

Sound Innovation Engine 1.0

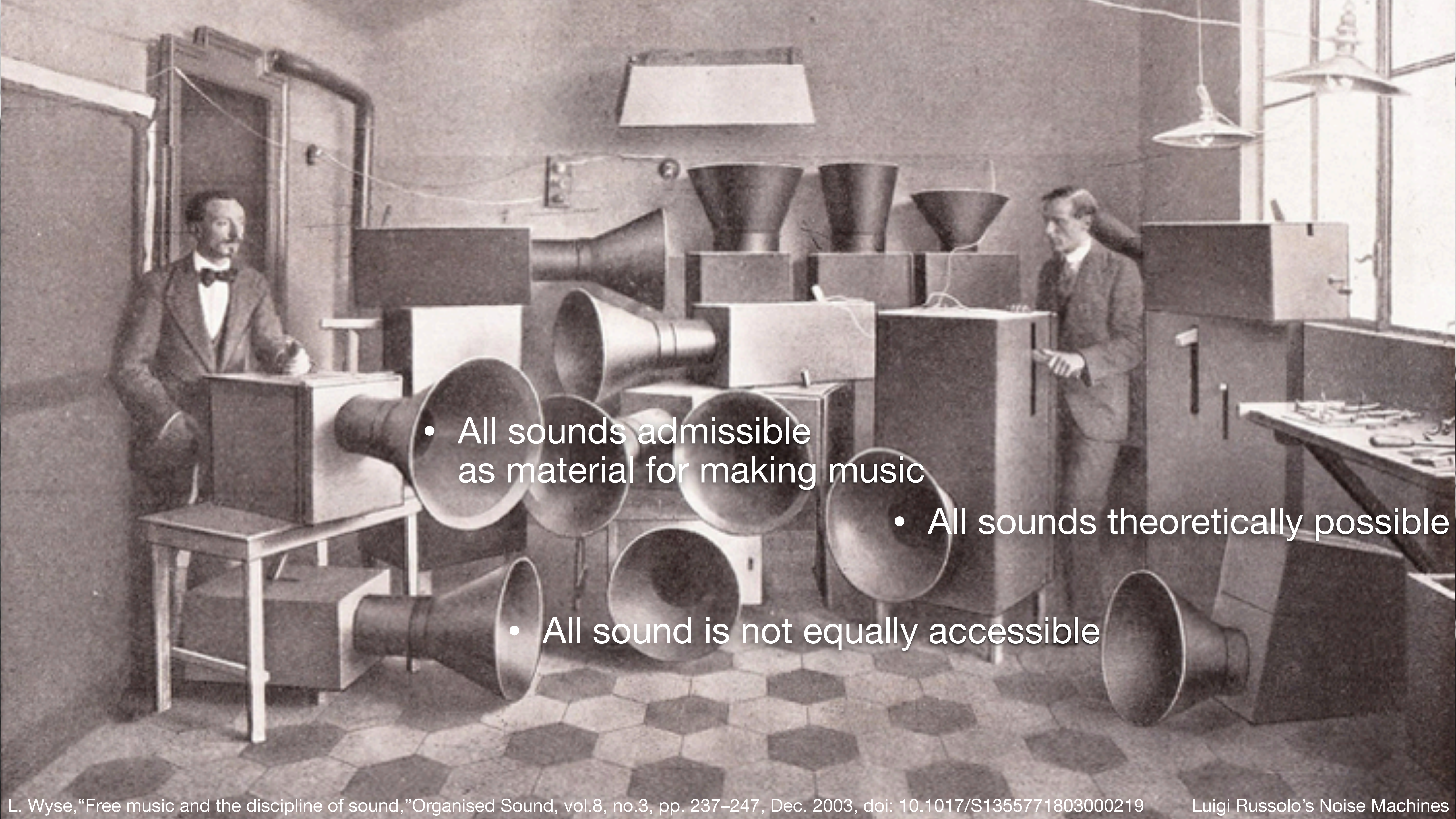
Towards Sound Innovation Engines

Using Pattern-Producing Networks and Audio Graphs

Björn Þór Jónsson, Çağrı Erdem, Stefano Fasciani and Kyrre Glette, January 2024



UiO • RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion
University of Oslo



- All sounds admissible as material for making music

- All sounds theoretically possible

- All sound is not equally accessible

**recognising sounds
you've never heard**

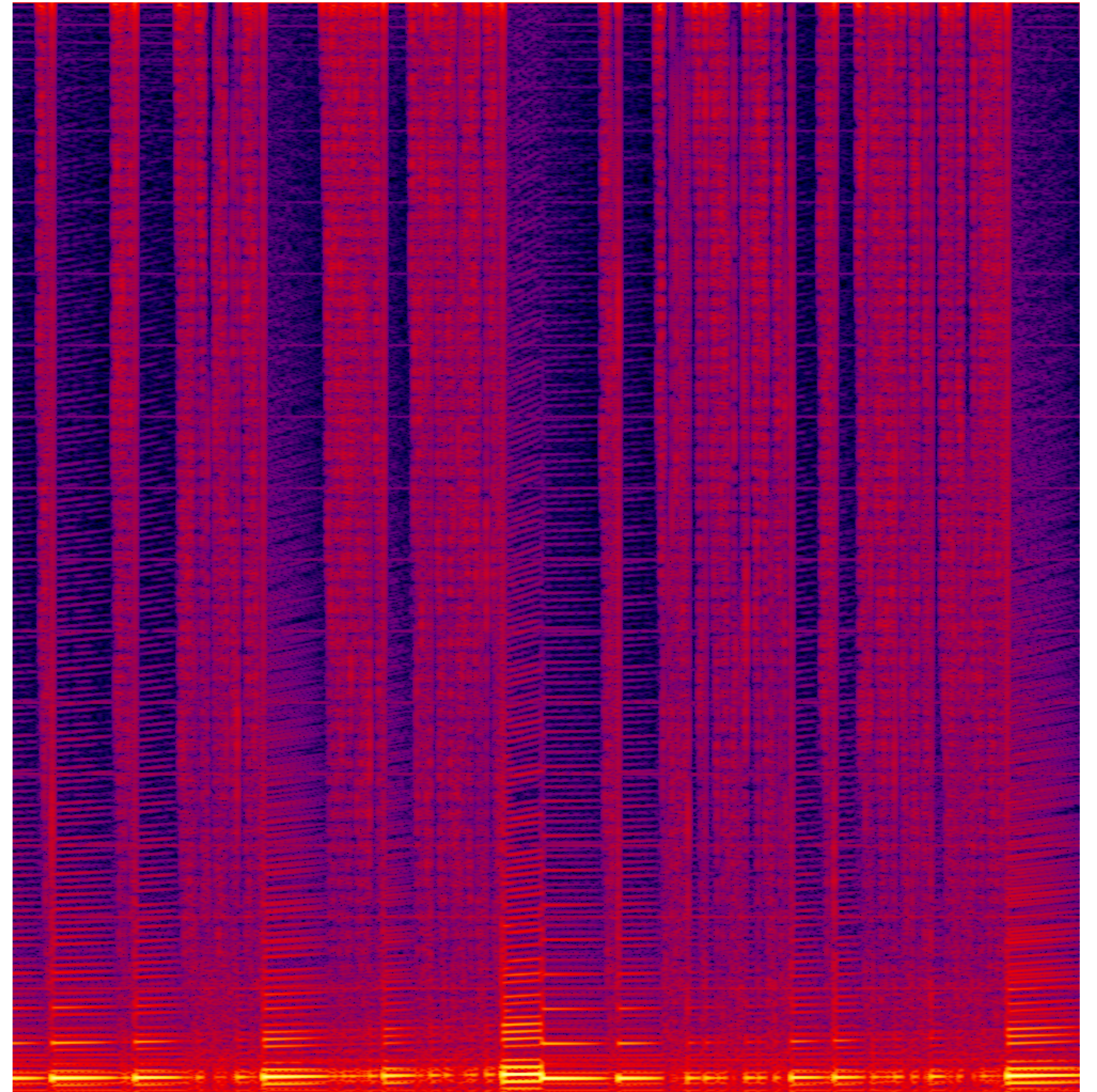
**finding sounds recognised as
pleasing but unfamiliar**

**Quality Diversity
to discover stepping stones
with goal switching
on a path to greatness**

Innovation Engines

Automate QD exploration with a model capable of:

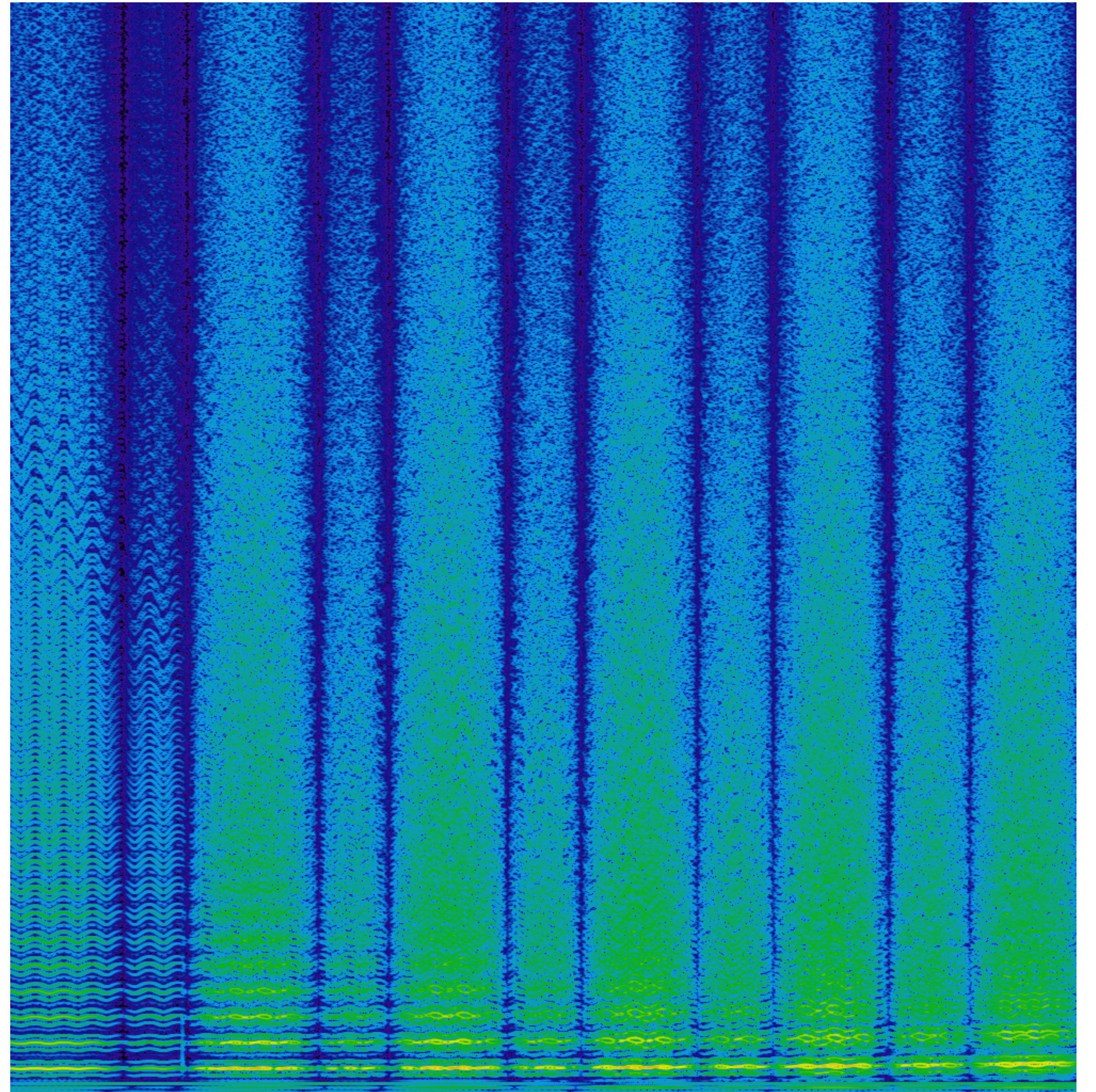
- distinguishing novelty
- evaluating quality



Innovation Engines

Ultimate goal:

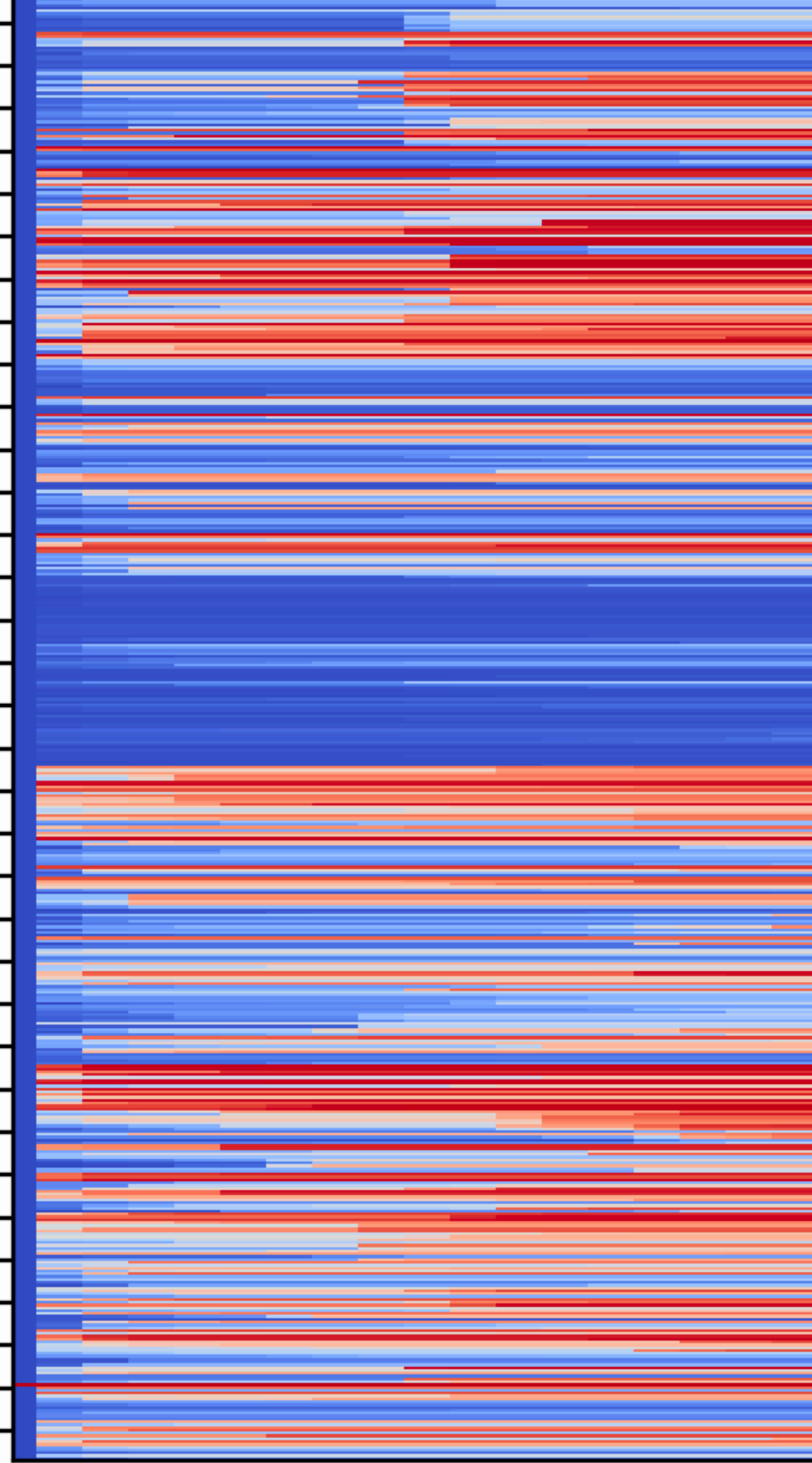
- Unsupervised classification
- Produce new types of things



Behavioural Descriptor

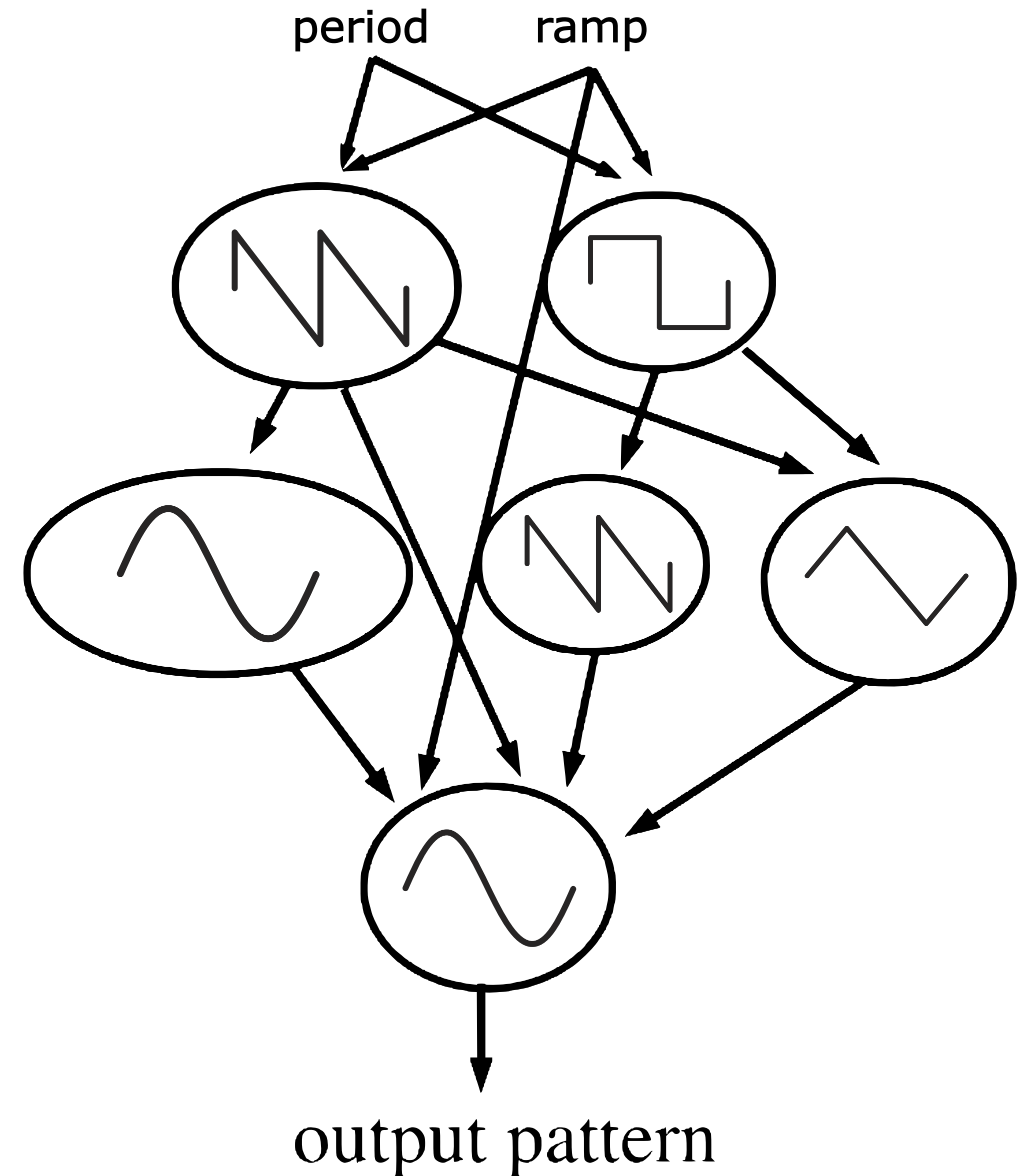
YAMNet

Giggle
Synthetic singing
Sniff
Heart murmur
Whimper (dog)
Goat
Roar
Patter
Guitar
Organ
Tabla
Brass instrument
Bell
Scratching (performance technique)
Swing music
Electronic dance music
Music of Asia
Wedding music
Rain on surface
Motor vehicle (road)
Bus
Aircraft engine
Engine starting
Chopping (food)
Scissors
Siren
Sewing machine
Explosion
Glass
Boiling
Bouncing
Creak
Sine wave
Mains hum

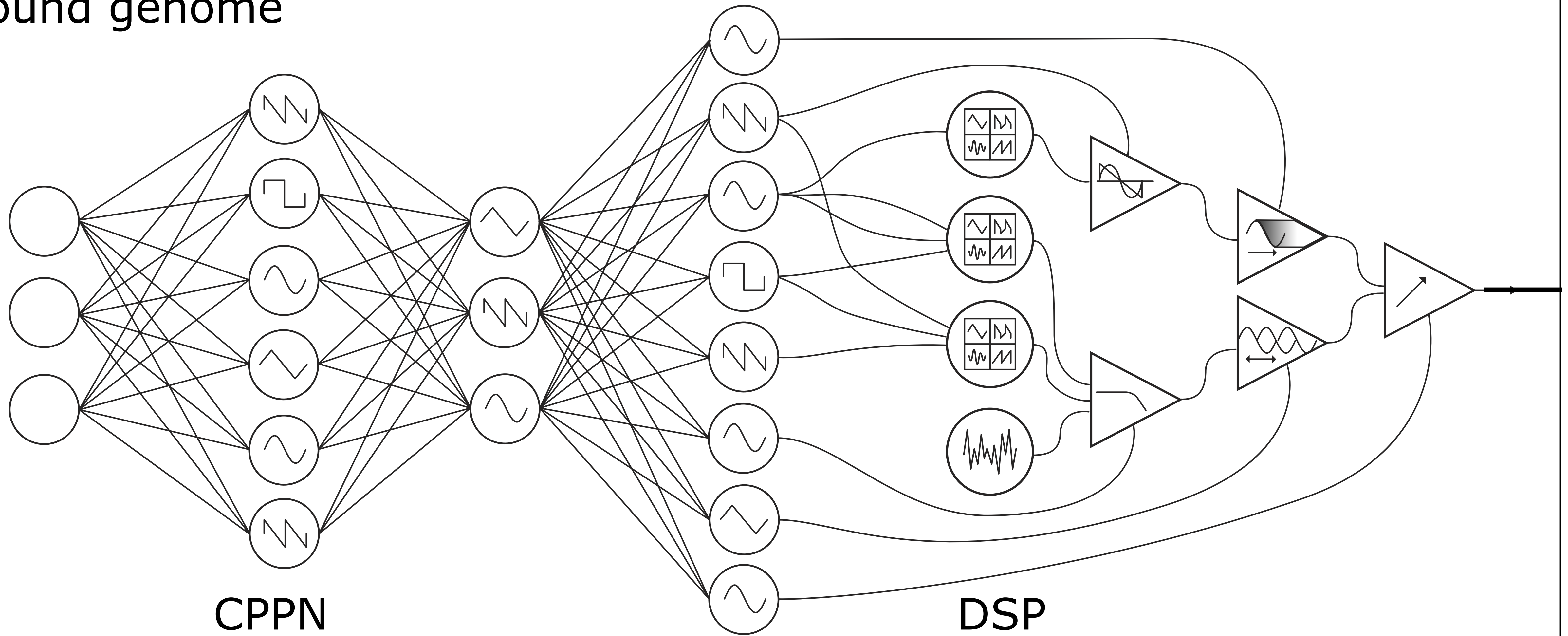


Compositional Pattern Producing Networks CPPNs

- Abstract unfolding development in evolutionary processes
- Applied to timbral development
 - Combined with DSP graphs



sound genome



CPPN

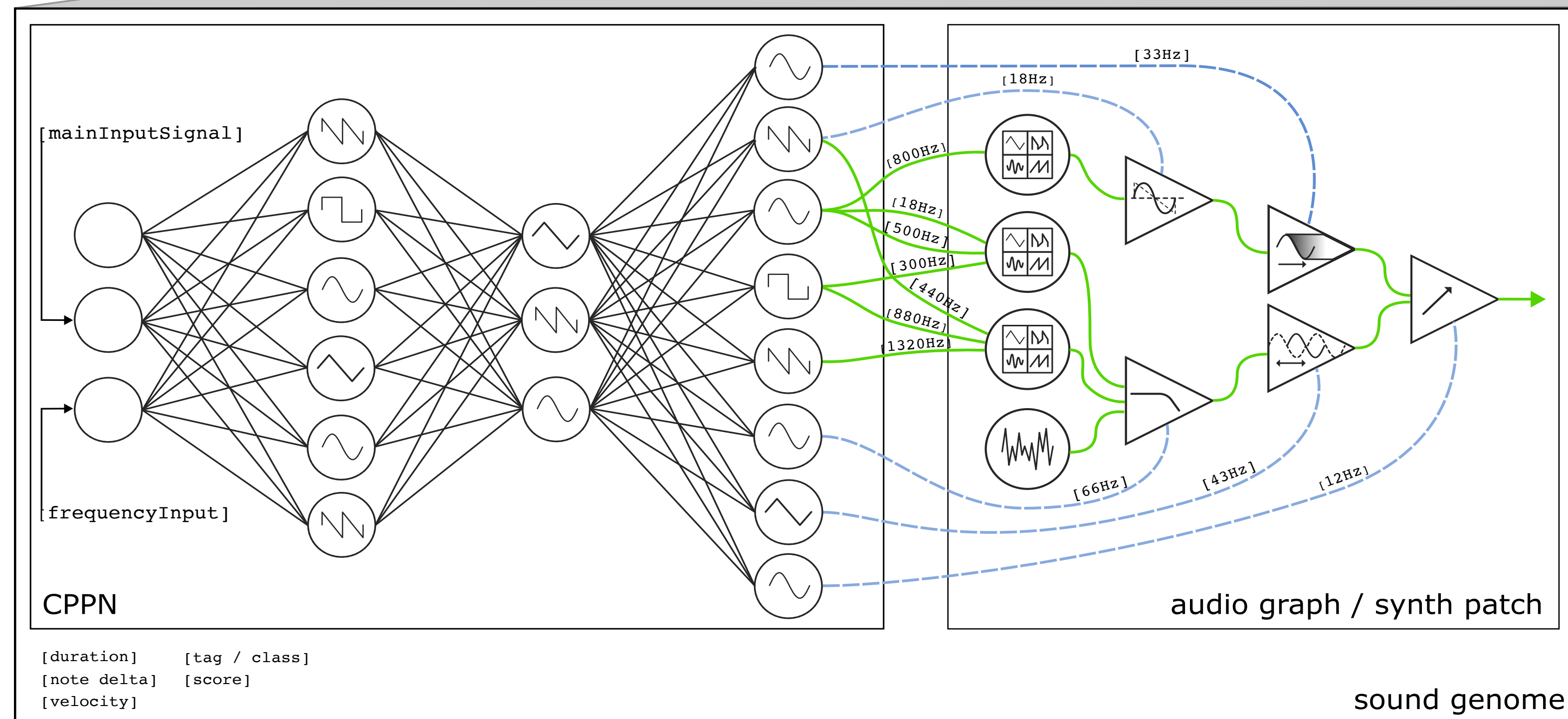
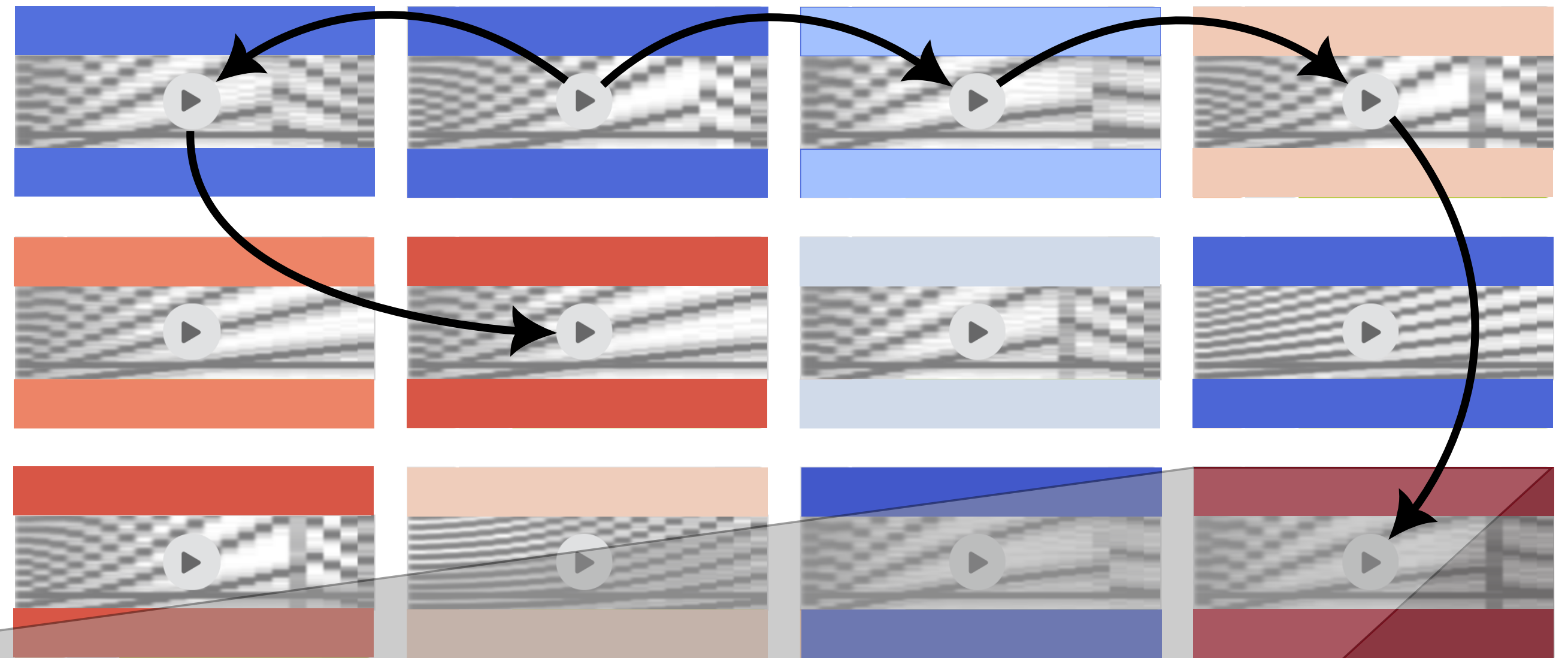
DSP

Signal Composition

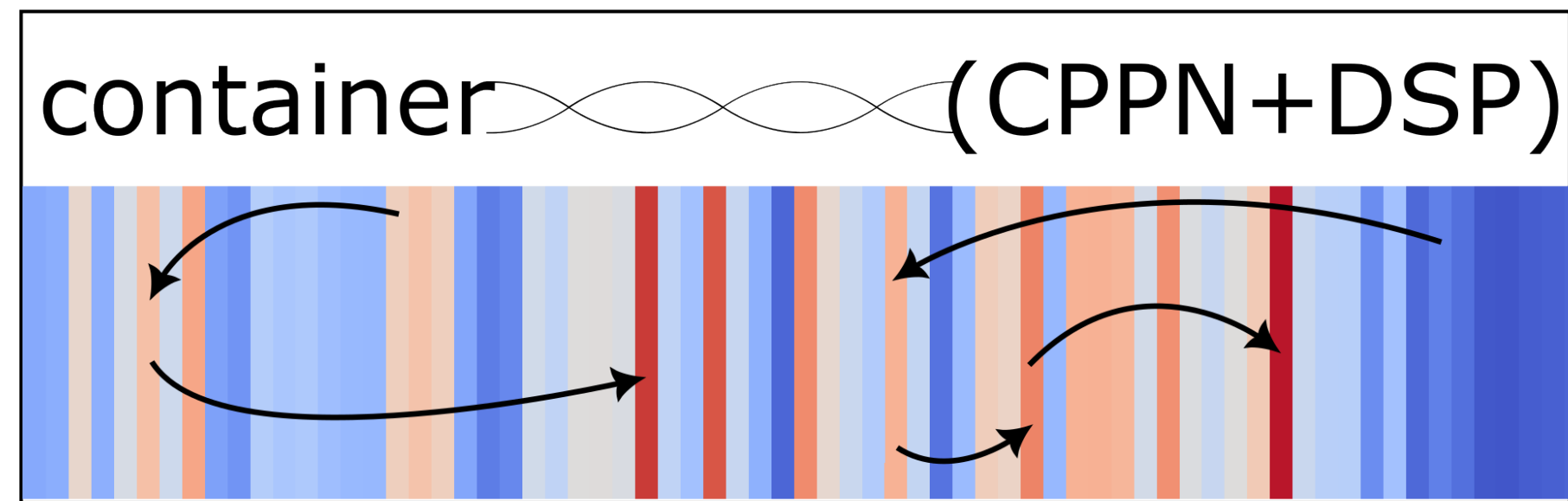
CPPN + DSP

QD algorithm

MAP-Elites

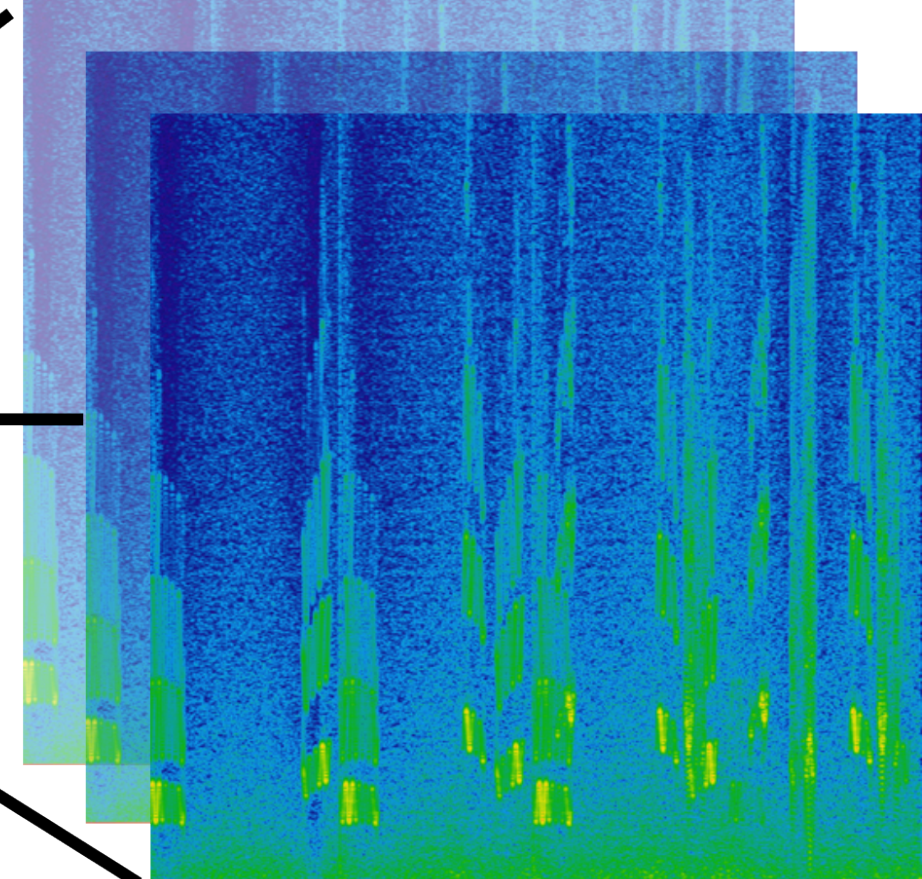
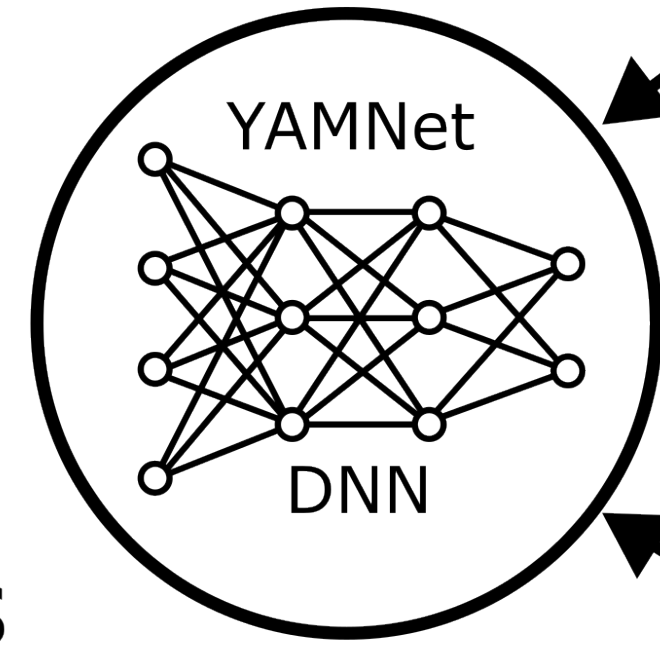


MAP-Elites



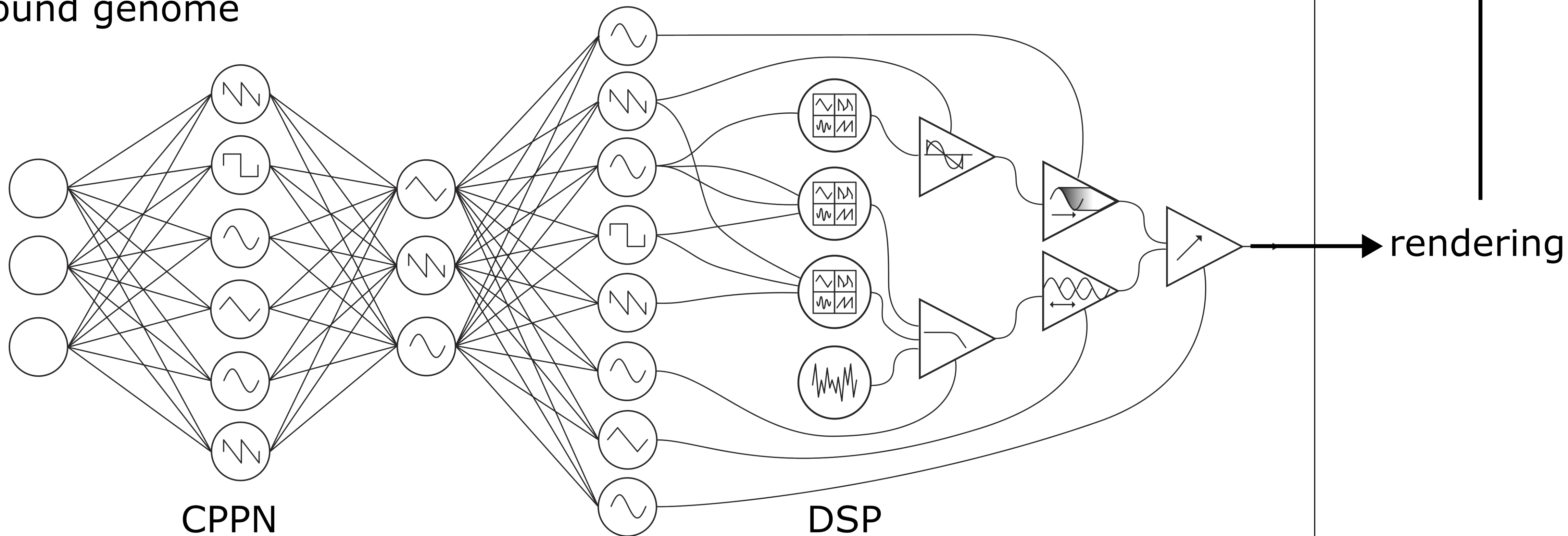
map and insert higher-scoring solutions

evaluation



selection + variation

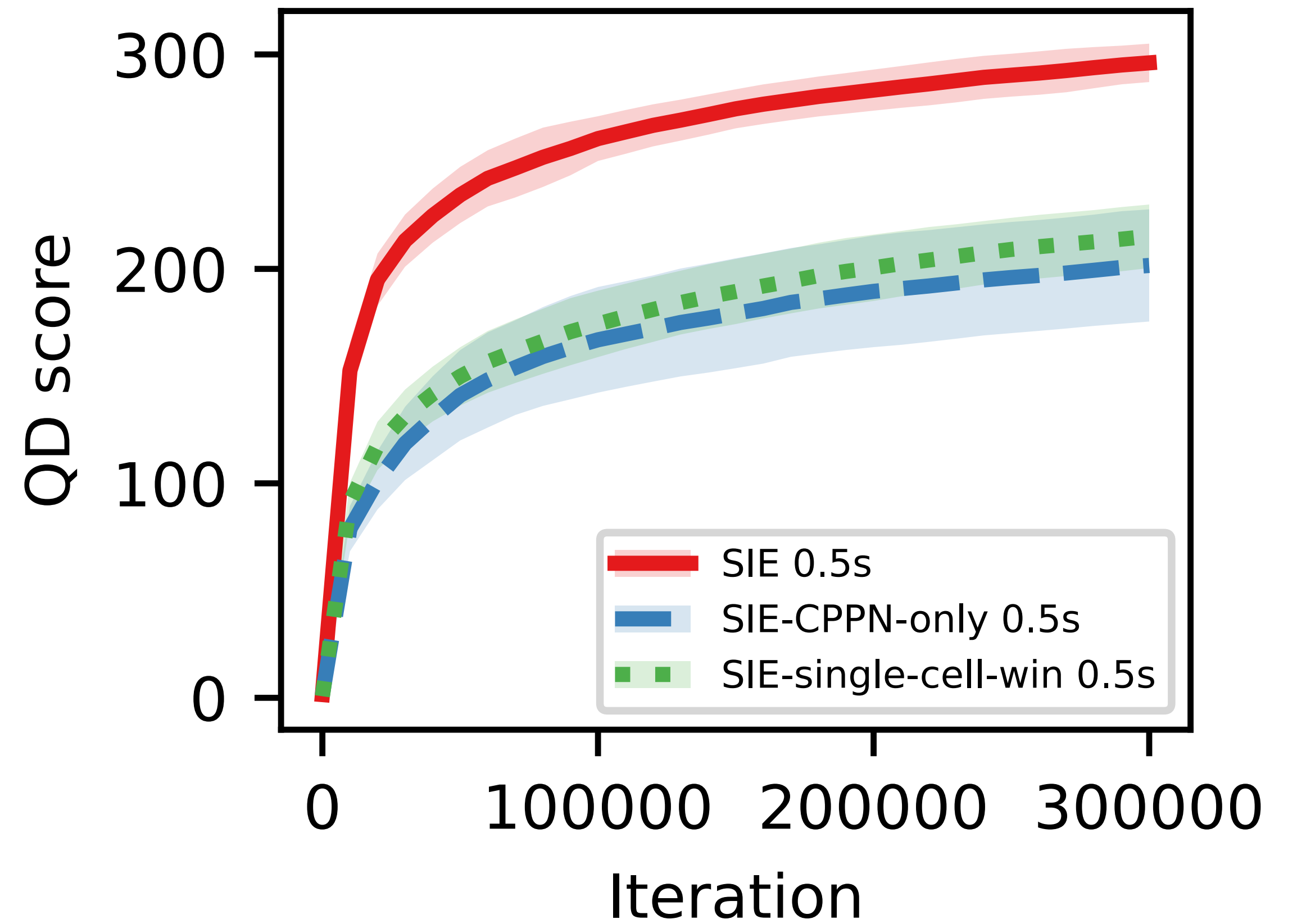
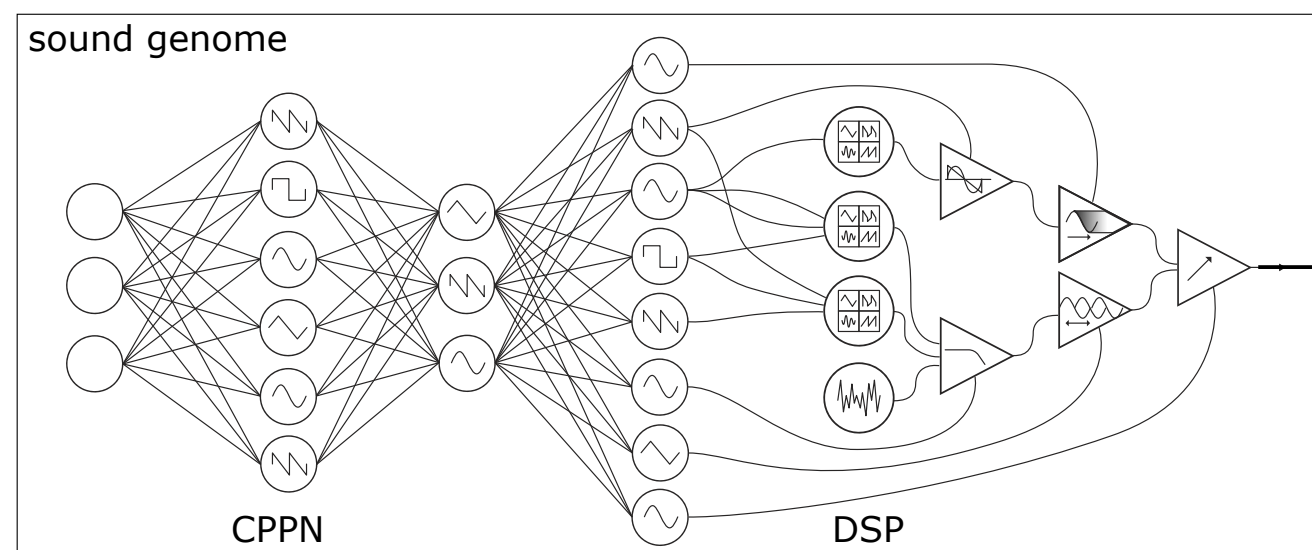
sound genome



Results

Signal Processing Graph

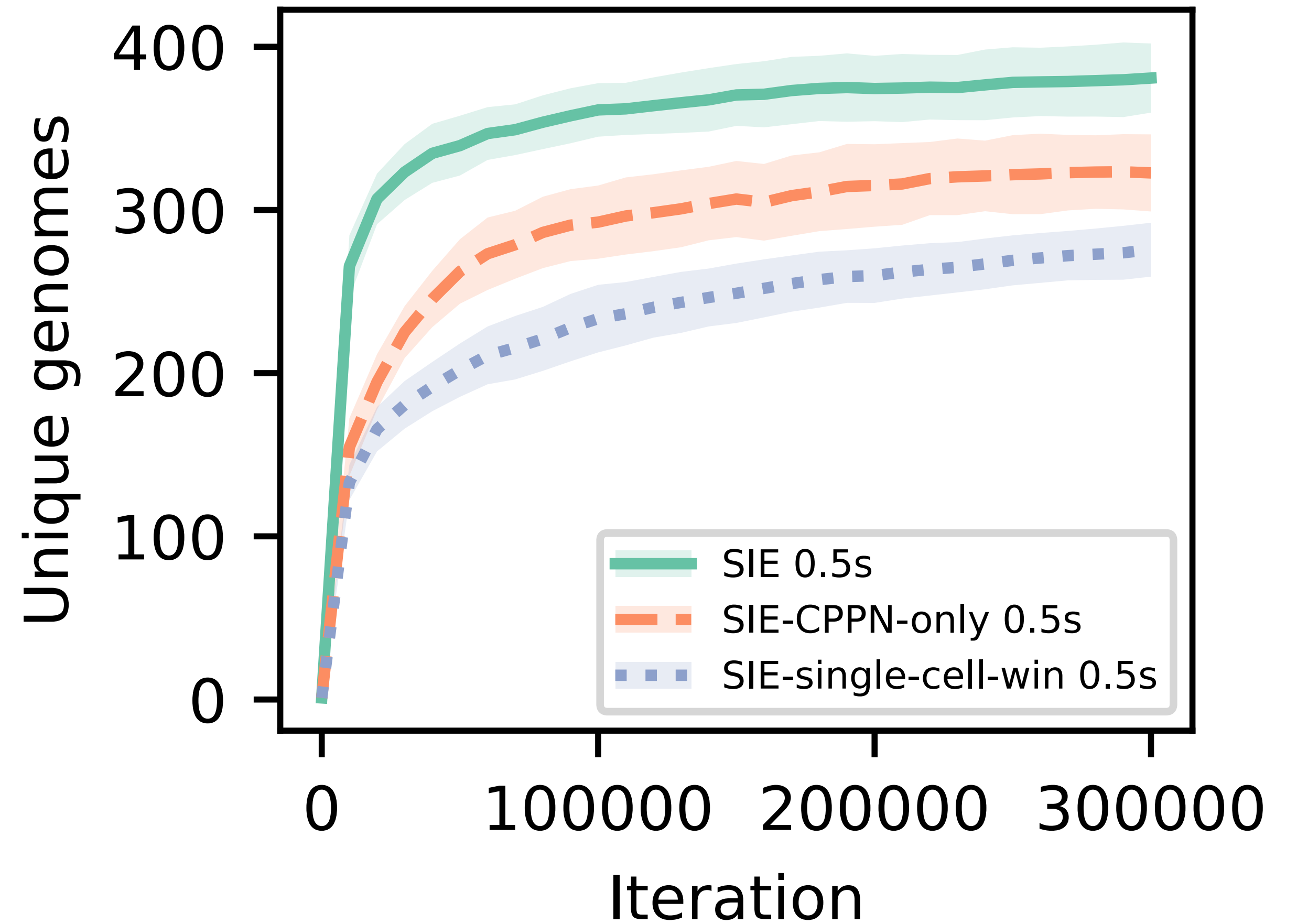
- QD-score:
 - CPPN + DSP
 - VS
 - CPPN only



Results

Elite Populations

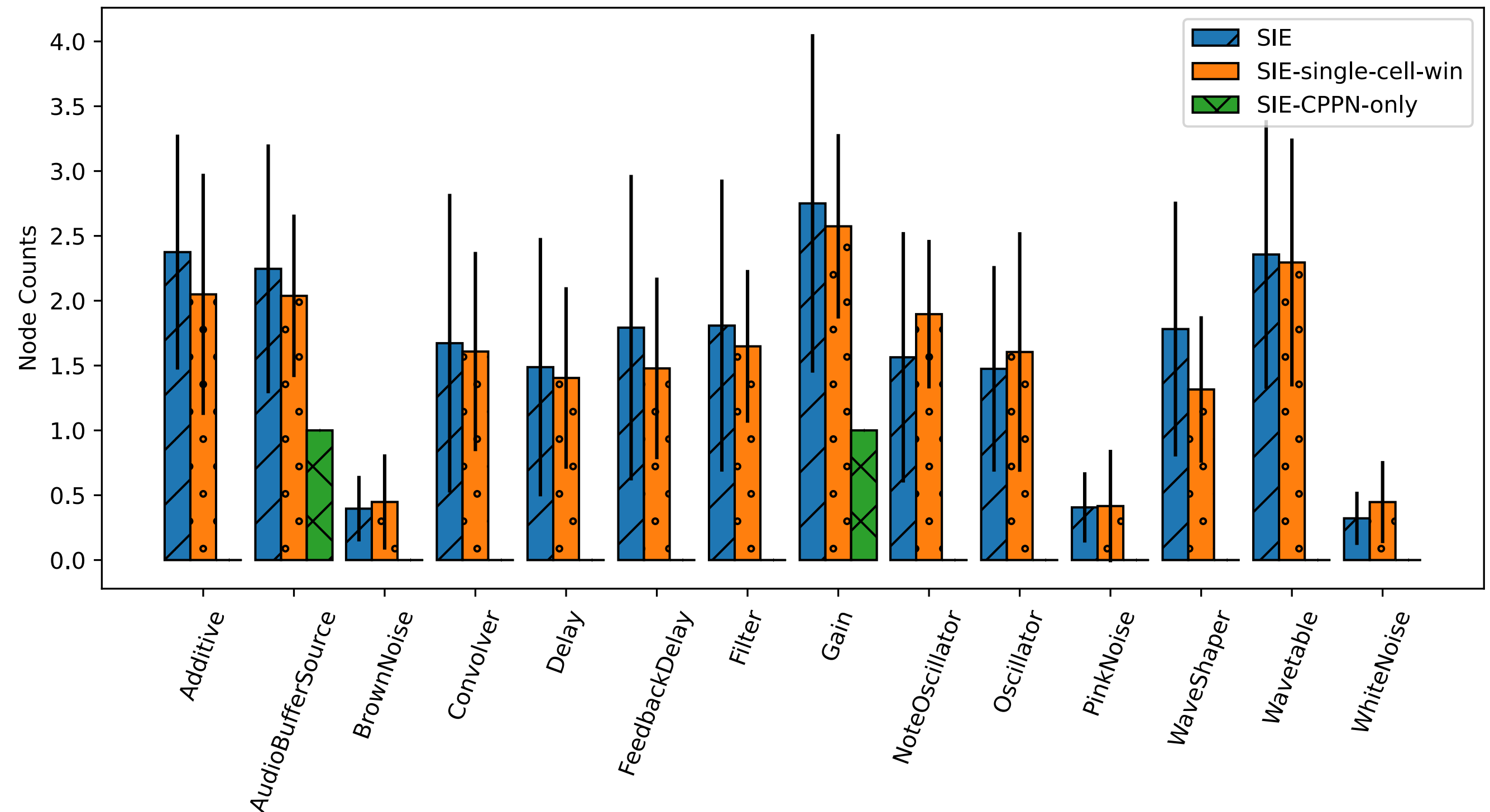
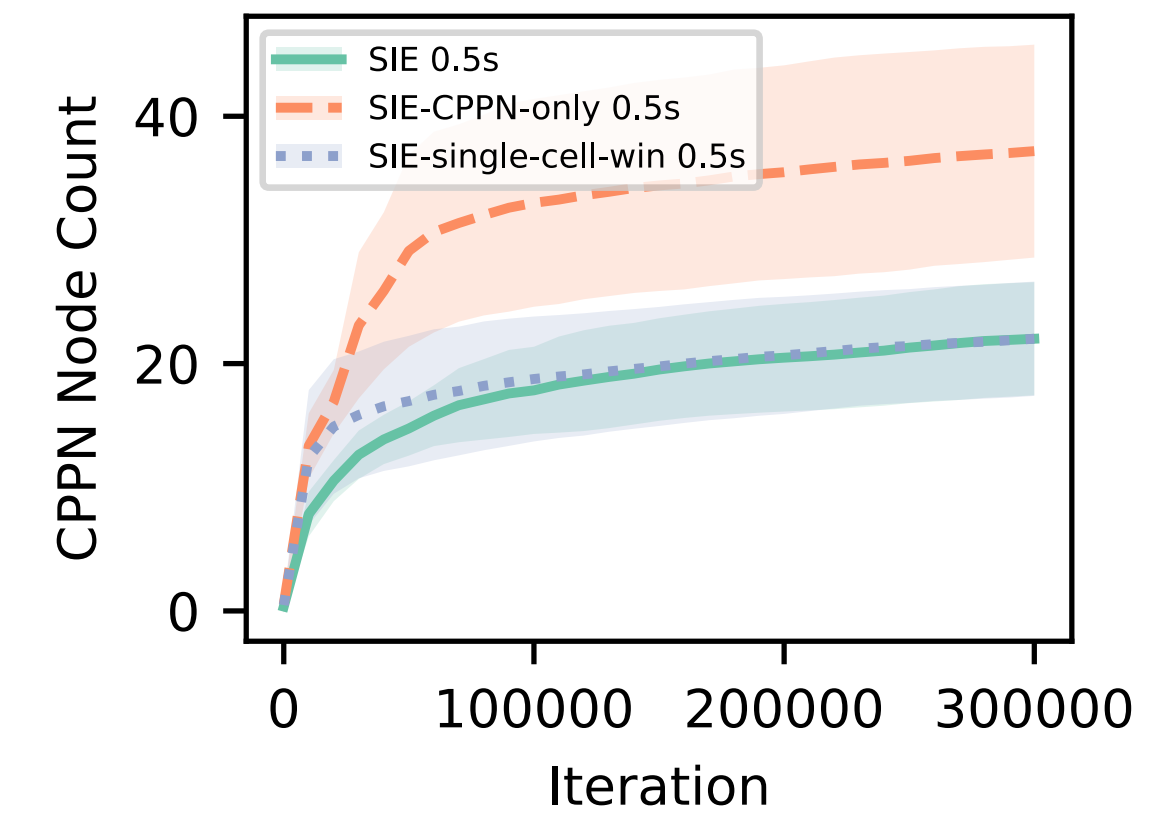
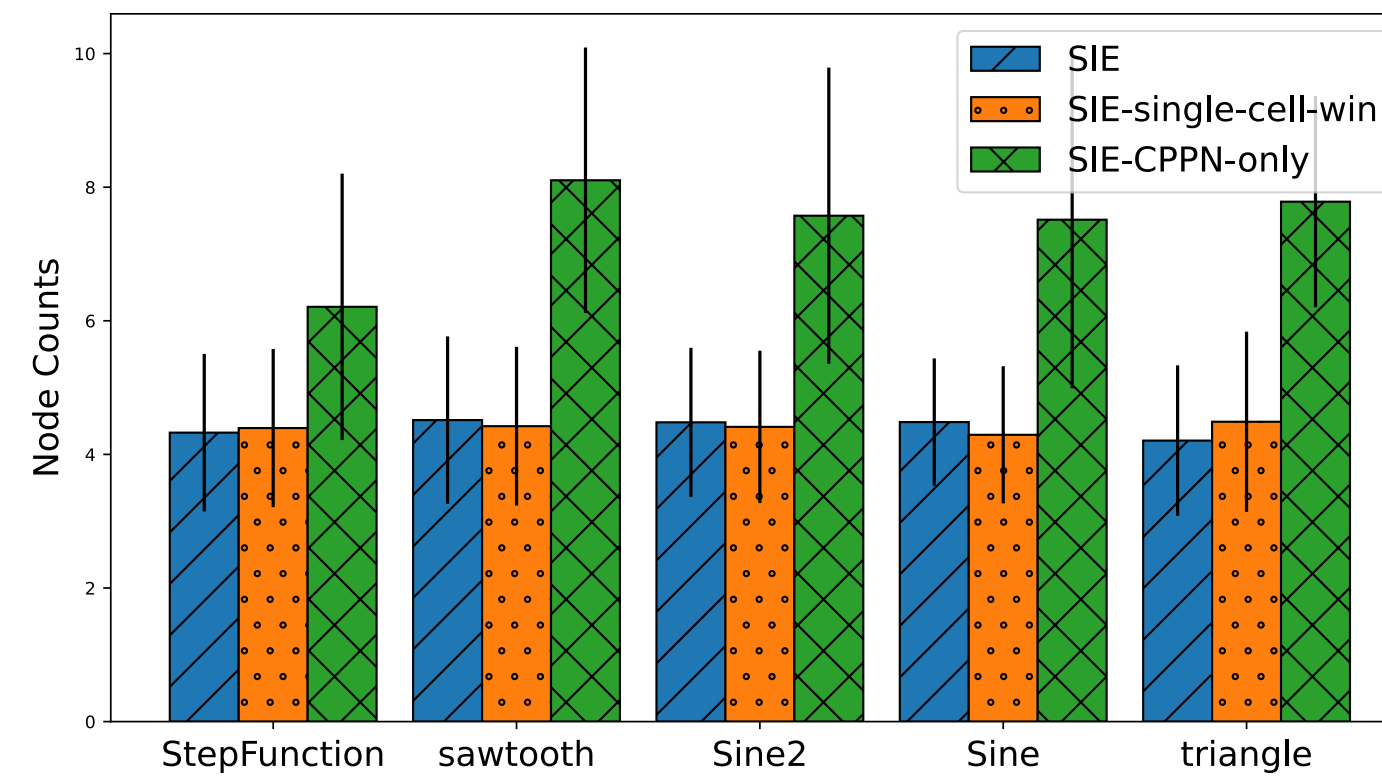
- More diversity when evaluating sound objects from CPPN + DSP genomes



Results

Genome Complexity

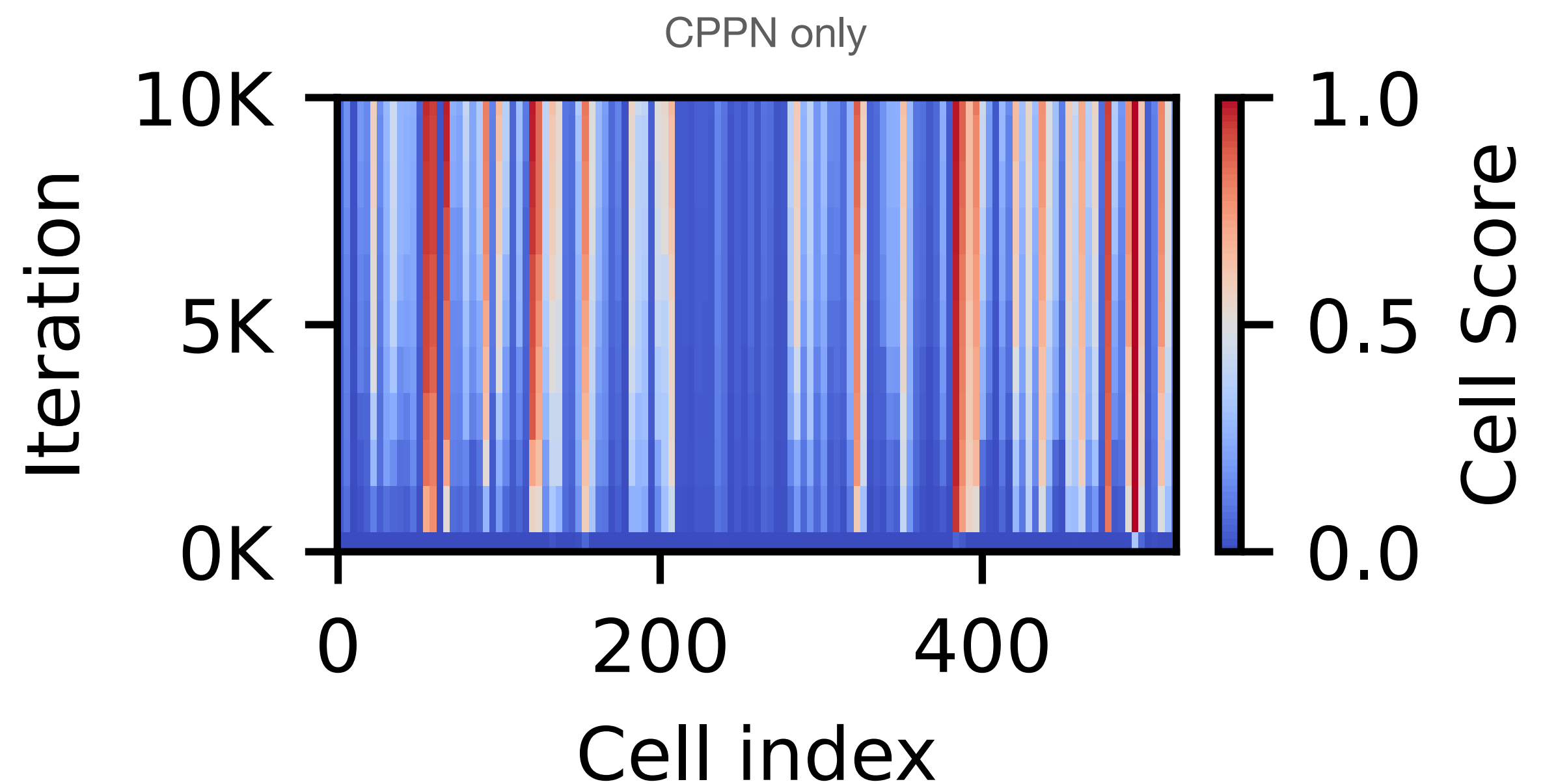
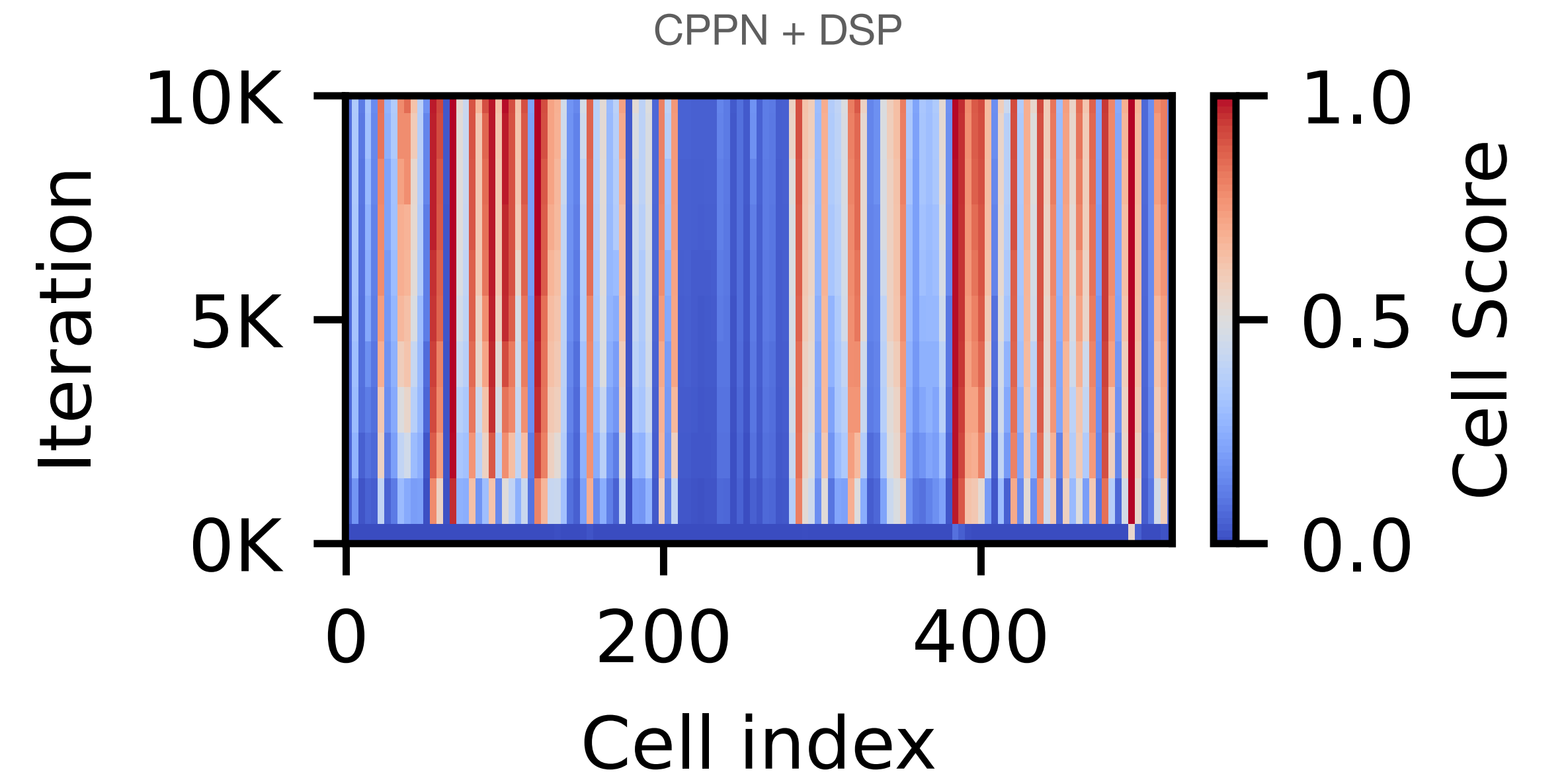
- Uniform set of CPPN activation functions
- Custom wavetable and additive synthesis DSP nodes prominent
- CPPN only genomes with higher CPPN node count
- Compensating for the lack of DSP?



Results

Performance Against Pre-trained Reward Signals

- High scores across most classes
- CPPN + DSP higher overall
- Synthesiser struggles with scoring high on musical classes
 - Understandably?

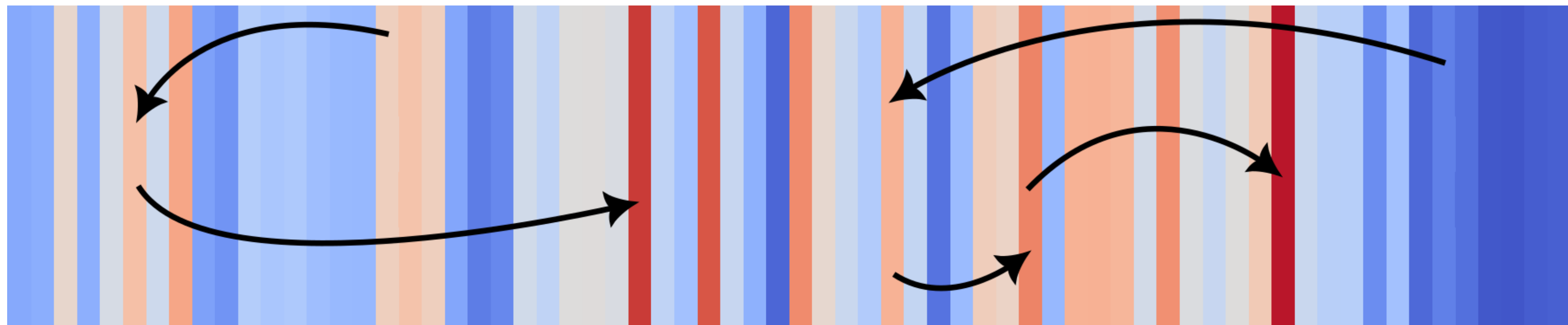


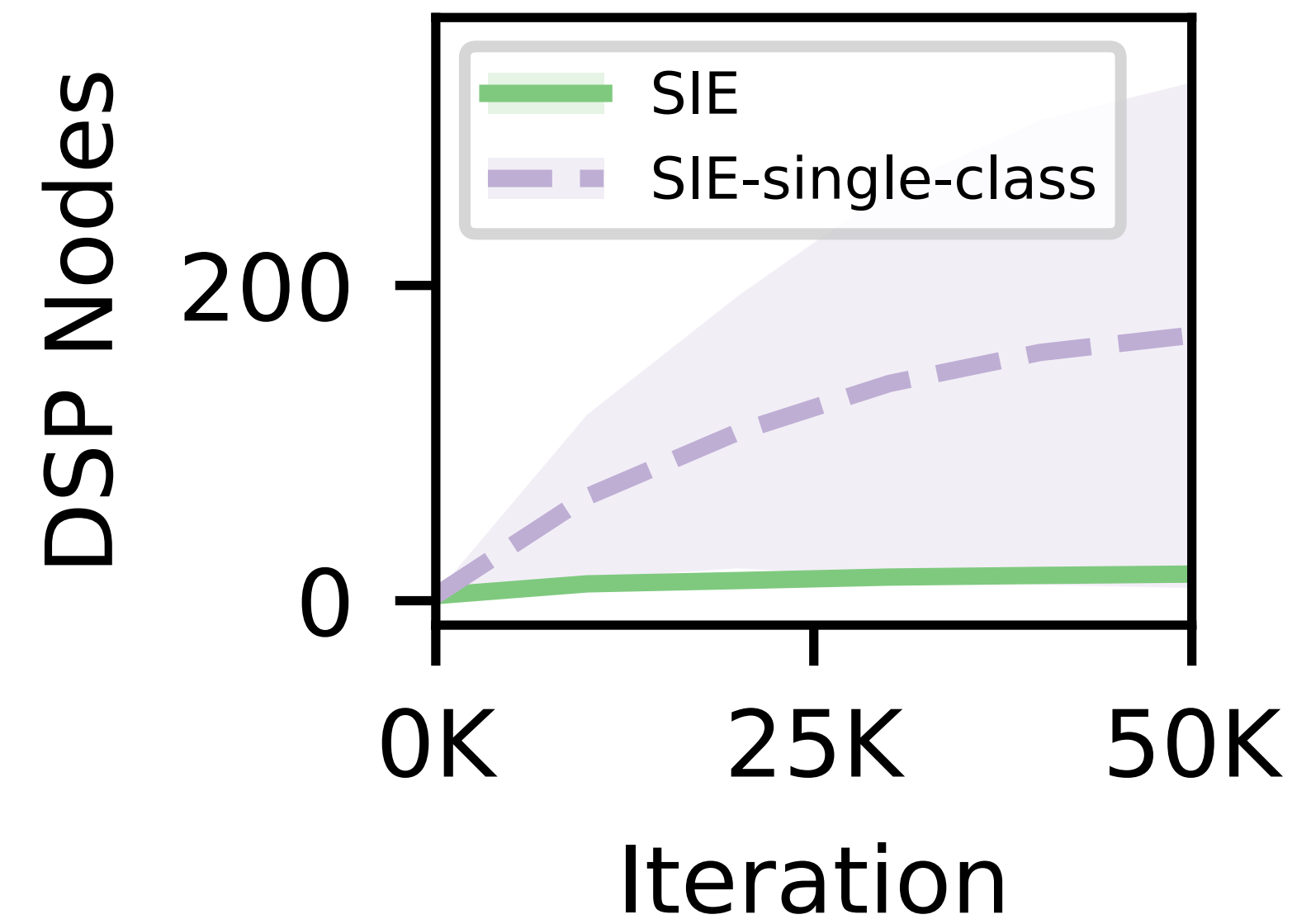
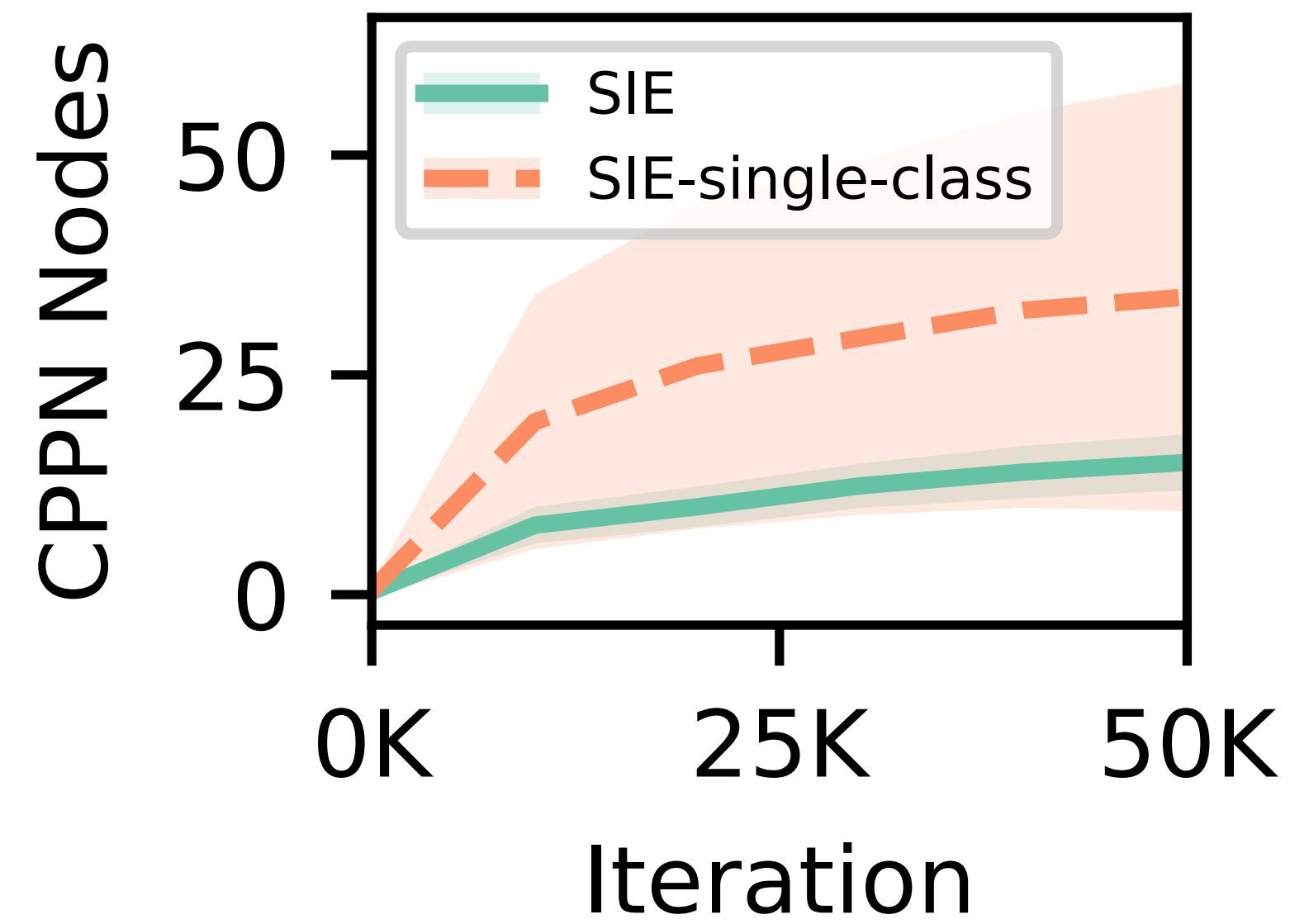
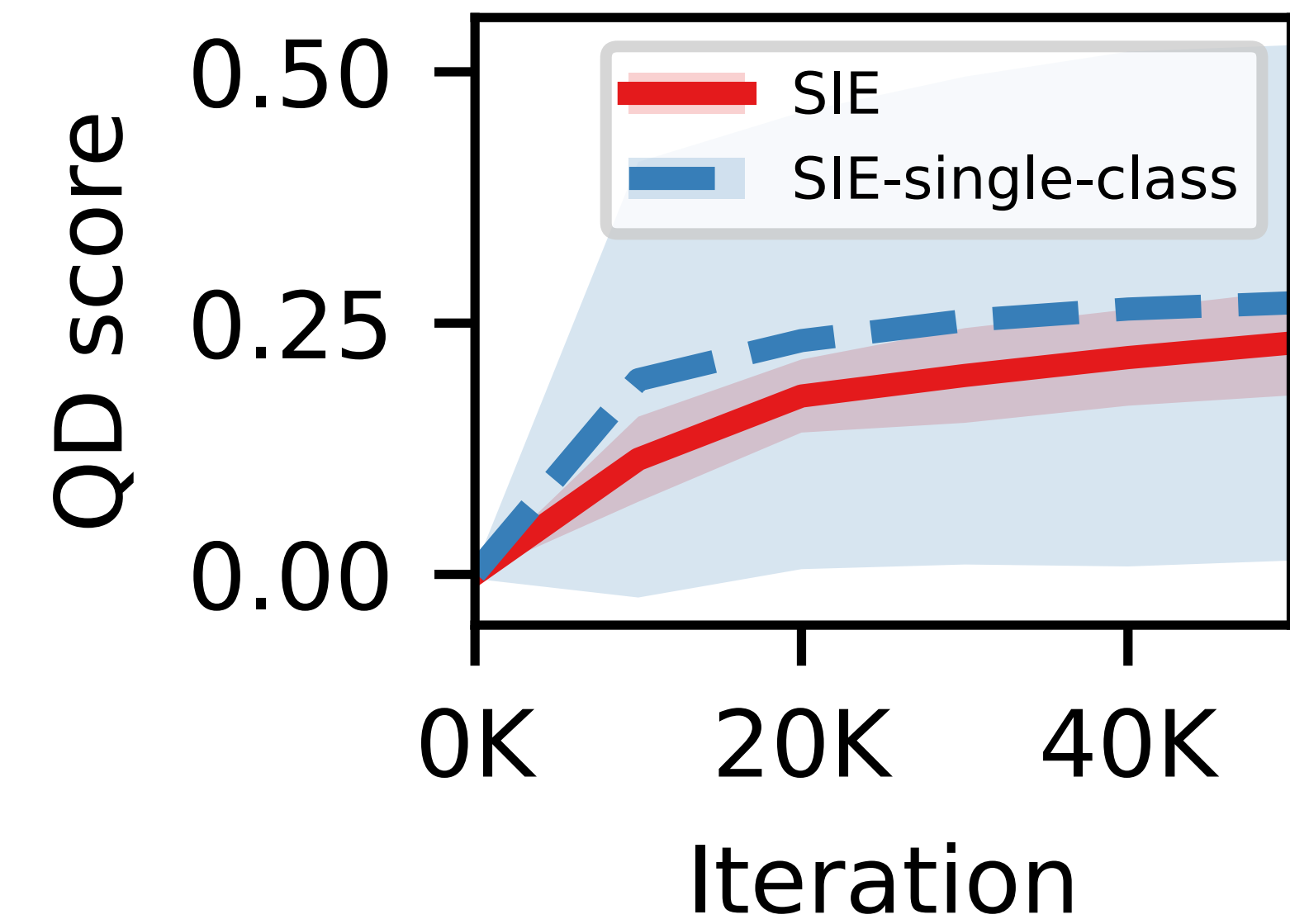
stepping
stones

21.7 ± 3.6

goal switches

63.2% of the 34.3 ± 4.5 mean new champions per class





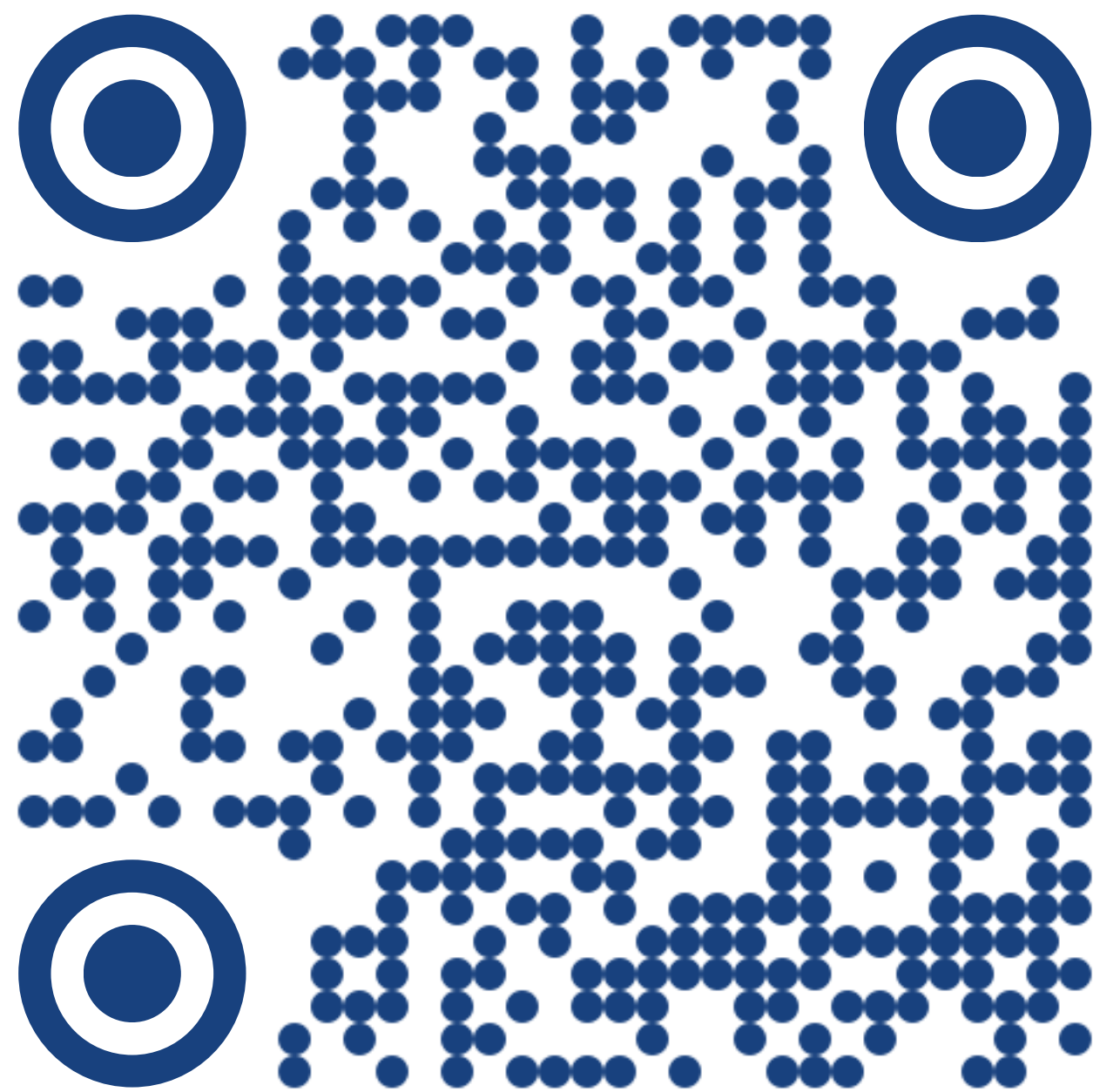
Abandoning Diversity

Similar performance at the cost of significantly higher complexity

Sound Objects and their Application

Evoruns Explorer

- Scrub through evolutionary runs, classes and generations



synth.is/exploring-evoruns?evoRunDirPath=/QD-Fox/evoruns/c

Evolutionary runs interface showing controls for scrubbing through runs, classes, and generations.

Evolutionary run: evorun 156 / 329: 01HCHSSVDBK0ATN8T36MZJCZX7_one_comb-dur_0.5

class: Fart (parent class: Fart) | score: 100% | Filter...

generation: 472968

automatic playback: < Auto >> Rnd 1C

manual interaction: Play, Refresh, Close, Heart, Download, Reverse, Anti-aliasing, Apply frequency updates to control signals

favourite 0 / 0

duration: 0.5

pitch: 0

velocity: 1



Evolving

Favorite

Live coding

Published

Exploring evolution runs

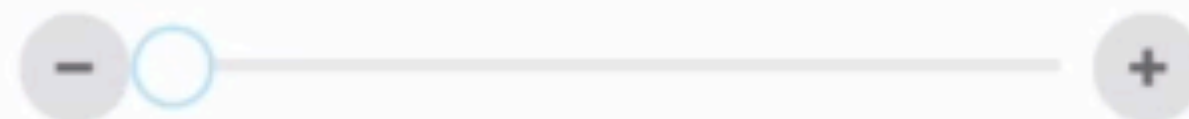
evorun 71 / 87:

...01H0C8NC7T8ZPNFFG0DNB4JTJ9_1dur-1nd



class: Violin, fiddle (parent class: Bowed string instrument) | score: 17% |

violin



generation: 10456



automatic playback



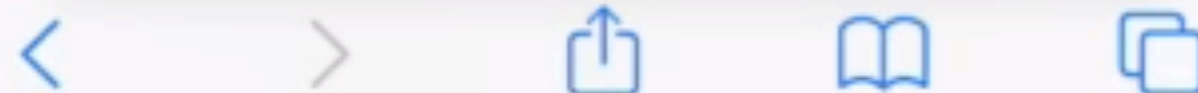
manual interaction



Reverse

AA

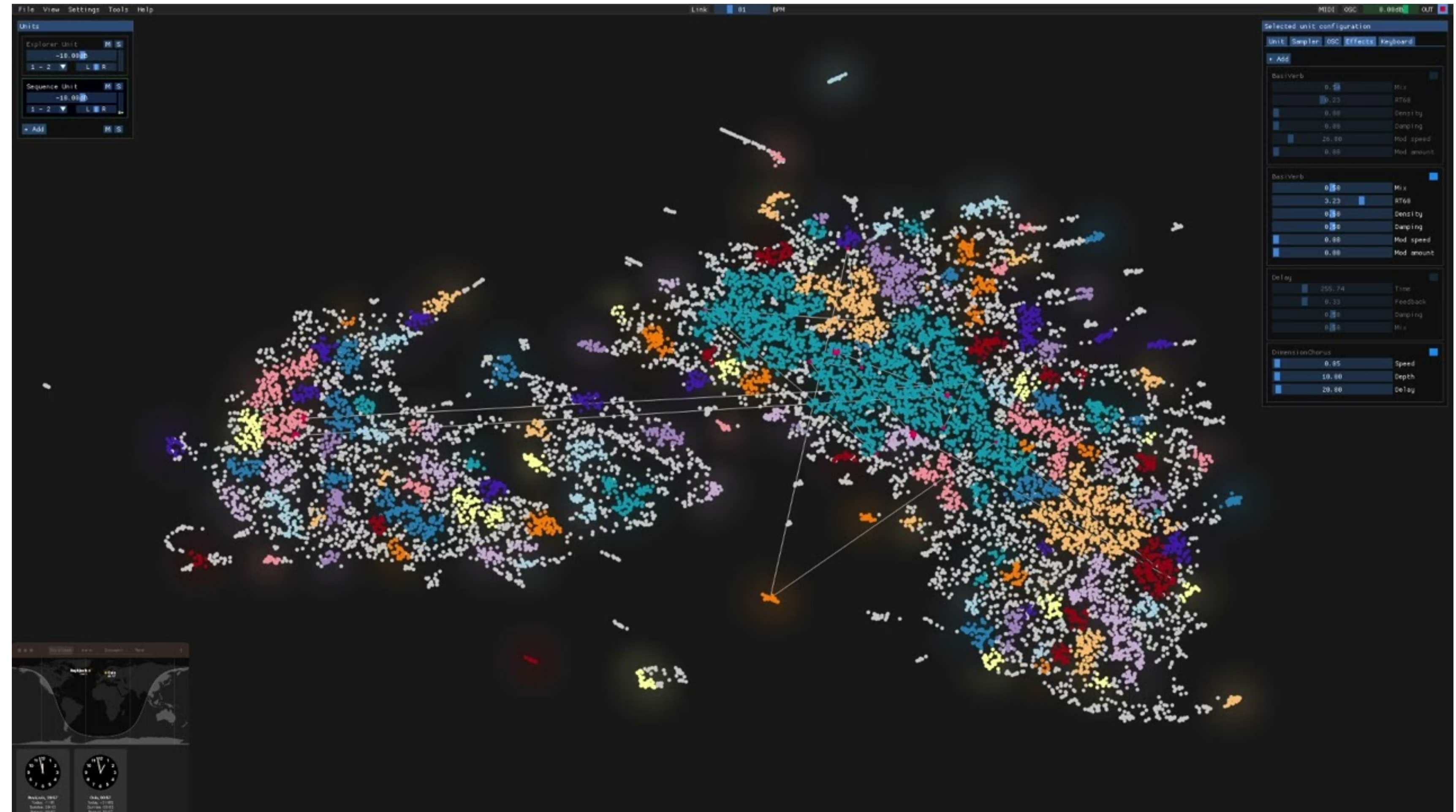
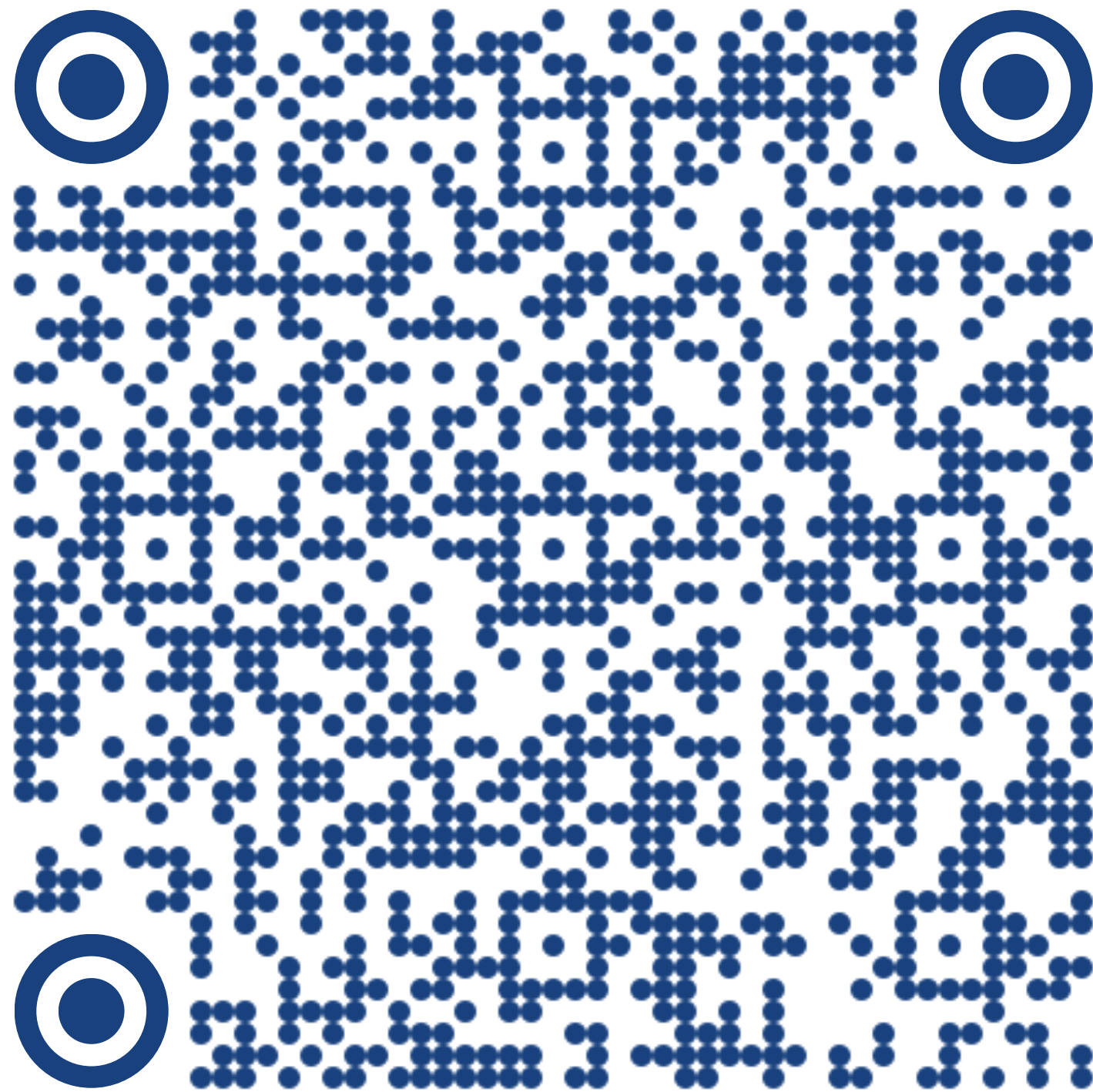
synth.is



Sound Objects and their Application

Evolutionary Sequences

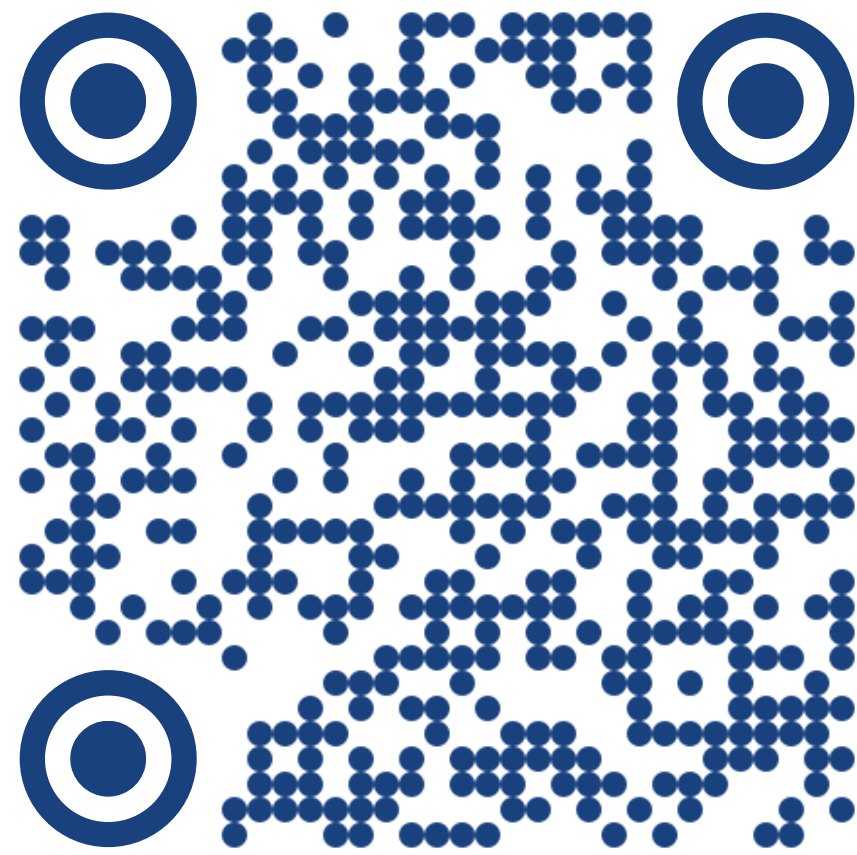
- Live-stream playlist:



Sound Objects and their Application

Evolutionary Sequences

- Some of this stuff is actually on main-streaming services (WIP):



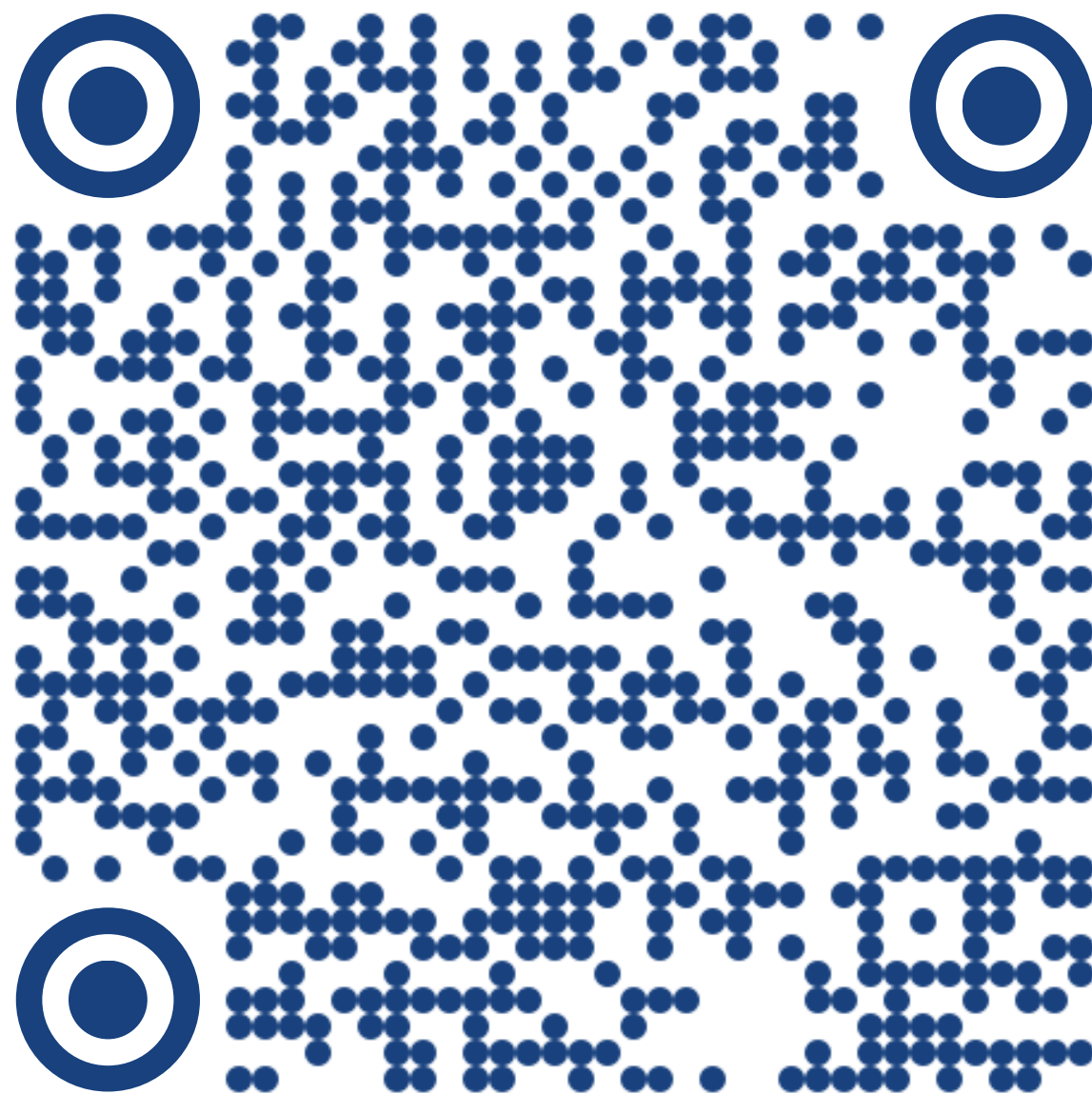
- IndieWeb publication planned

#	Title	Plays
1	Phytobenthos 1.5.00 aflands	15:19
2	Phytobenthos 1.5.01 aflands	3:20
3	Phytobenthos 1.5.02 aflands	2:25
4	Phytobenthos 1.5.03 aflands	18:45
5	Phytobenthos 1.5.04 aflands	9:04
6	Phytobenthos 1.5.05 aflands	1:08
7	Phytobenthos 1.5.06 aflands	7:37
8	Phytobenthos 1.5.07 aflands	5:54
9	Phytobenthos 1.5.08 aflands	1:52
10	Phytobenthos 1.5.09 aflands	4:31
11	Phytobenthos 1.5.10 aflands	15:02
12	Phytobenthos 1.5.11 aflands	3:50
13	Phytobenthos 1.5.12 aflands	2:42

Sound Objects and their Application

meat machines vs silicone machines

- in: 🦾 *organising sound*

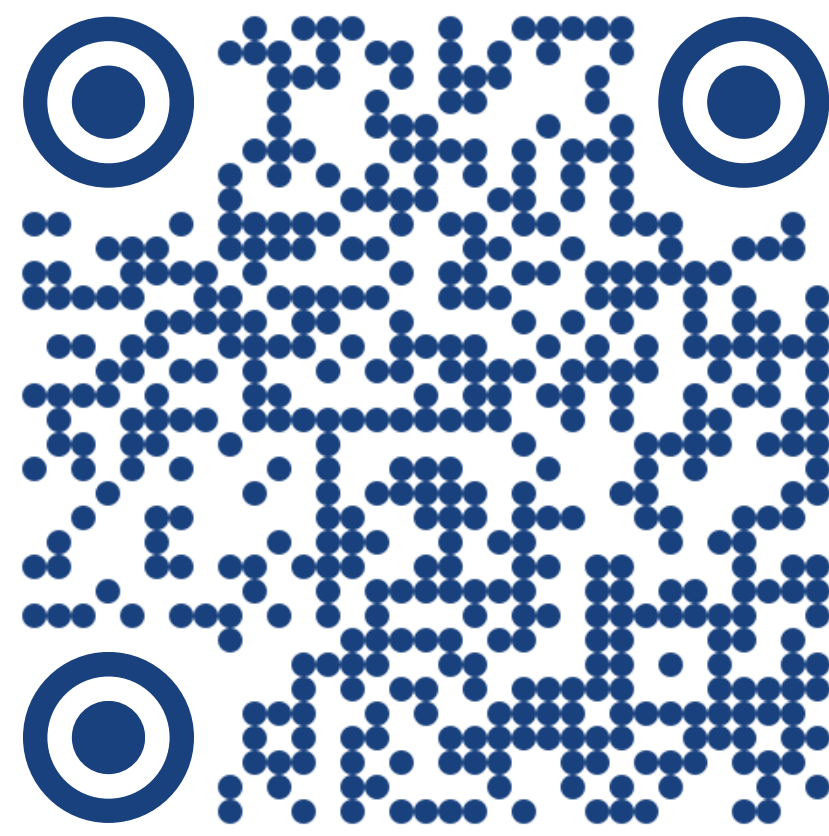


- mail results to: bthj@uio.no

A screenshot of the DataverseNO dataset viewer interface. The browser address bar shows the URL 'dataverse.no/dataset.xhtml?persistent...'. The page title is 'DataverseNO'. There are tabs for 'Files', 'Metadata', 'Terms', and 'Versions'. Below the tabs, there is a 'Change View' section with 'Table' and 'Tree' buttons. The 'Tree' button is selected, and a file tree is displayed. The tree structure is as follows:

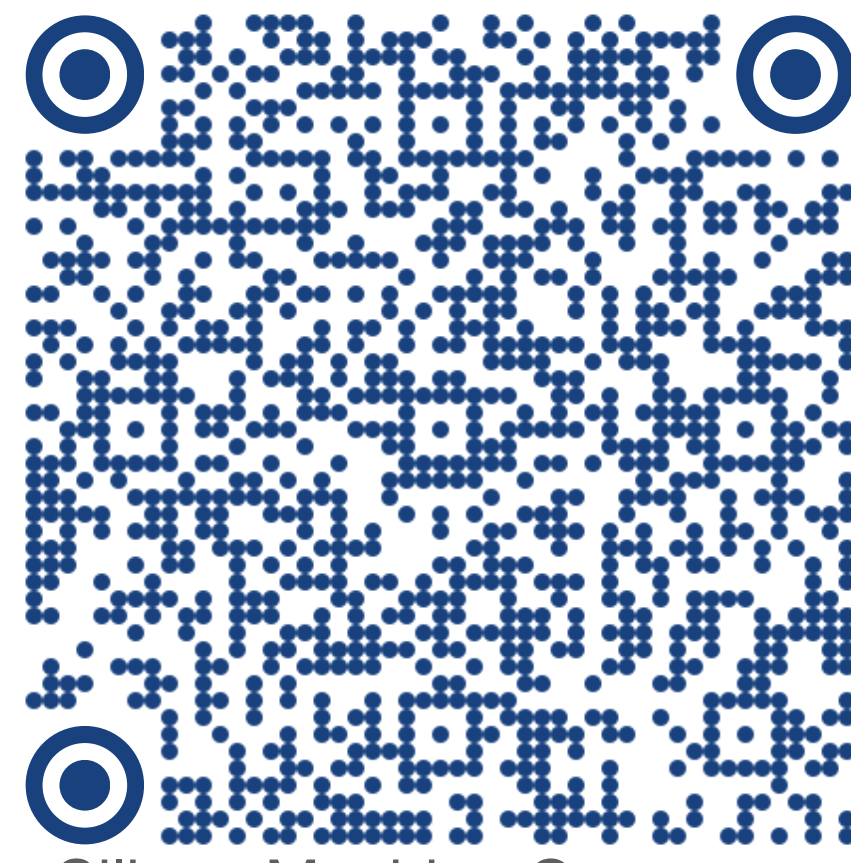
- evoruns-render
 - QD-Fox
 - conf-duration_delta_pitch_combinations-singleCellWin.tar.xz (151.2 MB)
 - conf-one_comb-CPPN_only.tar.xz (906.4 MB)
 - conf-one_comb-noNoise.tar.xz (3.9 GB)
 - conf-one_comb-single-class.tar.xz (1.6 MB)
 - conf-one_comb-singleCellWin.tar.xz (1.2 GB)
 - conf-one_comb.tar.xz (2.6 GB)
 - conf-single-class-runs.tar.xz (1.8 MB)
 - conf-single-class-runs_112-dur-pitch-vel-comb.tar.xz (3.3 MB)
 - conf-static_mutation_rate_combinations-singleCellWin.tar.xz (540.8 MB)
 - QD-nemur
 - QD-ROBIN-HPC
 - QD-ROBIN-workstations
- plots
- QD
- software
 - 00_README.md (4.9 KB)
 - 00_README.txt (4.9 KB)

con



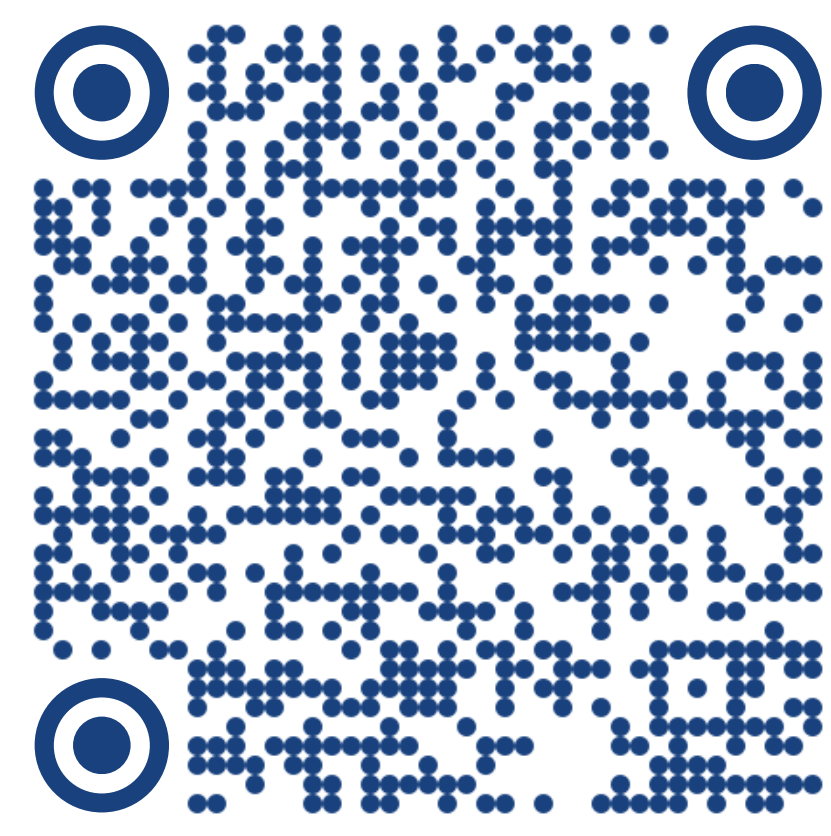
Evoruns Explorer

clu



Silicon Machine Composer

si



For Meat Machine Composers

on

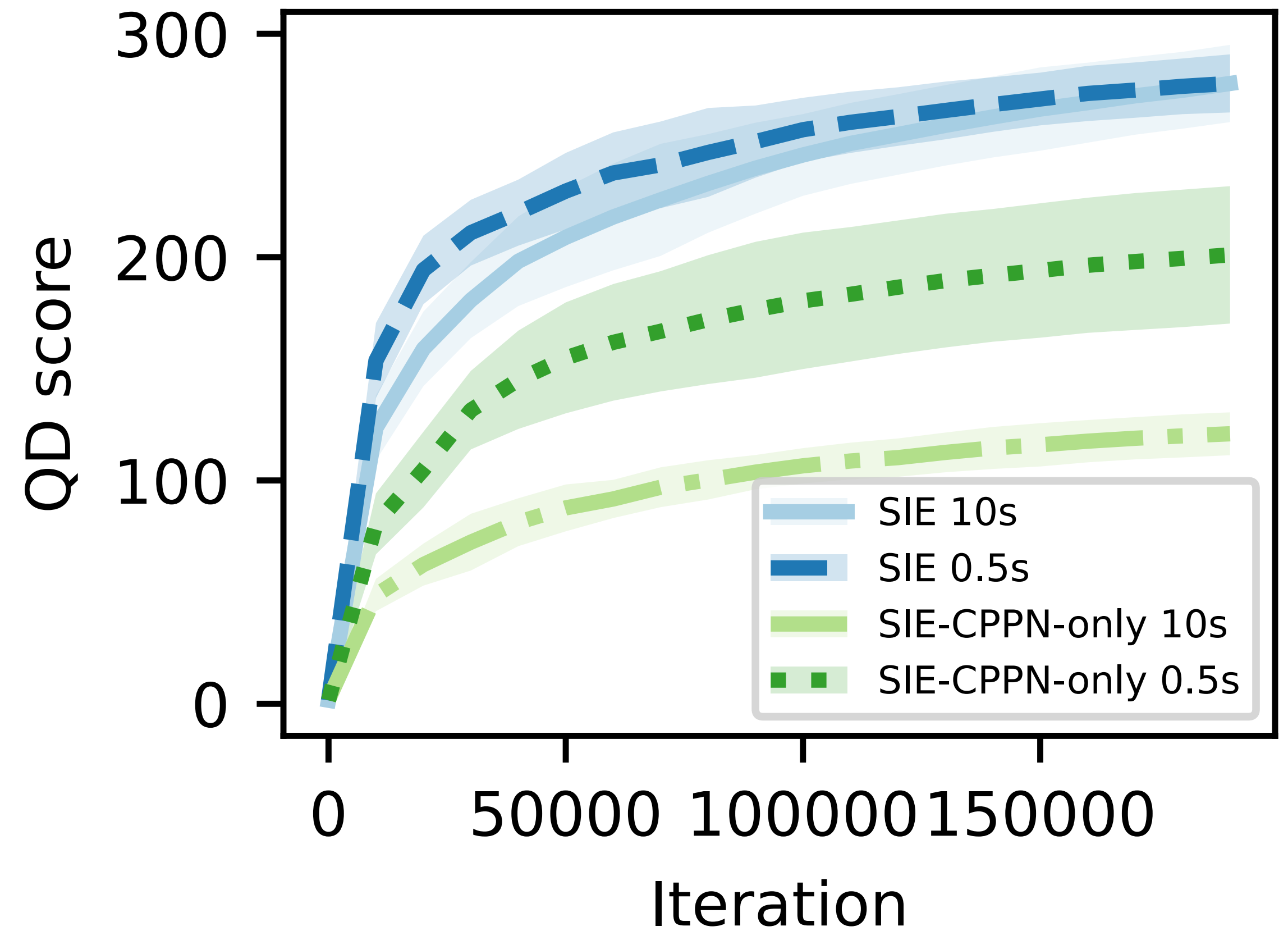
- Applying diversity-promoting algorithms with classifier reward signals is a viable approach for sound discovery
- The sound synthesis approach employed achieved high confidence from a DNN classifier
- Diverse and innovative sound objects were generated, suggesting further explorations in this system



Results

Temporal Pattern Revelation and Classifier Characteristics

- Longer time for sounds to develop doesn't benefit CPPN-only sounds
- Lack of DSP more apparent?
- CPPN + DSP duration agnostic



57.4% ± 3.4%

- map coverage, when incremental - otherwise immediately full coverage