MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment

Hao-Wen Dong*, Wen-Yi Hsiao*, Li-Chia Yang, Yi-Hsuan Yang
Research Center of IT Innovation, Academia Sinica

Demo Page **https://salu133445.github.io/musegan/**

**these authors contributed equally to this work*

Outline

- Goals & Challenges
- Data
- Proposed Model
- Results & Evaluation
- Future Works

Source Code **https://github.com/salu133445/musegan**
Demo Page **https://salu133445.github.io/musegan/**

Goals

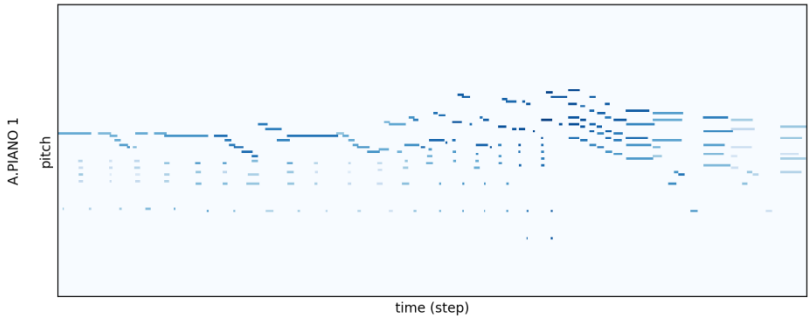
[Source Code]
https://github.com/salu133445/musegan

[Demo Page]
https://salu133445.github.io/musegan/

Generate pop music

- of **multiple tracks**

- in **piano-roll** format

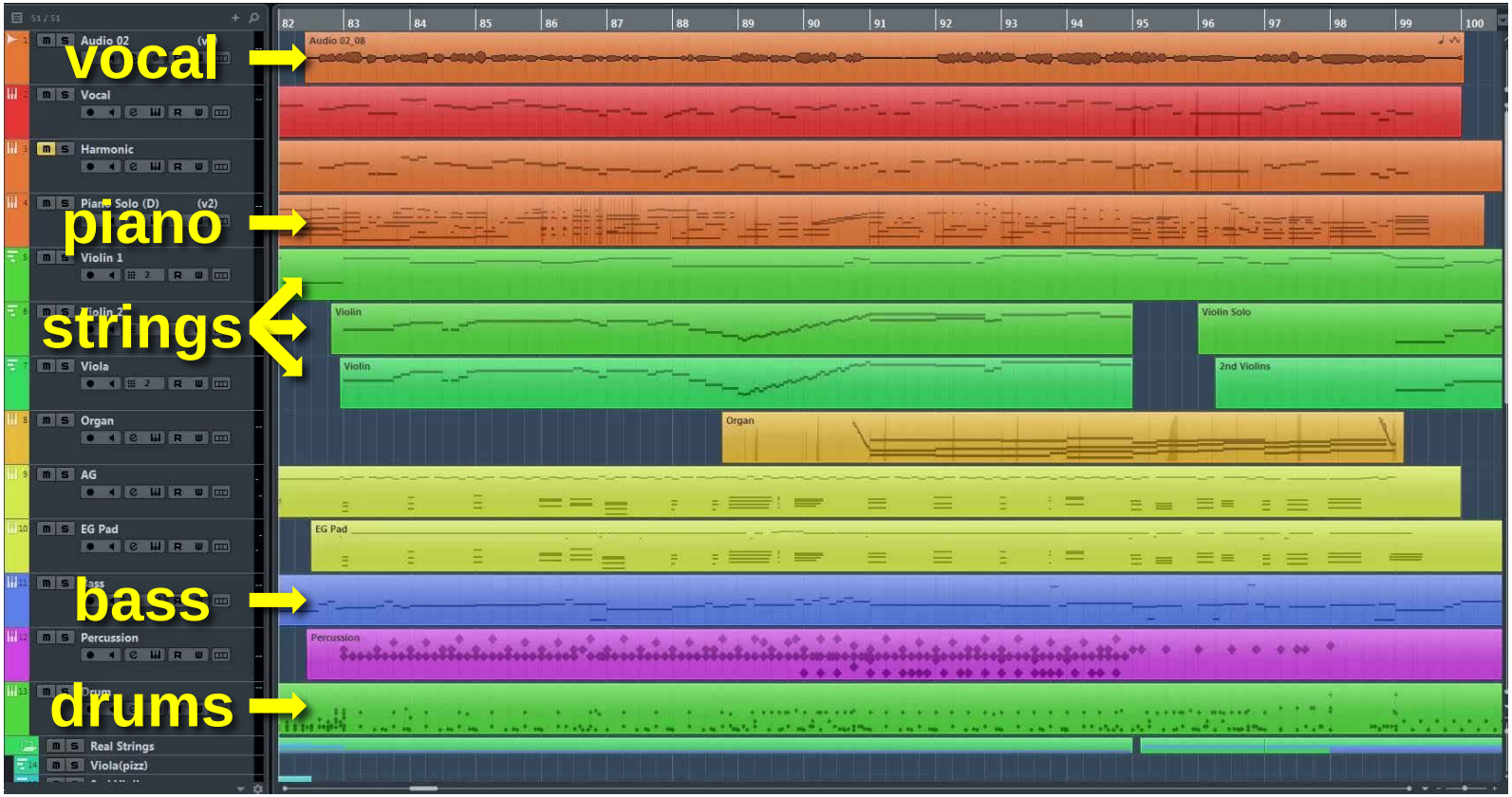


- using **GAN** with **CNNs**

Challenge I

Multi-track GAN

Multitrack Interdependency

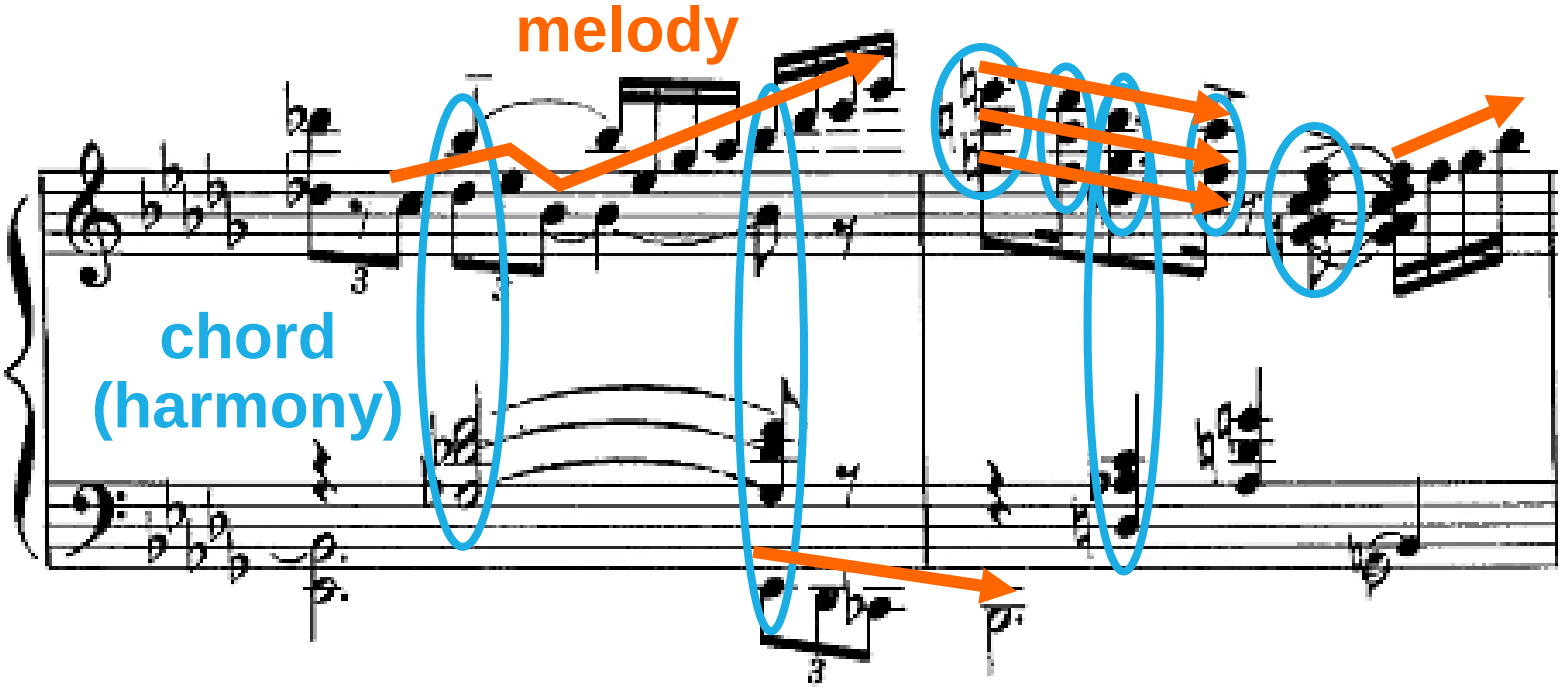


music & clip by ***phycause***

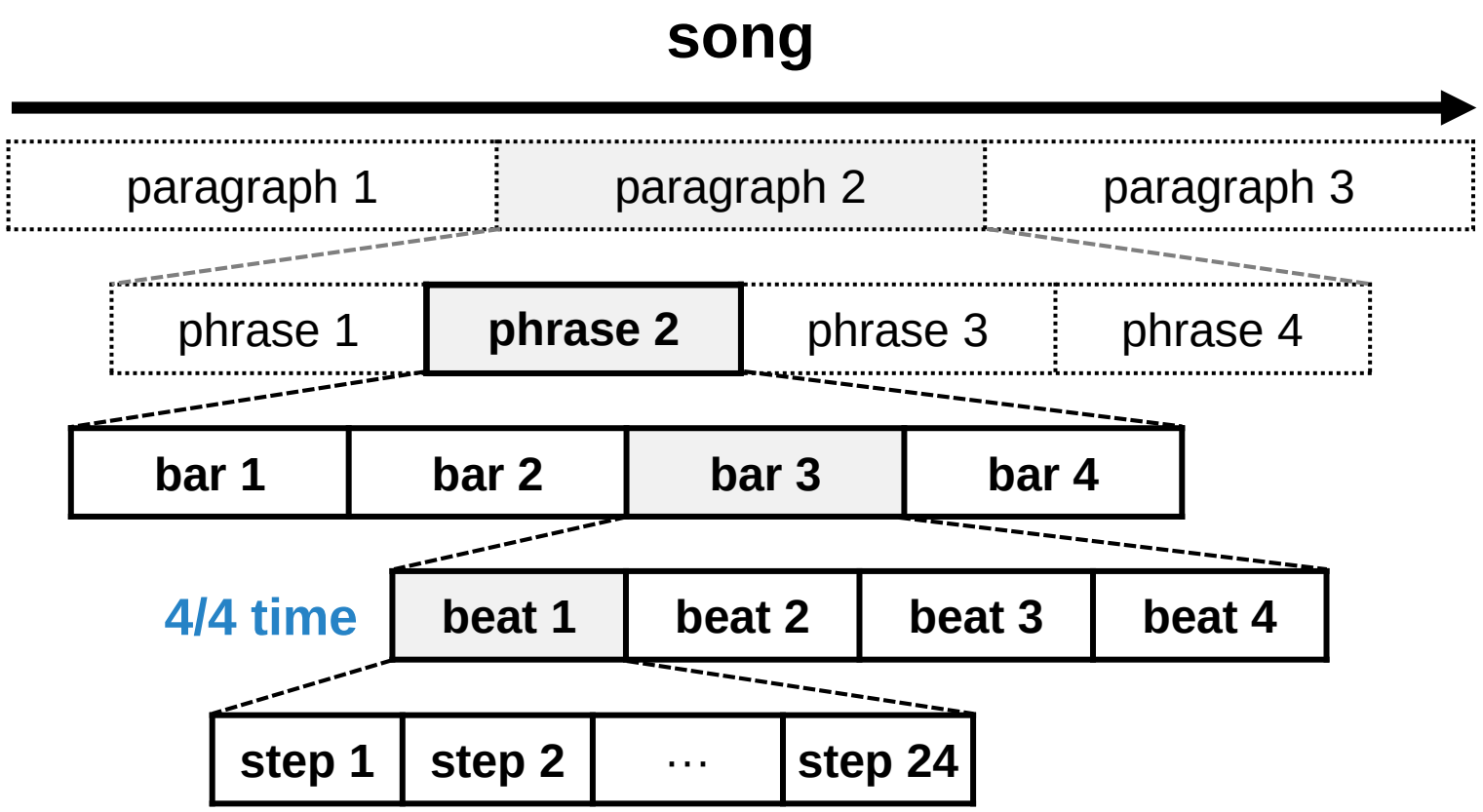
Challenge II

Music Texture

Convolutional Neural Networks

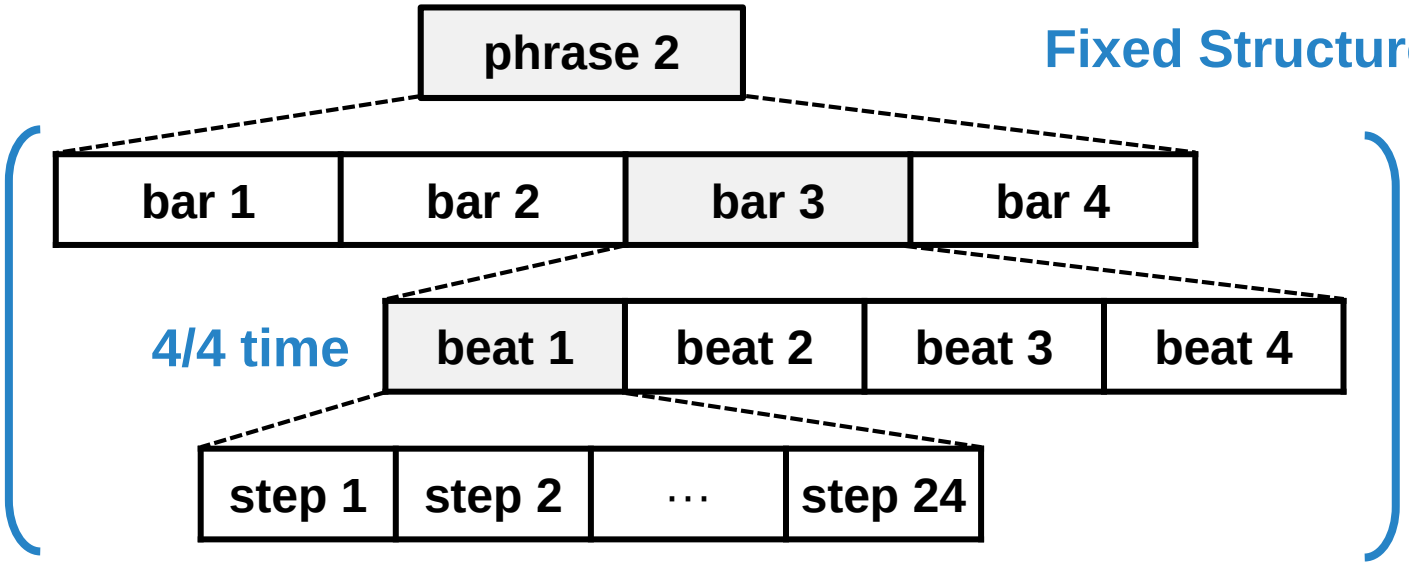


Challenge III
Temporal Structure



Challenge III
Temporal Structure

Convolutional Neural Networks

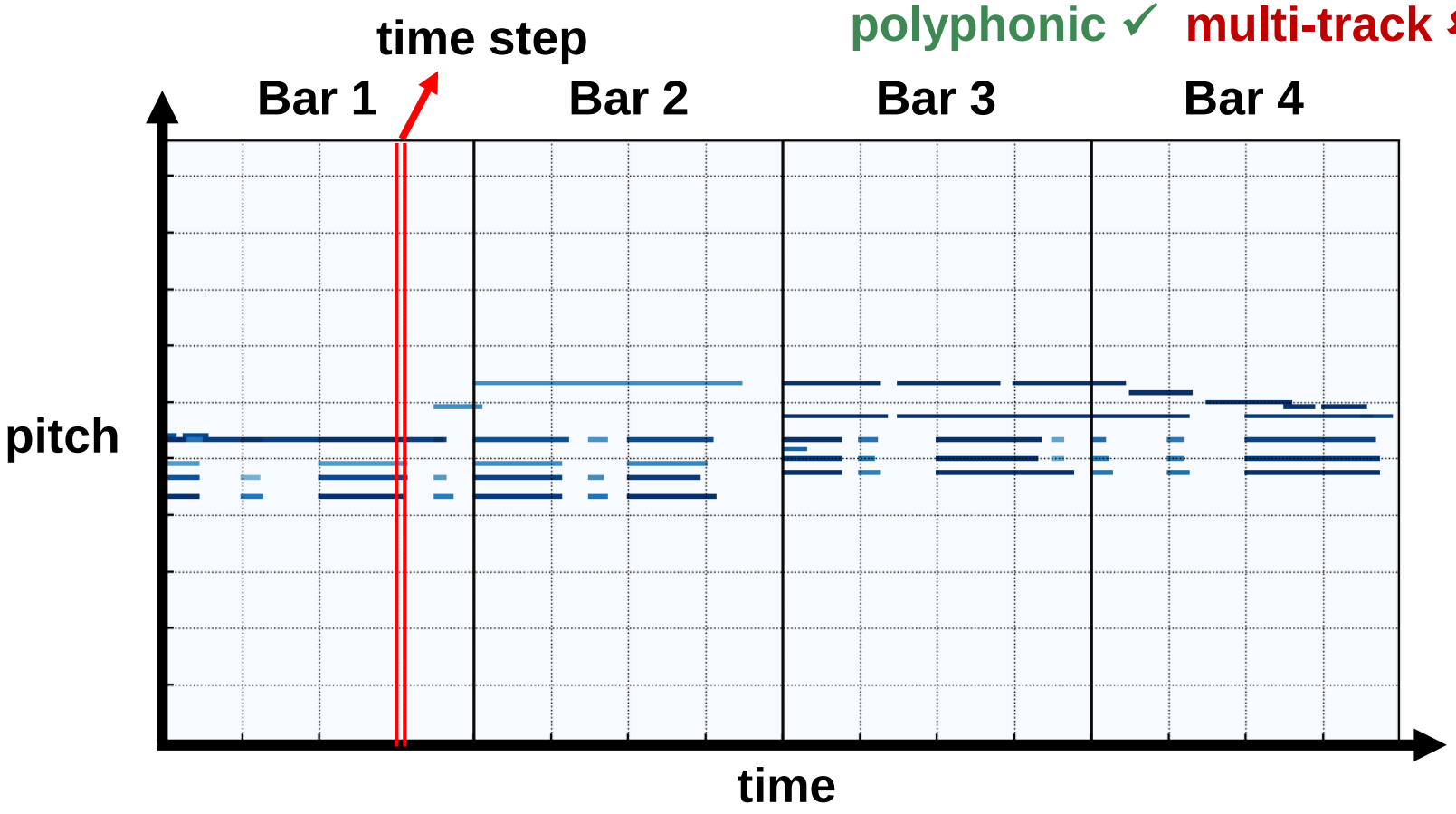


Data Representation

Piano-roll

(with symbolic timing)

polyphonic ✓ multi-track ✘

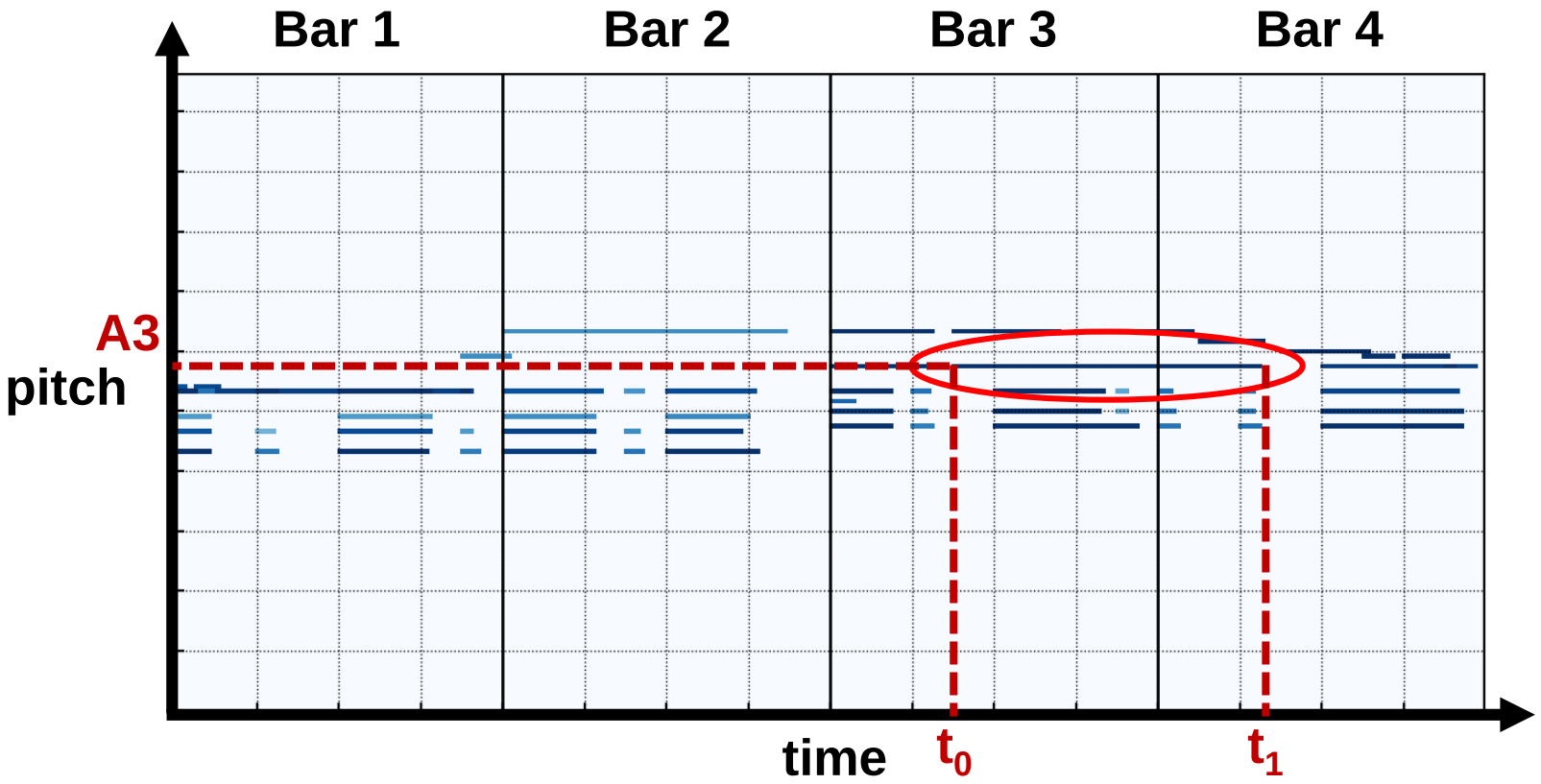


Data Representation

Piano-roll

(with symbolic timing)

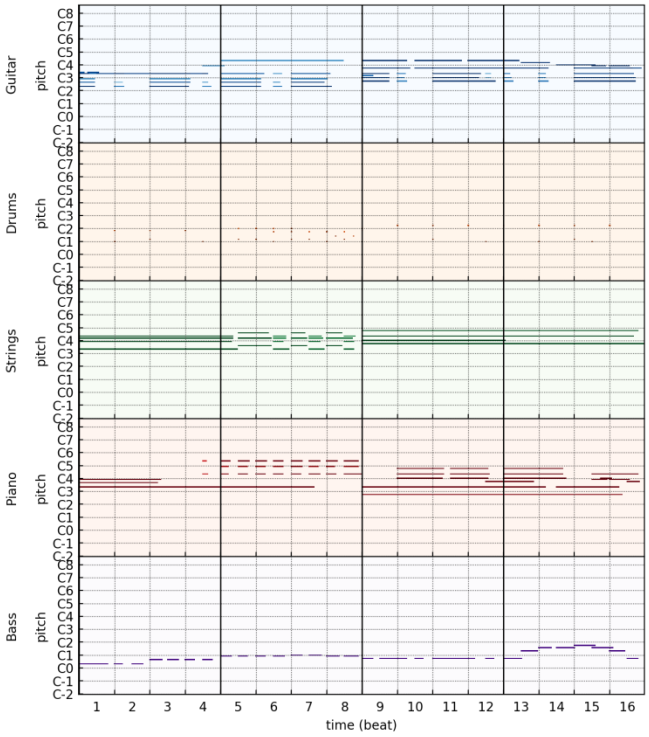
polyphonic ✓ multi-track ✘

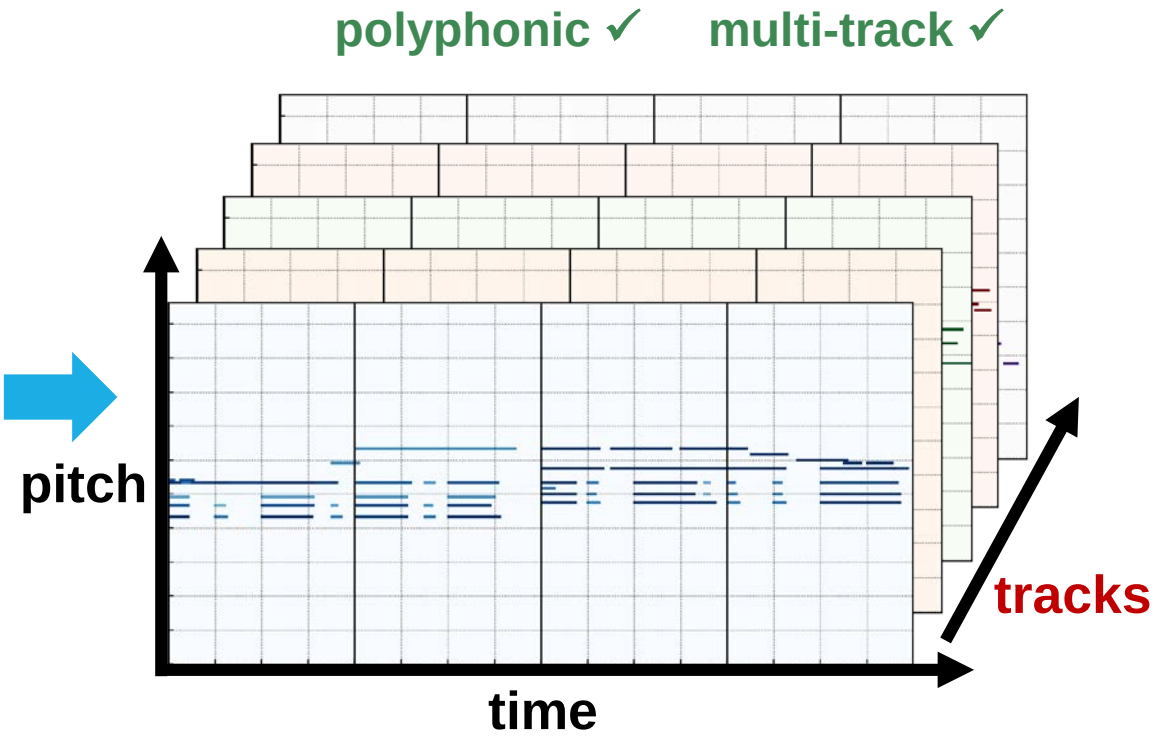


Multi-track Piano-roll

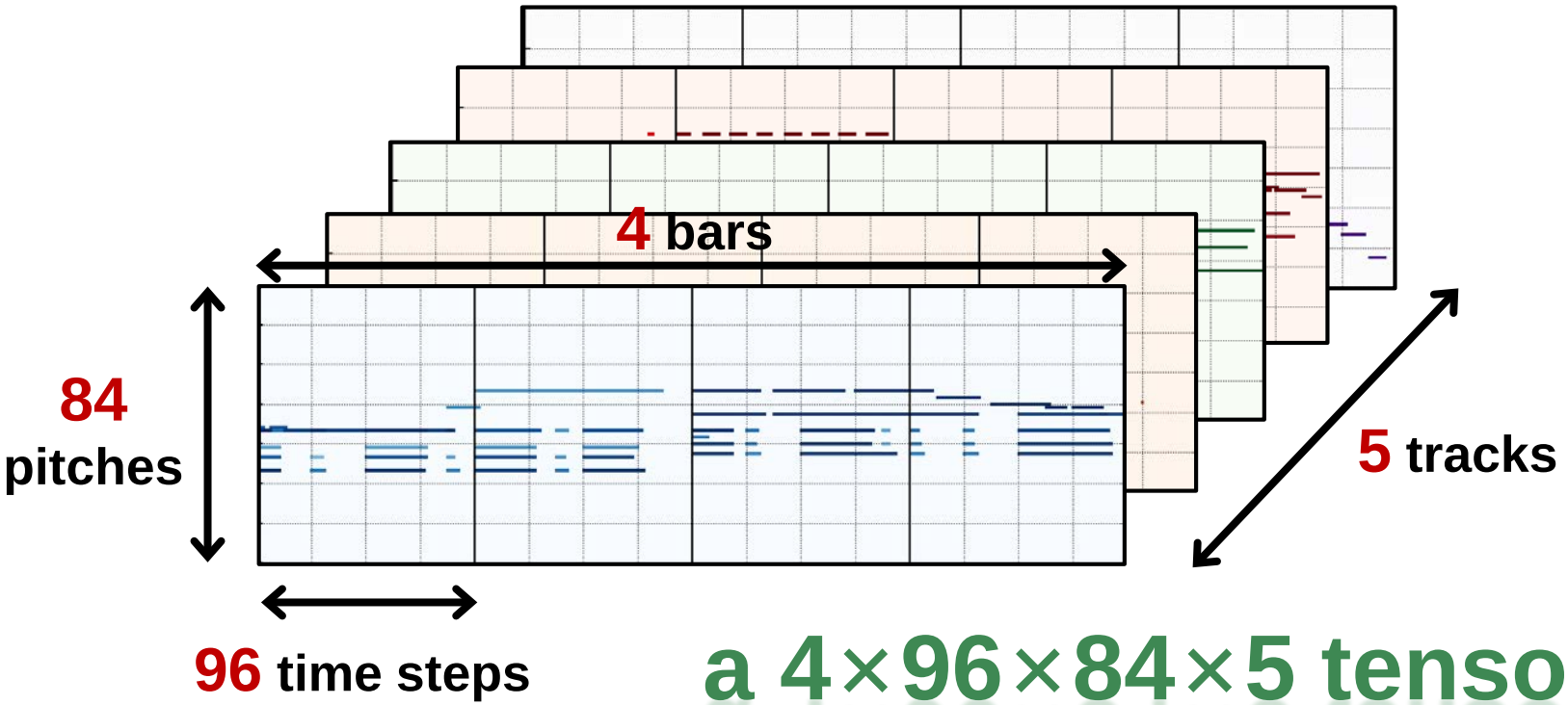
(with symbolic timing)

Data Representation





Data Representation



a 4×96×84×5 tensor

Data

[Dataset]
https://salu133445.github.io/musegan/dataset

[Pypianoroll]
https://salu133445.github.io/pypianoroll/

NEW!!

LPD (Lakh Pianoroll Dataset)

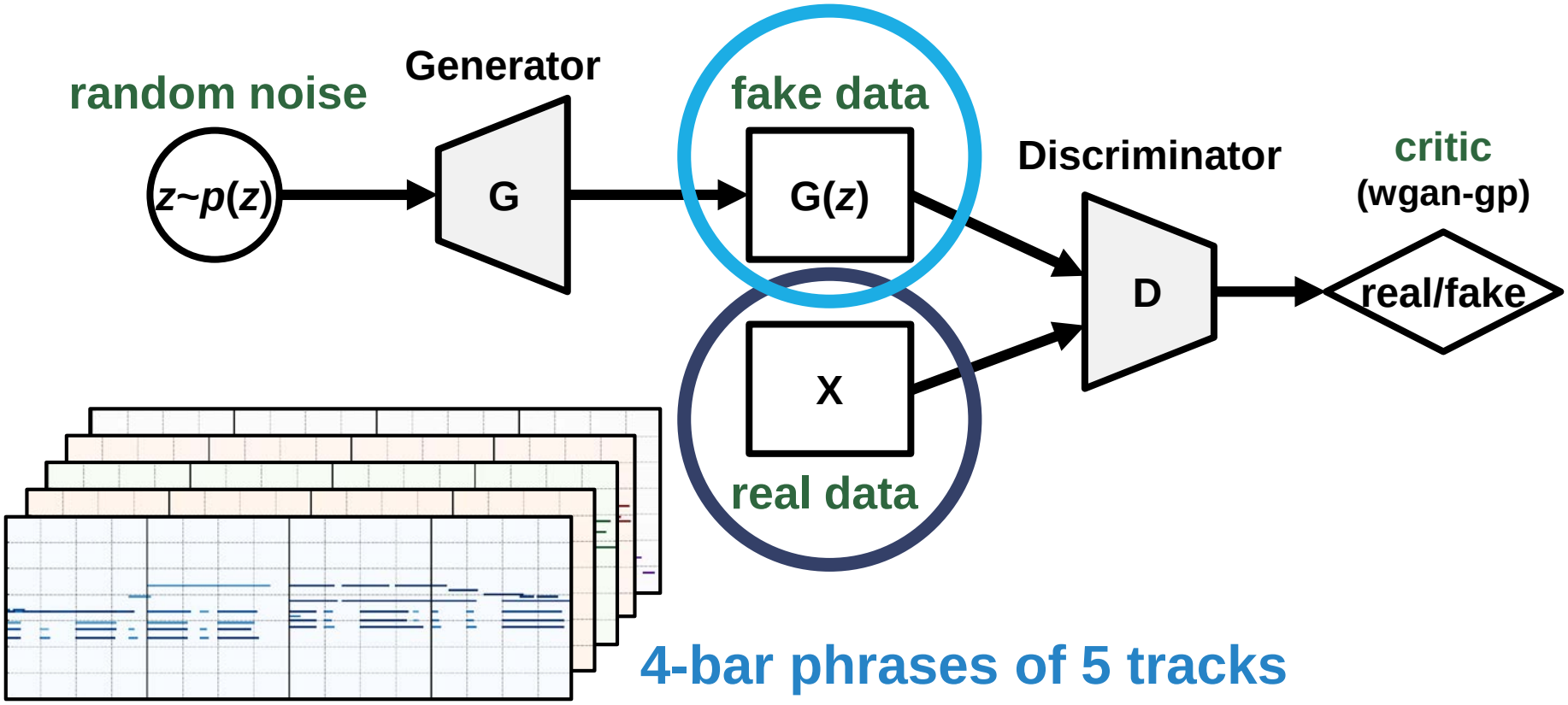
- **>170,000** multi-track piano-rolls
- Derived from Lakh MIDI Dataset
- Mainly pop songs

NEW!!

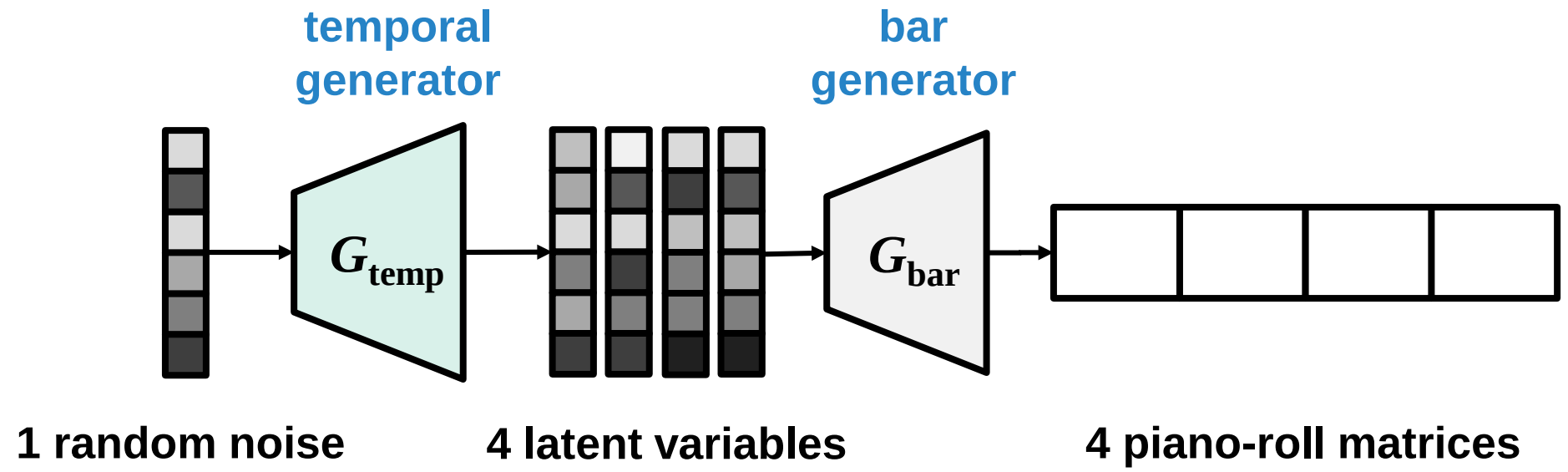
Pypianoroll (Python package)

- Manipulation & Visualization
- Efficient Save/Load
- Parse/Write MIDI files
- On **PYPI** (pip installable)

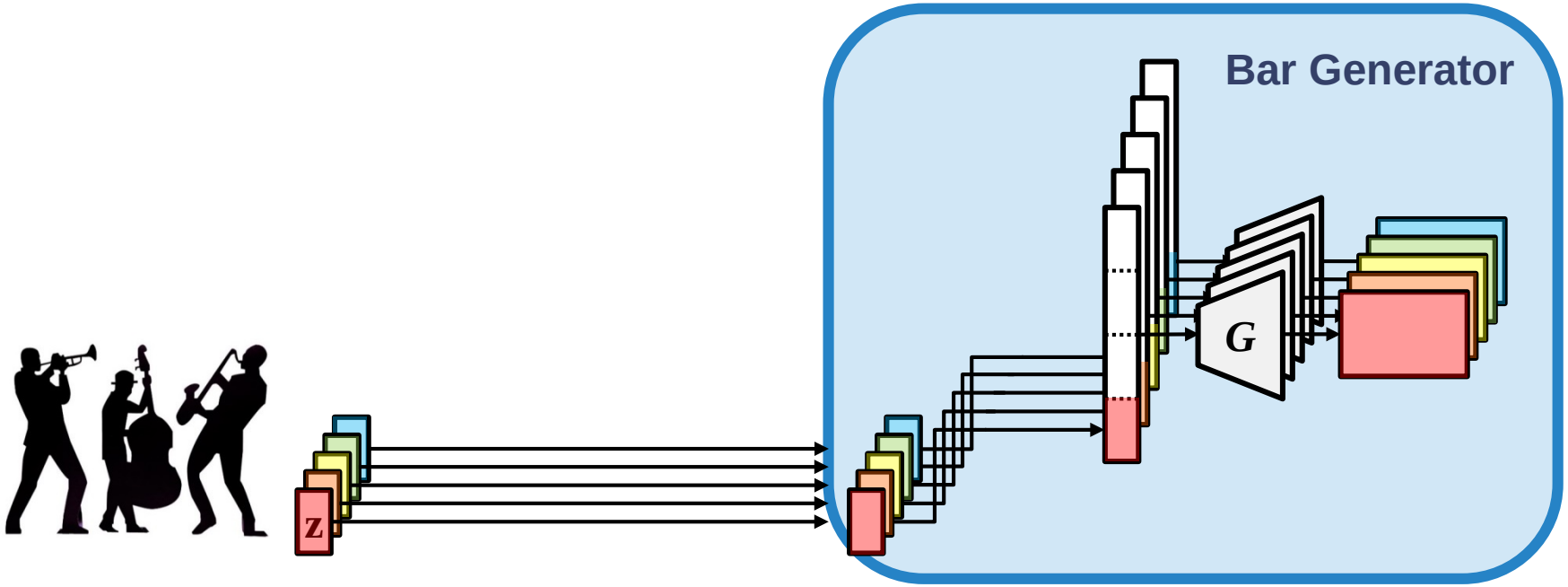
Generative Adversarial Networks



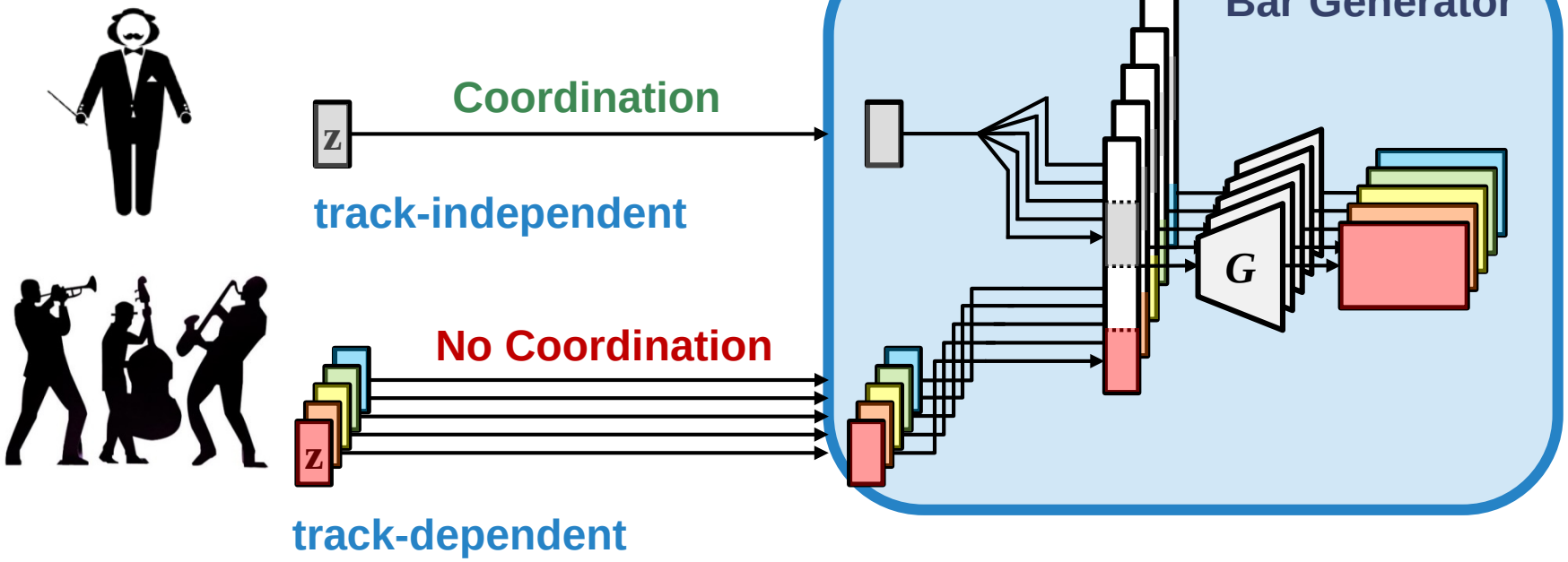
MuseGAN – An Overview



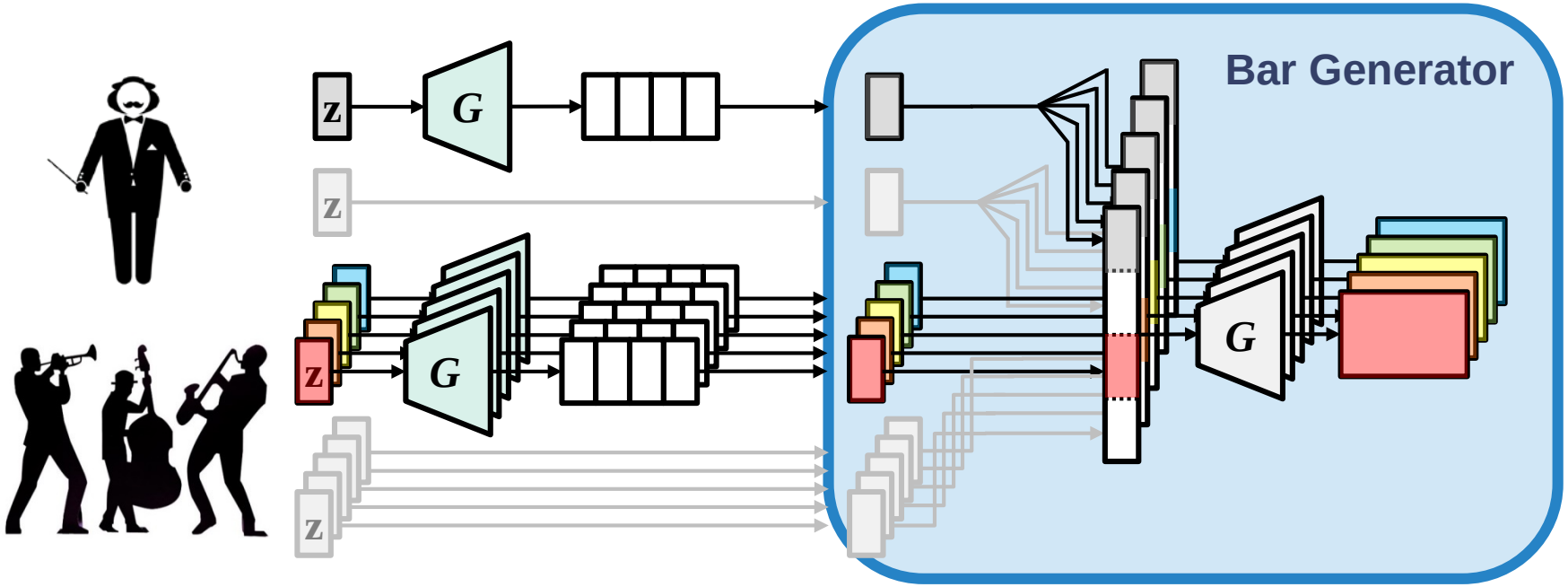
MuseGAN



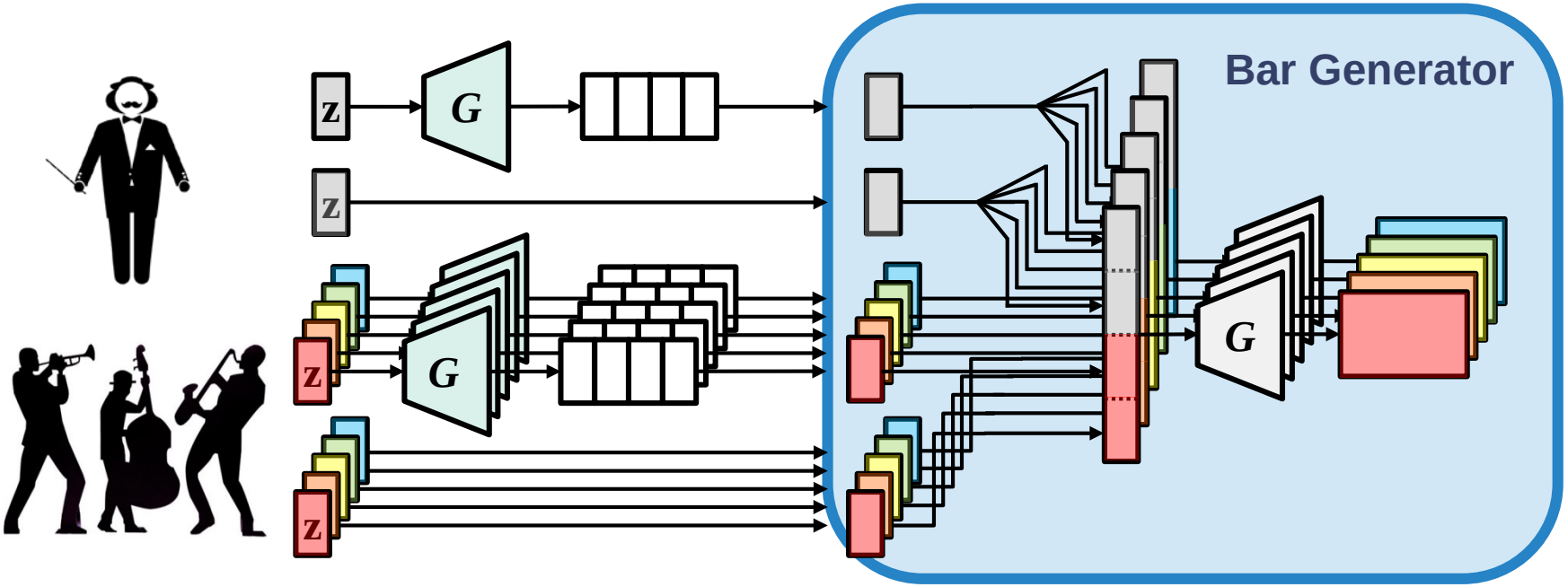
MuseGAN


Coordination
z
track-independent
No Coordination
z
track-dependent
Bar Generator
G

MuseGAN

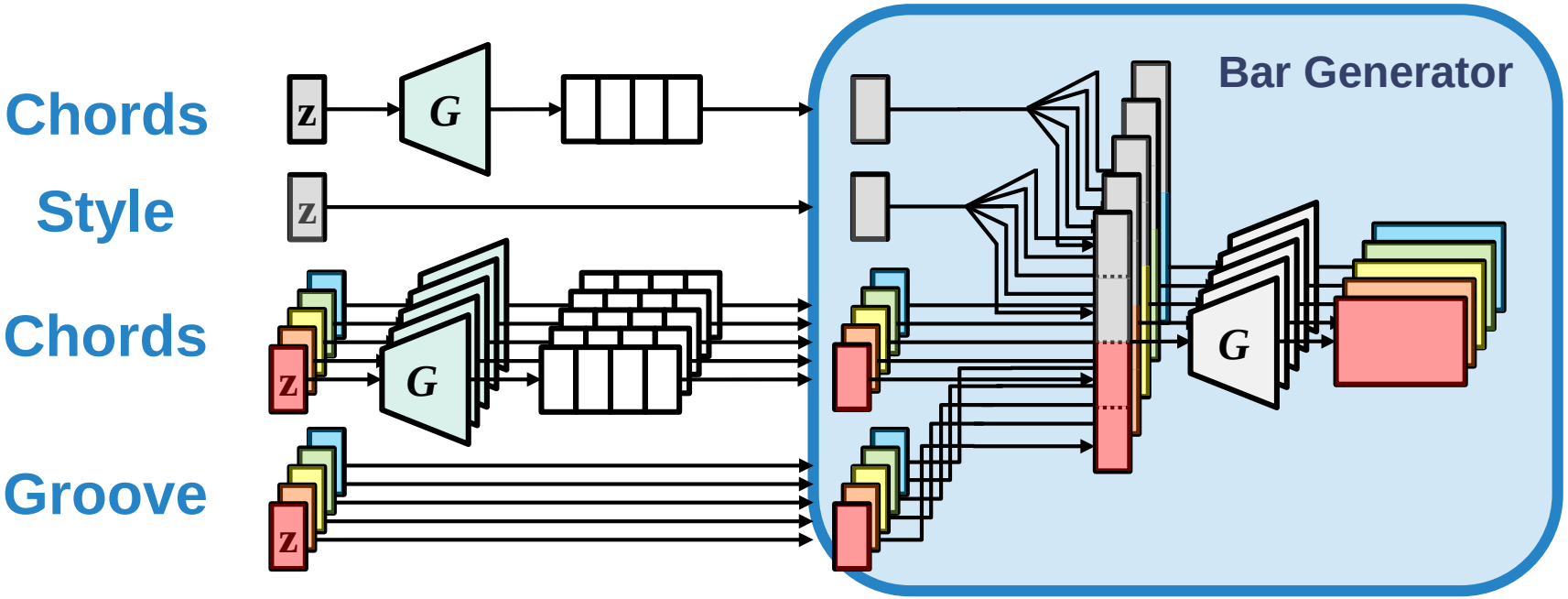


MuseGAN



MuseGAN

		Time	
		Dependent	Independent
Track	Dependent	**Melody**	**Groove**
	Independent	**Chords**	**Style**



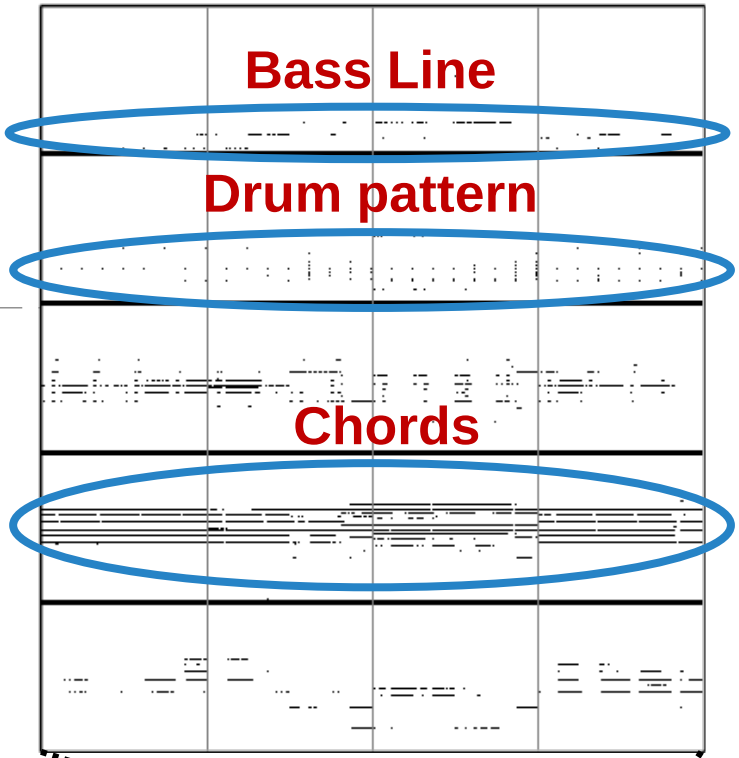
Results

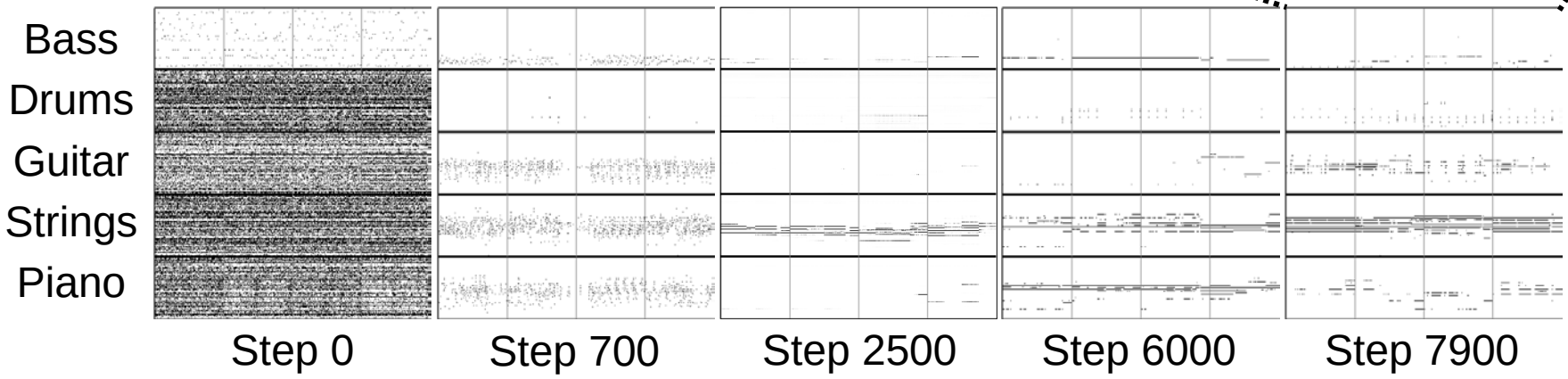
Sample 1

Sample 2

More Samples on Demo Page

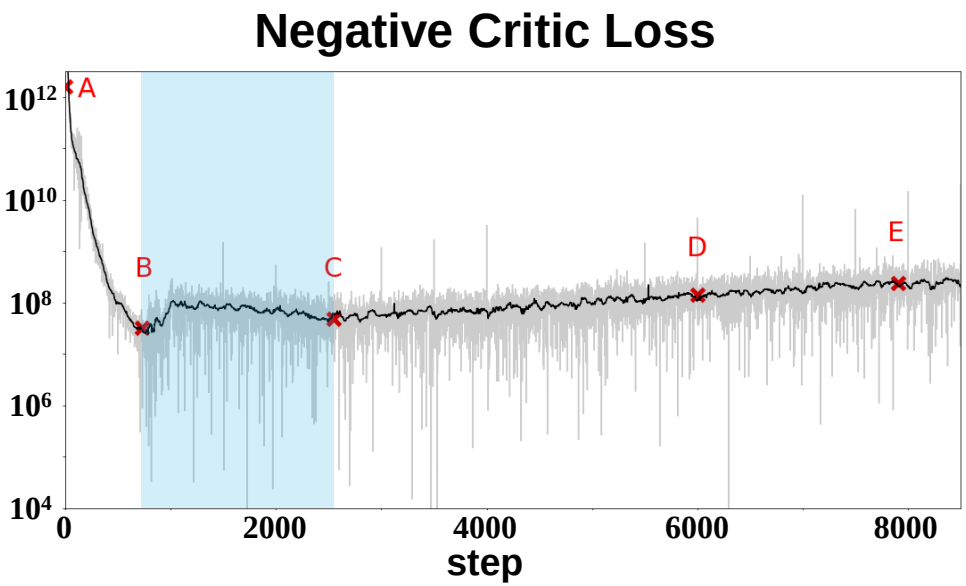
https://salu133445.github.io/musegan/

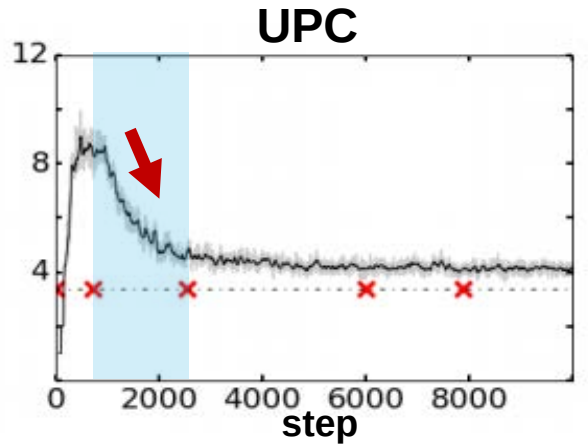


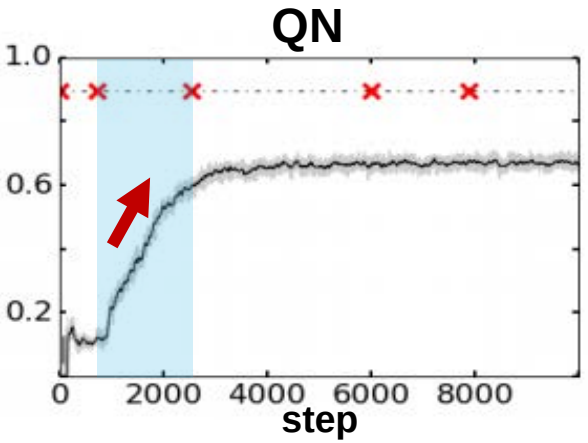


Monitor the Training

Objective Metrics







UPC	number of used pitch classes per bar
QN	ratio of qualified notes

User Study

from scratch		H	R	MS	C	OR
non-pro	jam	2.83	3.29	2.88	2.84	2.88
	comp	3.12	**3.36**	2.95	3.13	3.12
	hybrid	**3.15**	3.33	**3.09**	**3.30**	**3.16**
pro	jam	2.31	3.05	2.48	2.49	2.42
	comp	2.66	3.13	2.68	2.63	2.73
	hybrid	**2.92**	**3.25**	**2.81**	**3.00**	**2.93**

H: harmonious

R: rhythmic

MS: musically structured

C: coherent

OR: overall rating

composer

jamming

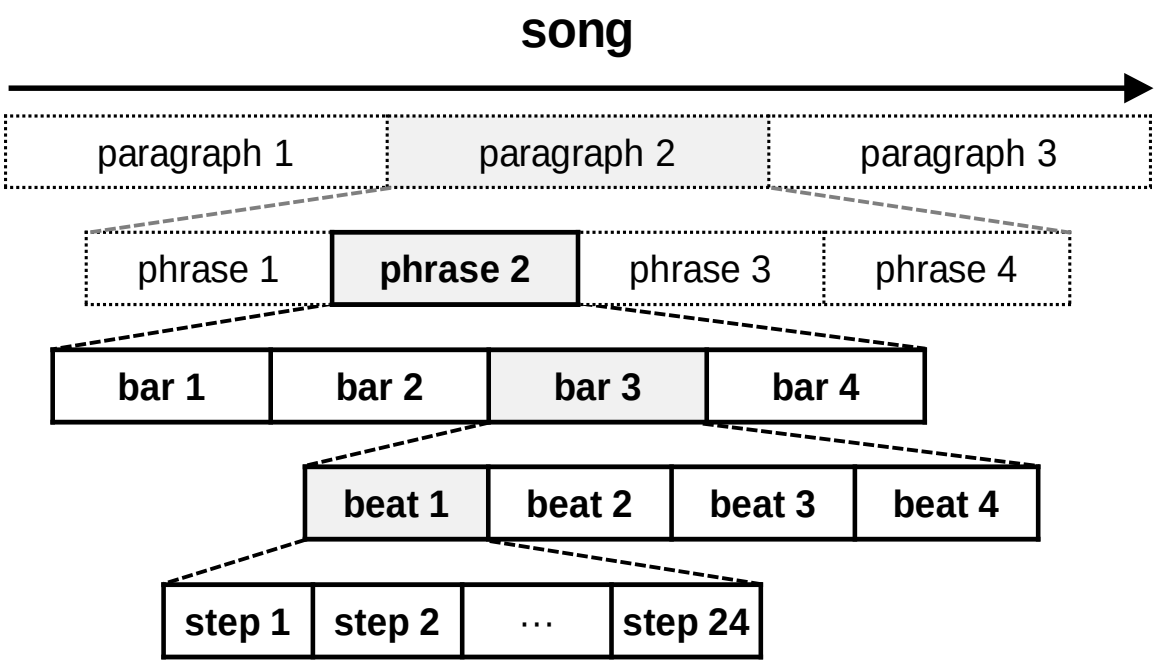
hybrid

Summary

- **MuseGAN**
  - a novel GAN for **multi-track sequence** generation
  - **multi-track**, **polyphonic** music
  - **human-AI cooperative** scenario (see the paper)
- **Lakh Pianoroll Dataset (LPD)** (**new dataset!!**)
- **Pypianoroll** (**new package!!**)

Future Works

Full Song Generation

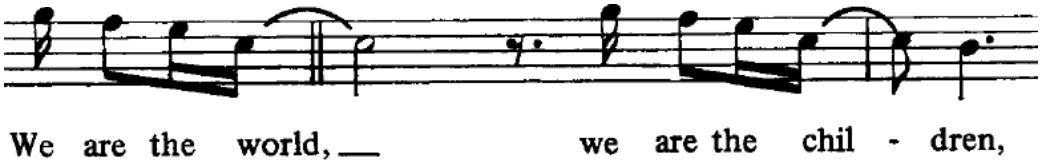


Hierarchical Temporal Structure

Future Works

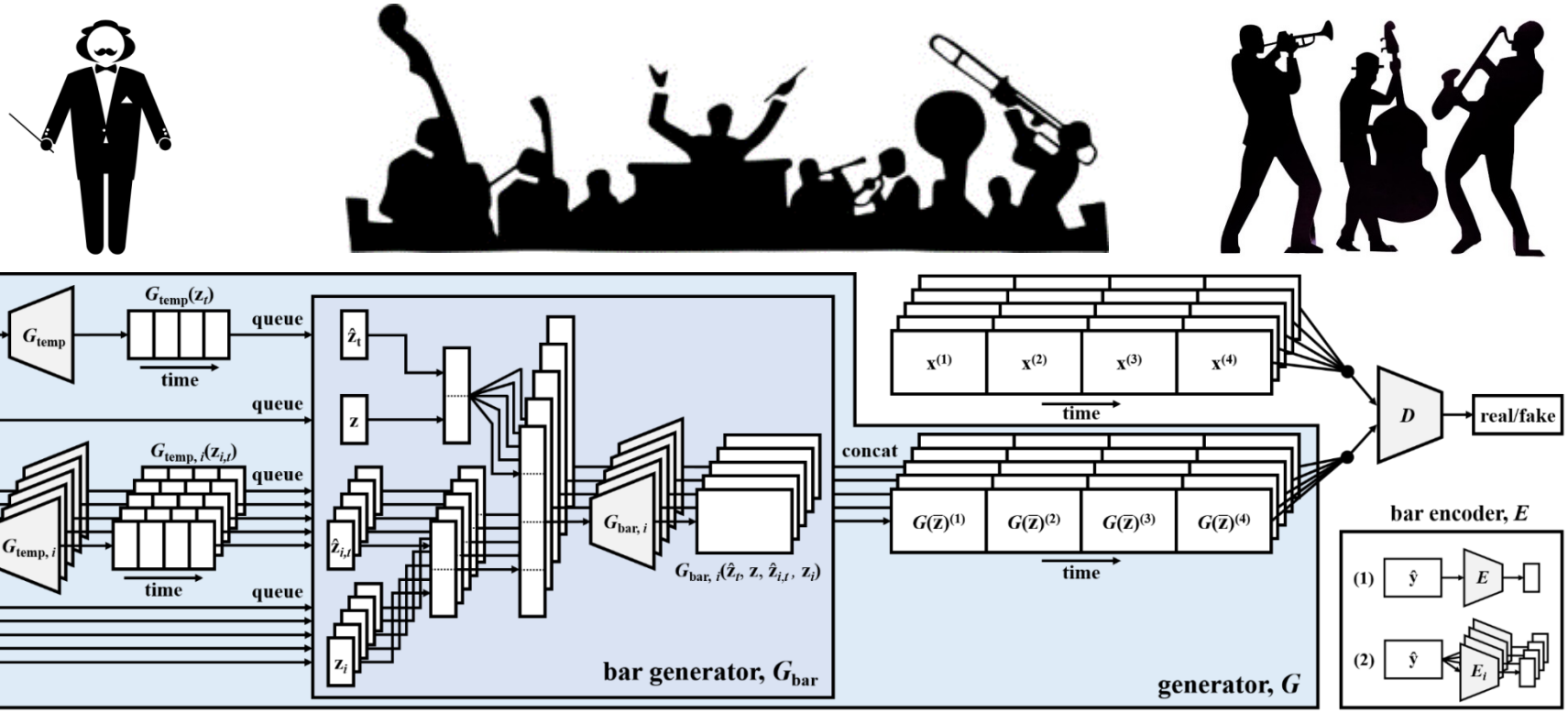
Cross-modal Generation

- Music + Video
- Music + Lyrics
- Video + Text



Source Code https://github.com/salu133445/musegan

Demo Page https://salu133445.github.io/musegan/



Q&A

MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment