

# Kohei SAIJO

E-mail: [saijo \[at\] pcl.cs.waseda.ac.jp](mailto:saijo[at]pcl.cs.waseda.ac.jp) | Office Phone: (+81)3-3203-4468 |

[🏠 Personal Website](#) | [🎓 Google Scholar](#) | [🐙 Github](#)

## Education

- 
- |   |                       |
|---|-----------------------|
| <b>Waseda University (Tokyo, Japan)</b> , Ph. D. in Computer Science and Engineering  | Apr. 2023 – Present   |
| <ul style="list-style-type: none"> <li>• Supervisor: Tetsuji Ogawa</li> <li>• Research topic: Universal and unified sound source separation</li> </ul>            |                       |
| <b>Waseda University (Tokyo, Japan)</b> , M.E. in Computer Science and Engineering  | Apr. 2021 – Mar. 2023 |
| <ul style="list-style-type: none"> <li>• Supervisor: Tetsuji Ogawa</li> <li>• Thesis title: Unsupervised Source Separation via Separation and Remixing</li> </ul> |                       |
| <b>Waseda University (Tokyo, Japan)</b> , B.E. in Computer Science and Engineering  | Apr. 2017 – Mar. 2021 |
| <ul style="list-style-type: none"> <li>• Supervisor: Tetsuji Ogawa</li> <li>• Thesis title: Neural speech enhancement robust to speaker movement</li> </ul>       |                       |

## Experience

- 
- |  |                       |
|--|-----------------------|
| <b>Reserach Assistant</b> , National Institute of Advanced Industrial Science and Technology – Tokyo, Japan  | Oct. 2024 – Present   |
| <ul style="list-style-type: none"> <li>• Mentor: Yoshiaki Bando</li> </ul>   |                       |
| <b>Research Internship</b> , Mitsubishi Electric Research Laboratories – Cambridge, MA, USA  | Nov. 2023 – Aug. 2024 |
| <ul style="list-style-type: none"> <li>• Mentor: Jonathan Le Roux</li> <li>• Worked on sound source separation</li> <li>• Major publication(s): <a href="#">[Saijo+, Interspeech'24]</a>, <a href="#">[Saijo+, IWAENC'24]</a>, <a href="#">[Saijo+, ICASSP'25]</a>, <a href="#">[Saijo+, ICASSP'25]</a></li> </ul> |                       |
| <b>Visiting Scholar</b> , Carnegie Mellon University – Pittsburgh, PA, USA   | Apr. 2023 – Jul. 2023 |
| <ul style="list-style-type: none"> <li>• Mentor: Shinji Watanabe</li> <li>• Worked on multi-task universal speech enhancement</li> <li>• Major publication(s): <a href="#">[Saijo+, ASRU'23]</a></li> </ul>  |                       |
| <b>Part-time Researcher</b> , LINE corporation – Tokyo, Japan  | Sep. 2021 – Apr. 2022 |
| <ul style="list-style-type: none"> <li>• Mentor: Robin Scheibler</li> <li>• Worked on unsupervised multi-channel source separation</li> <li>• Major publication(s): <a href="#">[Saijo+, Interspeech'22]</a></li> </ul>  |                       |
| <b>Research Internship</b> , LINE corporation – Tokyo, Japan   | Aug. 2021 – Sep. 2021 |
| <ul style="list-style-type: none"> <li>• Mentor: Robin Scheibler</li> <li>• Worked on multi-channel joint source separation and dereverberation</li> <li>• Major publication(s): <a href="#">[Saijo+, Interspeech'22]</a></li> </ul>   |                       |

## External Research Funding

- 
- |  |                       |
|--|-----------------------|
| <b>Research Fellowship for Young Scientists (DC2)   1,600,000 JPY</b><br><i>from Japan Society for the Promotion of Science (JSPS)</i>                             | Apr. 2024 - Mar. 2026 |
| <ul style="list-style-type: none"> <li>• The goal of this research project is to establish a framework that enables us to build an on-demand separation</li> </ul> |                       |

system in any target domain of interest. To this end, I am working on (i) improving the stability and performance of unsupervised sound source separation and (ii) constructing a lightweight but high-performing source separation model.

#### **Super Global University (Visit to Carnegie Mellon University) | 800,000 JPY**

Apr. 2023 - Jul. 2023

*from Top Global University Japan Project*

- In this study abroad program, I have participated in joint research with Prof. Shinji Watanabe at the Language Technologies Institute, Carnegie Mellon University. The primary goal of our project was to build a single speech enhancement model that can handle all the major speech enhancement tasks.
- Major publications: [\[Saijo+, ASRU'23\]](#), [\[Zhang+, ASRU'23\]](#)

#### **Support for Pioneering Research Initiated by the Next Generation (SPRING) | 500,000 JPY**

Apr. 2023 - Mar. 2024

*from Japan Science and Technology Agency (JST)*

- This research project aimed to establish a universal unsupervised source separation method to build an on-demand separation system in any target domain of interest. To this end, I have established a novel unsupervised source separation framework which works without assuming the sound source types.
- Major publications: [\[Saijo+, ICASSP'23\]](#), [\[Saijo+, Interspeech'23\]](#),

### **Awards**

---

#### **ISS Young Researcher's Award in Speech Field**

Dec. 2023

*from the Institute of Electronics, Information and Communication Engineers (IEICE)*

#### **Repayment Exemption for Graduate Students with Excellent Achievements (full-exemption)**

Mar. 2023

*from Japan Student Services Organization (JASSO)*

#### **Best Student Paper Award**

Mar. 2022

*from the Acoustical Society of Japan (ASJ)*

### **Reviewing Activity**

---

- ICASSP: 2024, 2025
- Interspeech: 2024, 2025
- WASPAA: 2025
- IEEE/ACM TASLP: 2024, 2025
- Speech communication: 2023, 2024

### **Skills**

---

**Language:** Japanese (native), English (conversant)

**Programming:** Python, C, LaTeX

**Software & Tools:** PyTorch, TensorFlow, ESPnet, Asteroid

### **Publications**

---

#### **Journal Paper (First Author)**

1. Kohei Saijo, Tetsuji Ogawa, "Stabilizing and Enhancing Remixing-based Unsupervised Sound Source Separation," *APSIPA Transactions on Signal and Information Processing*, Vol. 14 (1), 2025.

#### **International Conference (First Author)**

1. Kohei Saijo, Yoshiaki Bando, “**Is MixIT Really Unsuitable for Correlated Sources? Exploring MixIT for Unsupervised Pre-training in Music Source Separation**,” *2025 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, October 2025.
2. Kohei Saijo, Tetsuji Ogawa, “**A Comparative Study on Positional Encoding for Time-frequency Domain Dual-path Transformer-based Source Separation Models**,” *2025 33rd European Signal Processing Conference (EUSIPCO)*, September 2025.
3. Kohei Saijo, Wangyou Zhang, Samuele Cornell, Robin Scheibler, Chenda Li, Zhaoheng Ni, Anurag Kumar, Marvin Sach, Yihui Fu, Wei Wang, Tim Fingscheidt, and Shinji Watanabe, “**Interspeech 2025 URGENT Speech Enhancement Challenge**,” *2025 26th Annual Conference of International Speech Communication Association (INTERSPEECH)*, August 2025.
4. Kohei Saijo, Janek Ebbers, François G Germain, Gordon Wichern, Jonathan Le Roux, “**Task-Aware Unified Source Separation**,” *2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2025.
5. Kohei Saijo, Janek Ebbers, François G Germain, Sameer Khurana, Gordon Wichern, Jonathan Le Roux, “**Leveraging Audio-Only Data for Text-Queried Target Sound Extraction**,” *2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2025.
6. Kohei Saijo, Gordon Wichern, François G. Germain, Zexu Pan, Jonathan Le Roux, “**TF-LoCoformer: Transformer with Local Modeling by Convolution for Speech Separation and Enhancement**,” *International Workshop on Acoustic Signal Enhancement (IWAENC)*, September 2024 [**Best student paper award finalist**].
7. Kohei Saijo, Gordon Wichern, François G. Germain, Zexu Pan, Jonathan Le Roux, “**Enhanced Reverberation as Supervision for Unsupervised Speech Separation**,” *2024 25th Annual Conference of International Speech Communication Association (INTERSPEECH)*, September 2024.
8. Kohei Saijo, Wangyou Zhang, Zhong-Qiu Wang, Shinji Watanabe, Tetsunori Kobayashi, and Tetsuji Ogawa, “**A Single Speech Enhancement Model Unifying Dereverberation, Denoising, Speaker Counting, Separation, and Extraction**,” *2023 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, December 2023.
9. Kohei Saijo, Tetsuji Ogawa, “**Remixing-based Unsupervised Source Separation from Scratch**,” *2023 24th Annual Conference of International Speech Communication Association (INTERSPEECH)*, August 2023.
10. Kohei Saijo, Tetsuji Ogawa, “**Self-Remixing: Unsupervised Speech Separation via Separation and Remixing**,” *2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, June 2023.
11. Kohei Saijo, Tetsuji Ogawa, “**Unsupervised Training of Sequential Neural Beamformer Using Coarsely-separated and Non-separated Signal**,” *2022 23rd Annual Conference of International Speech Communication Association (INTERSPEECH)*, September 2022.
12. Kohei Saijo, Robin Scheibler, “**Spatial Loss for Unsupervised Multi-channel Source Separation**,” *2022 23rd Annual Conference of International Speech Communication Association (INTERSPEECH)*, September 2022.
13. Kohei Saijo, Robin Scheibler, “**Independence-based Joint Dereverberation and Separation with Neural Source Model**,” *2022 23rd Annual Conference of International Speech Communication Association (INTERSPEECH)*, September 2022.
14. Kohei Saijo, Tetsuji Ogawa, “**Remix-Cycle-Consistent Learning on Adversarially Learned Separator for Accurate and Stable Unsupervised Speech Separation**,” *2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2022.
15. Kohei Saijo, Kazuhiro Katagiri, Masaru Fujieda, Tetsunori Kobayashi, Tetsuji Ogawa, “**Comparative Study on DNN-based Minimum Variance Beamforming Robust to Small Movements of Sound Sources**,” *2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, December 2021.

#### International Conference (Co-author)

1. Francesco Paissan, Gordon Wichern, Yoshiki Masuyama, Ryo Aihara, François G. Germain, Kohei Saijo, Jonathan Le Roux, “**FasTUSS: Faster Task-Aware Unified Source Separation**,” *2025 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, October 2025.

2. Wangyou Zhang, Kohei Saijo, Samuele Cornell, Robin Scheibler, Chenda Li, Zhaoheng Ni, Anurag Kumar, Marvin Sach, Wei Wang, Yihui Fu, Shinji Watanabe, Tim Fingscheidt, and Yanmin Qian, **“Lessons Learned from the URGENT 2024 Speech Enhancement Challenge,”** 2025 26th Annual Conference of International Speech Communication Association (INTERSPEECH), August 2025.
3. Tomohiro Hayashi, Riku Ogino, Kohei Saijo, Tetsuji Ogawa, **“What to Refer and How? - Exploring Handling of Auxiliary Information in Target Speaker Extraction,”** 2024 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), December 2024.
4. Zexu Pan, Gordon Wichern, François G. Germain, Kohei Saijo, Jonathan Le Roux, **“PARIS: Pseudo-AutoRegressive Siamese Training for Online Speech Separation,”** 2024 25th Annual Conference of International Speech Communication Association (INTERSPEECH), September 2024.
5. Wangyou Zhang, Robin Scheibler, Kohei Saijo, Samuele Cornell, Chenda Li, Zhaoheng Ni, Anurag Kumar, Jan Pirklbauer, Marvin Sach, Shinji Watanabe, Tim Fingscheidt, and Yanmin Qian, **“URGENT Challenge: Universality, Robustness, and Generalizability For Speech Enhancement,”** 2024 25th Annual Conference of International Speech Communication Association (INTERSPEECH), September 2024.
6. Wangyou Zhang, Kohei Saijo, Jee-weon Jung, Chenda Li, Shinji Watanabe, and Yanmin Qian, **“Beyond Performance Plateaus: A Comprehensive Study on Scalability in Speech Enhancement,”** 2024 25th Annual Conference of International Speech Communication Association (INTERSPEECH), September 2024.
7. Xuankai Chang, Brian Yan, Kwanghee Choi, Jee-weon Jung, Yichen Lu, Soumi Maiti, Roshan Sharma, Jiatong Shi, Jinchuan Tian, Shinji Watanabe, Yuya Fujita, Takashi Maekaku, Pengcheng Guo, Yao-Fei Cheng, Pavel Denisov, Kohei Saijo, and Hsiu-Hsuan Wang, **“Exploring Speech Recognition, Translation, and Understanding with Discrete Speech Units: A Comparative Study,”** 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), April 2024.
8. Wangyou Zhang, Kohei Saijo, Zhong-Qiu Wang, Shinji Watanabe, and Yanmin Qian, **“Toward Universal Speech Enhancement For Diverse Input Conditions,”** 2023 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), December 2023.
9. Riku Ogino, Kohei Saijo, Tetsuji Ogawa, **“Design of Discriminators in GAN-Based Unsupervised Learning of Neural Post-Processors for Suppressing Localized Spectral Distortion,”** 2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), November 2024.