

# Adversarial Generation of Time-Frequency Features

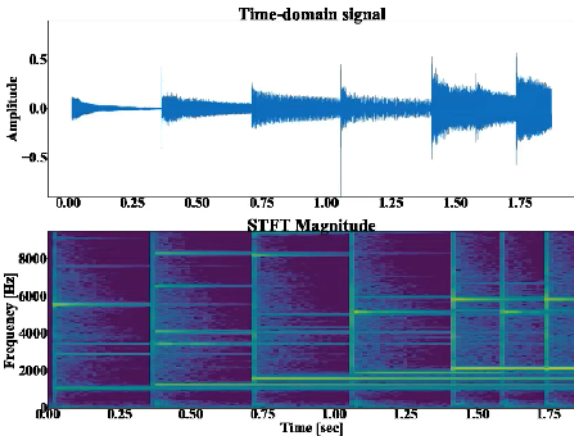
with application in audio synthesis

Speaker: Andrés Marafioti

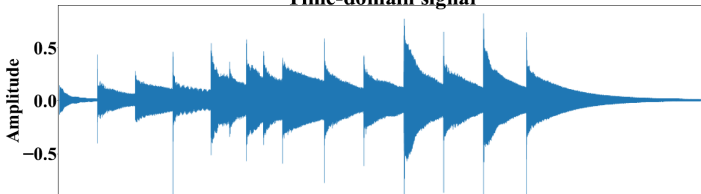
Co-Authors: Nathanaël Perraudin, Nicki Holighaus, Piotr Majdak

Acoustics Research Institute, Vienna  
Austrian Academy of Sciences

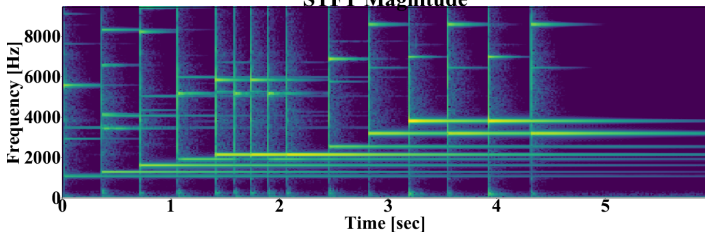
International Conference on Machine Learning  
Long Beach, California, June 11th, 2019



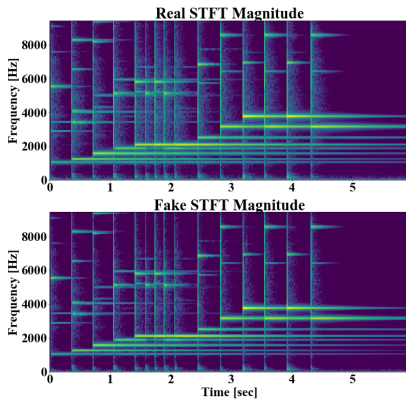
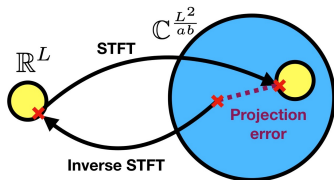
Time-domain signal

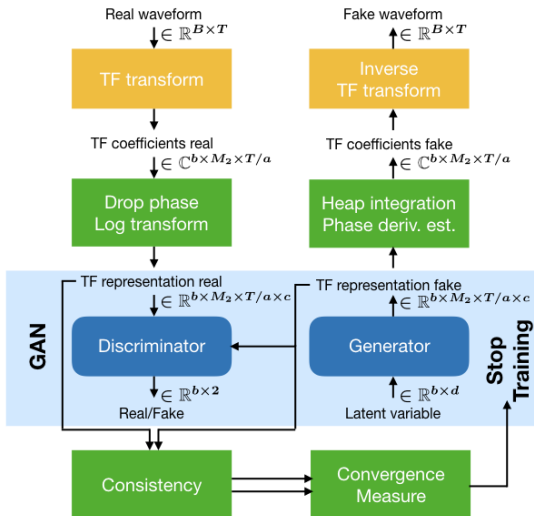


STFT Magnitude



# Is it consistent?





- We trained on a dataset of spoken English digits [0-9].
- We evaluated our results with perceptual tests.
- Audio examples and implementations are available at `tifgan.github.io`

	WaveGAN digits	TiFGAN-M digits
	vs TiFGAN	vs WaveGAN
Real	86%	94%
TiFGAN	–	75%
WaveGAN	25%	–

Thank you for your attention!  
Supported by the Austrian Science Fund  
(FWF; MERLIN, I 3067-N30).