Table 1: Detection results under no perturbations on FMA.

(a) AudioSeal

Tau	**0.01**	**0.02**	**0.05**	**0.1**	**0.15**
FPR	0.210	0.130	0.060	0.030	0.020
FNR	0.000	0.000	0.000	0.000	0.000

(b) AudioSeal-B

Tau	**0.625**	**0.6875**	**0.75**	**0.8125**	**0.875**
FPR	0.080	0.010	0.010	0.000	0.000
FNR	0.000	0.000	0.010	0.040	0.110

(c) Timbre

Tau	**0.625**	**0.6875**	**0.75**	**0.8125**	**0.875**
FPR	0.100	0.000	0.000	0.000	0.000
FNR	0.000	0.000	0.000	0.010	0.030

(d) WavMark

Tau	**0.0**	**0.0625**	**0.125**	**0.1875**	**0.25**
FPR	0.000	0.000	0.000	0.000	0.000
FNR	0.000	0.000	0.000	0.000	0.000

Table 2: Detection results under background noise and time stretch on FMA.

(a) AudioSeal: Background noise

SNR	**FNR (AudioSeal-B)**	**FNR (AudioSeal)**	**FPR (AudioSeal-B)**	**FPR (AudioSeal)**
5	0.44	0.03	0.00	0.01
10	0.26	0.00	0.00	0.02
20	0.04	0.00	0.00	0.01
30	0.02	0.00	0.01	0.01
40	0.01	0.00	0.00	0.01

(b) AudioSeal: Time stretch

Stretch Factor	**FNR (AudioSeal-B)**	**FNR (AudioSeal)**	**FPR (AudioSeal-B)**	**FPR (AudioSeal)**
0.7	0.24	0.60	0.01	0.01
0.9	0.19	0.45	0.00	0.02
1.1	0.10	0.25	0.01	0.02
1.3	0.17	0.57	0.01	0.01
1.5	0.31	0.73	0.02	0.00

(c) Timbre: Background noise

SNR	**FNR**	**FPR**
5	0.52	0.00
10	0.25	0.00
20	0.04	0.00
30	0.02	0.00
40	0.01	0.00

(d) Timbre: Time stretch

Stretch Factor	**FNR**	**FPR**
0.7	0.08	0.00
0.9	0.04	0.00
1.1	0.04	0.00
1.3	0.12	0.00
1.5	0.13	0.00

(e) WavMark:Background noise

SNR	**FNR**	**FPR**
5	0.96	0.00
10	0.68	0.00
20	0.07	0.00
30	0.00	0.00
40	0.00	0.00

(f) WavMark: Time stretch

Stretch Factor	**FNR**	**FPR**
0.7	0.41	0.00
0.9	0.20	0.00
1.1	0.24	0.00
1.3	0.57	0.00
1.5	0.84	0.00

Table 3: Results under black-box Square attack on FMA.

(a) AudioSeal

ℓ_∞ **Bound**	**FNR**	**SNR**	**ViSQOL**
0.05	0.0	24.85	4.57
0.1	0.0	18.84	4.33
0.15	0.0	15.33	4.14
0.2	0.0	12.83	3.96

(b) AudioSeal-B

ℓ_∞ **Bound**	**FNR**	**SNR**	**ViSQOL**
0.05	0.15	24.85	4.55
0.1	0.25	18.83	4.29
0.15	0.45	15.33	4.08
0.2	0.55	12.85	3.91

(c) Timbre

ℓ_∞ **Bound**	**FNR**	**SNR**	**ViSQOL**
0.05	0.0	25.68	4.79
0.1	0.07	19.66	4.58
0.15	0.14	16.14	4.39
0.2	0.21	13.66	4.22

(d) WavMark

ℓ_∞ **Bound**	**FNR**	**SNR**	**ViSQOL**
0.05	0.0	25.96	4.81
0.1	0.67	19.63	4.70
0.15	1.0	16.18	4.58
0.2	1.0	14.52	4.52

Table 4: Detection results under white-box removal attack on FMA.

(a) AudioSeal

SNR	**20**	**30**	**40**	**50**	**60**
FNR	1.00	0.85	0.35	0.00	0.00

(b) AudioSeal-B

SNR	**20**	**30**	**40**	**50**	**60**
FNR	1.00	1.00	0.95	0.50	0.00

(c) Timbre

SNR	**20**	**30**	**40**	**50**	**60**
FNR	1.00	0.95	0.85	0.45	0.10

(d) WavMark

SNR	**20**	**30**	**40**	**50**	**60**
FNR	1.00	1.00	1.00	0.50	0.40

Table 5: Detection results under white-box forgery attack on FMA.

(a) AudioSeal

SNR	**20**	**30**	**40**	**50**	**60**
FPR	1.00	1.00	1.00	1.00	1.00

(b) AudioSeal-B

SNR	**20**	**30**	**40**	**50**	**60**
FPR	1.00	0.95	0.90	0.40	0.30

(c) Timbre

SNR	**20**	**30**	**40**	**50**	**60**
FPR	1.00	1.00	0.90	0.50	0.20

(d) WavMark

SNR	**20**	**30**	**40**	**50**	**60**
FPR	1.00	1.00	1.00	1.00	1.00

Table 6: Results for I-FGSM on AudioMarkData.

(a) I-FGSM watermark removal: AudioSeal

SNR	20	30	40	50	60
FNR	1.00	0.75	0.25	0.00	0.00

(b) I-FGSM watermark removal: AudioSeal-B

SNR	20	30	40	50	60
FNR	1.00	1.00	0.90	0.40	0.15

(c) I-FGSM watermark removal: Timbre

SNR	20	30	40	50	60
FNR	1.00	1.00	0.85	0.45	0.05

(d) I-FGSM watermark removal: WavMark

SNR	20	30	40	50	60
FNR	1.00	1.00	0.50	0.50	0.15

(e) I-FGSM watermark forgery: AudioSeal

SNR	20	30	40	50	60
FPR	1.00	1.00	1.00	1.00	1.00

(f) I-FGSM watermark forgery: AudioSeal-B

SNR	20	30	40	50	60
FPR	1.00	1.00	0.75	0.20	0.00

(g) I-FGSM watermark forgery: Timbre

SNR	20	30	40	50	60
FPR	1.00	1.00	1.00	0.90	0.50

(h) I-FGSM watermark forgery: WavMark

SNR	20	30	40	50	60
FPR	1.00	1.00	1.00	1.00	1.00

Table 7: Results for composited no-box attacks.

(a) EnCodeC (24kHz) followed by MP3 (16kbps)

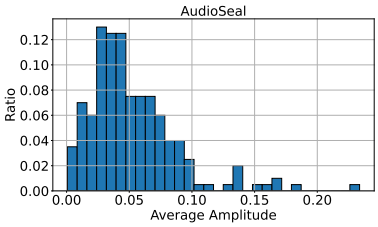
Method	FNR	FPR
AudioSeal	0.99	0.00
AudioSeal-B	1.00	0.00
Timbre	0.99	0.00
WavMark	1.00	0.00

(b) MP3 (16kbps) followed by EnCodeC (24kHz)

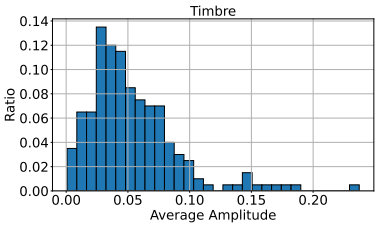
Method	FNR	FPR
AudioSeal	0.95	0.00
AudioSeal-B	1.00	0.00
Timbre	0.96	0.00
WavMark	1.00	0.00

(c) Gaussian noise (SNR 20dB) followed by MP3 (16kbps)

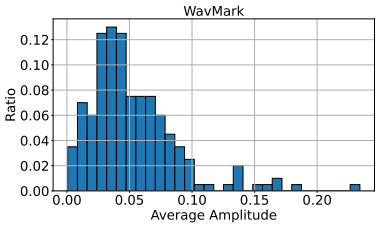
Method	FNR	FPR
AudioSeal	0.09	0.00
AudioSeal-B	0.65	0.00
Timbre	0.38	0.00
WavMark	1.00	0.00



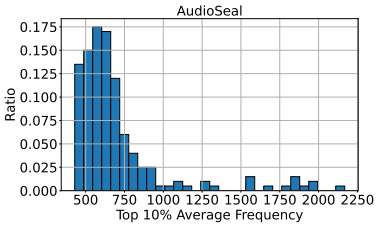
(a) Amplitude - AudioSeal



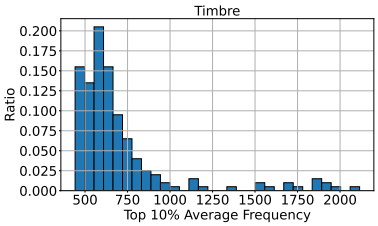
(b) Amplitude - Timbre



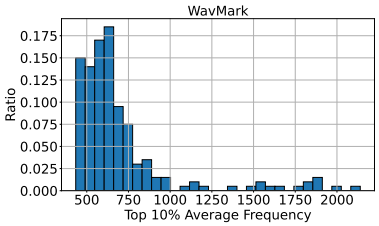
(c) Amplitude - WavMark



(d) Frequency - AudioSeal



(e) Frequency - Timbre



(f) Frequency - WavMark

Figure 1: Distribution of the watermarked audio's average amplitude and the average frequency of the top 10% amplitude after applying a fast Fourier transform (FFT).