

# EXPLORING SAMPLING TECHNIQUES FOR GENERATING MELODIES WITH A TRANSFORMER

Mathias Rose Bjare<sup>1</sup>, Stefan Lattner<sup>2</sup> and Gerhard Widmer<sup>1,3</sup>

<sup>1</sup>Institute of Computational Perception (CP-JKU), Johannes Kepler University Linz, Austria

<sup>2</sup>Sony Computer Science Laboratories (CSL), Paris, France

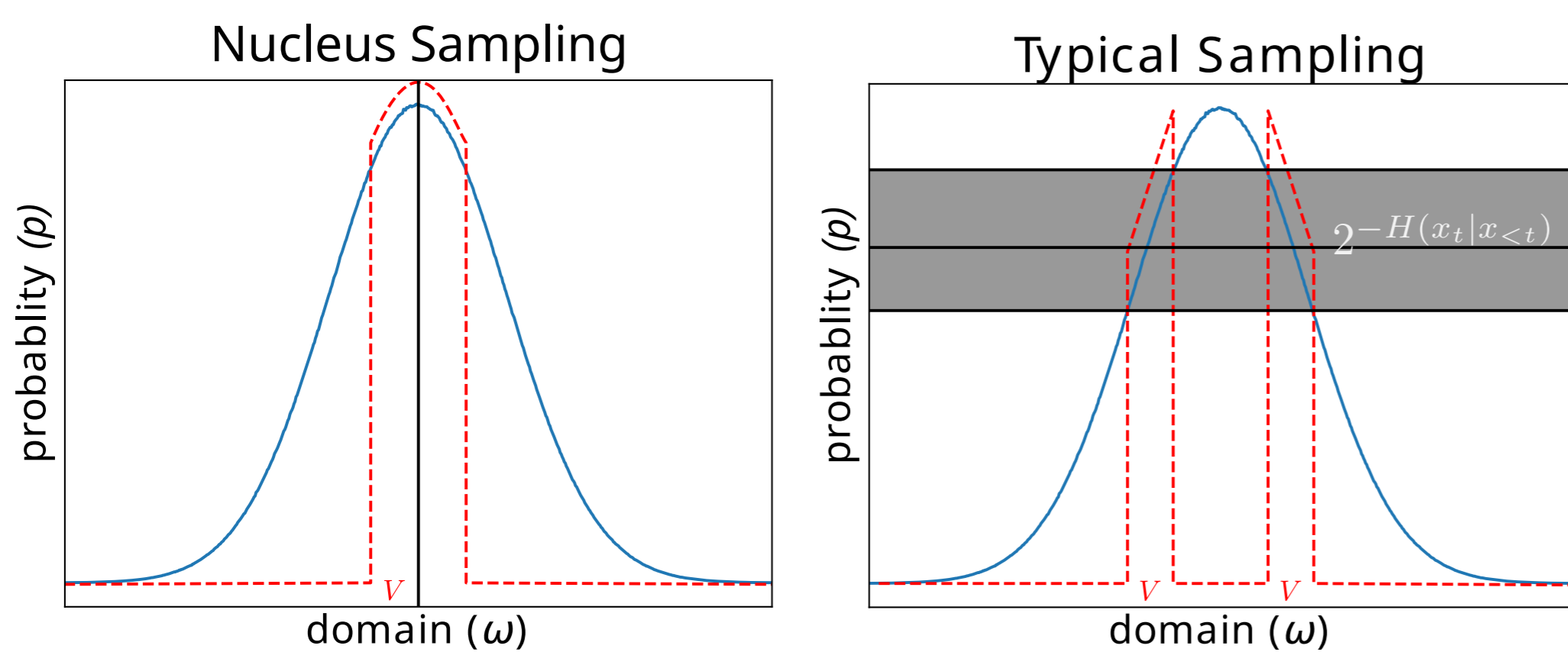
<sup>3</sup>LIT AI Lab, Linz Institute of Technology, Austria

## INTRODUCTION

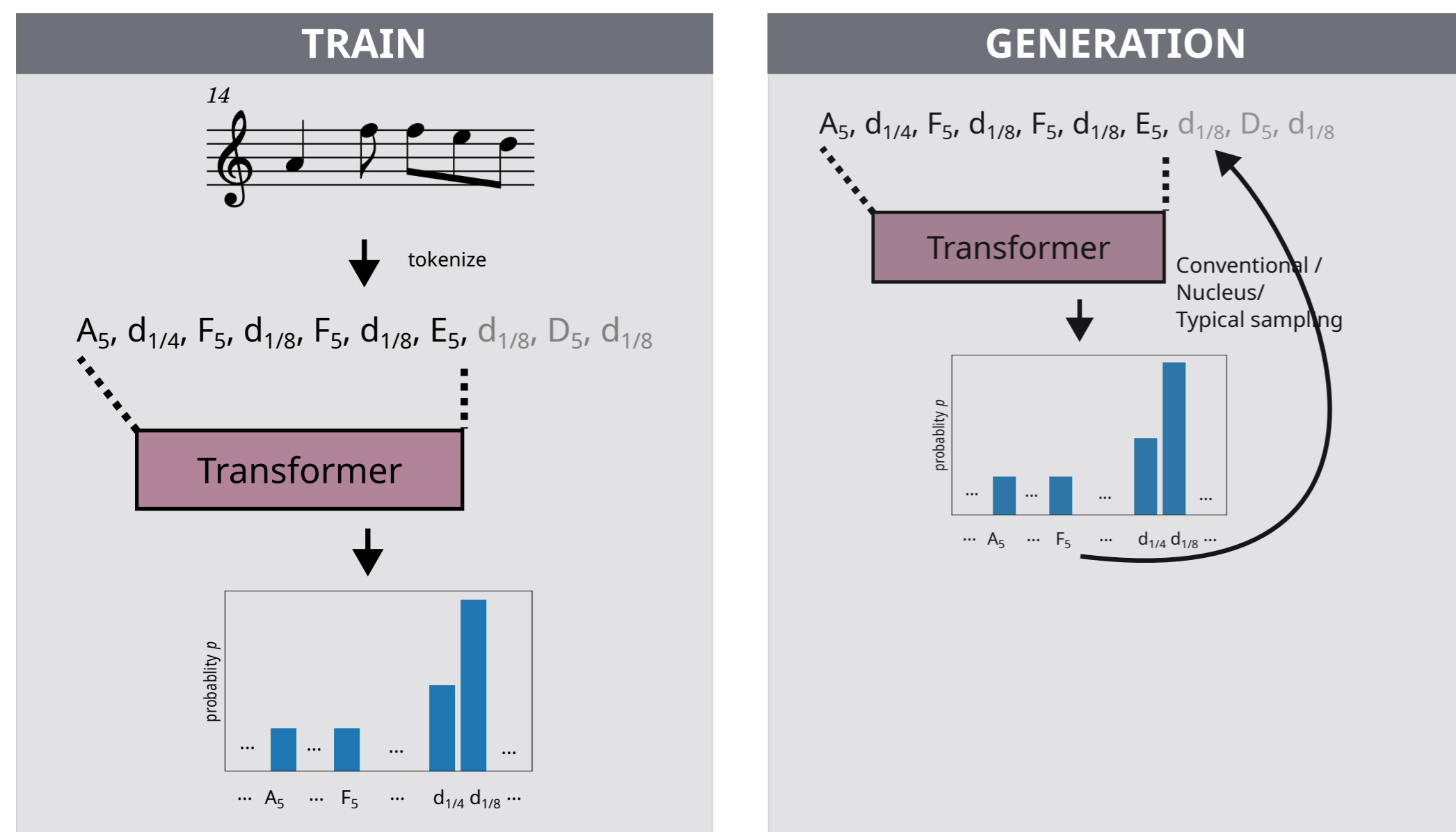
- We investigate the impact of distribution sampling techniques on musical qualities such as diversity and structure.
- We evaluate the effect of the sampling techniques in optimal circumstances and suboptimal circumstances.

## DISTRIBUTION TRUNCATION

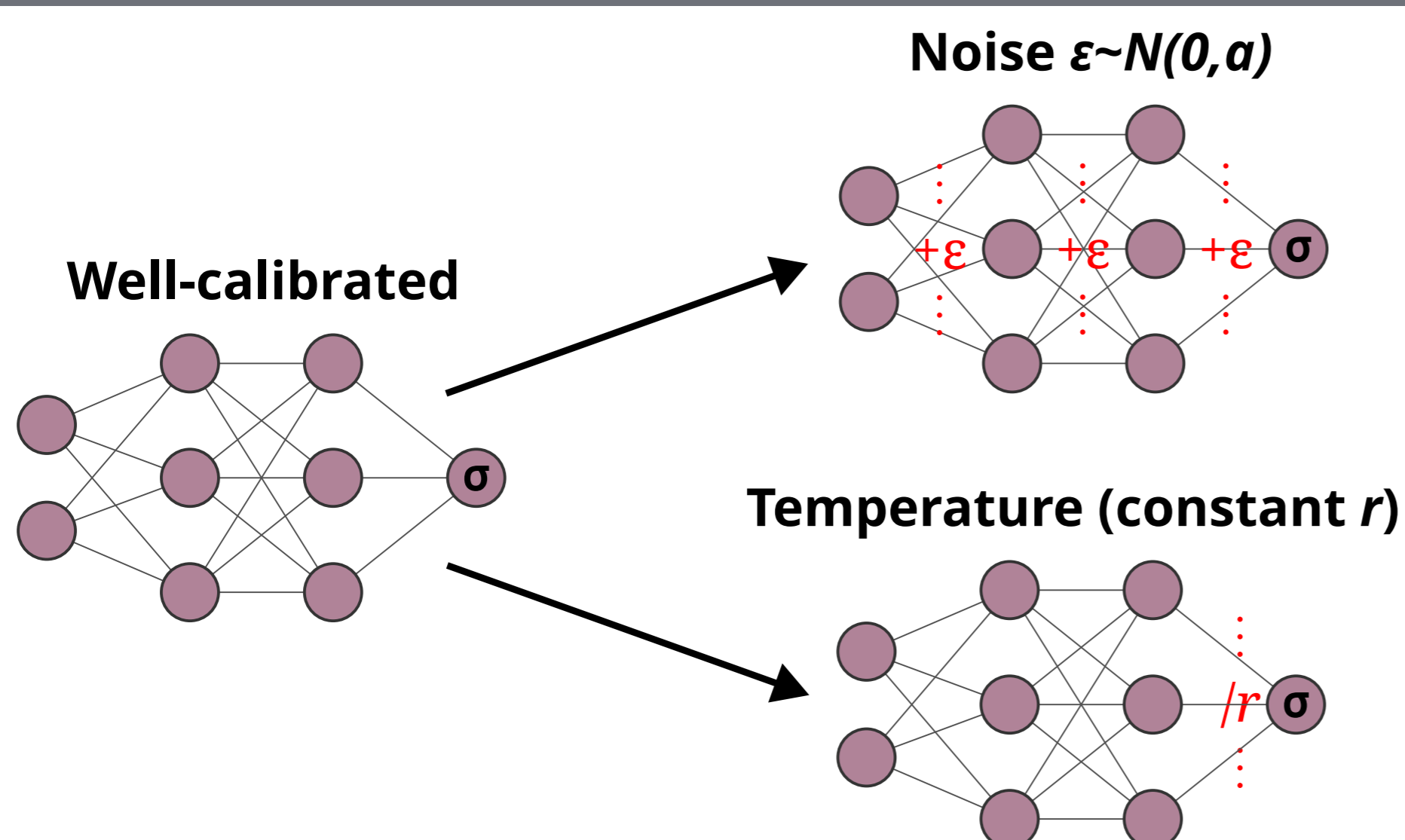
- **Nucleus sampling**: removes the largest set of events with a probability that sums to a threshold.
- **Typical sampling**: sorts events by "typicality" (deviation of event's information content to entropy) and removes the least typical events with a probability that sums to a threshold.



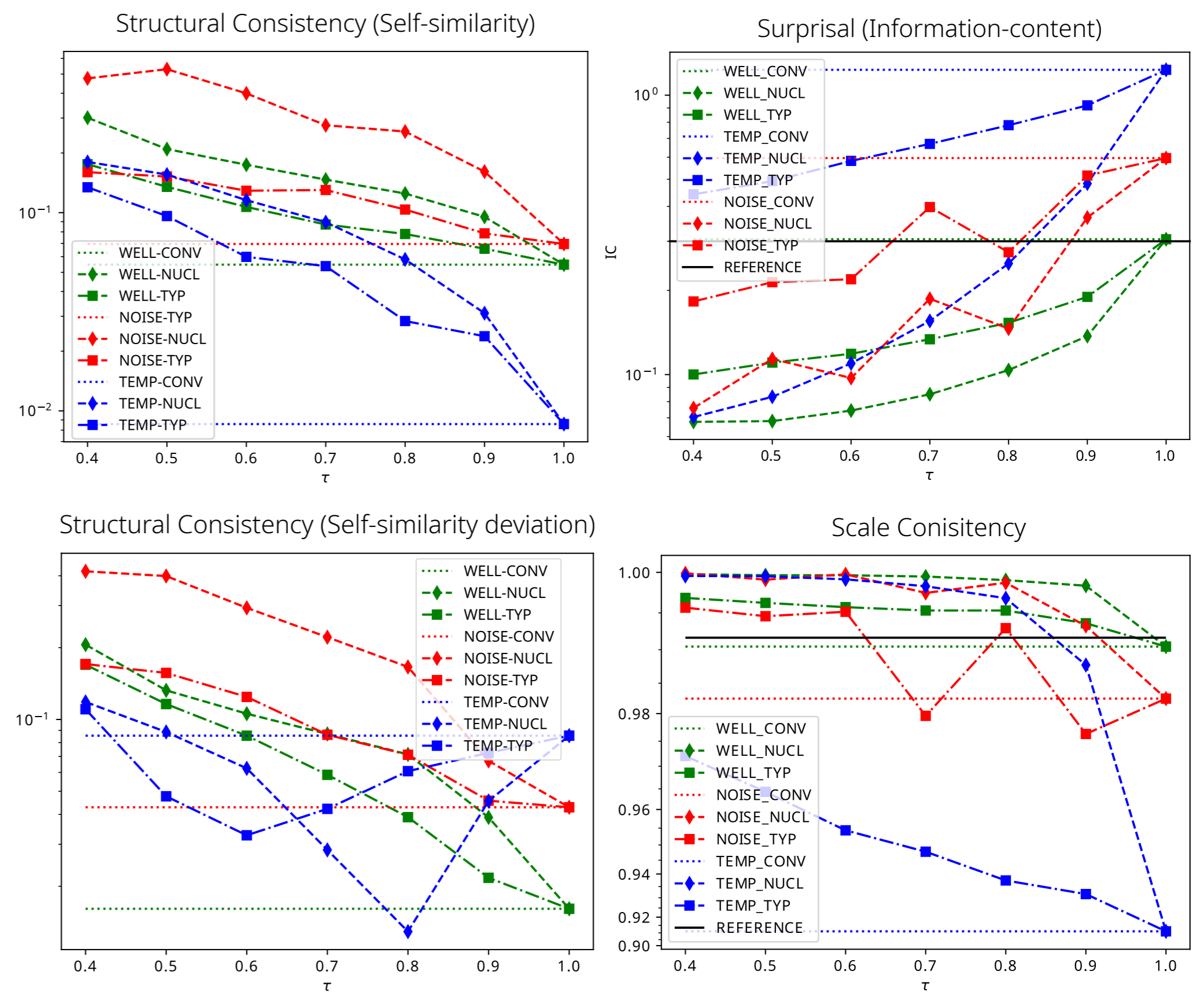
## EXPERIMENTS



## MODEL DEGRADATIONS



## OBJECTIVE EVALUATIONS



## SUBJECTIVE EVALUATIONS

| Method     | QULT    | ST_STR         | LT_STR         | CPLX           |
|------------|---------|----------------|----------------|----------------|
| REFERENCE  | 3.7±1.0 | 3.8±1.0        | 3.7±1.1        | 3.6±0.8        |
| WELL_CONV  | 3.2±1.1 | 3.7±0.9        | 3.5±1.2        | 3.3±1.0        |
| WELL_NUCL  | 3.6±1.1 | <b>3.9±1.1</b> | <b>3.7±1.1</b> | 2.8±1.0        |
| WELL_TYP   | 3.4±1.2 | 3.6±0.9        | <b>3.7±1.0</b> | 3.3±1.0        |
| NOISE_CONV | 2.7±1.0 | 3.2±0.9        | 3.0±1.0        | 2.8±0.9        |
| NOISE_NUCL | 2.6±1.3 | 3.2±1.4        | 2.8±1.5        | 2.5±1.2        |
| NOISE_TYP  | 2.7±1.1 | 3.2±1.1        | 3.1±1.2        | 2.4±1.0        |
| TEMP_CONV  | 2.1±1.3 | 2.7±1.1        | 2.1±1.1        | <b>3.7±1.0</b> |
| TEMP_NUCL  | 3.4±1.2 | 3.6±0.9        | 3.4±1.3        | 3.4±1.1        |
| TEMP_TYP   | 2.2±1.1 | 2.7±0.9        | 2.4±1.0        | 3.3±0.8        |

Listener study mean-opinion score  $\pm$  one standard deviation. The measured attributes are overall quality (**QULT**), perceived short-term structure (**ST\_STR**), long-term structure (**LT\_STR**), and complexity (**CPLX**).

## CONCLUSION

- Higher truncation strength leads to **increased structural and tonal consistency**.
- The truncation techniques improved musical qualities in suboptimal circumstances but not in optimal.

