Adaptive Frequency Neural Networks for Dynamic Pulse and Metre Perception

CITY UNIVERSITY LONDON
EST 1894

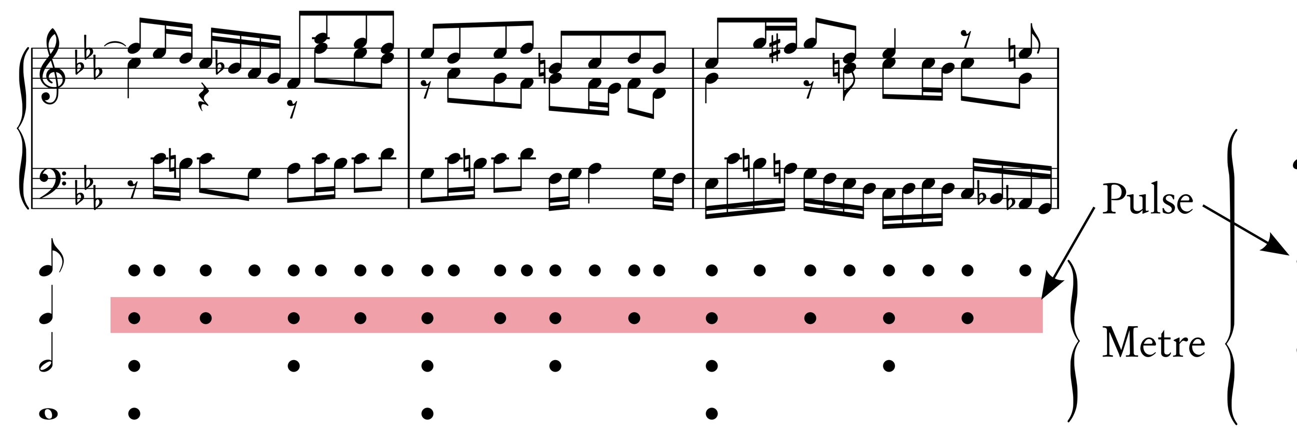
Andrew Lambert [1 2], Tillman Weyde [1], Newton Armstrong [2]
{andrew.lambert.1; t.e.weyde; newton.armstrong.1}@city.ac.uk
[1] Department of Computer Science, [2] Department of Music

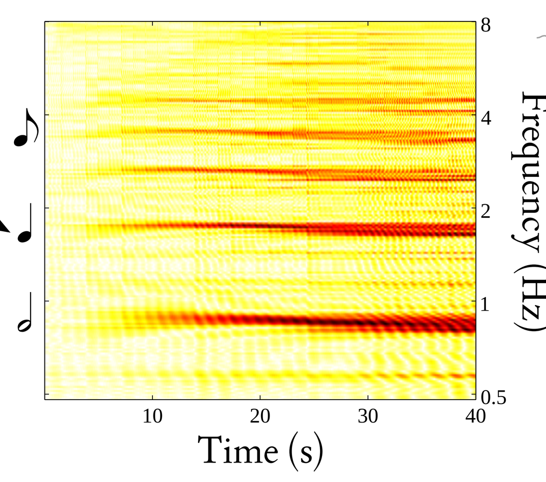
Beat Induction

A perceptual and cognitive process by which a steady **pulse** is perceived when listening to a rhythm.

Metre

The multi-layered divisions of time present in music, of which the referent layer is the **pulse**. Other layers in music **divide** the pulse into the smallest subdivisions of time, and **extend** it towards larger measures, phrases, periods, and even higher order forms.





Nonlinear Resonance

A model of the way our entire nervous system resonates to rhythms. A population of neurons is represented as a canonical **nonlinear oscillator**.

Expressive Timing

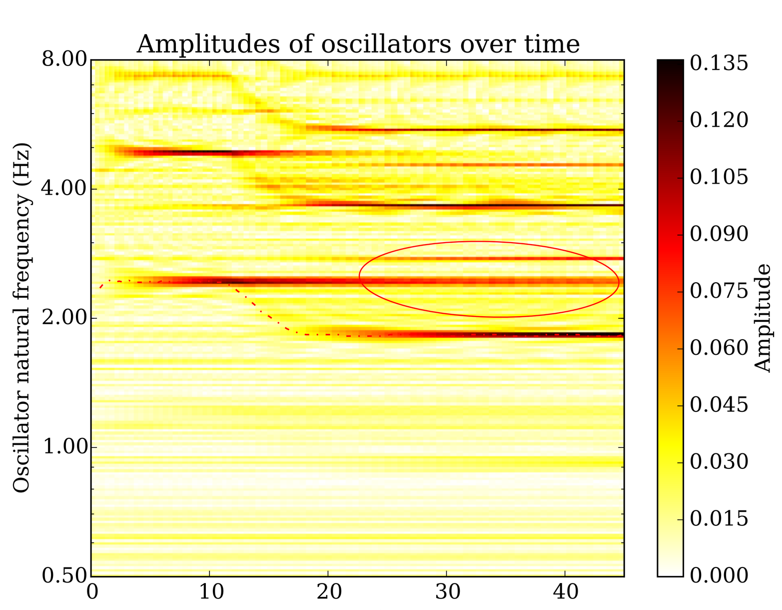
Expressive shaping of the music including **tempo change**, rubato, and groove that affects the perception of the **pulse** and **metre**.

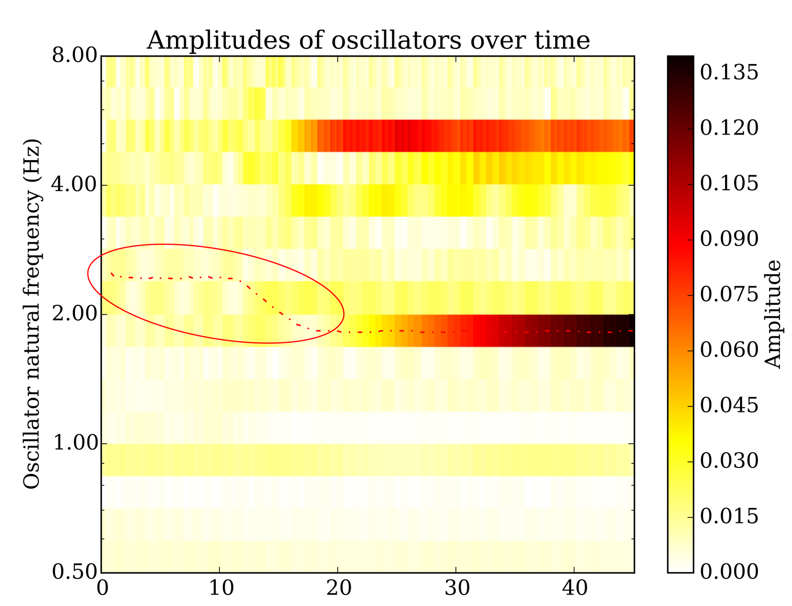
Gradient Frequency Neural Network (GFNN)

A network of canonical **oscillators** distributed across a frequency gradient. Rhythm-harmonic frequency information is added to the signal, which can be interpreted as a perception of **metre** (Large et al., 2010).

Interference

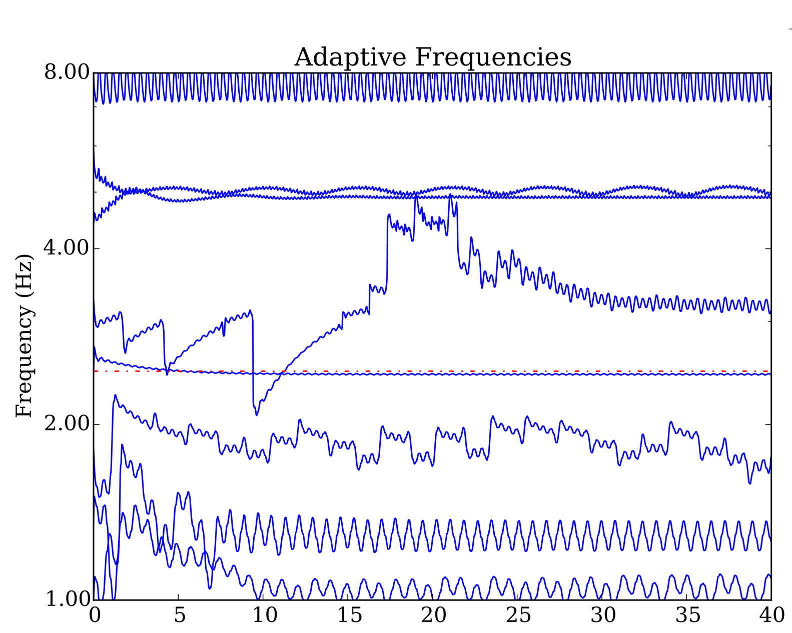
If tempo changes, the memory of the original tempo can persist, causing interference in the output.





GFNN Density

A low densities can reduce interference, but lead to missing important frequencies, especially if tempo is changing.



Adaptive Frequency Neural Network (AFNN)

Applies Righetti et al.'s (2006) Hebbian learning rule to the oscillator frequencies in the network. The frequencies adapt to the stimulus through an **attraction** to local areas of resonance.

$$\frac{d\omega_i}{dt} = \frac{-\epsilon_f}{r} x(t)\sin(\phi_i) - \frac{\epsilon_h}{r}\left(\frac{\omega_i - \omega_{i0}}{\omega_{i0}}\right)$$

A secondary **elasticity** rule is also added, which attracts the frequencies back to their original value.

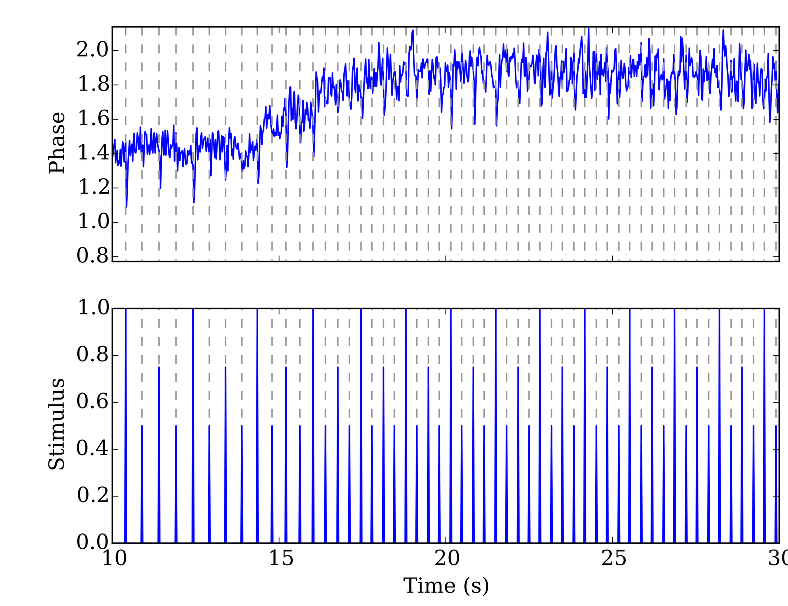
Benefits

+ Reduced density and interference
+ More adaptation for dynamic pulse

Phase Based Evaluation

So far, evaluation of GFNNs has not considered **phase** information. This is important as, in relation to pulse, it constitutes the difference between **on-beat** and **off-beat** prediction.

We propose a **weighted phase output**, shown in (1), by taking the magnitude ($r = |z|$) and angle ($\varphi = arg(z)$) of the oscillators.



$$\Phi = \sum_{i=0}^{N} r_i \varphi_i \qquad (1)$$

Experiment

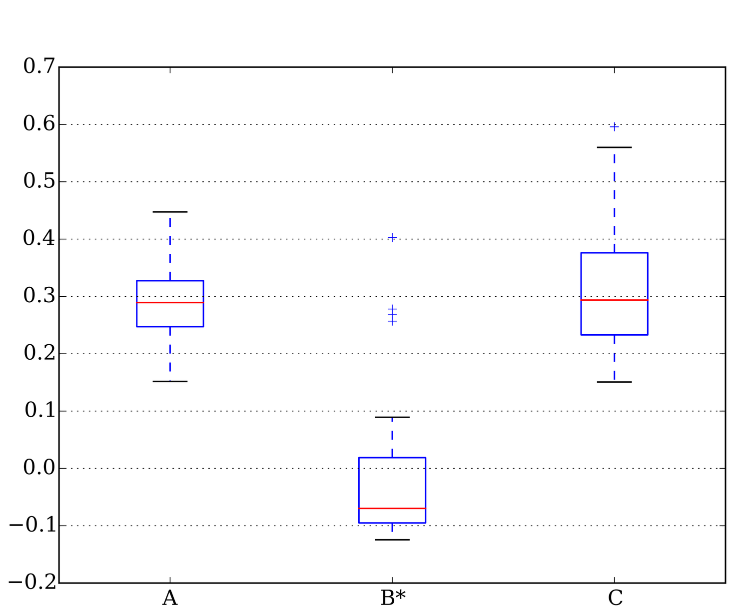
Aims:
+ **Compare** AFNN with GFNN presented in Velasco and Large (2011), and Large et. al (2015)
+ Test AFNN on **dynamic** pulses

Method:
+ 5 stimulus categories
  - Isochronous and Son Clave (Velasco and Large 2011)
  - Large et al.'s (2015) rhythm syncopation levels 1-4
  - Accelerando and Ritardando dynamic pulses

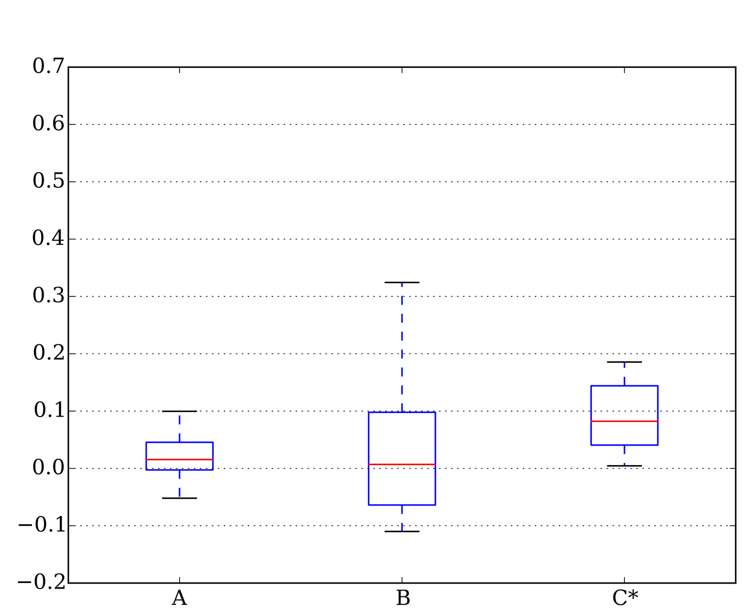
Evaluation

+ Compare **weighted phase output** with a **ground truth phase** signal.
+ Ground truth signal resembles an **inverted beat-pointer** model and represents phase growing from 0 to 2π in an oscillation.
+ Quantitative comparison via Pearson product-moment correlation coefficient (PCC)

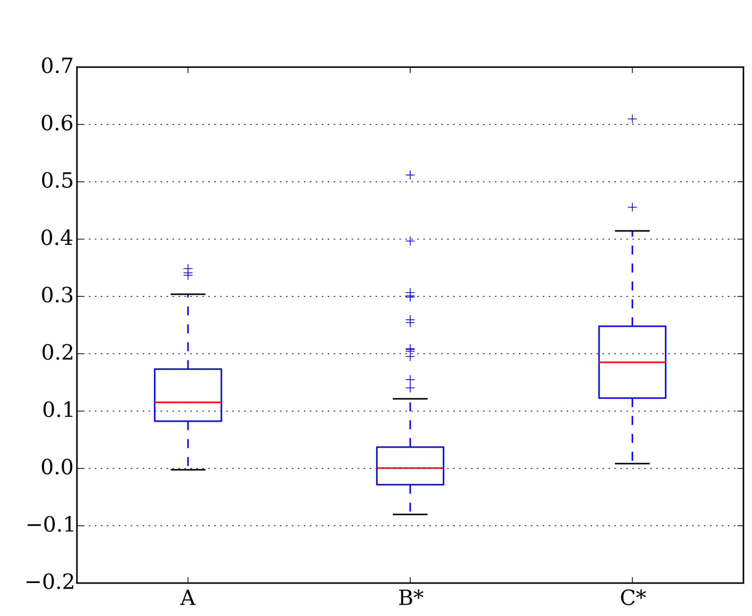
AFNNs and GFNNs operate on more than one metrical level, so **small positive correlations** indicate good frequency and phase response.



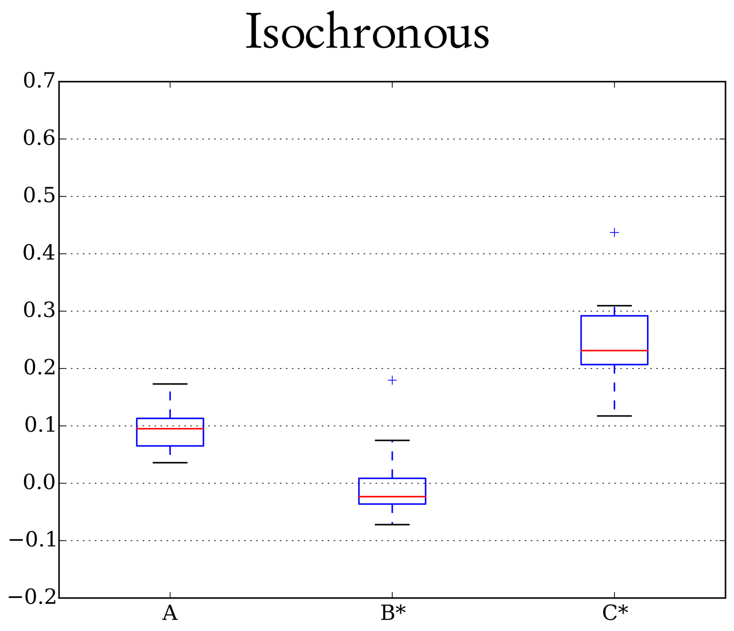
Isochronous



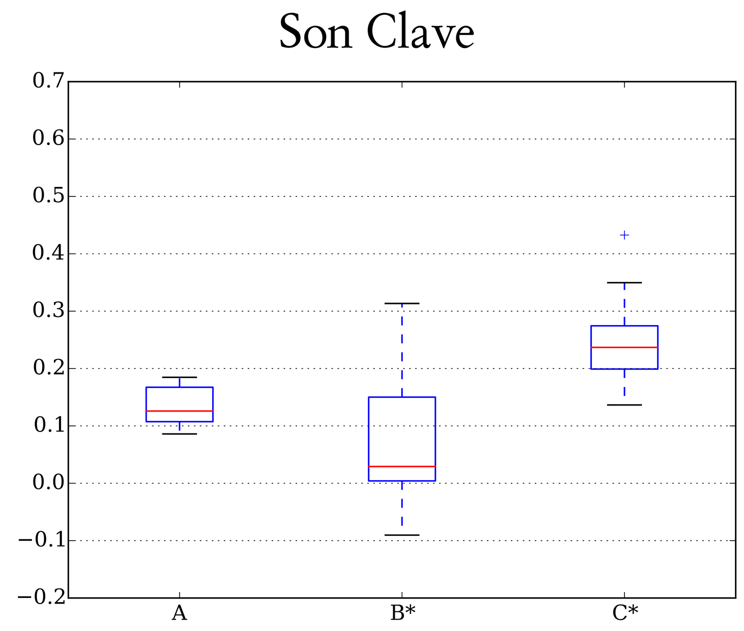
Son Clave



Levels 1-4



Accel.



Rit.

A) GFNN
B) Low density GFNN
C) AFNN
*Denotes significance in a Wilcoxon signed rank test ($p < 0.05$)

Results

+ **Isochronous:** GFNN is effective, AFNN matches performance, despite low density.
+ **Son Clave:** All networks perform poorly, AFNN performs significantly better.
+ **Levels 1-4:** AFNN is significantly improved but still a poor correlation.
+ **Accel. and Rit.:** GFNN performs poorly, AFNN shows acceptable correlation and is a significant improvement over the GFNN.

When compared with GFNNs, we showed an **improved** response by AFNNs to rhythmic stimuli with both **steady** and **varying** pulse.

Edward W. Large. Neurodynamics of Music. In Mari R. Jones, Richard R. Fay, and Arthur N. Popper, editors, Music Perception, number 36 in Springer Handbook of Auditory Research, pages 201–31. Springer New York, 2010.

Edward W. Large, Jorge A. Herrera, and Marc J. Velasco. Neural networks for beat perception in musical rhythm. Frontiers in Systems Neuroscience, 9(159), 2015.

Ludovic Righetti, Jonas Buchli, and Auke J. Ijspeert. Dynamic hebbian learning in adaptive frequency oscillators. Physica D: Nonlinear Phenomena, 216(2):269–81, 2006.

Marc J. Velasco and Edward W. Large. Pulse Detection in Syncopated Rhythms using Neural Oscillators. In 12th International Society for Music Information Retrieval Conference, pages 185–90, Miami, FL, 2011.