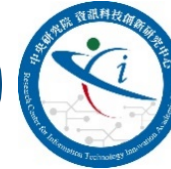


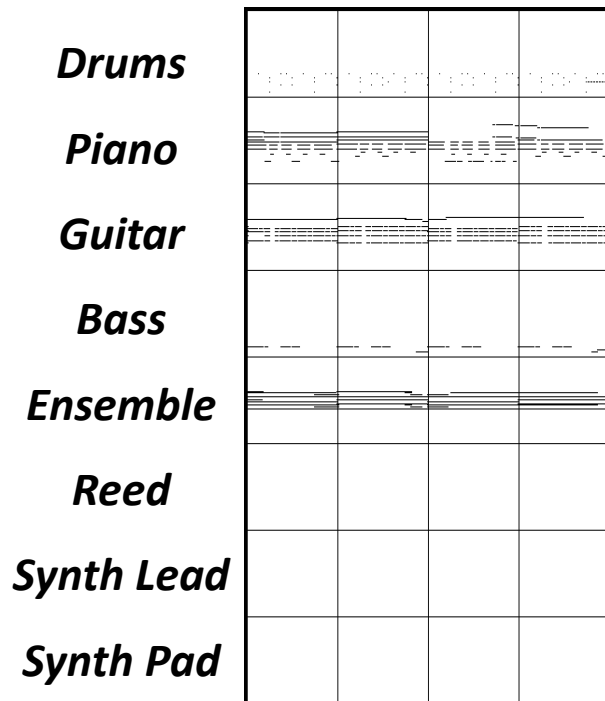
# Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation

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



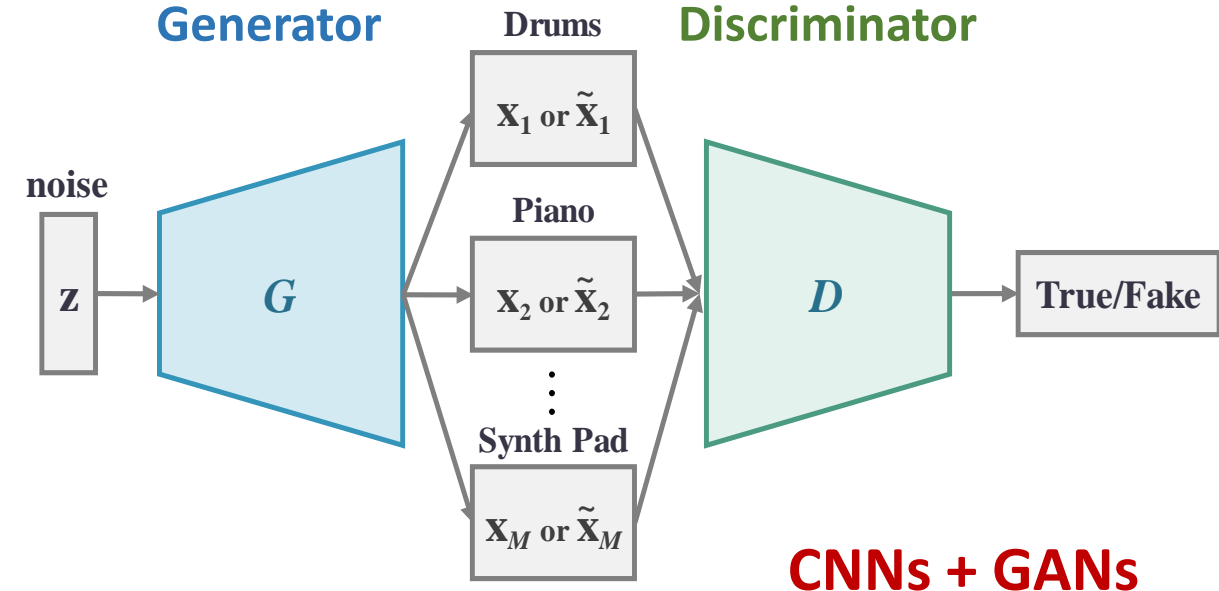
**stand B-7**

## Target outputs - Multi-track piano-rolls



**multi-track  
polyphonic**

## Convolutional Generative Adversarial Networks



# Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation

Hao-Wen Dong and Yi-Hsuan Yang

stand B-7

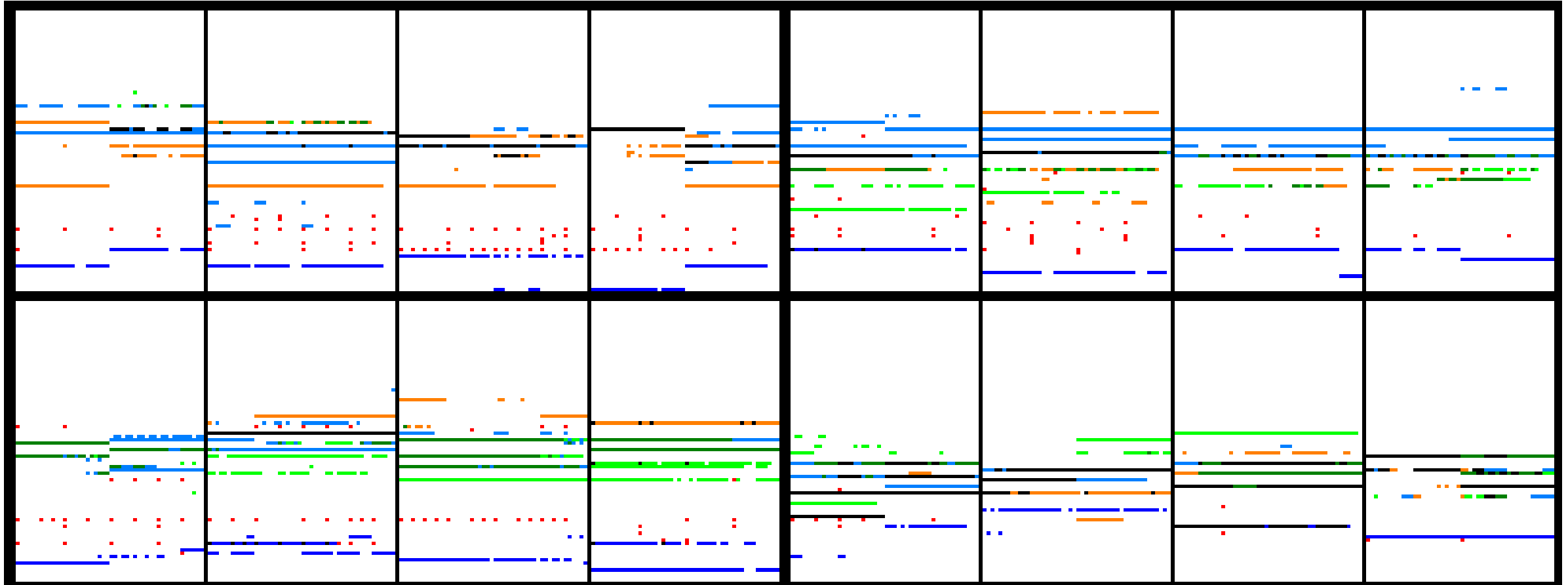
*Drums*

*Piano*

*Guitar*

*Bass*

*Strings*



# Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation

Hao-Wen Dong and Yi-Hsuan Yang

stand B-7

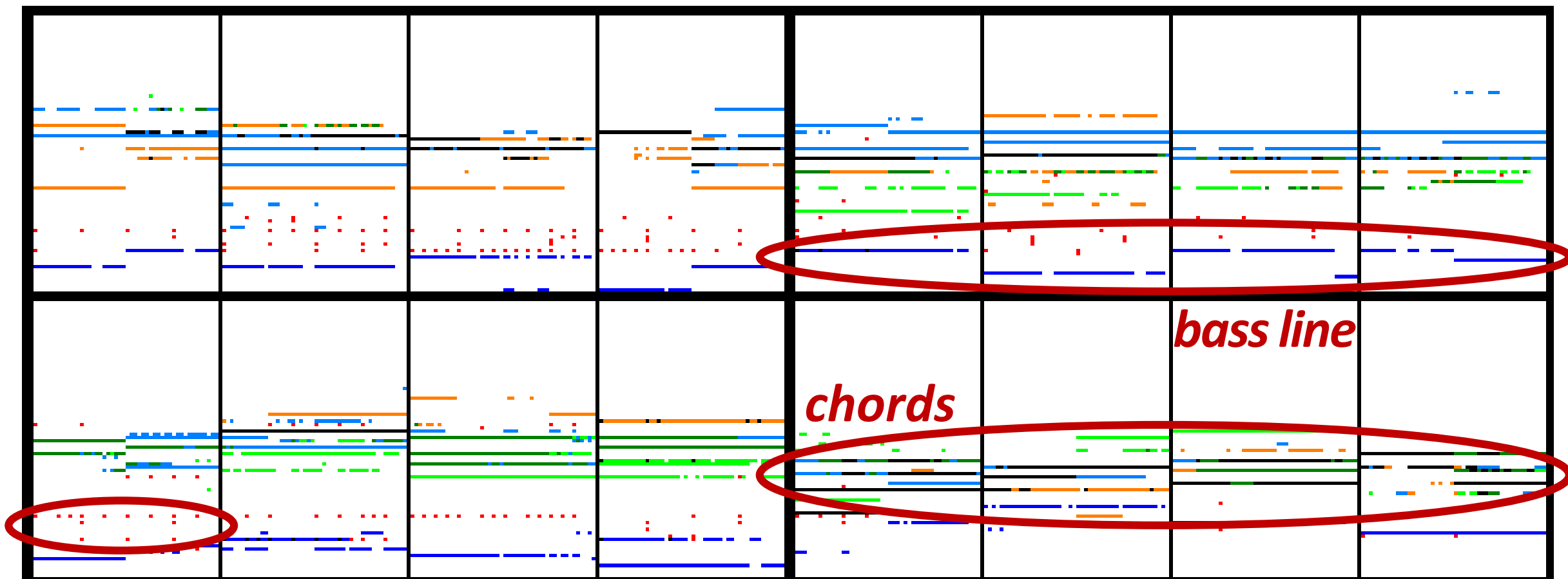
*Drums*

*Piano*

*Guitar*

*Bass*

*Strings*



*drum patterns*

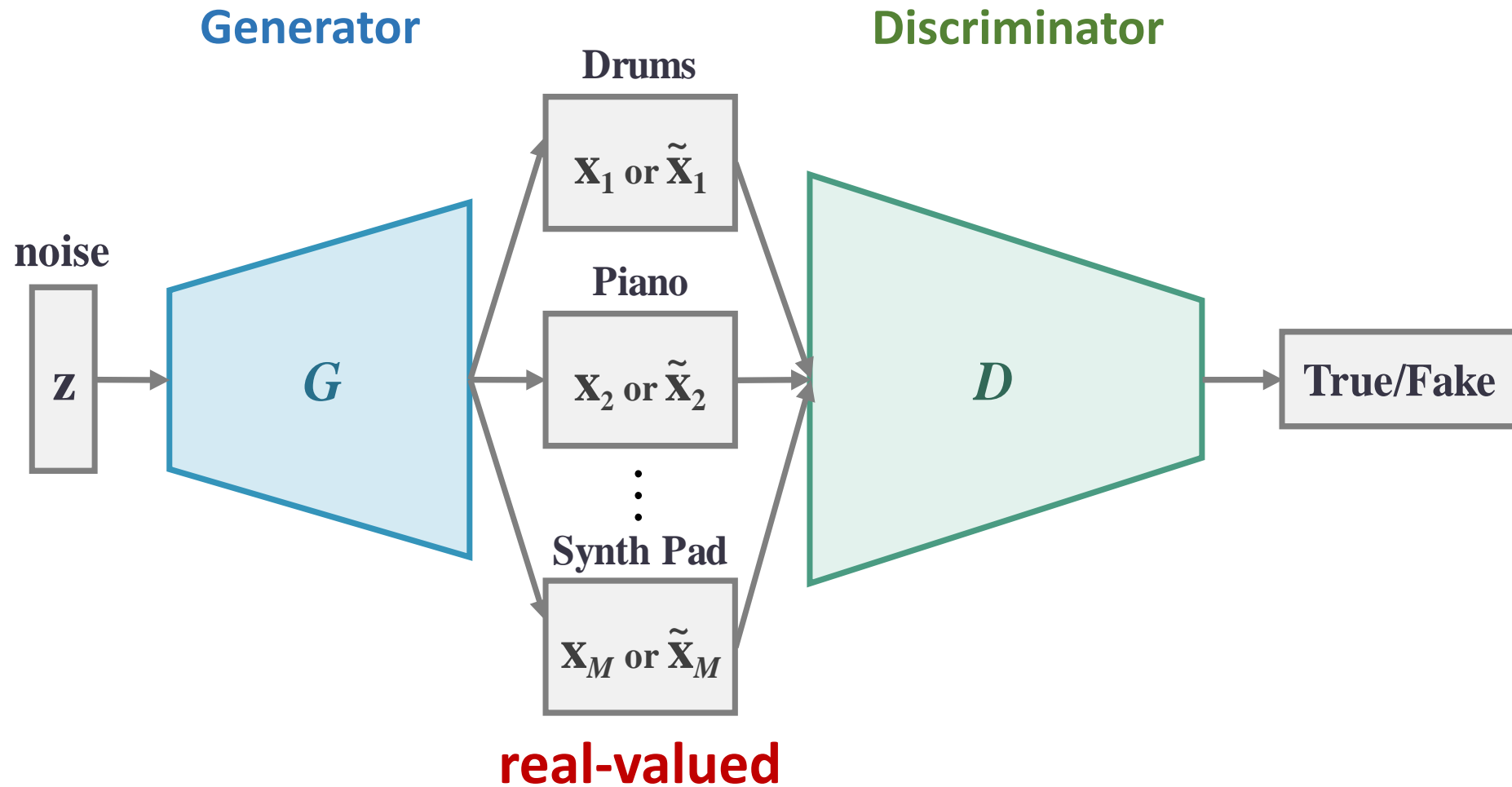
*chords*

*bass line*

# Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation

Hao-Wen Dong and Yi-Hsuan Yang

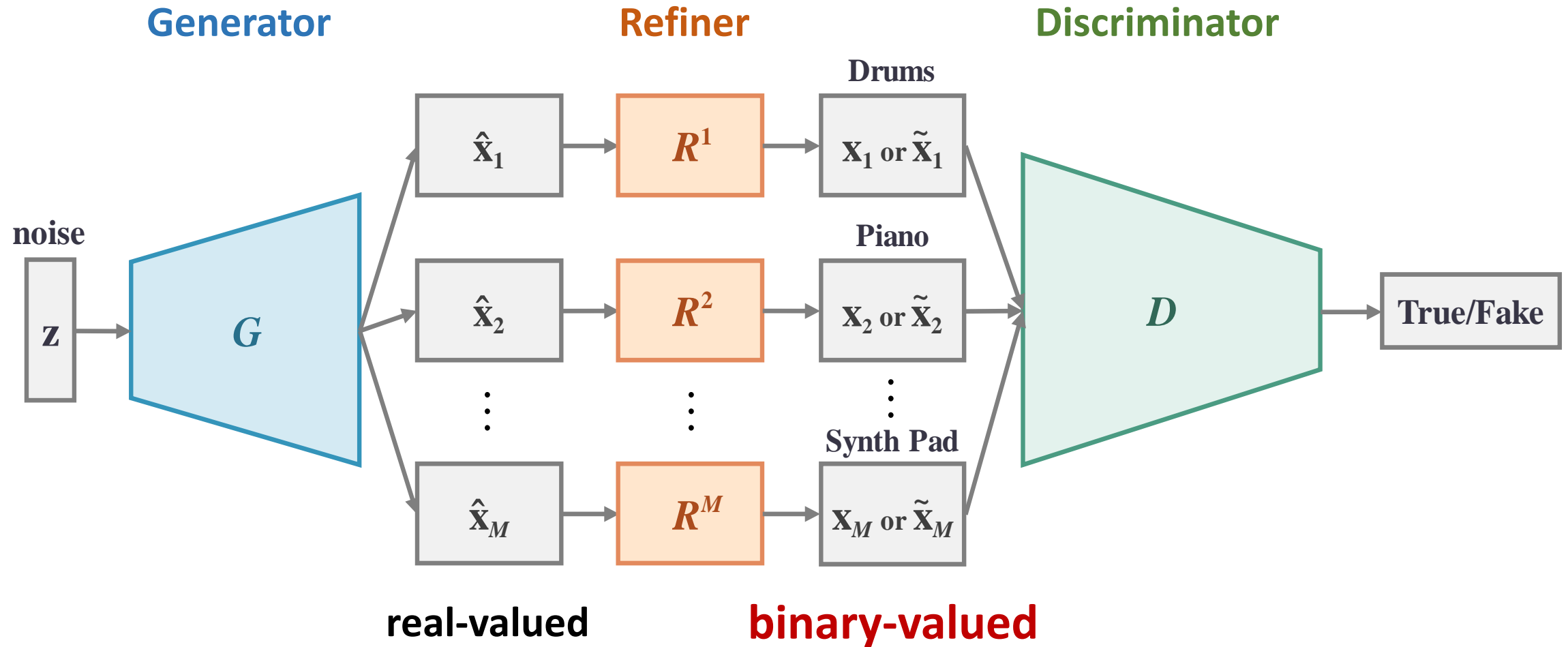
stand B-7



# Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation

Hao-Wen Dong and Yi-Hsuan Yang

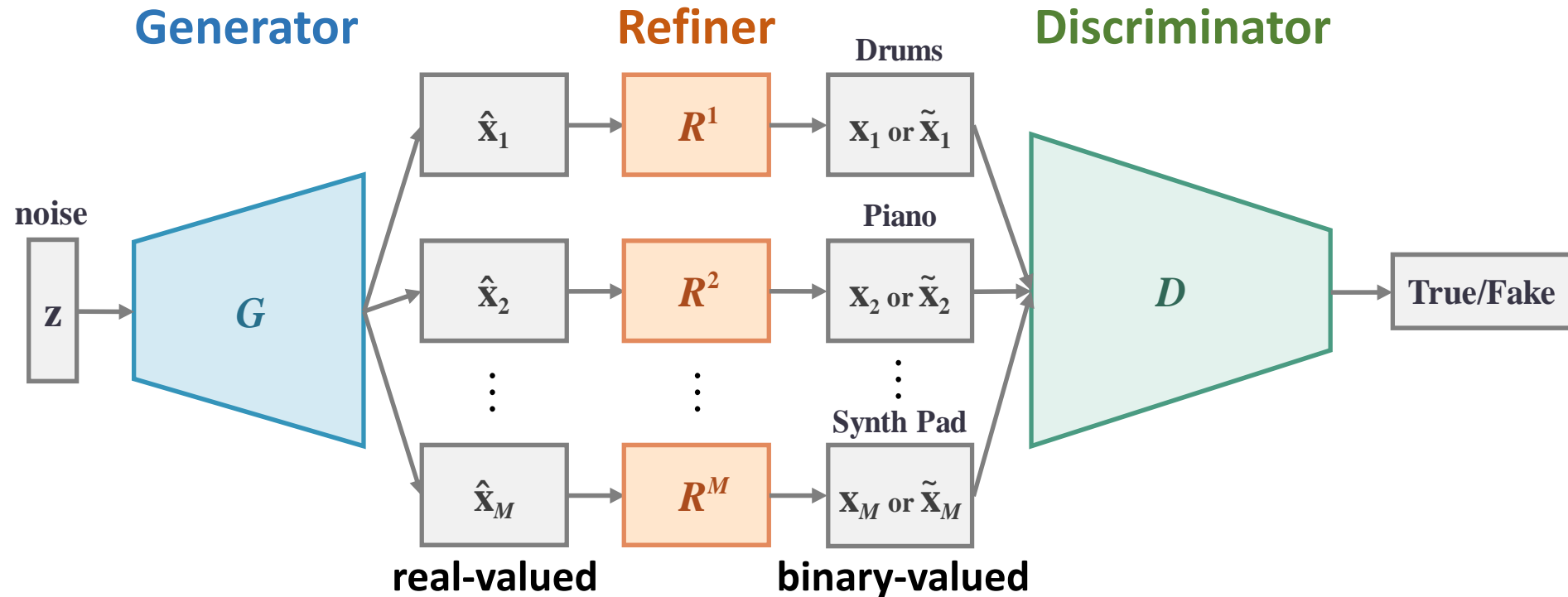
stand B-7



# Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation

Hao-Wen Dong and Yi-Hsuan Yang

stand B-7



|               | Generator's outputs  | Real data     |
|---------------|----------------------|---------------|
| MuseGAN       | real-valued          | binary-valued |
| BinaryMuseGAN | <b>binary-valued</b> | binary-valued |

# Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation

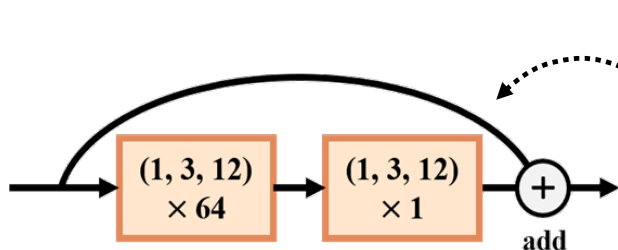
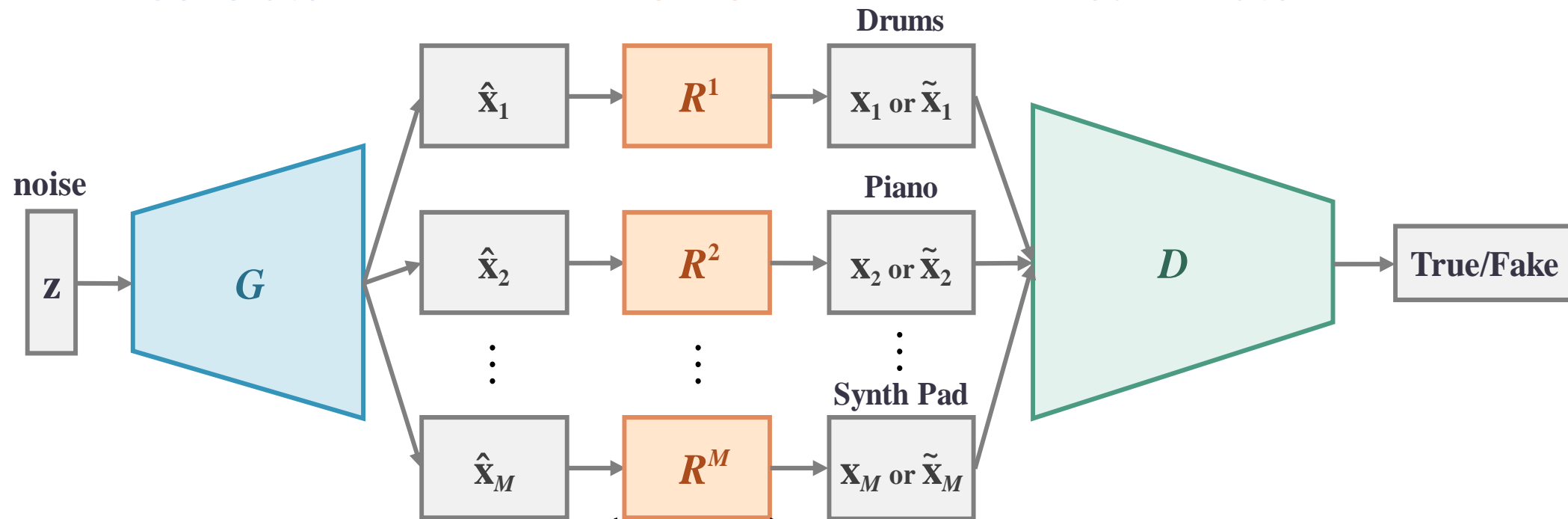
Hao-Wen Dong and Yi-Hsuan Yang

stand B-7

Generator

Refiner

Discriminator



**Deterministic Binary Neurons (DBNs)**

$$DBN(x) = u(\sigma(x) - 0.5)$$

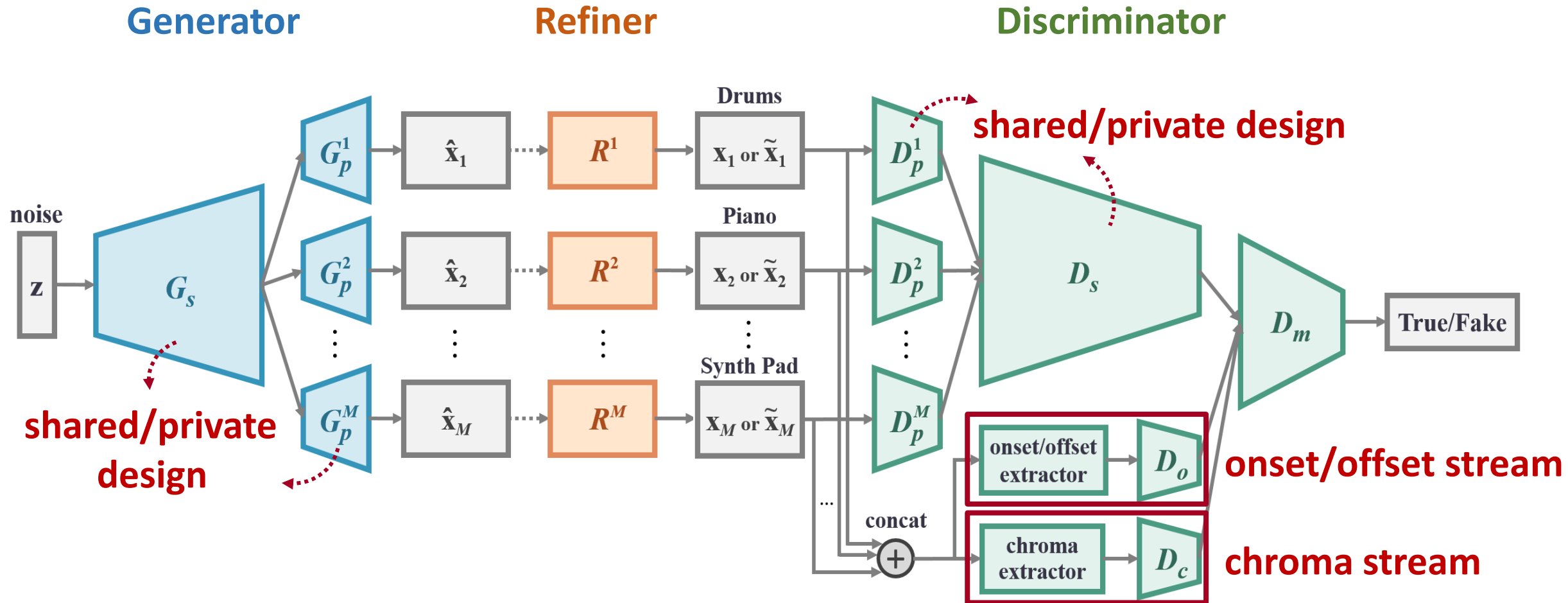
**Stochastic Binary Neurons (SBNs)**

$$SBN(x) = u(\sigma(x) - v), v \sim U[0, 1]$$

# Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation

Hao-Wen Dong and Yi-Hsuan Yang

stand B-7

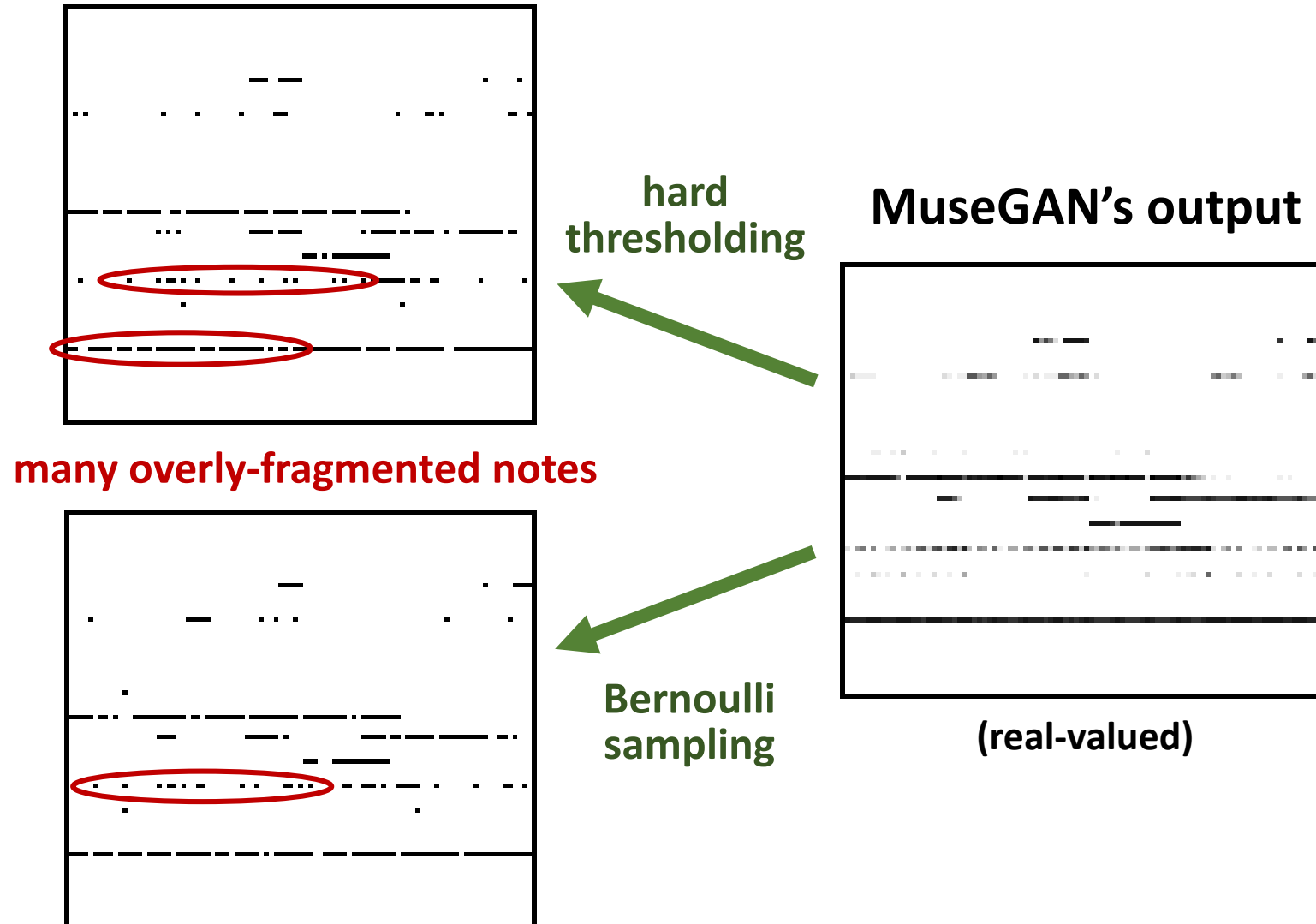




# Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation

Hao-Wen Dong and Yi-Hsuan Yang

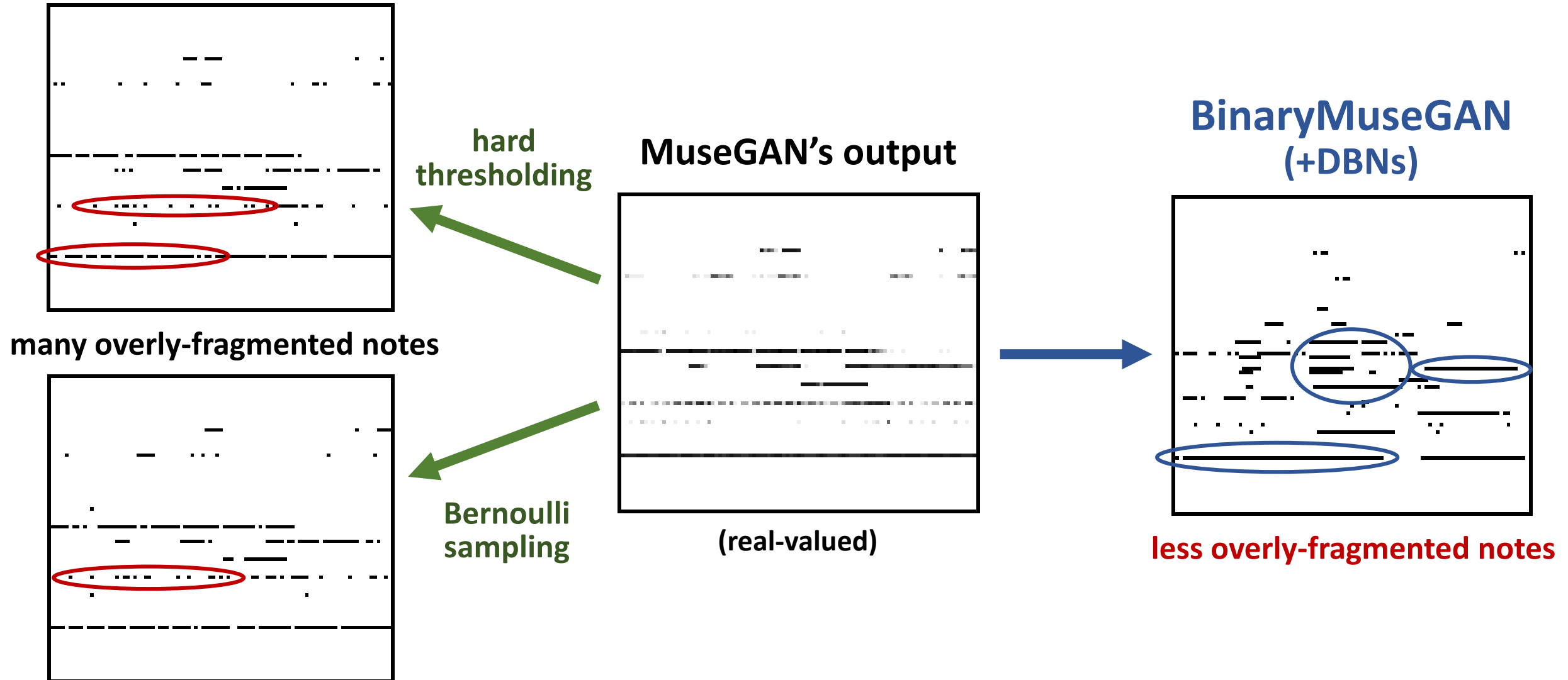
stand B-7



# Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation

Hao-Wen Dong and Yi-Hsuan Yang

stand B-7



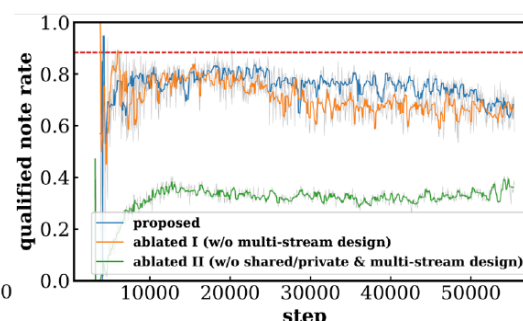
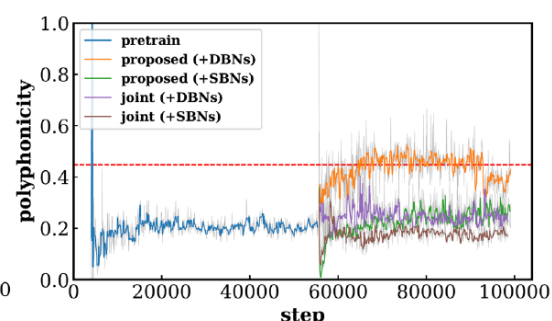
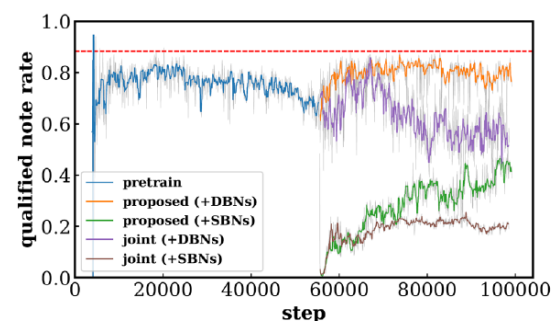
# Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation

Hao-Wen Dong and Yi-Hsuan Yang

stand B-7

MuseGAN BinaryMuseGAN (different training strategies)

|    | training data | pretrained  |             | proposed    |             | joint       |      | end-to-end  |             |
|----|---------------|-------------|-------------|-------------|-------------|-------------|------|-------------|-------------|
|    |               | BS          | HT          | SBNs        | DBNs        | SBNs        | DBNs | SBNs        | DBNs        |
| QN | 0.88          | <b>0.67</b> | <b>0.72</b> | 0.42        | <b>0.78</b> | 0.18        | 0.55 | <b>0.67</b> | 0.28        |
| PP | 0.48          | 0.20        | 0.22        | <b>0.26</b> | <b>0.45</b> | 0.19        | 0.19 | 0.16        | <b>0.29</b> |
| TD | 0.96          | <b>0.98</b> | 1.00        | <b>0.99</b> | <b>0.87</b> | <b>0.95</b> | 1.00 | 1.40        | 1.10        |



Come to learn more and listen to the demos!

## Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation

Hao-Wen Dong and Yi-Hsuan Yang

Research Center for IT Innovation, Academia Sinica, Taipei, Taiwan

[Demo Website] <https://salu133445.github.io/bmusegan/>



### >> Introduction

# MuseGAN [1] shows the promise of using GANs [2] with CNNs to generate *multitrack pianorolls*. But it requires further postprocessing at test time to binarize the generator's (G) output

# BinaryMuseGAN (proposed) adopts *binary neurons* [3] to binarize G's output during training

### >> Data

# Lakh Pianoroll Dataset (LPD) — *LPD-cleaned* subset

# Consider only songs with an *alternative* tag to make the training data cleaner

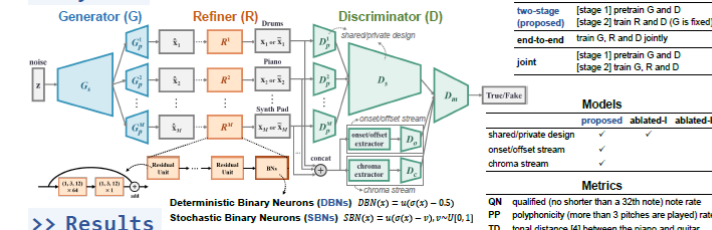
# 13,746 4-bar phrases from 2,291 songs

# 96 time steps in a bar, 84 possible pitches (C1 to B7)

# 8 tracks — Drums, Piano, Guitar, Bass, Ensemble, Reed, Synth Lead and Synth Pad

# Target output tensor shape — (4, 96, 84, 8)

### >> System



### >> Results

|    | training data | pretrained  |      | proposed    |             | joint       |      | end-to-end |             | ablated-I |      | ablated-II |      |
|----|---------------|-------------|------|-------------|-------------|-------------|------|------------|-------------|-----------|------|------------|------|
|    |               | BS          | HT   | SBNs        | DBNs        | SBNs        | DBNs | SBNs       | DBNs        | BS        | HT   | BS         | HT   |
| QN | 0.88          | 0.67        | 0.72 | 0.42        | <b>0.78</b> | 0.18        | 0.55 | 0.67       | 0.28        | 0.61      | 0.64 | 0.35       | 0.37 |
| PP | 0.48          | 0.20        | 0.22 | <b>0.26</b> | <b>0.45</b> | 0.19        | 0.19 | 0.16       | <b>0.29</b> | 0.19      | 0.20 | 0.14       | 0.14 |
| TD | 0.96          | <b>0.98</b> | 1.00 | <b>0.99</b> | <b>0.87</b> | <b>0.95</b> | 1.00 | 1.40       | 1.10        | 1.00      | 1.00 | 1.30       | 1.40 |

### >> Conclusions

# While the generated results appear preliminary and lack musicality, we showed the potential of adopting binary neurons in a music generation system

# Using DBNs leads to better objective scores than hard thresholding, Bernoulli sampling and SBNs

# It might also be interesting to use binary neurons in music transcription (binary-valued outputs as well)

### >> References

- [1] Hao-Wen Dong, Wen-Yi Hsiao, Li-Chia Yang, and Yi-Hsuan Yang. MuseGAN: Symbolic-domain music generation and accompaniment with multi-track sequential generative adversarial networks. In *Proc. AAAI*, 2018.
- [2] Ian J. Goodfellow et al. Generative adversarial nets. In *Proc. NIPS*, 2014.
- [3] Yoshua Bengio, Nicholas Leonard, and Aaron C. Courville. Estimating or propagating gradients through stochastic neurons for conditional computation. *arXiv preprint arXiv:1308.3432*, 2013.
- [4] Christopher Hart, Mark Sandler, and Martin Gasser. Detecting harmonic change in musical audio. In *Proc. ACM Workshop on Audio and Music Computing Multimedia*, 2006.