The VoicePrivacy 2020 Challenge

Objective evaluation-ZEBRA

Odyssey 2020

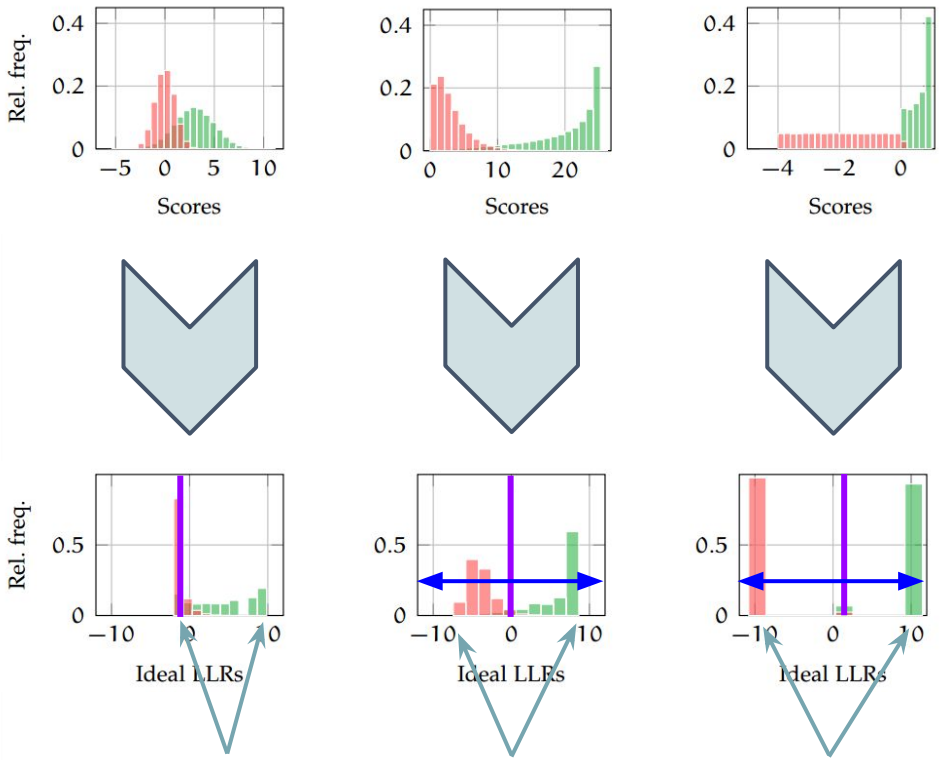
Presenter: Andreas Nautsch

Natalia Tomashenko [1]
Brij M.L. Srivastava [2]
Xin Wang [3]
Emmanuel Vincent [4]
Andreas Nautsch [5]
Junichi Yamagishi [3,6]
Nicholas Evans [5]
Jose Patino [5]
Jean-François Bonastre [1]
Paul-Gauthier Noé [1]
Massimiliano Todisco [5]
Mohamed Maouche [2]
Benjamin O'Brien [7]
Anais Chanclu [1]

[1] LIA – University of Avignon – France
[2] Inria – France
[3] NII – Tokyo – Japan
[4] Inria – France
[5] Audio Security and Privacy Group, EURECOM – France
[6] University of Edinburgh – UK
[7] Aix-Marseille University – France

4th November 2020

Recap: EER, min Cllr & ZEBRA's "worst case"



ZEBRA worst-case:
max(abs(LLR))

Ideal score calibration:

1. identify bins with same ratios of errors
2. map scores to unified scale

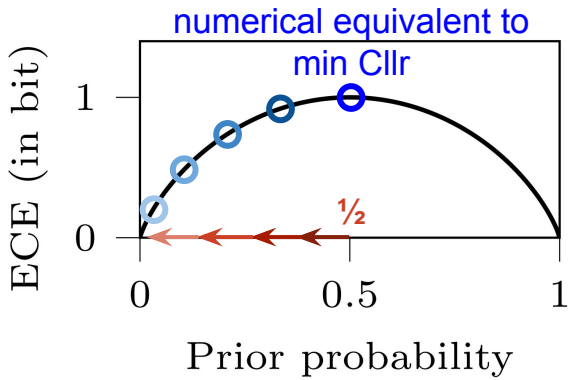
EER: one threshold, varies on "unified scale"

min Cllr: generalised class discrimination

Figures from Nautsch: "Speaker Recognition in Unconstrained Environments," PhD thesis, TU Darmstadt, 2019.

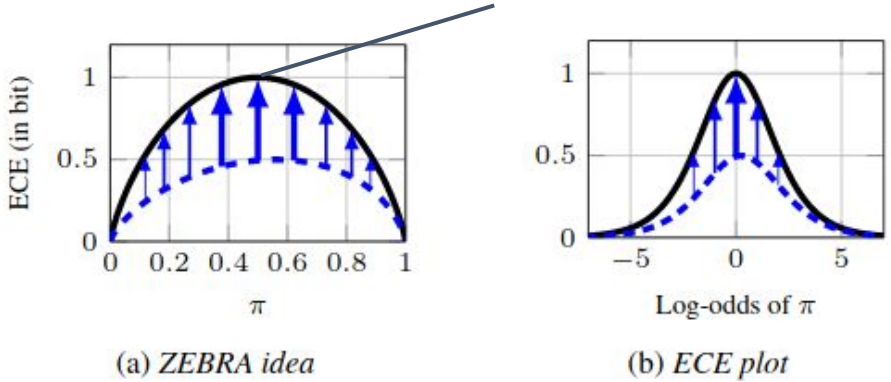
Recap: min Cllr & ZEBRA's "expectation"

$$\frac{\pi}{|\mathcal{S}_{\mathcal{A}}|}\sum_{a\in\mathcal{S}_{\mathcal{A}}}\log_2\left(1+\frac{1-\pi}{a\,\pi}\right)+\frac{1-\pi}{|\mathcal{S}_{\mathcal{B}}|}\sum_{b\in\mathcal{S}_{\mathcal{B}}}\log_2\left(1+\frac{b\,\pi}{1-\pi}\right)$$



If perfect privacy, then must:

a) min Cllr is upper bound
b) profile is symmetric



(a) *ZEBRA idea* (b) *ECE plot*

Prior: chosen by adversary

⇒ *inaccessible to us,*
but we can simulate :)

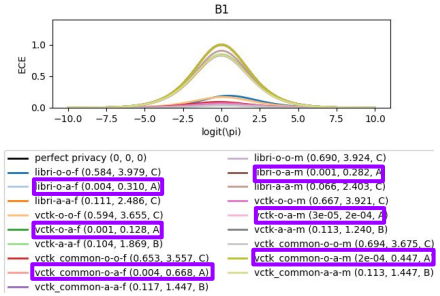
Empirical observations:

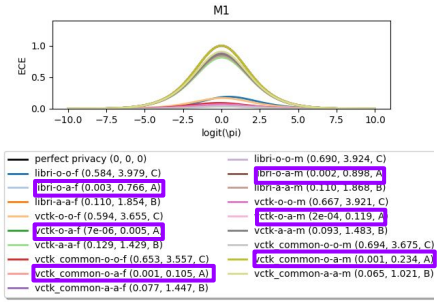
i) min Cllr often close to upper bound
ii) lots of symmetries

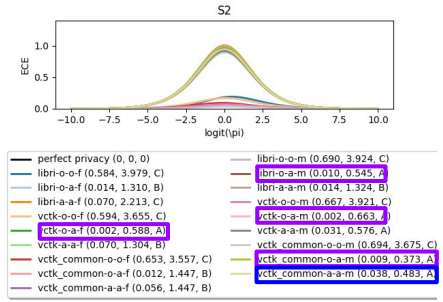
ZEBRA: by system for all VoicePrivacy test data sets (primary)

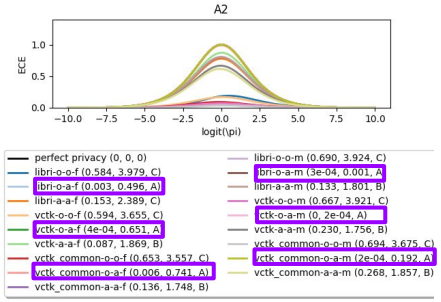
Categorical tags of worst-case privacy disclosure

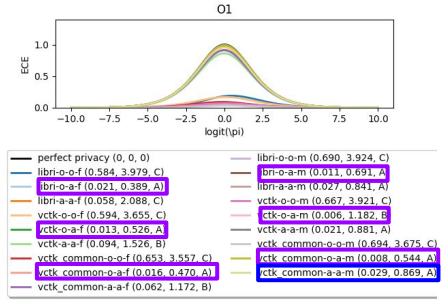
Tag	Category	Posterior odds ratio (flat prior)
0	$l = 1 = 10^0$	50 : 50 (flat posterior)
A	$10^0 < l < 10^1$	more disclosure than 50 : 50
B	$10^1 \leq l < 10^2$	one wrong in 10 to 100
C	$10^2 \leq l < 10^4$	one wrong in 100 to 10 000
D	$10^4 \leq l < 10^5$	one wrong in 10 000 to 100 000
E	$10^5 \leq l < 10^6$	one wrong in 100 000 to 1 000 000
F	$10^6 \leq l$	one wrong in at least 1 000 000

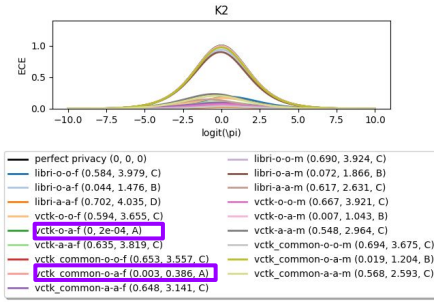


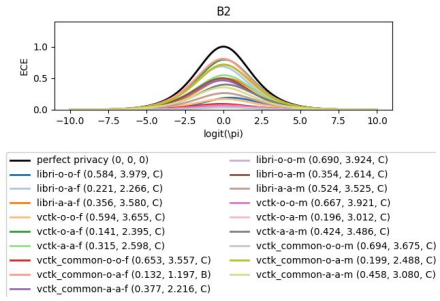


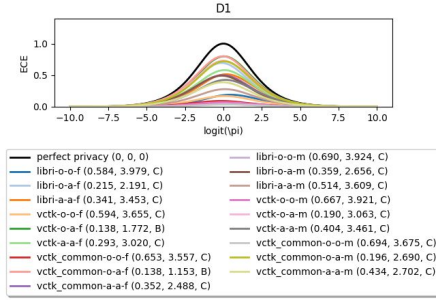


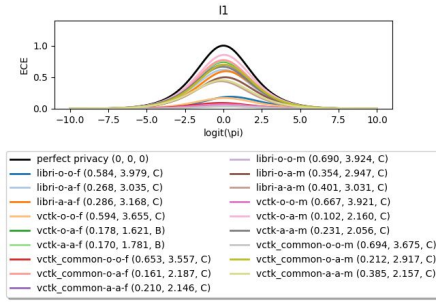




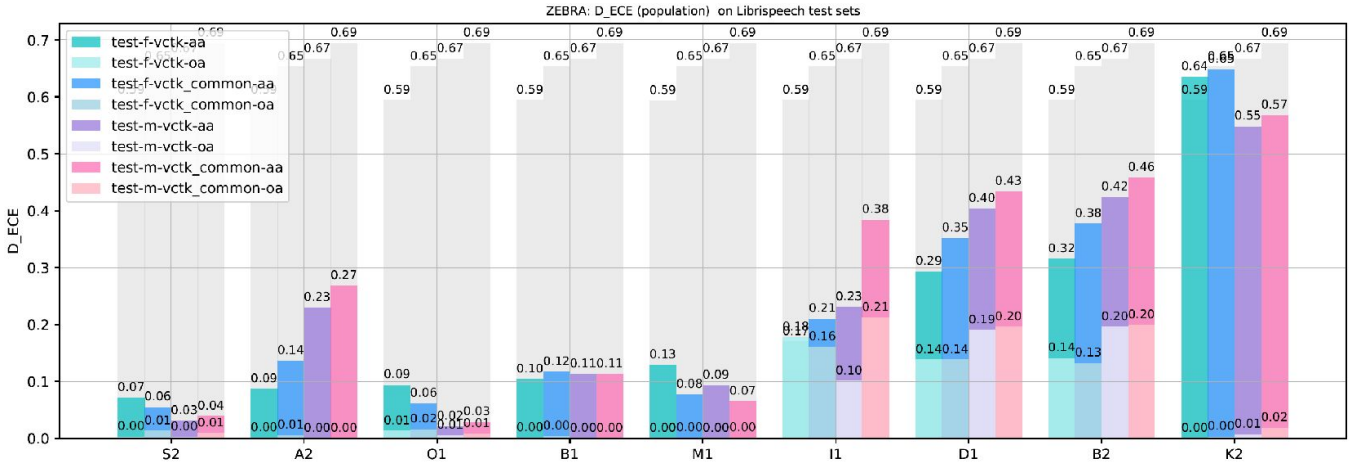




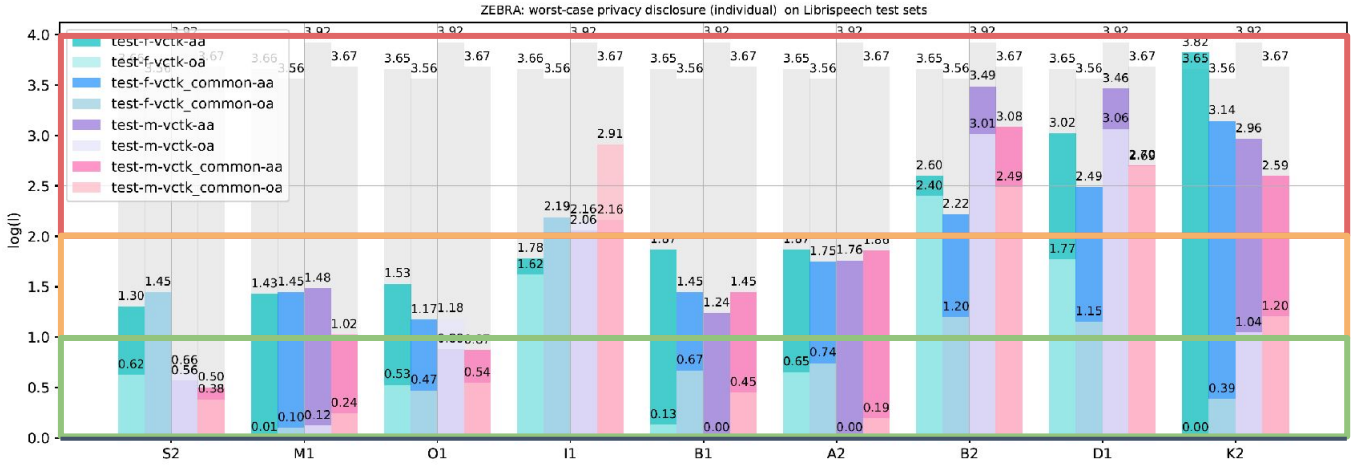




ZEBRA: D_ECE (population) on VCTK-test

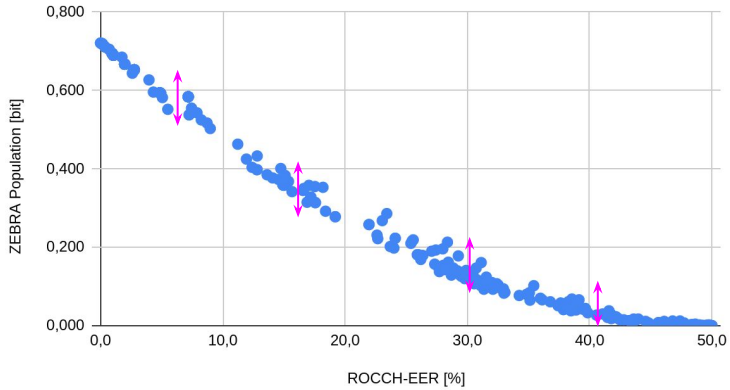


ZEBRA: worst-case privacy disclosure (individual) on VCTK-test

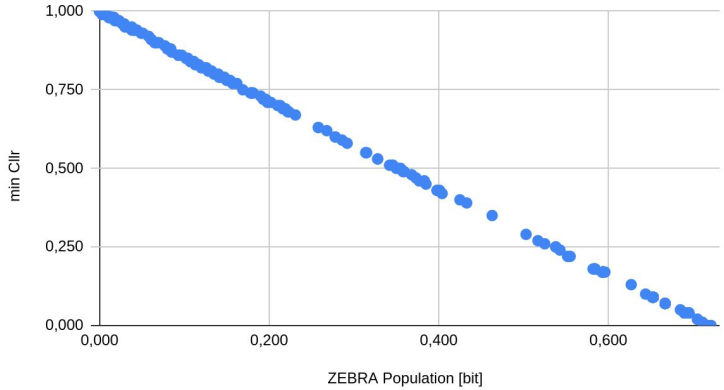


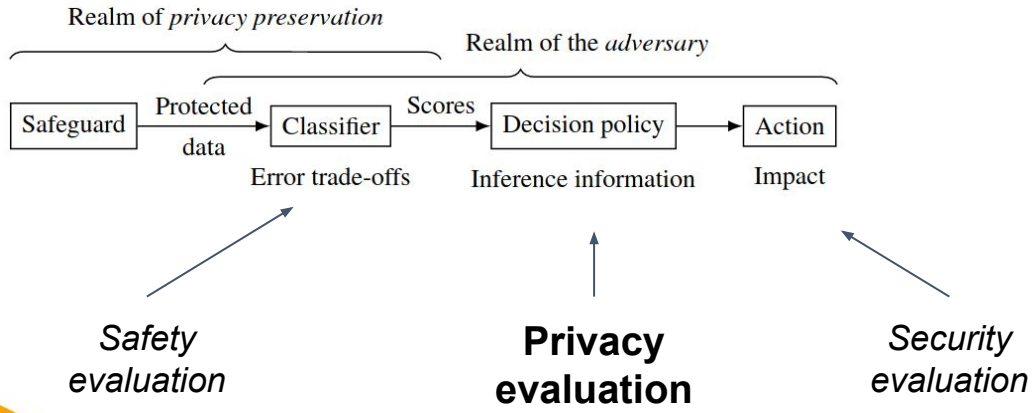
Metric correlation — NOT CAUSALITY

ZEBRA Population [bit] vs. ROCCH-EER [%]

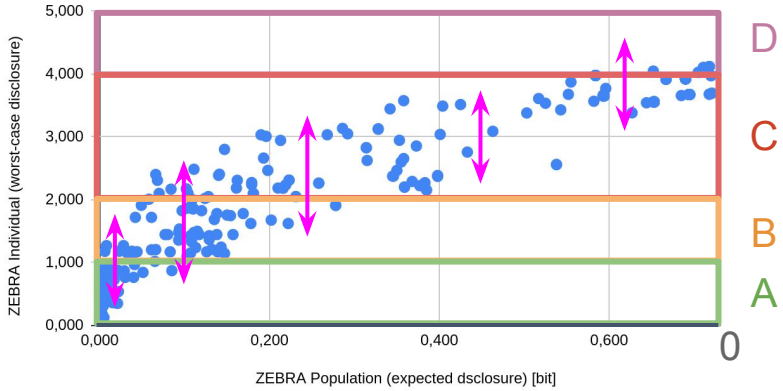


ZEBRA Population [bit] vs. min Cllr

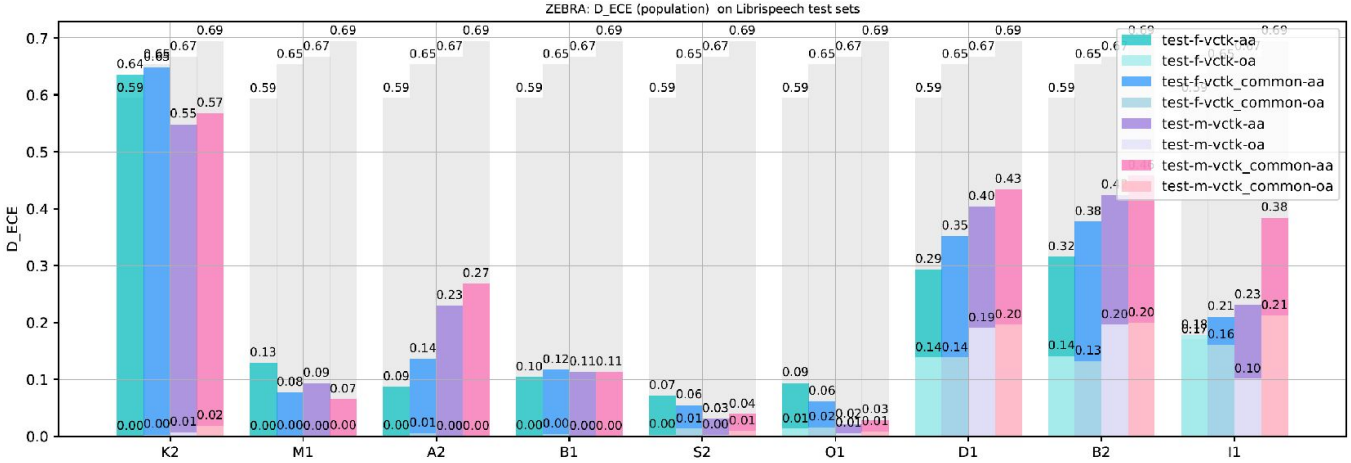




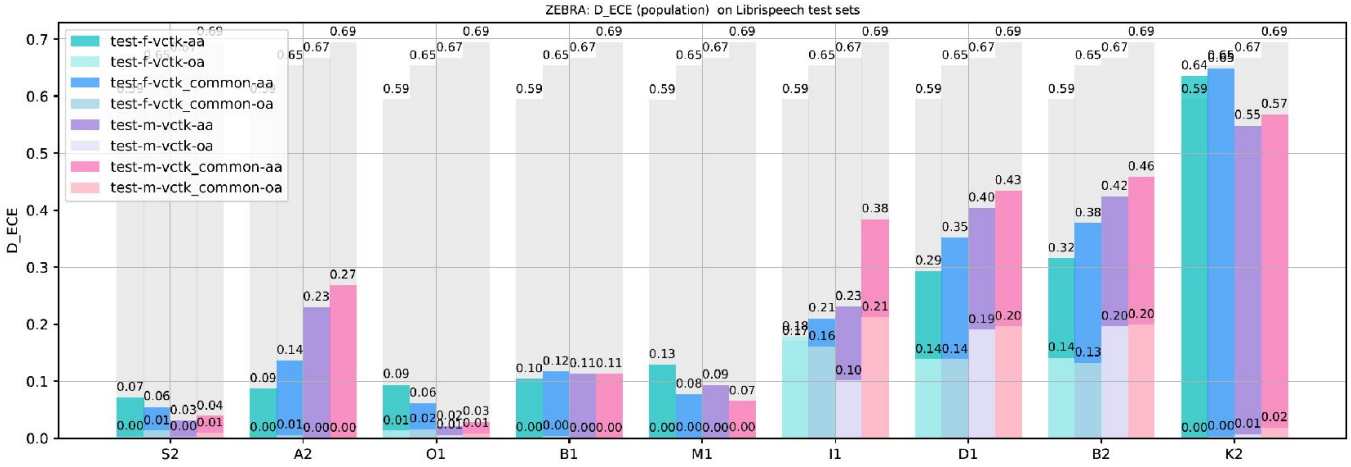
ZEBRA expected disclosue [bit] vs. worst-case disclosure



ZEBRA: D_ECE (population) on VCTK-test

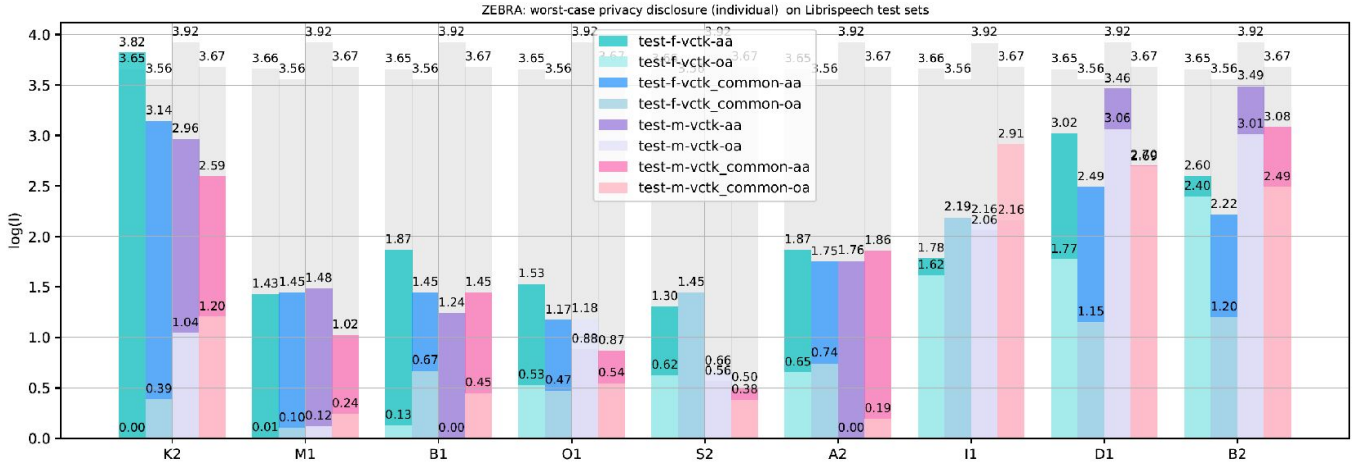


Sorted by oa f-test-vctk

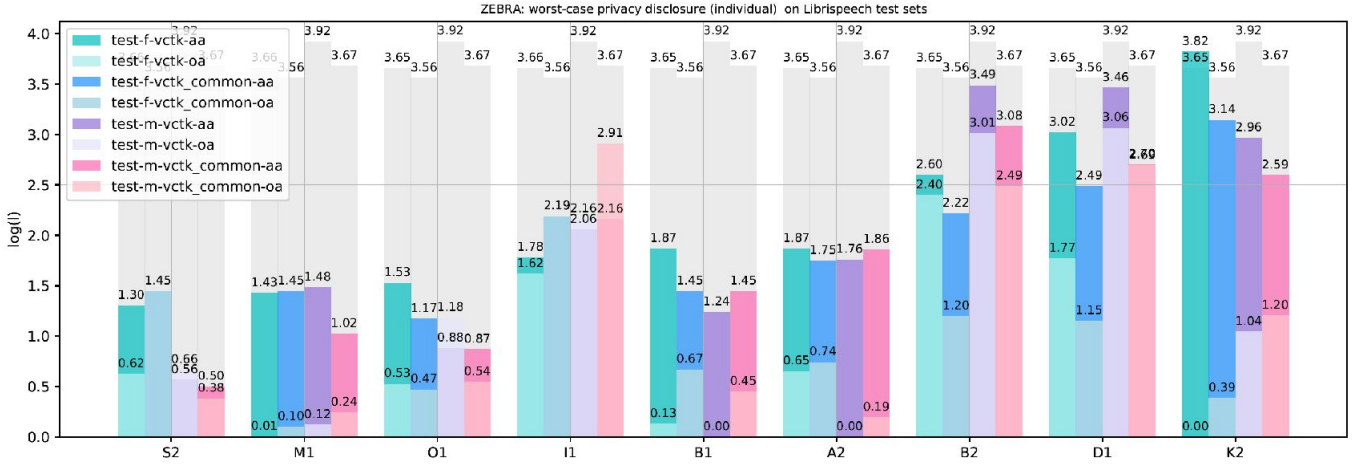


Sorted by aa f-test-vctk

ZEBRA: worst-case privacy disclosure (individual) on VCTK-test

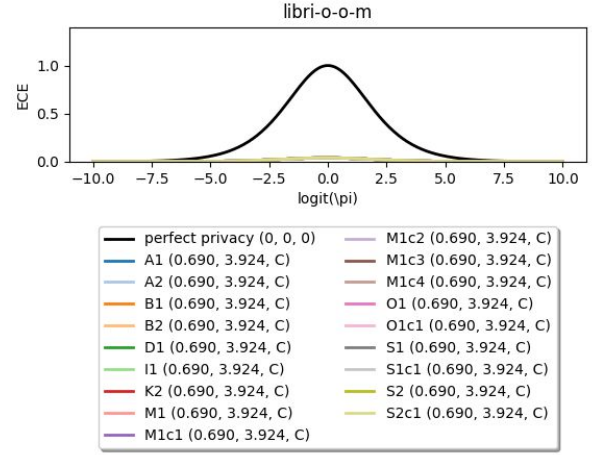


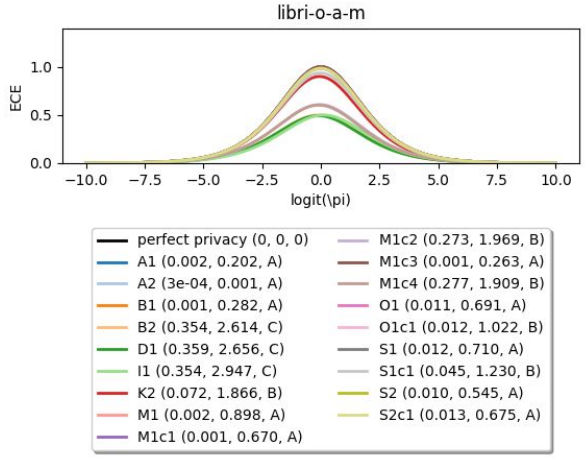
Sorted by oa f-test-vctk

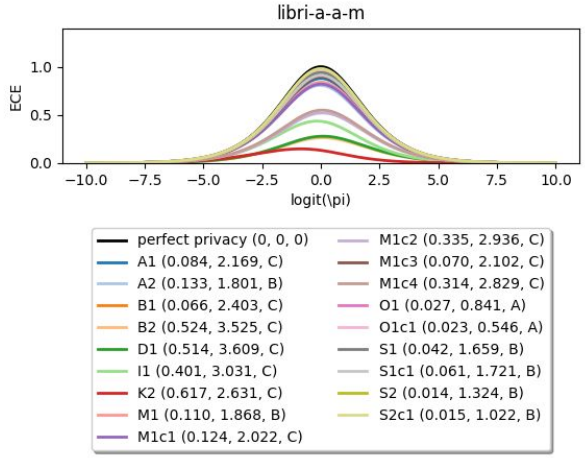


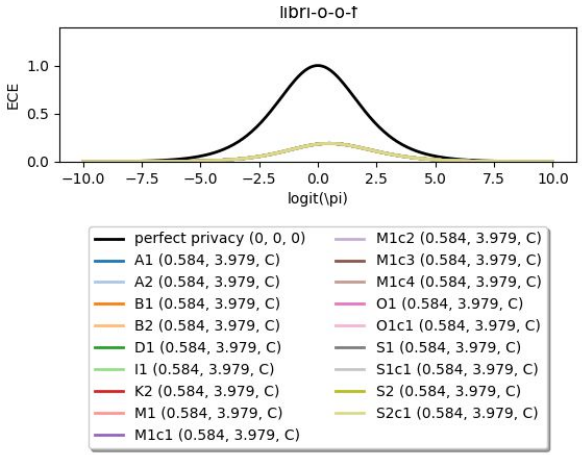
Sorted by aa f-test-vctk

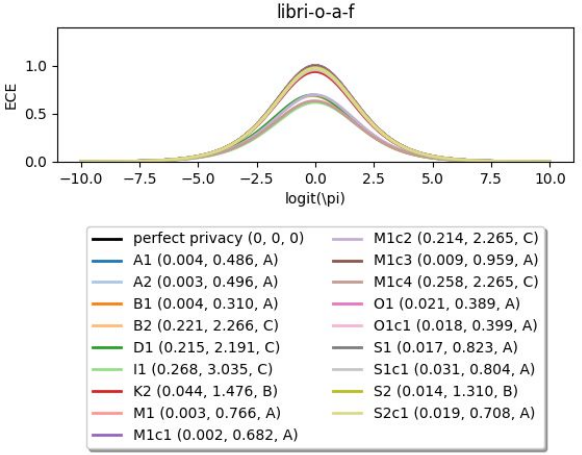
ZEBRA: LibriSpeech (all systems)

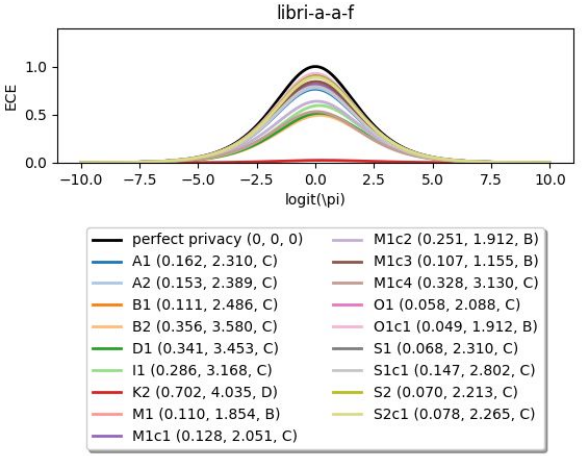




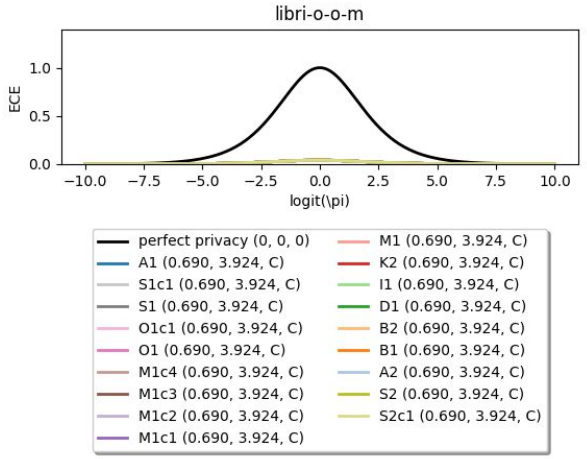


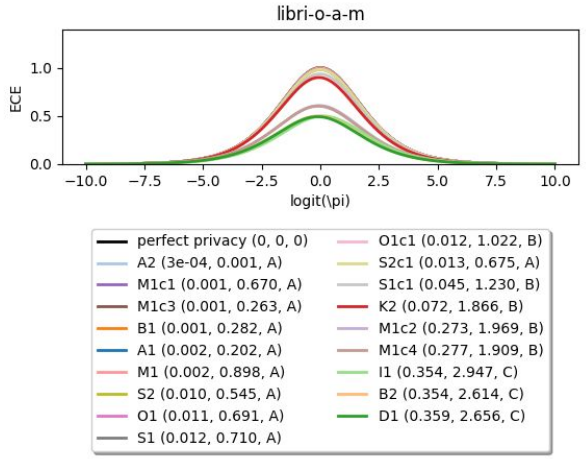


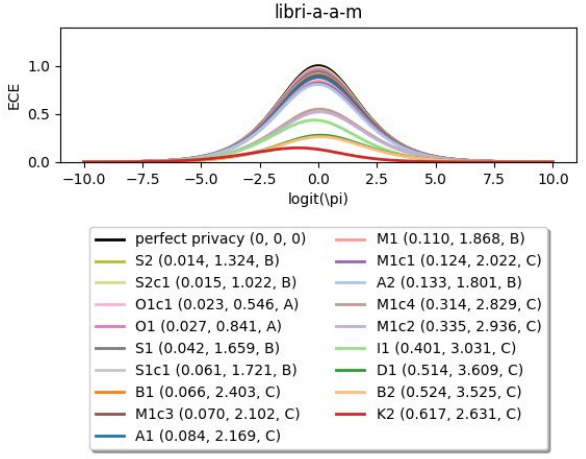


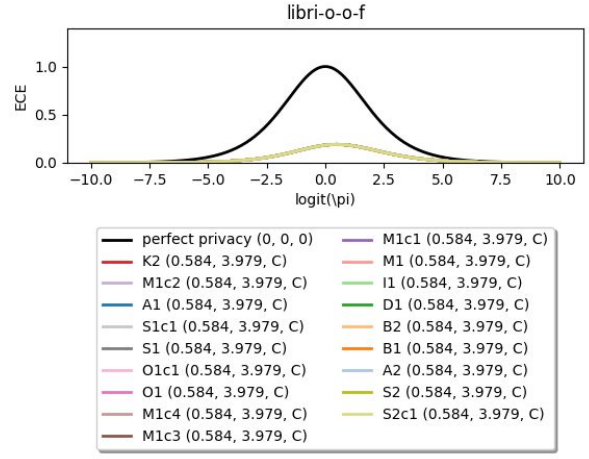


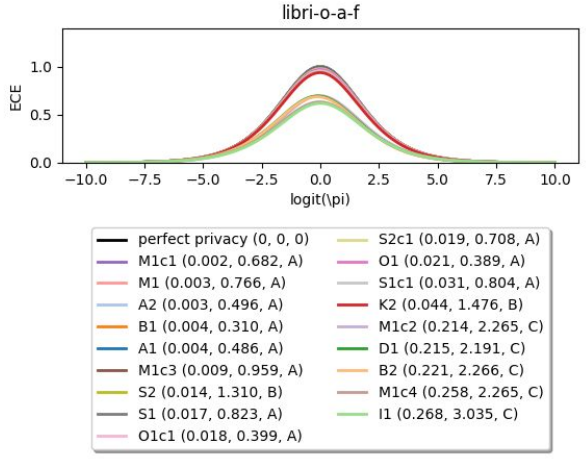
ZEBRA: LibriSpeech (all systems) ordered by performance

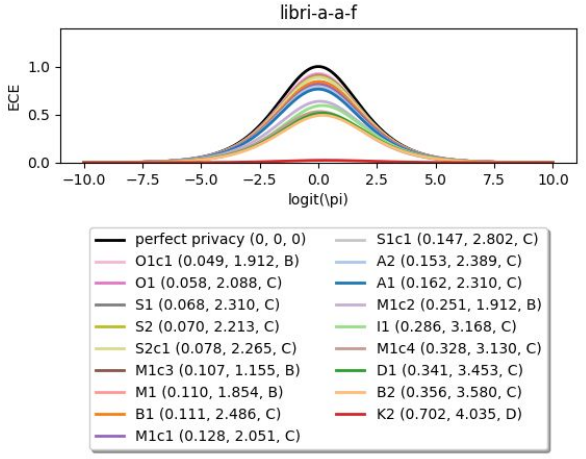




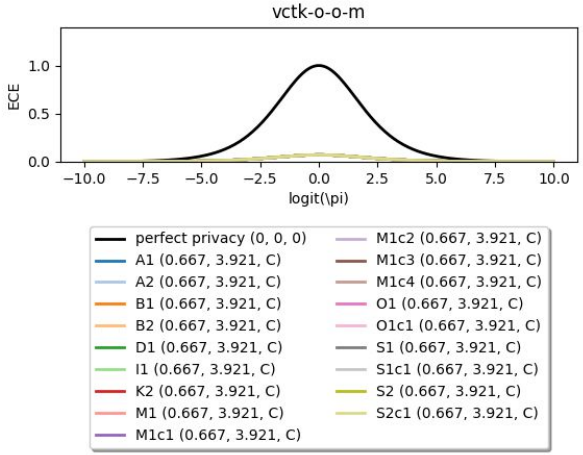


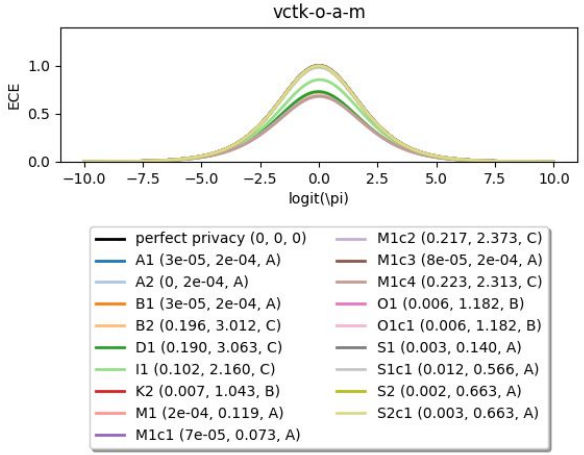


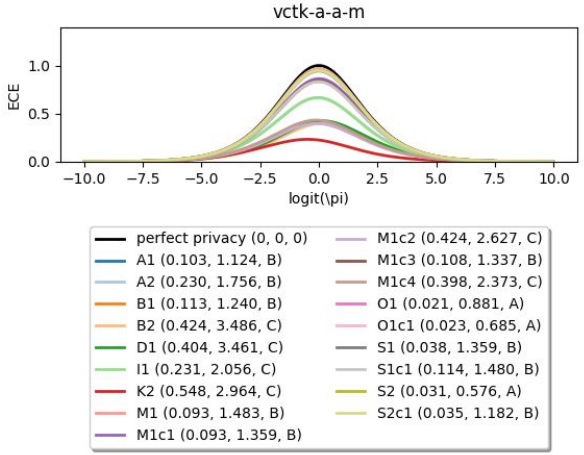


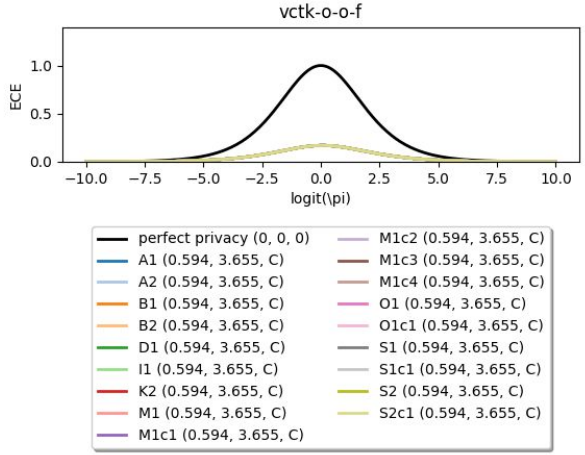


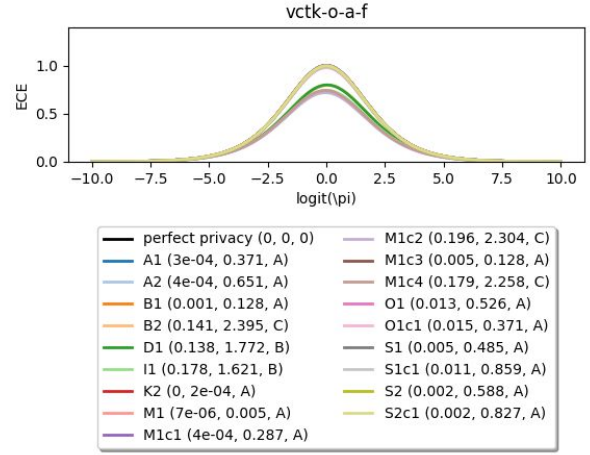
ZEBRA: VCTK-different (all systems)

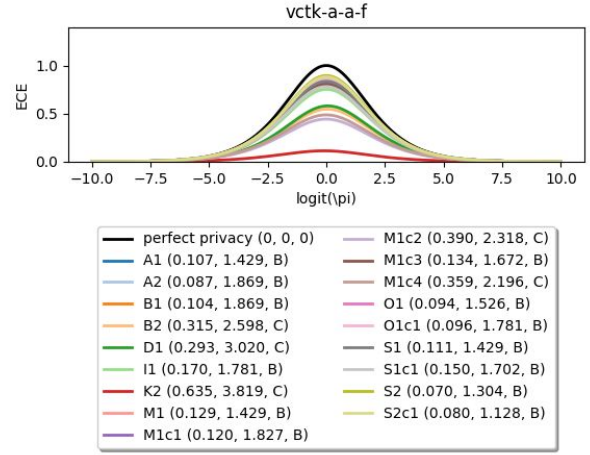




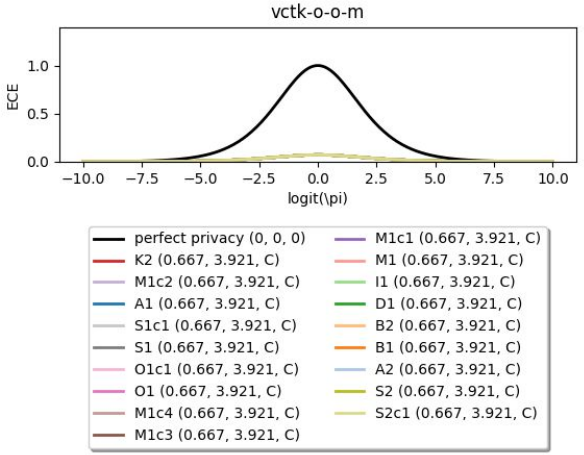


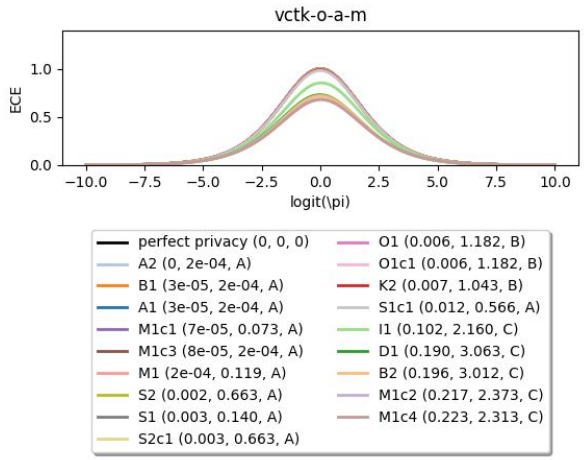


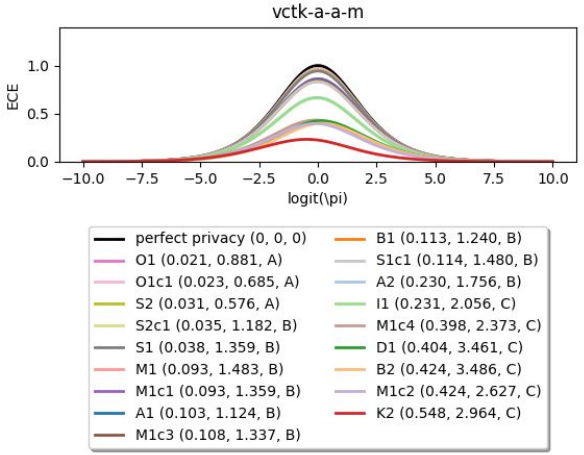


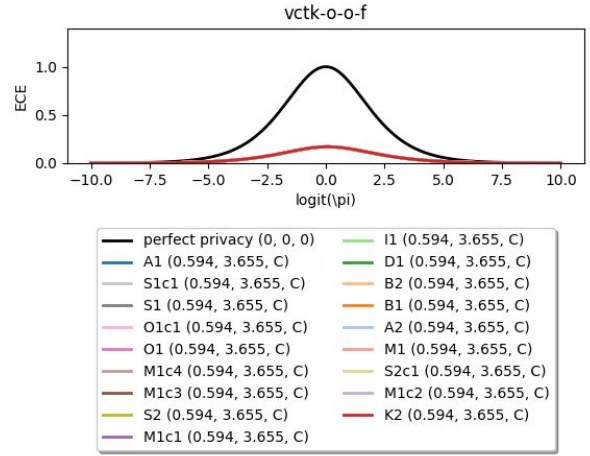


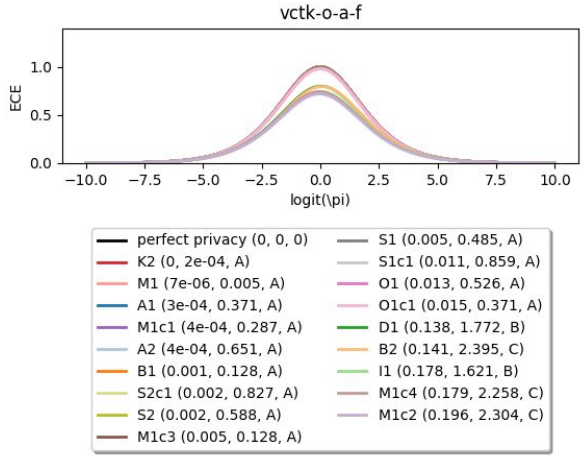
ZEBRA: VCTK-different (all systems) ordered by performance

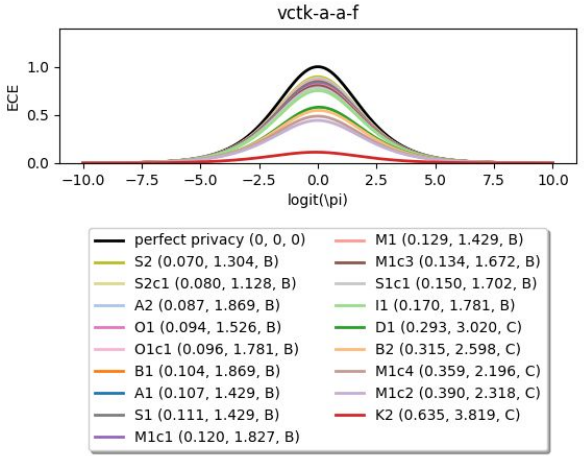




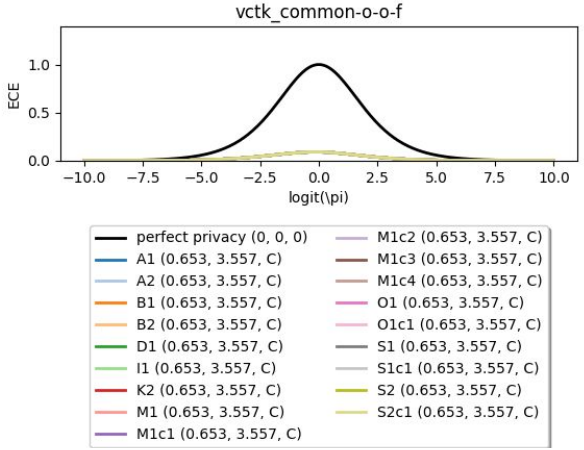


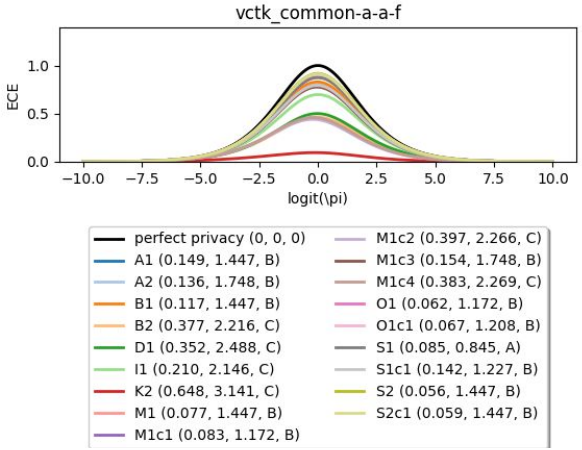


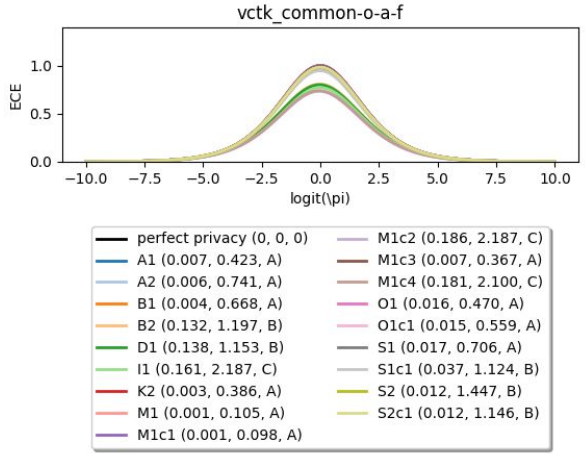


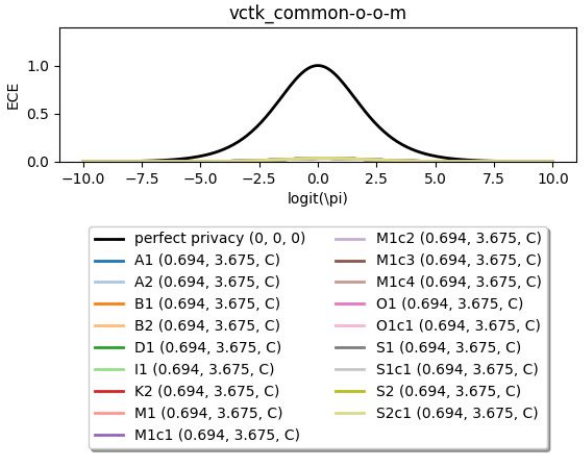


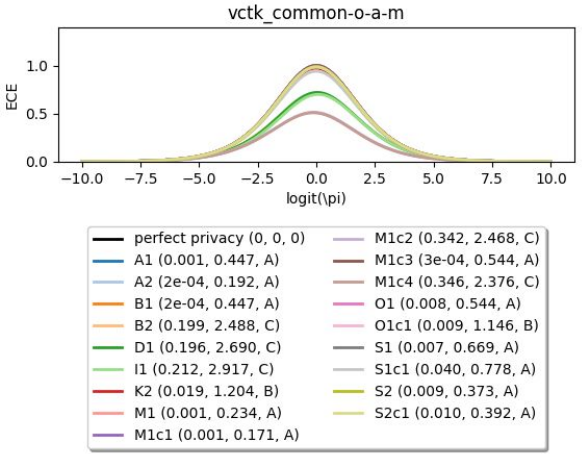
ZEBRA: VCTK-common (all systems)

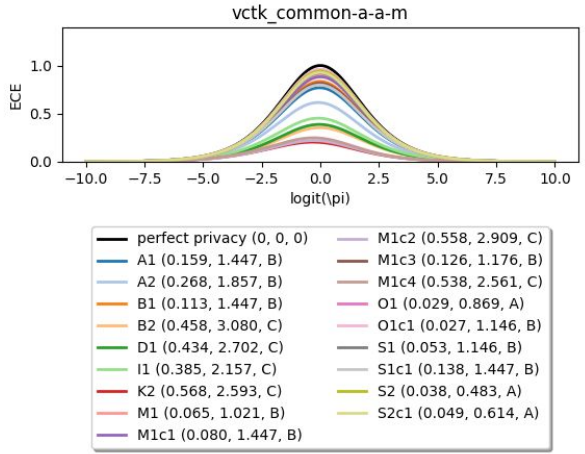












ZEBRA: VCTK-common (all systems) ordered by performance