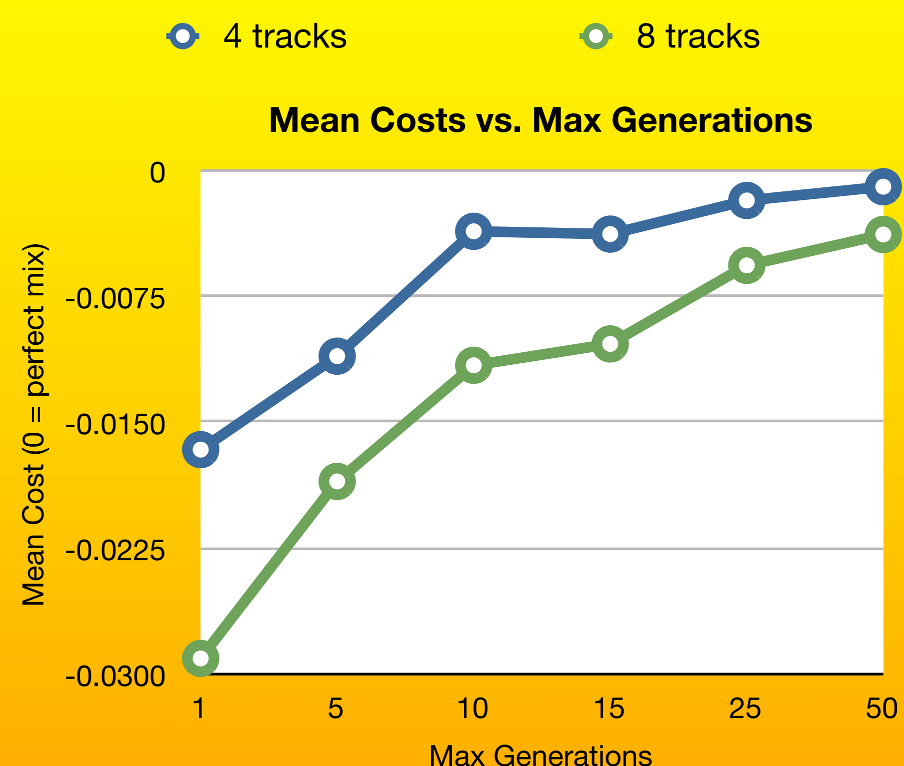


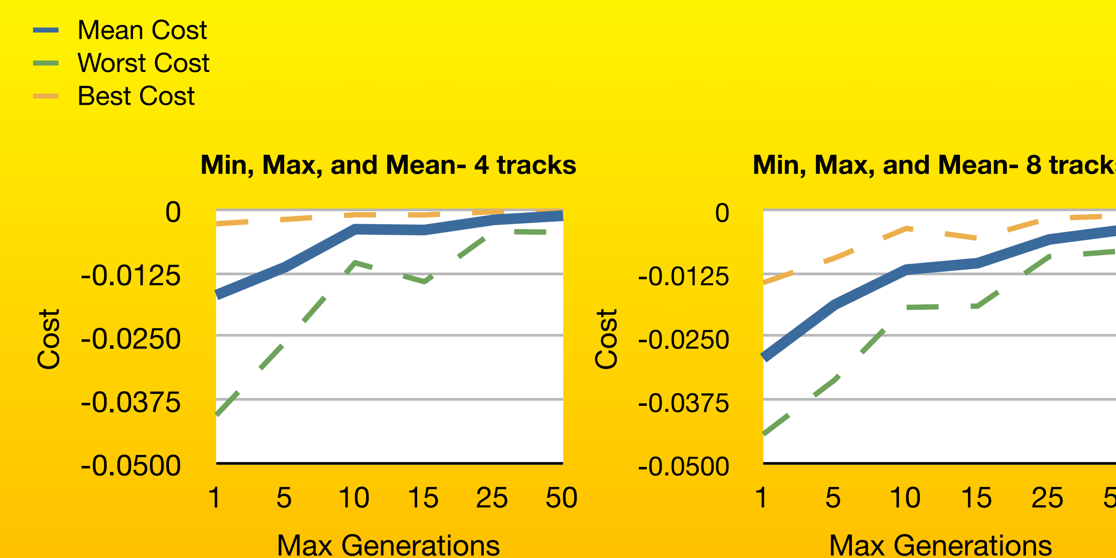
Generation: 1 5 6 24 39 50

Why did track 5 miss the target by such a large margin?
Track 5 is a snare drum track whose gain was randomly assigned to be -43dB in this test. The final effect of this track on the overall mix may have been too quiet to be perceptible by the cost function due to its low gain or its percussive nature, which may indicate a shortcoming of this approach.

System Performance

- tested the ability of the system to recreate known mixes over a varying number of tracks in the mix and number of generations that the genetic optimization runs
- as can be expected, the system performs better on collections of fewer tracks and when allowed to run for more generations





Mean Cost vs. Maximum and Minimum Costs found over 20 Runs

due to the random initialization of the genetic optimization algorithm, performance can vary widely, especially in shorter runs (fewer generations). a more intelligent initialization system could greatly help this.

Future Work

- Deal with more than just gain adjustments: panning, compression, effects...
- Combine this with a robust track-by-track initialization approach for a more complete automatic mixing system
- Subjective testing of system's performance versus human-made mixes
- Develop a collaborative repository of multitrack recordings to make the 'intelligent recording studio' a widespread effort

More Info

- The code uses two free MATLAB toolboxes:
  - **Elias Pampalk's MA toolbox** for sone and SH calculation: http://www.pampalk.at/ma/
  - **Genetic Algorithm Optimization Toolbox (GAOT)**: http://www.ise.ncsu.edu/mirage/GAToolBox/gaot/
- Code to be released as a toolbox on NYU Music Tech Research site www.nyu.edu/projects/mtr/ and on author's homepage www.bennettk.com

Works Cited

1. Pampalk, Elias et al. "On the Evaluation of Perceptual Similarity Measures for Music." Proceedings of the 6th International Conference on Digital Audio Effects (DAFx'03) 8 (2003).
2. Houck, Christopher et al. "A Genetic Algorithm for Function Optimization: A Matlab Implementation." (1996).
3. Horner, Andrew et al. "Machine Tongues XVI: Genetic Algorithms and Their Application to FM Matching Synthesis." Computer Music Journal 17.4 (2003): 17-29.
4. Tan, B.G. and Lim, S.M. "Automated Parameter Optimization for Double Frequency Modulation Synthesis Using the Genetic Annealing Algorithm". Journal of the Audio Engineering Society 44.1/2 (1996): 3-15.