



Inter-University Research Institute Corporation /
Research Organization of Information and Systems

National Institute of Informatics

ICASSP 2024
SLP.L20.2

Can Large-scale Vocoded Spoofed Data Improve Speech Spoofing Countermeasure with a Self-supervised Front End?

/wʌn/ /ʃɪn/

Wang Xin, Junichi Yamagishi

National Institute of Informatics

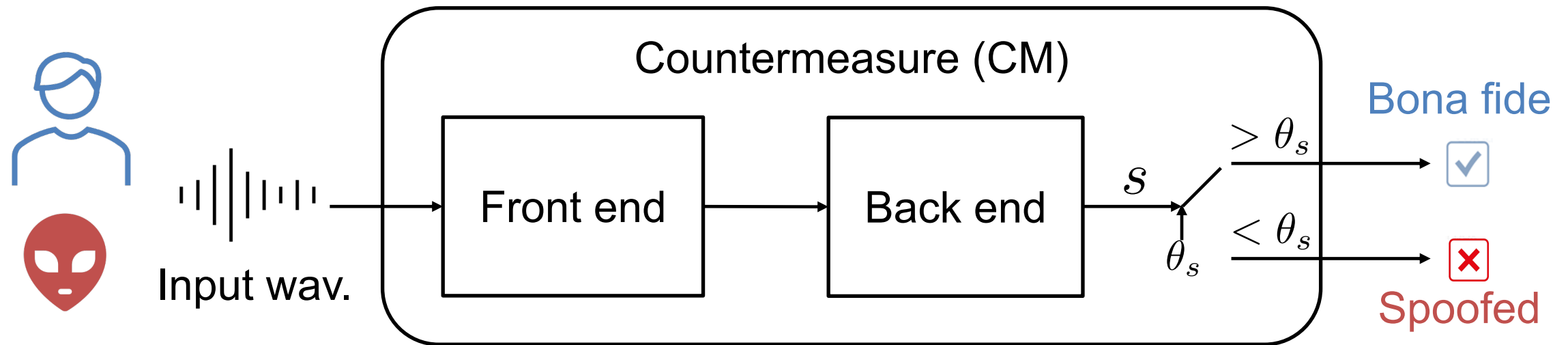
Summary in one slide

- ❑ Extension of previous work (Wang 2023)
- ❑ Motivation: not use any spoofed training data from text-to-speech or voice conversion
- ❑ Method
 - Upstream SSL training, using **vocoded VoxCeleb2**
 - Downstream SSL fine-tuning, using vocoded ASVspoof19
- ❑ Our best overall results

Introduction

□ A binary classification task

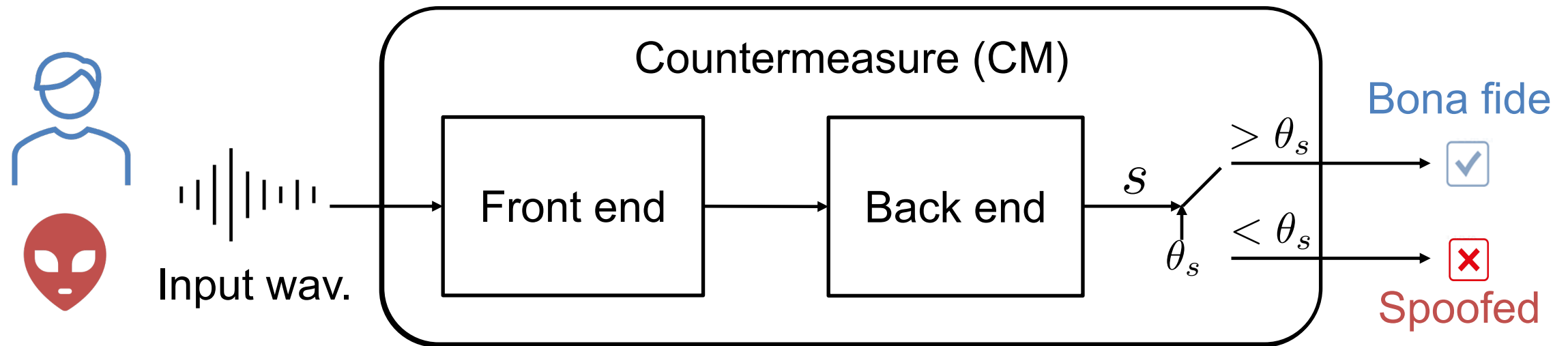
- Bona fide: human voice
- Spoofed: text-to-speech (TTS) or voice conversion (VC) voice
- Metric: equal error rate (EER)



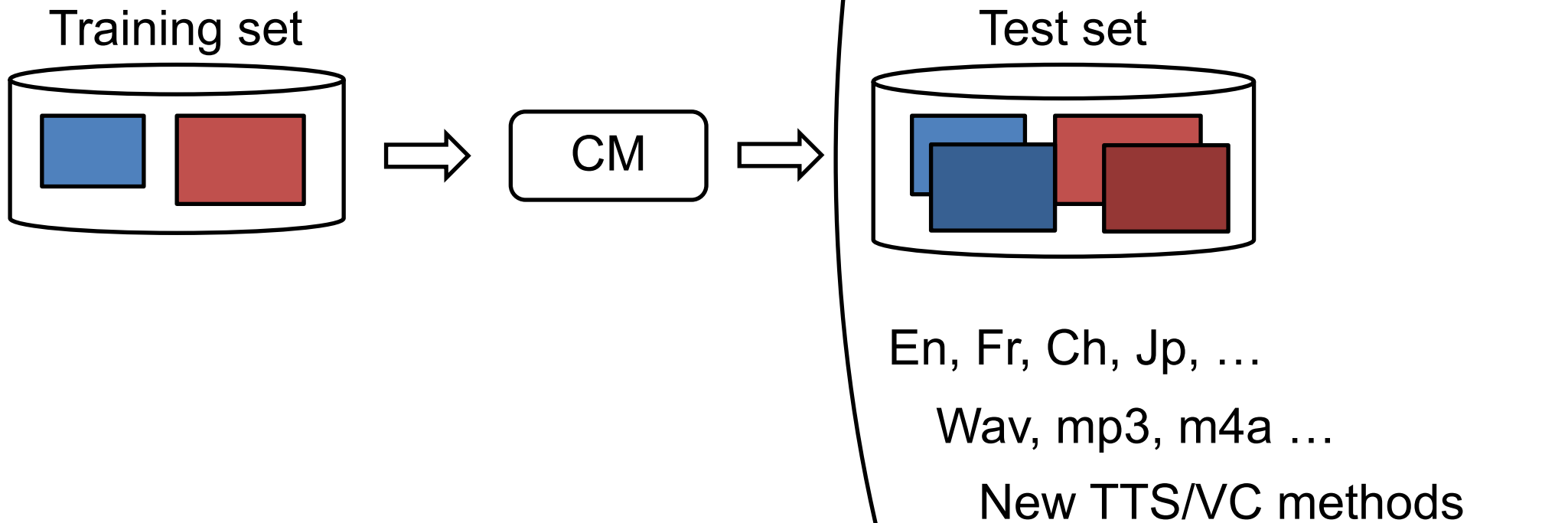
Introduction

□ CM architecture in our studies (Wang 2022, Tak 2022)

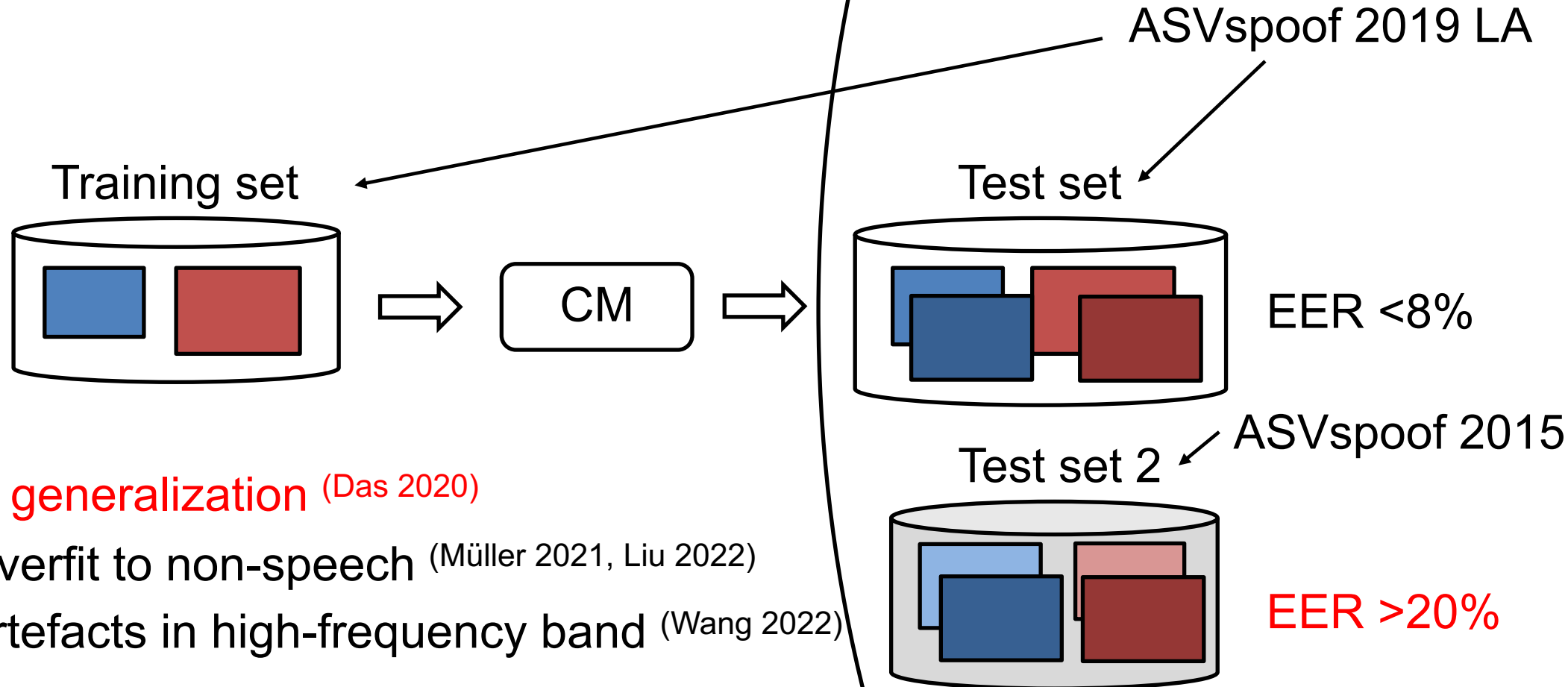
- **Front end:** self-supervised learning (SSL) model
 - wav2vec 2.0 **XLSR-53** (Conneau 2021)
- **Back end:** global average pooling + 4-layer neural network



Generalization is desired

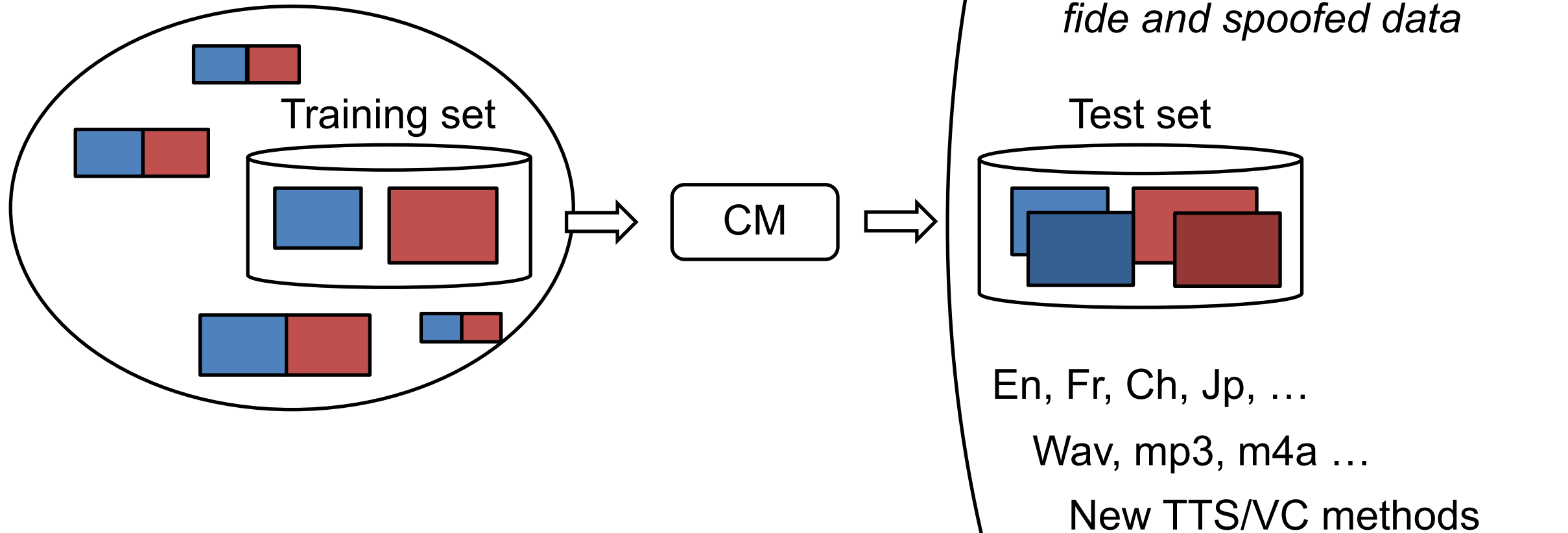


Generalization is challenging

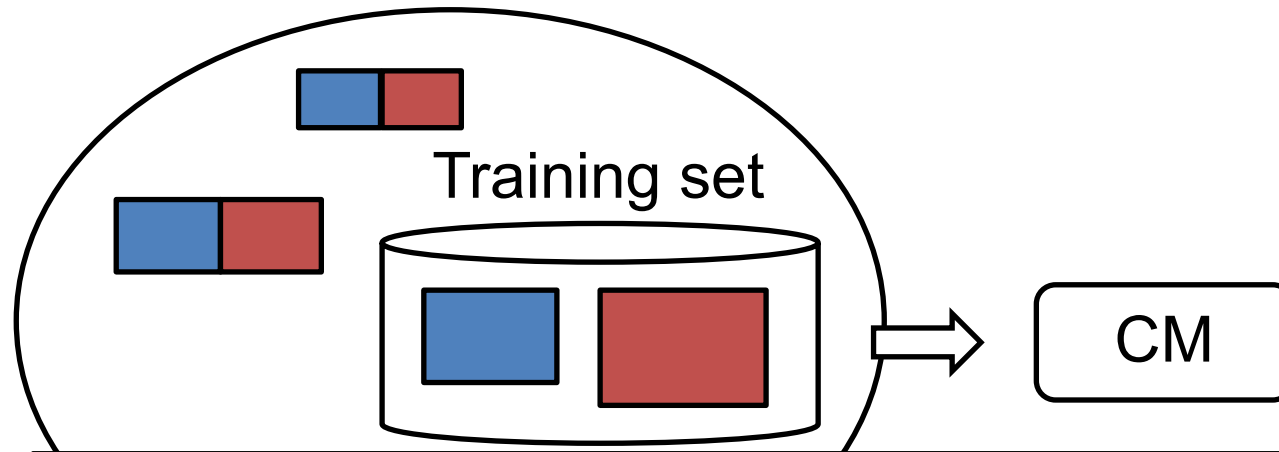


- **Poor generalization** (Das 2020)
 - Overfit to non-speech (Müller 2021, Liu 2022)
 - Artefacts in high-frequency band (Wang 2022)
 - ...

Generalization may need more data



Generalization may need more data



ASVspoof 2019: A large-scale public database of synthesized, converted and replayed speech

Xin Wang^{a,*}, Junichi Yamagishi^{a,b,**}, Massimiliano Todisco^{c,**}, Héctor Delgado^{c,**}, Andreas Nautsch^{c,**}, Nicholas Evans^{c,**}, Md Sahidullah^{d,**}, Ville Vestman^{c,**}, Tomi Kinnunen^{c,**}, Kong Aik Lee^{e,**}, Lauri Juvela^g, Paavo Alku^g, Yu-Huai Peng^h, Hsin-Te Hwang^h, Yu Tsao^h, Hsin-Min Wang^h, Sébastien Le Maguerⁱ, Markus Becker^j, Fergus Henderson^j, Rob Clark^j, Yu Zhang^j, Quan Wang^j, Ye Jia^j, Kai Onuma^k, Koji Mushika^k, Takashi Kaneda^k, Yuan Jiang^l, Li-Juan Liu^l, Yi-Chiao Wu^m, Wen-Chin Huang^m, Tomoki Toda^m, Kou Tanakaⁿ, Hirokazu Kameokaⁿ, Ingmar Steiner^o, Driss Matrouf^p, Jean-François Bonastre^p, Avashna Govender^b, Srikanth Ronanki^q, Jing-Xuan Zhang^r, Zhen-Hua Ling^r

^aNational Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan

^bCentre for Speech Technology Research, University of Edinburgh, UK

^cEURECOM, Campus SophiaTech, 450 Route des Chappes, 06410 Biot, France

^dUniversité de Lorraine, CNRS, Inria, LORIA, F-54000, Nancy, France

^eUniversity of Eastern Finland, Joensuu campus, Länsskatu 15, FI-80110 Joensuu, Finland

^fNEC Corp., 7-1, Shiba 5-chome Minato-ku, Tokyo 108-8001, Japan

^gAalto University, Rakentajanaukio 2 C, 00076 Aalto, Finland

^hAcademia Sinica, 128, Sec. 2, Academia Road, Nankang, Taipei, Taiwan

ⁱGoogle Inc., 1600 Amphitheatre Parkway, Mountain View, CA 94043, USA

^jHQV, Shinjuku Park Tower 25F, 4-7-1 Nishi-Shinjuku, Shinjuku-ku, Tokyo 163-1055 Japan

^kPhysical Research, High-tech Development Zone, No. 960 Wuyangang West Road, Hefei, 230086, China

^lNagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan

^mNTT Communication Science Laboratories, 2-1, Morinomiya Wakusaba, Atsugi-shi, Kanagawa, 243-0292 Japan

ⁿand FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^oand FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^pand FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^qand FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^rand FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^sand FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^tand FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^uand FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^vand FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^wand FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^xand FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^yand FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^zand FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{aa}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ab}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ac}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ad}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ae}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{af}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ag}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ah}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ai}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{aj}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ak}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{al}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{am}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{an}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ao}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ap}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{aq}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ar}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{as}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{at}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{au}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{av}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{aw}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ax}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ay}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{az}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ba}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bb}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bc}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bd}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{be}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bf}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bg}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bh}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bi}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bj}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bk}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bl}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bm}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bn}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bo}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bp}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bq}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{br}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bs}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bt}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bu}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bv}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bw}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bx}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{by}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{bz}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ca}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cb}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cc}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cd}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ce}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cf}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cg}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ch}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ci}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cj}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ck}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cl}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cm}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cn}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{co}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cp}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cq}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cr}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cs}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ct}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cu}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cv}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cw}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cx}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cy}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{cz}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{da}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{db}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dc}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dd}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{de}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{df}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dg}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dh}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{di}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dj}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dk}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dl}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dm}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dn}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{do}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dp}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dq}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dr}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ds}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dt}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{du}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dv}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dw}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dx}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dy}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{dz}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ea}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{eb}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ec}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ed}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ee}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ef}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{eg}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{eh}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ei}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ej}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ek}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{el}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{em}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{en}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{eo}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ep}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{eq}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{er}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{es}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{et}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{eu}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ev}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ew}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ex}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ey}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ez}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{fa}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{fb}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{fc}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{fd}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{fe}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{ff}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{fg}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

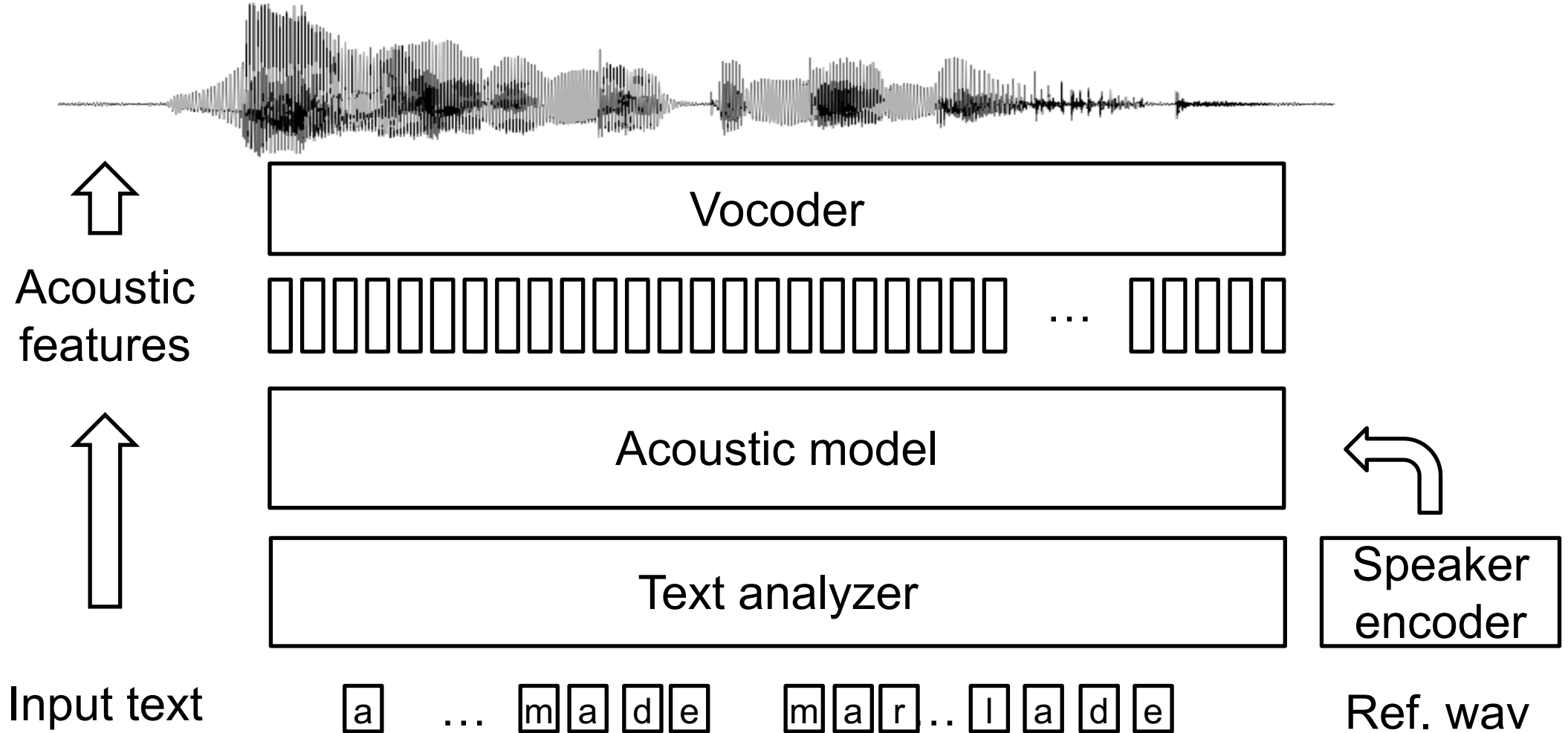
^{fh}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{fi}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

^{fj}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3, 98205 Göttingen, Germany

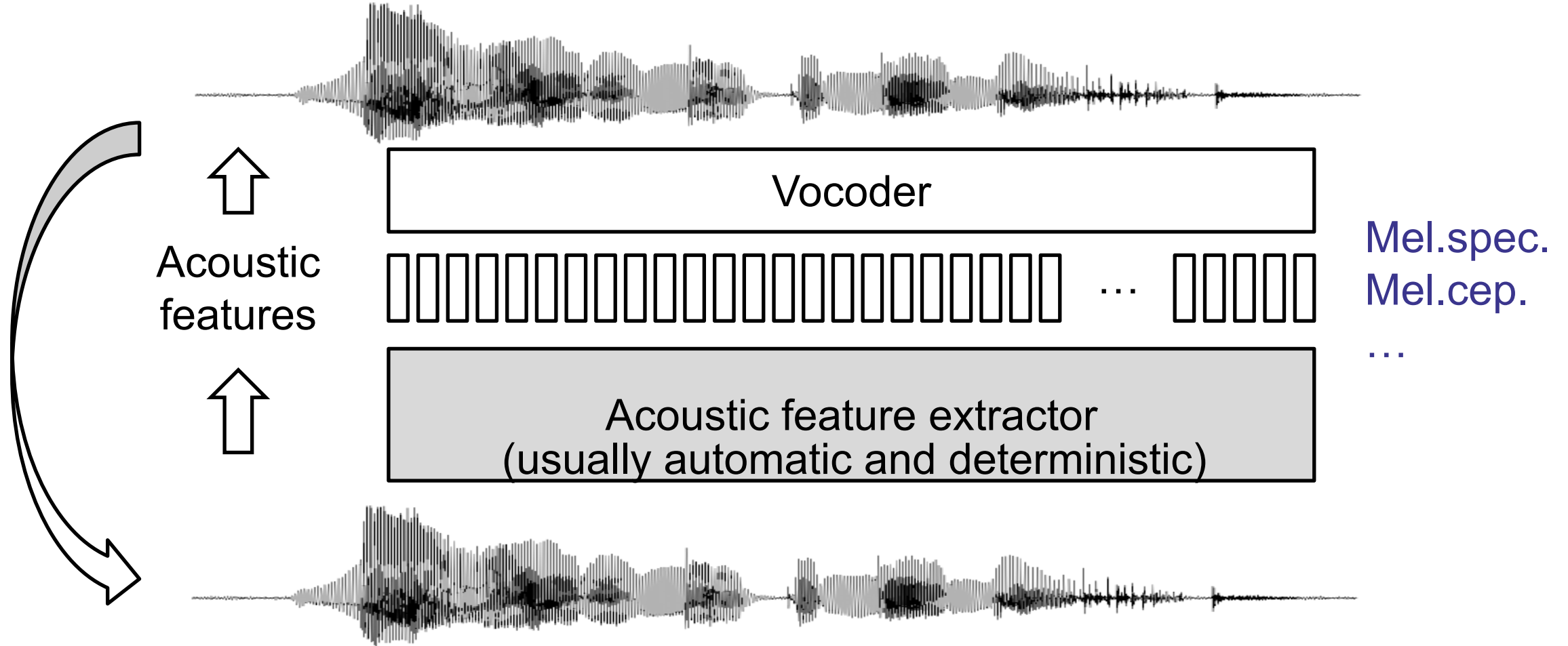
^{fk}and FRIIS, GmbH, Friedrich-Schiller-Str. 1-3,

Idea: creating spoofed training data by vocoding



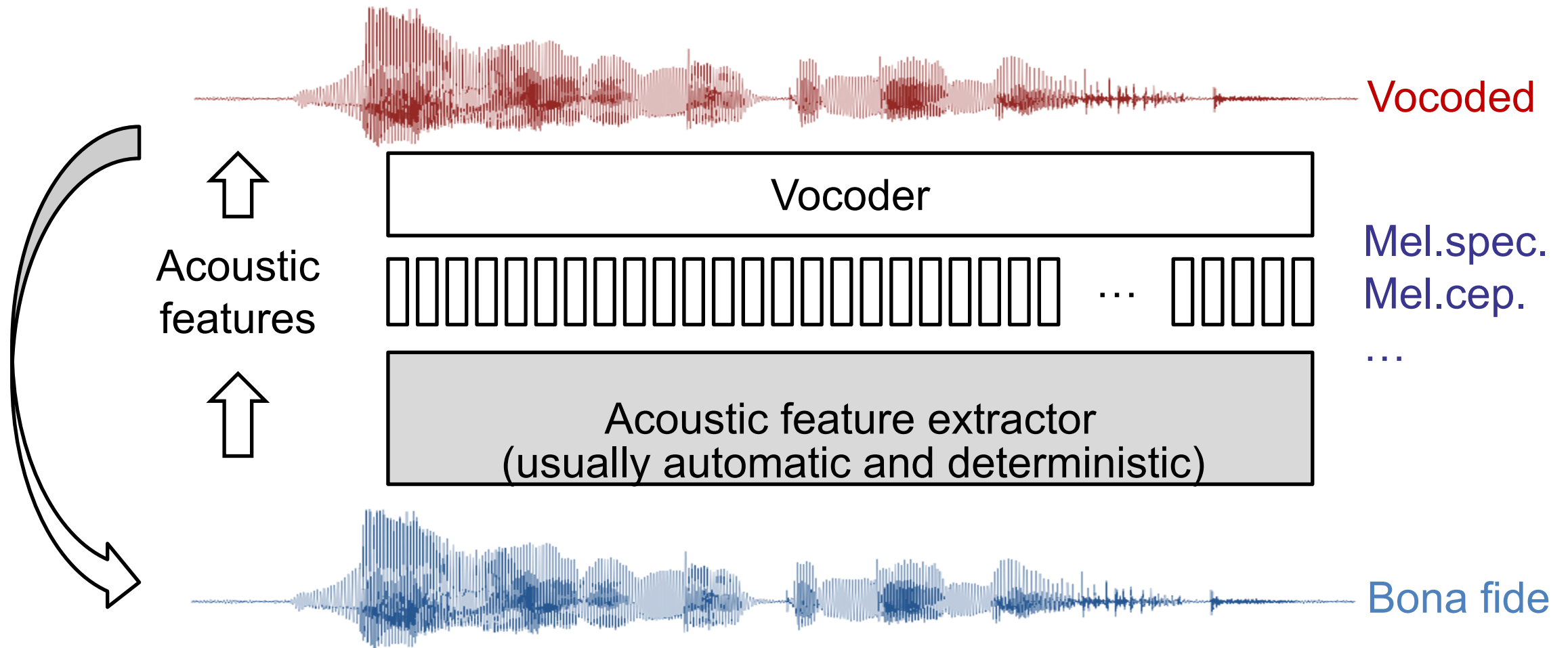
Example of full-fledged TTS

Idea: creating spoofed training data by vocoding



Vocoding is TTS using a perfect acoustic model

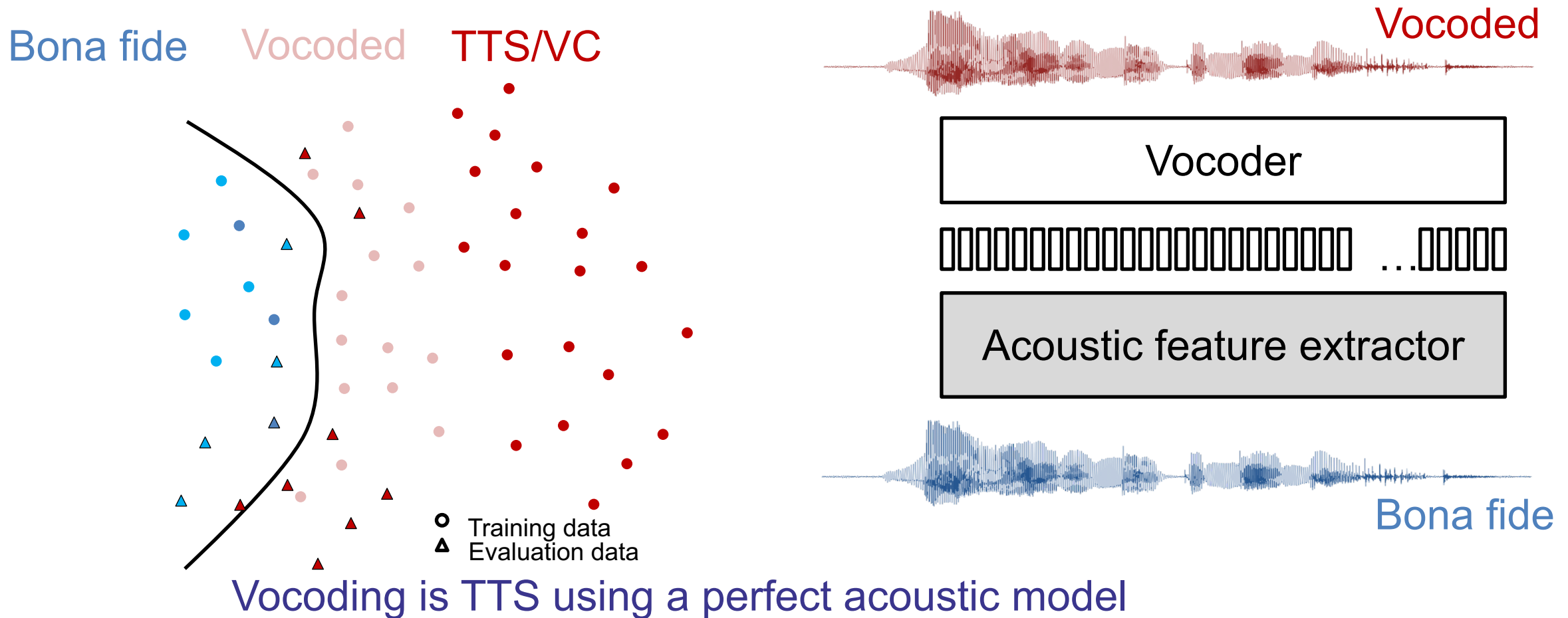
Idea: creating spoofed training data by vocoding



Vocoding is TTS using a perfect acoustic model

Idea: creating spoofed training data by vocoding

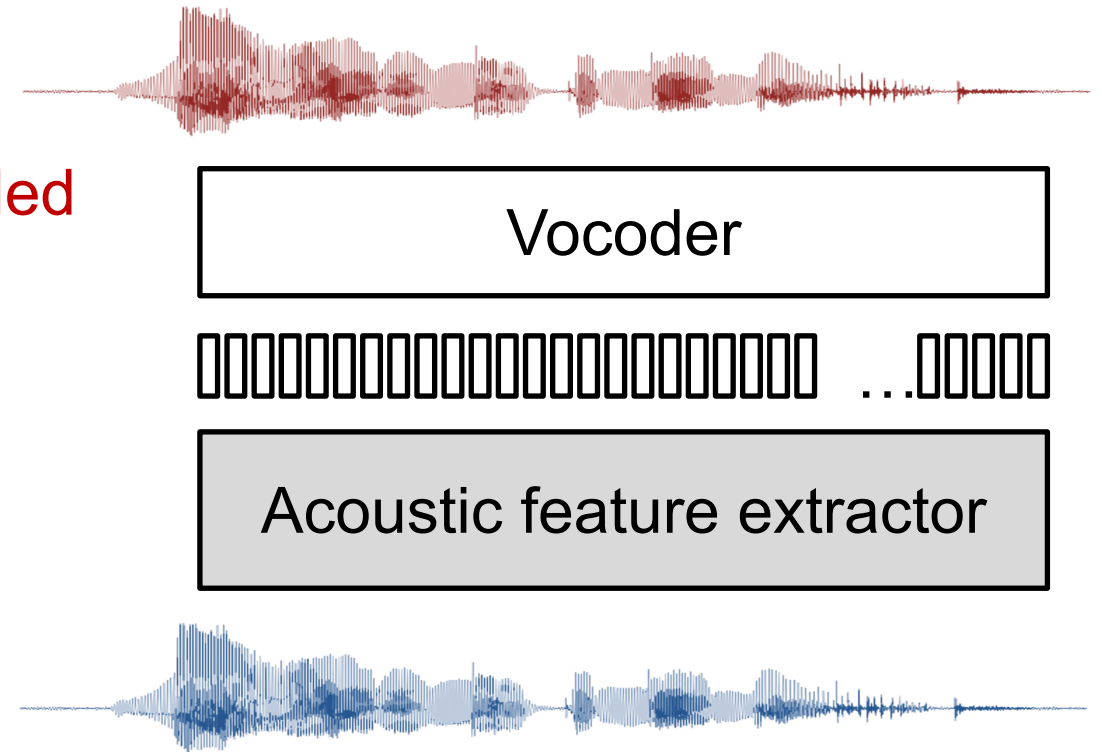
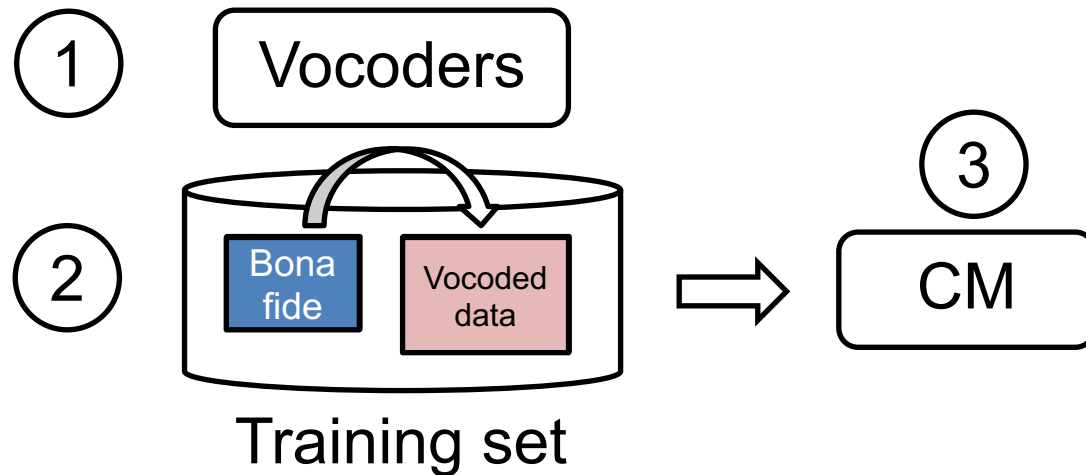
□ Assumption (ideally)



Idea: creating spoofed training data by vocoding

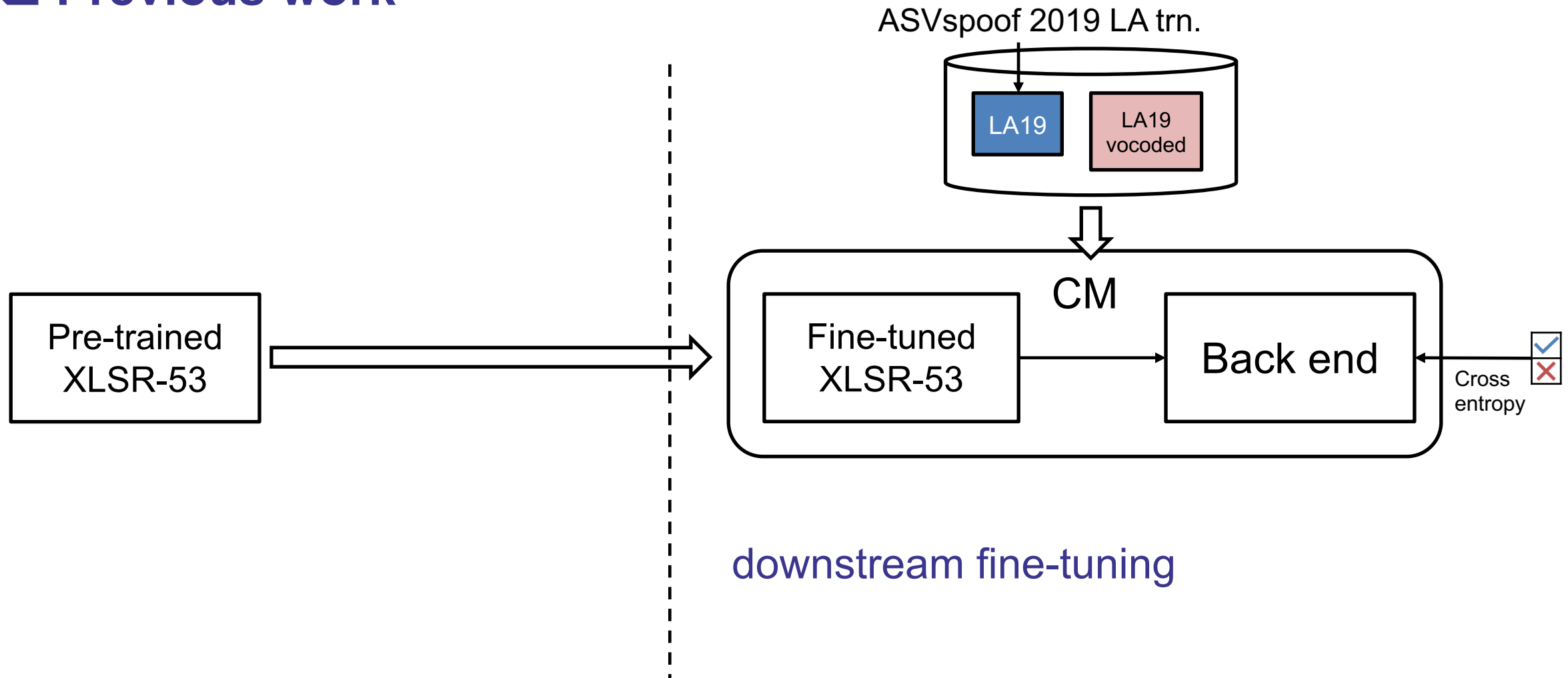
□ Procedure

1. Prepare (or download) vocoders
2. Vocode the bona fide data
3. Train CM using **bona fide** and **vocoded** data



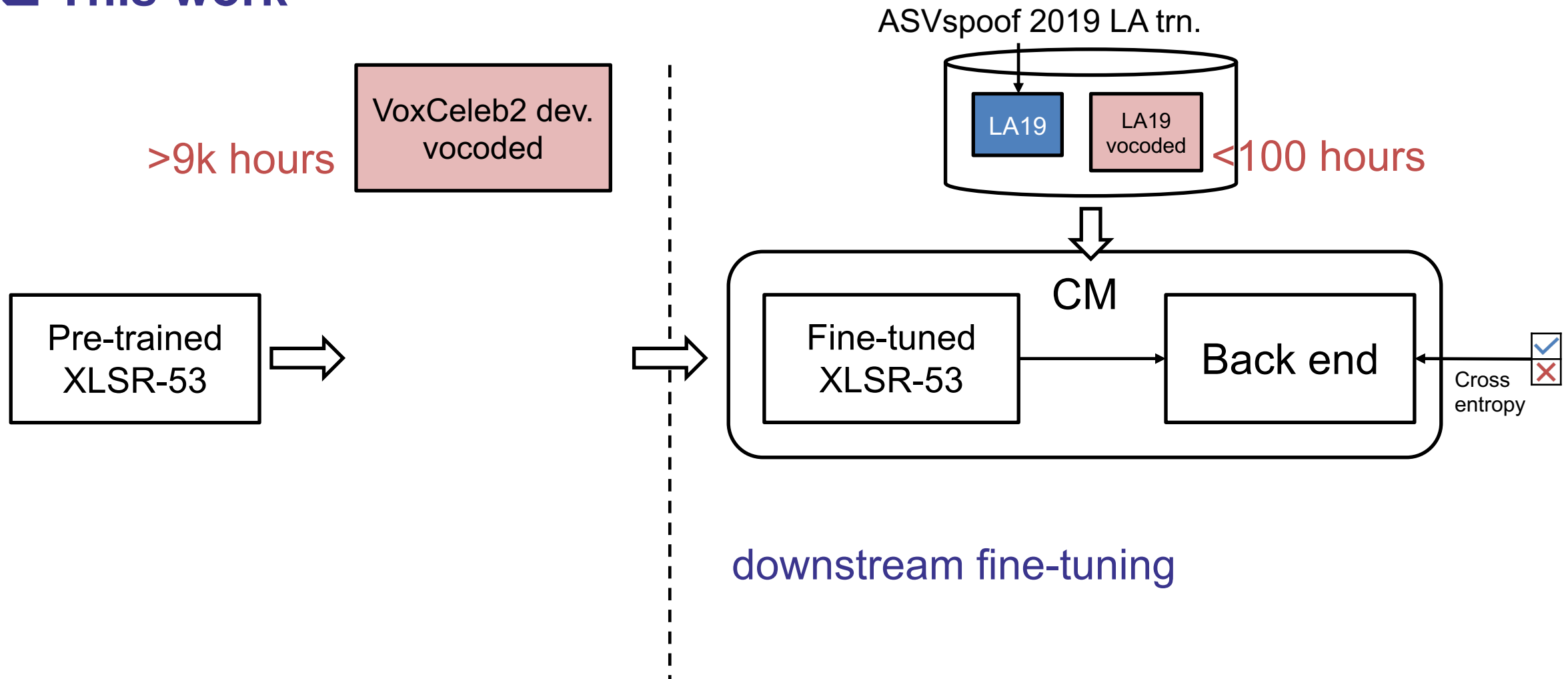
Method: CM training using bona fide & vocoded data

❑ Previous work (Wang 2023)



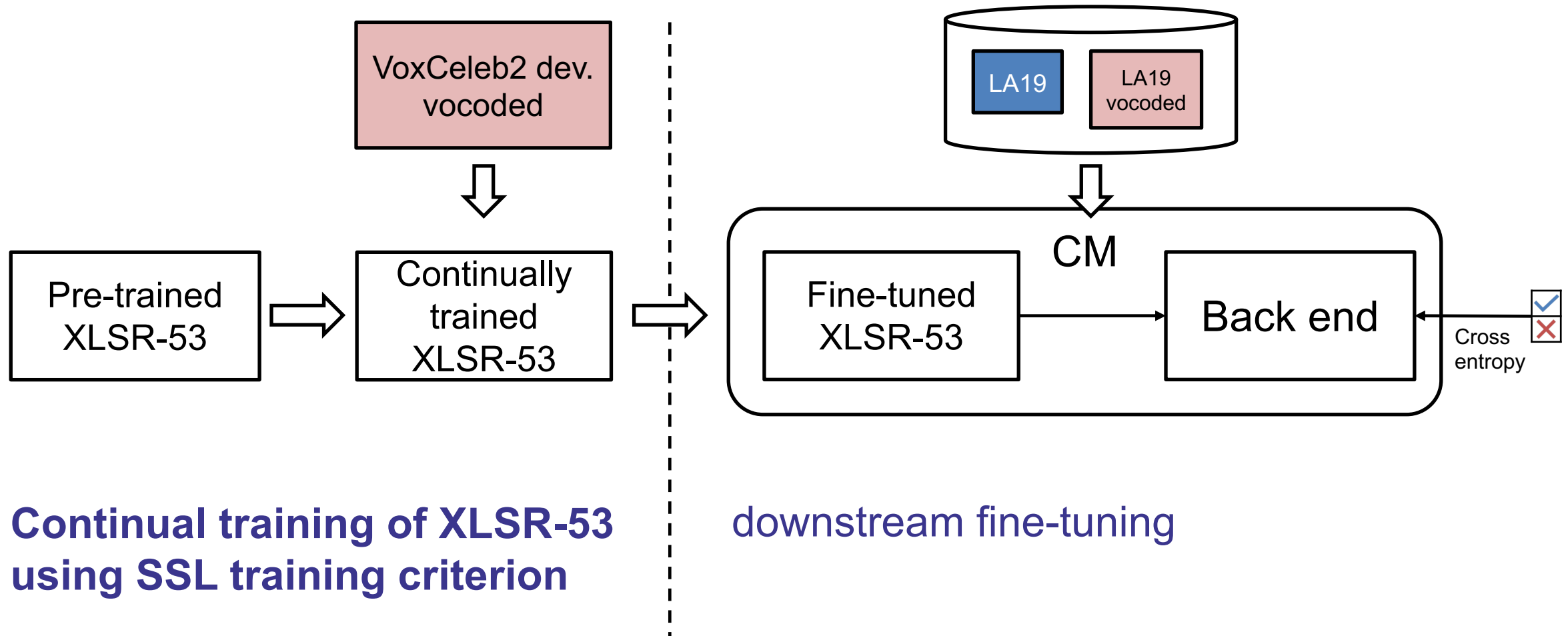
Method: CM training using bona fide & vocoded data

□ This work



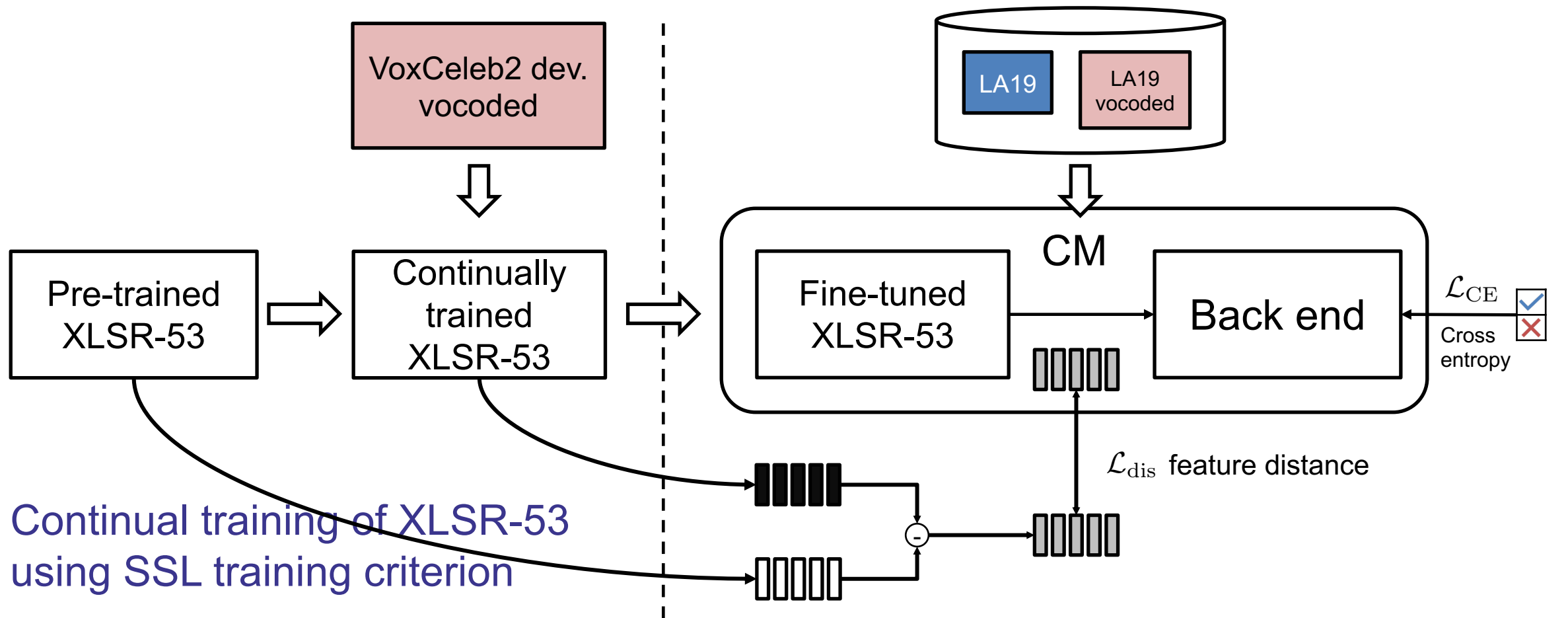
Method: CM training using bona fide & vocoded data

□ This work – method 1



Method: CM training using bona fide & vocoded data

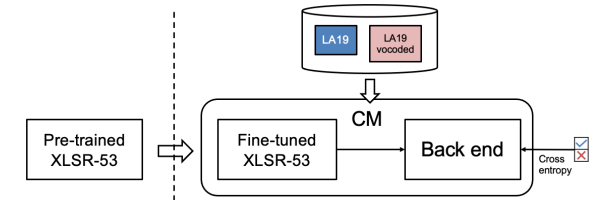
□ This work – method 2



Method: CM training using bona fide & vocoded data

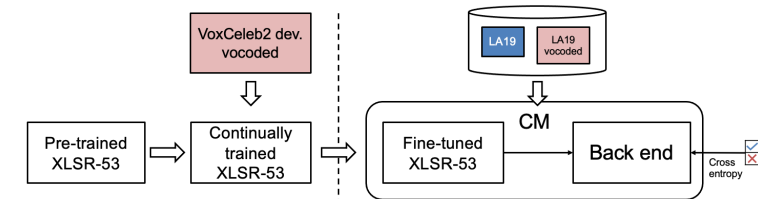
❑ Previous work (Wang 2023)

- Vocoded ASVspoof 2019 LA trn.
- downstream fine-tuning of SSL model

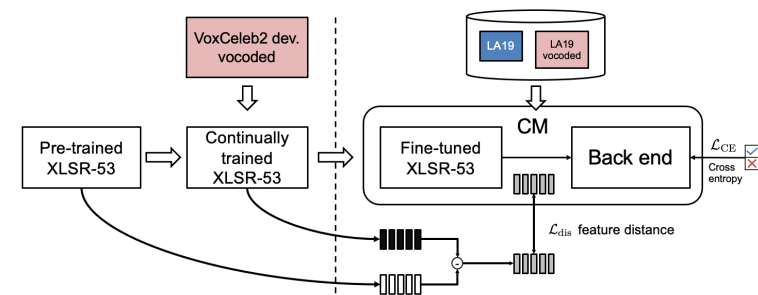


❑ This work

- Vocoded VoxCeleb2 dev.
- Upstream training of SSL
- Downstream training + distillation

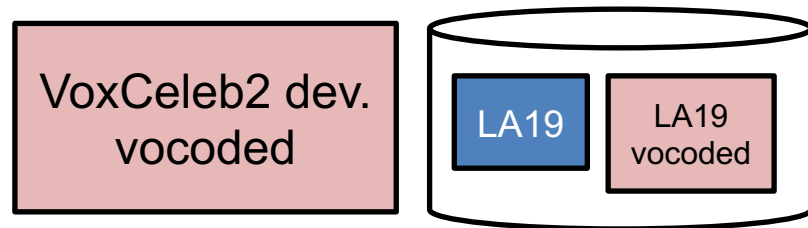


❑ Related studies using DSP-based vocoders (Wu 2013, Khoury 2014, Sizov 2015, Saratxaga 2016, Pal 2018)

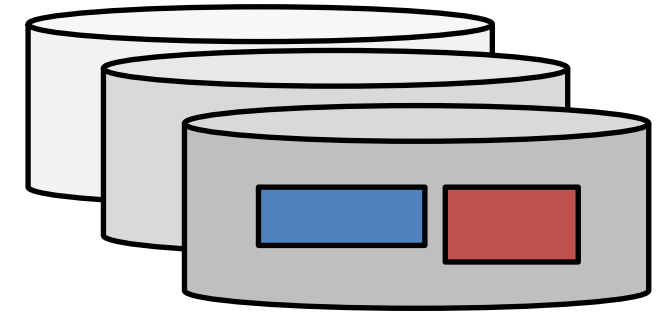


Experiment

Training data



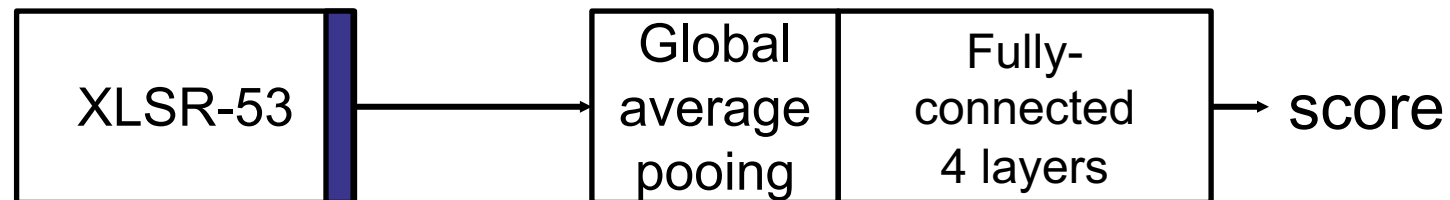
Evaluation data



□ Training data

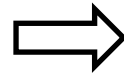
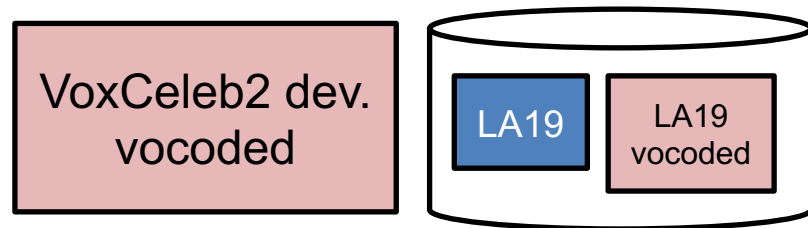
- SSL upstream training: vocoded VoxCeleb2 dev.
- Downstream fine-tuning: bonafide + vocoded ASVspoof 2019 LA trn.
- Vocoders: Hifi-GAN (Kong 2020), NSF (Wang 2019), NSF-GAN, WaveGlow (Prenger 2019)

□ CM

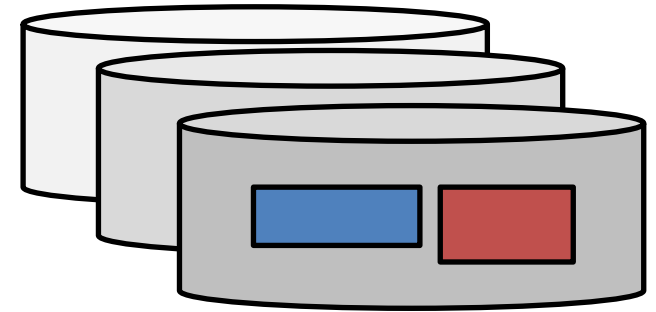


Experiment

Training data



Evaluation data



□ Evaluation data

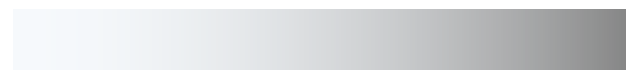
- ASVspoof 2019 LA test set, 2021 LA & DF eval sets
- ASVspoof 2019 LA test set non-speech removed, 2021 LA & DF hidden track
- WaveFake (Frank 2021) , In-the-Wild (Müller 2022)
- three independent training-evaluation rounds
- averaged EERs

More challenging due to domain mismatch

Experiment results



Low EER



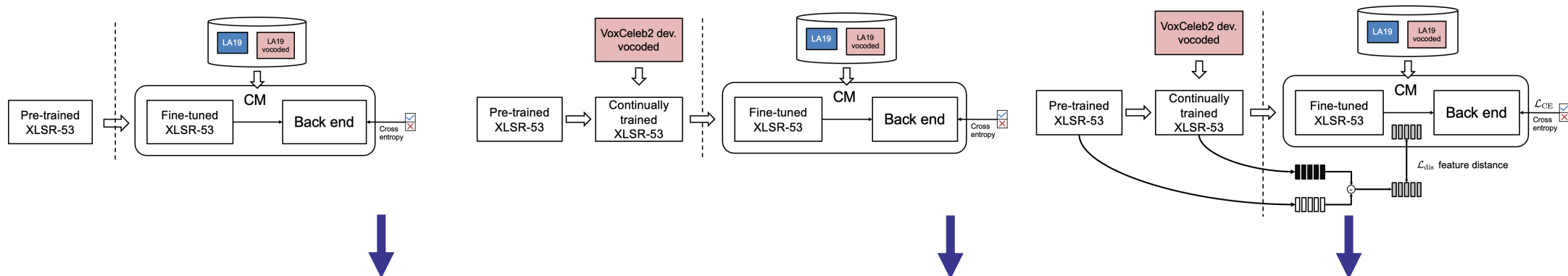
High EER

Systems using different training configurations

CM	ID	B1			B2			B3			P1		P2		P3	
	Front end SSL(s) SSL distilling	xlsr	xlsr,	w2v	xlsr,	w2v		xlsr,	w2v		v.vox	xlsr,	v.vox	xlsr,	v.vox	
		-		×		✓					-		×		✓	
	Data for fine-tune CM	voc.LA									voc.LA					
EER on each test set	LA19eval	3.45		1.97		1.26		2.09		2.01		1.91				
	LA21eval	17.59		13.94		21.09		16.88		14.94		15.92				
	DF21eval	6.53		4.04		14.72		4.34		5.28		5.67				
	LA19etrim	2.69		2.80		3.74		3.33		2.79		3.28				
	LA21hid	13.93		14.05		20.03		16.02		13.95		14.97				
	DF21hid	8.89		9.10		15.27		7.71		8.40		8.84				
	WaveFake	7.33		1.48		5.88		1.94		0.89		1.30				
	InWild	6.78		4.25		13.20		5.84		4.07		6.10				
	Pooled	11.13		12.95		14.06		10.54		9.07		9.98				
	Pooled	11.13		12.95		14.06		10.54		9.07		9.98				

Pooled EER
single threshold

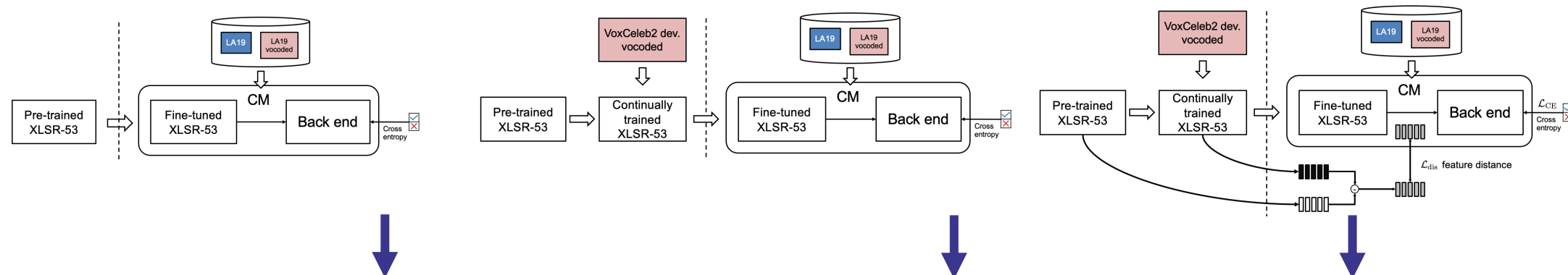
Experiment results



ID		B1	B2	B3	P1	P2	P3
Test sets	LA19eval	3.45	1.97	1.26	2.09	2.01	1.91
	LA21eval	17.59	13.94	21.09	16.88	14.94	15.92
	DF21eval	6.53	4.04	14.72	4.34	5.28	5.67
	LA19etrim	2.69	2.30	3.74	3.33	2.79	3.28
	LA21hid	13.93	9.16	20.03	16.02	13.95	14.97
	DF21hid	8.89	8.16	11.27	7.71	8.40	8.84
	WaveFake	7.33	1.48	5.88	1.94	0.89	1.30
	InWild	6.78	4.25	13.20	5.84	4.07	6.10
	Pooled	11.13	12.33	11.06	10.54	9.07	9.98

B1 vs P1:
continually
trained SSL is
not useless

Experiment results



Test sets	ID	B1	B2	B3	P1	P2	P3
	LA19eval	3.45	1.97	1.26	2.09	2.01	1.91
	LA21eval	17.59	13.94	21.09	16.88	14.94	15.92
	DF21eval	6.53	4.04	14.72	4.34	5.28	5.67
	LA19etrim	2.69	2.30	3.74	3.33	3.74	3.28
	LA21hid	13.93	10.16	20.03	16.02	13.95	14.97
	DF21hid	8.89	8.16	11.27	7.71	8.19	8.84
	WaveFake	7.33	1.48	5.88	1.94	0.89	1.30
	InWild	6.78	4.25	13.20	5.84	4.07	6.10
	Pooled	11.13	12.33	12.06	10.54	8.74	9.98

B1 vs P1:
continually
trained SSL is
not useless

Merging
two
SSLs is
helpful



Experiment results



ID	B1	B2	B3	P1	P2	P3
LA19eval	3.45	1.07	1.26	2.09	2.01	1.91
LA21eval	17.59	13.94	21.09	16.88	14.94	15.92
DF21eval	6.53	1.01	11.72	4.34	5.28	5.67
LA19etrim	2.69	2.80	3.74	3.33	2.79	3.28
LA21hid	13.93	11.39	11.00	16.68	13.95	14.97
DF21hid	8.89	1.00	5.27	7.71	8.40	8.84
WaveFake	7.33	1.48	5.88	1.94	0.89	1.30
InWild	6.78	4.25	13.20	5.84	4.07	6.10
Pooled	11.13	12.83	21.88	18.34	9.07	9.98

B1 is the best model in our previous work (Wang 2023)

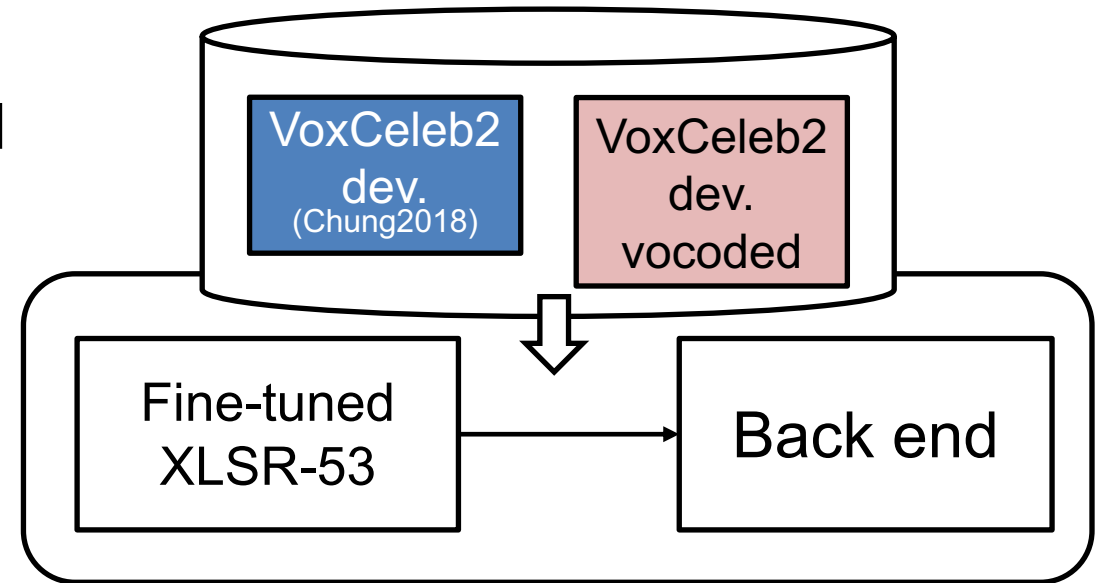
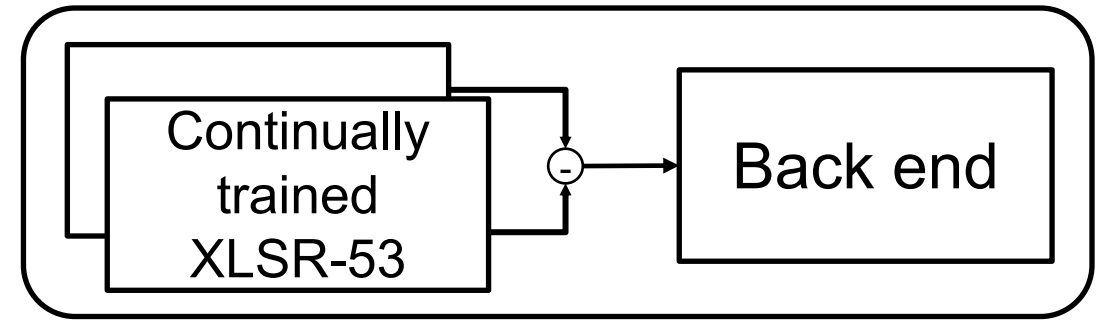
It is better than using TTS/VC spoofed data



Experiment results

❑ Other results in the paper

- Using two SSLs without distillization?
- Downstream fine-tuning using vocoded voxceleb2?
- ...



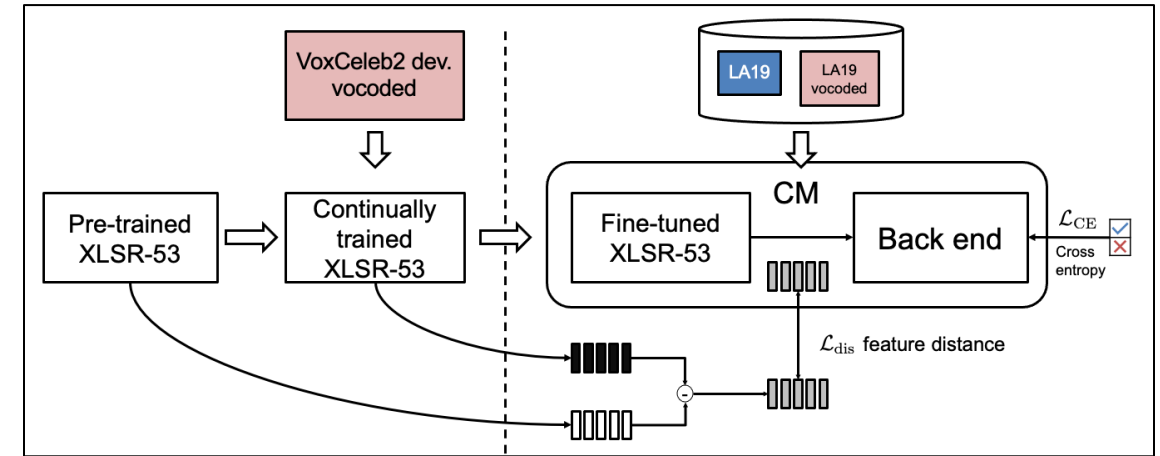
Summary

□ Method

- Large scale vocoded VoxCeleb2
- Upstream SSL training
- Downstream fine-tuning + distilling

□ Results

- Slightly outperformed previous work (pooled EER)



Thank you



project/[10-asvspoof-vocoded-trn-ssl](https://github.com/nii-yamagishilab/project-NN-Pytorch-scripts/tree/master/project/10-asvspoof-vocoded-trn-ssl)

Appendix

CM	ID	B1	B2	B3	P1	P2	P3	B1-b	P3-b	B1-c	P3-c
	Front end SSL(s) SSL distilling	xlsr -	xlsr, w2v ×	xlsr, w2v ✓	v.vox -	xlsr, v.vox ×	xlsr, v.vox ✓	xlsr -	xlsr, v.vox ✓	xlsr -	xlsr, v.vox ✓
Data for fine-tune CM		voc.LA			voc.LA			LA19trn		voc.VoxCel	
Test sets	LA19eval	3.45	1.97	1.26	2.09	2.01	1.91	0.22	0.13	3.59	3.71
	LA21eval	17.59	13.94	21.09	16.88	14.94	15.92	2.69	3.29	15.22	12.37
	DF21eval	6.53	4.04	14.72	4.34	5.28	5.67	4.27	3.45	5.99	3.31
	LA19etrim	2.69	2.80	3.74	3.33	2.79	3.28	7.37	7.37	2.74	3.63
	LA21hid	13.93	14.05	20.03	16.02	13.95	14.97	15.56	24.23	10.14	9.53
	DF21hid	8.89	9.10	15.27	7.71	8.40	8.84	9.16	13.95	9.03	7.77
	WaveFake	7.33	1.48	5.88	1.94	0.89	1.30	23.75	15.44	13.41	24.17
	InWild	6.78	4.25	13.20	5.84	4.07	6.10	13.52	12.32	6.90	7.00
Pooled		11.13	12.95	14.06	10.54	9.07	9.98	12.76	12.50	10.92	12.26

Reference

DSP-based vocoders

- Xingming Wang, Xiaoyi Qin, Tinglong Zhu, Chao Wang, Shilei Zhang, and Ming Li. The DKU-CMRI System for the ASVspoof 2021 Challenge: Vocoder Based Replay Channel Response Estimation. In *Proc. ASVspoof challenge workshop*, 16–21. 2021.
- Monisankha Pal, Dipjyoti Paul, and Goutam Saha. Synthetic Speech Detection Using Fundamental Frequency Variation and Spectral Features. *Computer Speech & Language* 48. Elsevier: 31–50. 2018.
- Ibon Saratxaga, Jon Sanchez, Zhizheng Wu, Inma Hernaez, and Eva Navas. Synthetic Speech Detection Using Phase Information. *Speech Communication* 81 (July): 30–41. doi:10.1016/j.specom.2016.04.001. 2016.
- Aleksandr Sizov, Elie Khoury, Tomi Kinnunen, Zhizheng Wu, and Sébastien Marcel. Joint Speaker Verification and Antispoofing in the I-Vector Space. *IEEE Transactions on Information Forensics and Security* 10 (4). IEEE: 821–832. doi:10.1109/TIFS.2015.2407362. 2015.
- Elie Khoury, Tomi Kinnunen, Aleksandr Sizov, Zhizheng Wu, and Sébastien Marcel. Introducing I-Vectors for Joint Anti-Spoofing and Speaker Verification. In *Proc. Interspeech*, 61–65. 2014.
- Jon Sanchez, Ibon Saratxaga, Inma Hernaez, Eva Navas, and Daniel Erro. A Cross-Vocoder Study of Speaker Independent Synthetic Speech Detection Using Phase Information. In *Proc. Interspeech*. 2014.
- Zhizheng Wu, Xiong Xiao, Eng Siong Chng, and Haizhou Li. Synthetic Speech Detection Using Temporal Modulation Feature. In *Proc. ICASSP*, 7234–7238. 2013.

neural vocoders

- Joel Frank, and Lea Schönherr. WaveFake: A Data Set to Facilitate Audio DeepFake Detection. In *Proc. NeurIPS Datasets and Benchmarks 2021*. 2021.
- Chengzhe Sun, Shan Jia, Shuwei Hou, Ehab AlBadawy, and Siwei Lyu. Exposing AI-Synthesized Human Voices Using Neural Vocoder Artifacts. ArXiv Preprint ArXiv:2302.09198. 2023.