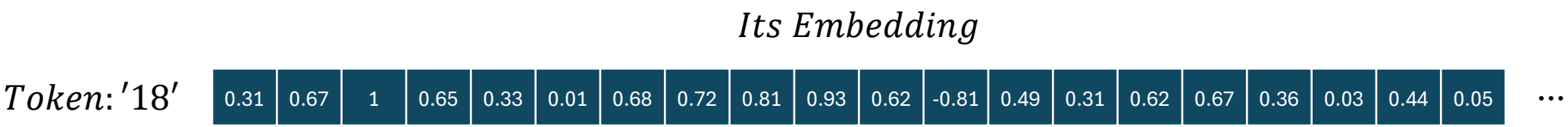
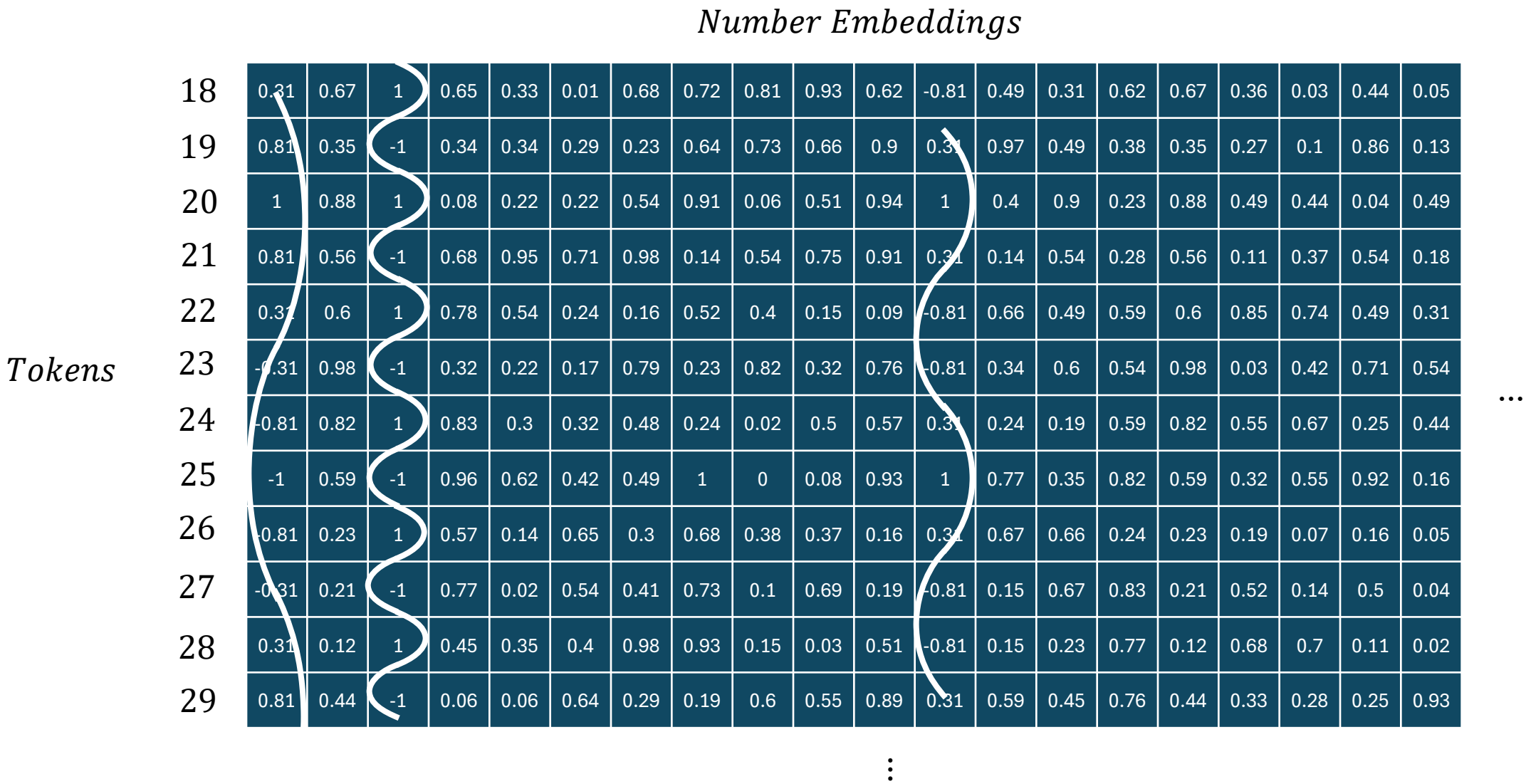

Its Embedding
Token: '18'
0.31
0.67
1
0.65
0.33
0.01
0.68
0.72
0.81
0.93
0.62
-0.81
0.49
0.31
0.62
0.67
0.36
0.03
0.44
0.05
...

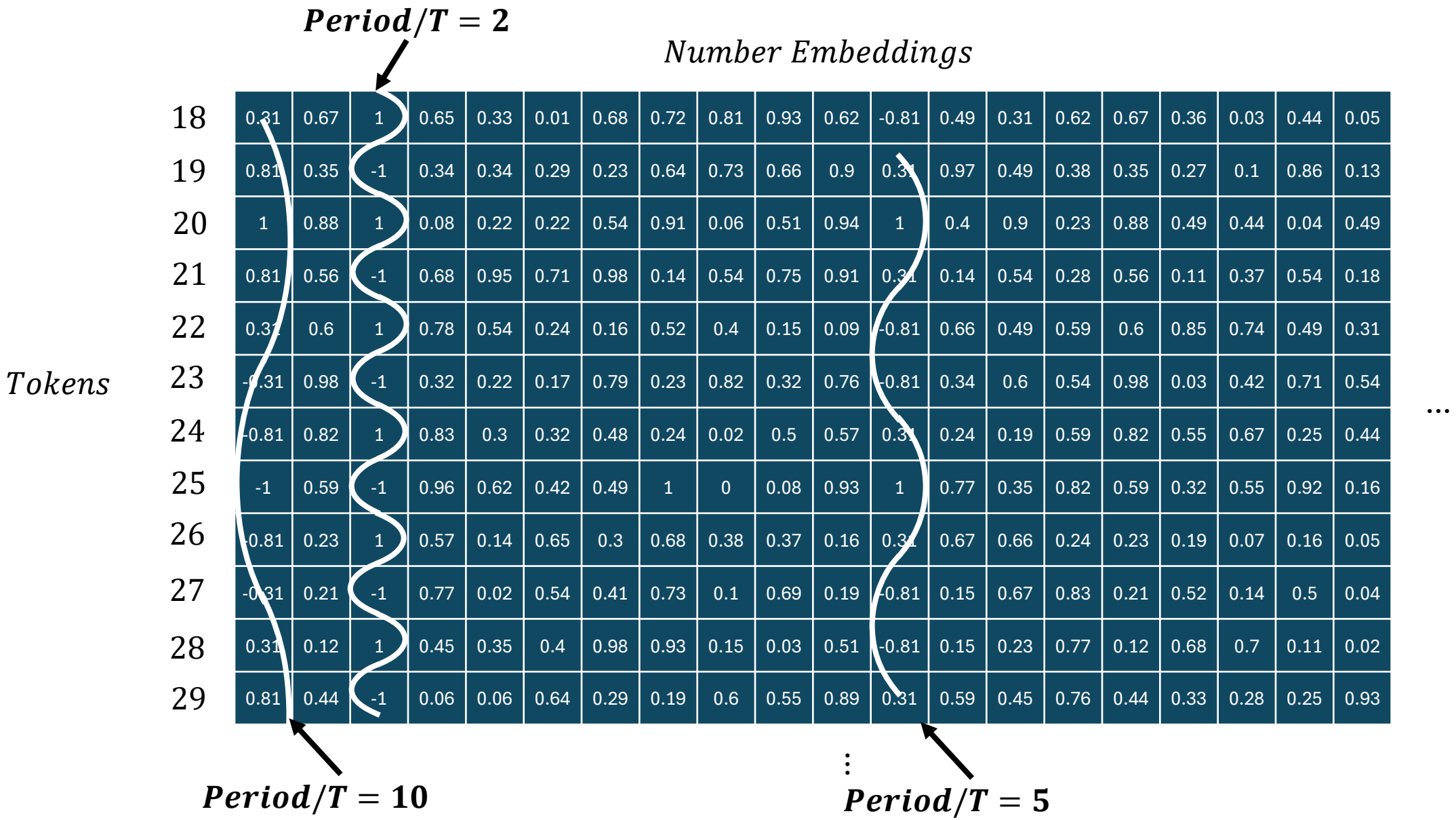
Number Embeddings

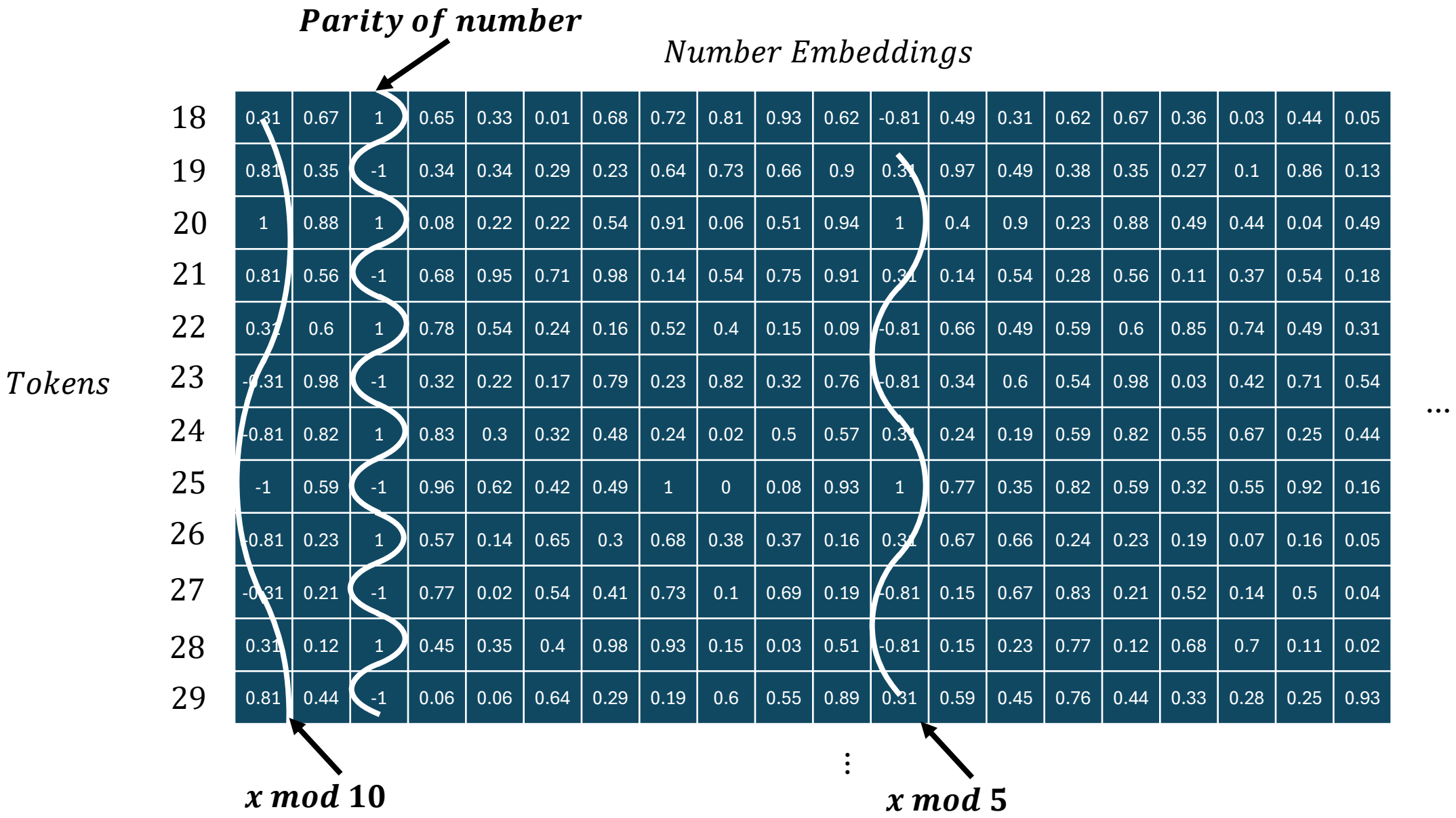
Tokens																				
18	0.31	0.67	1	0.65	0.33	0.01	0.68	0.72	0.81	0.93	0.62	-0.81	0.49	0.31	0.62	0.67	0.36	0.03	0.44	0.05
19	0.81	0.35	-1	0.34	0.34	0.29	0.23	0.64	0.73	0.66	0.9	0.31	0.97	0.49	0.38	0.35	0.27	0.1	0.86	0.13
20	1	0.88	1	0.08	0.22	0.22	0.54	0.91	0.06	0.51	0.94	1	0.4	0.9	0.23	0.88	0.49	0.44	0.04	0.49
21	0.81	0.56	-1	0.68	0.95	0.71	0.98	0.14	0.54	0.75	0.91	0.31	0.14	0.54	0.28	0.56	0.11	0.37	0.54	0.18
22	0.31	0.6	1	0.78	0.54	0.24	0.16	0.52	0.4	0.15	0.09	-0.81	0.66	0.49	0.59	0.6	0.85	0.74	0.49	0.31
23	-0.31	0.98	-1	0.32	0.22	0.17	0.79	0.23	0.82	0.32	0.76	-0.81	0.34	0.6	0.54	0.98	0.03	0.42	0.71	0.54
24	-0.81	0.82	1	0.83	0.3	0.32	0.48	0.24	0.02	0.5	0.57	0.31	0.24	0.19	0.59	0.82	0.55	0.67	0.25	0.44
25	-1	0.59	-1	0.96	0.62	0.42	0.49	1	0	0.08	0.93	1	0.77	0.35	0.82	0.59	0.32	0.55	0.92	0.16
26	-0.81	0.23	1	0.57	0.14	0.65	0.3	0.68	0.38	0.37	0.16	0.31	0.67	0.66	0.24	0.23	0.19	0.07	0.16	0.05
27	-0.31	0.21	-1	0.77	0.02	0.54	0.41	0.73	0.1	0.69	0.19	-0.81	0.15	0.67	0.83	0.21	0.52	0.14	0.5	0.04
28	0.31	0.12	1	0.45	0.35	0.4	0.98	0.93	0.15	0.03	0.51	-0.81	0.15	0.23	0.77	0.12	0.68	0.7	0.11	0.02
29	0.81	0.44	-1	0.06	0.06	0.64	0.29	0.19	0.6	0.55	0.89	0.31	0.59	0.45	0.76	0.44	0.33	0.28	0.25	0.93

...

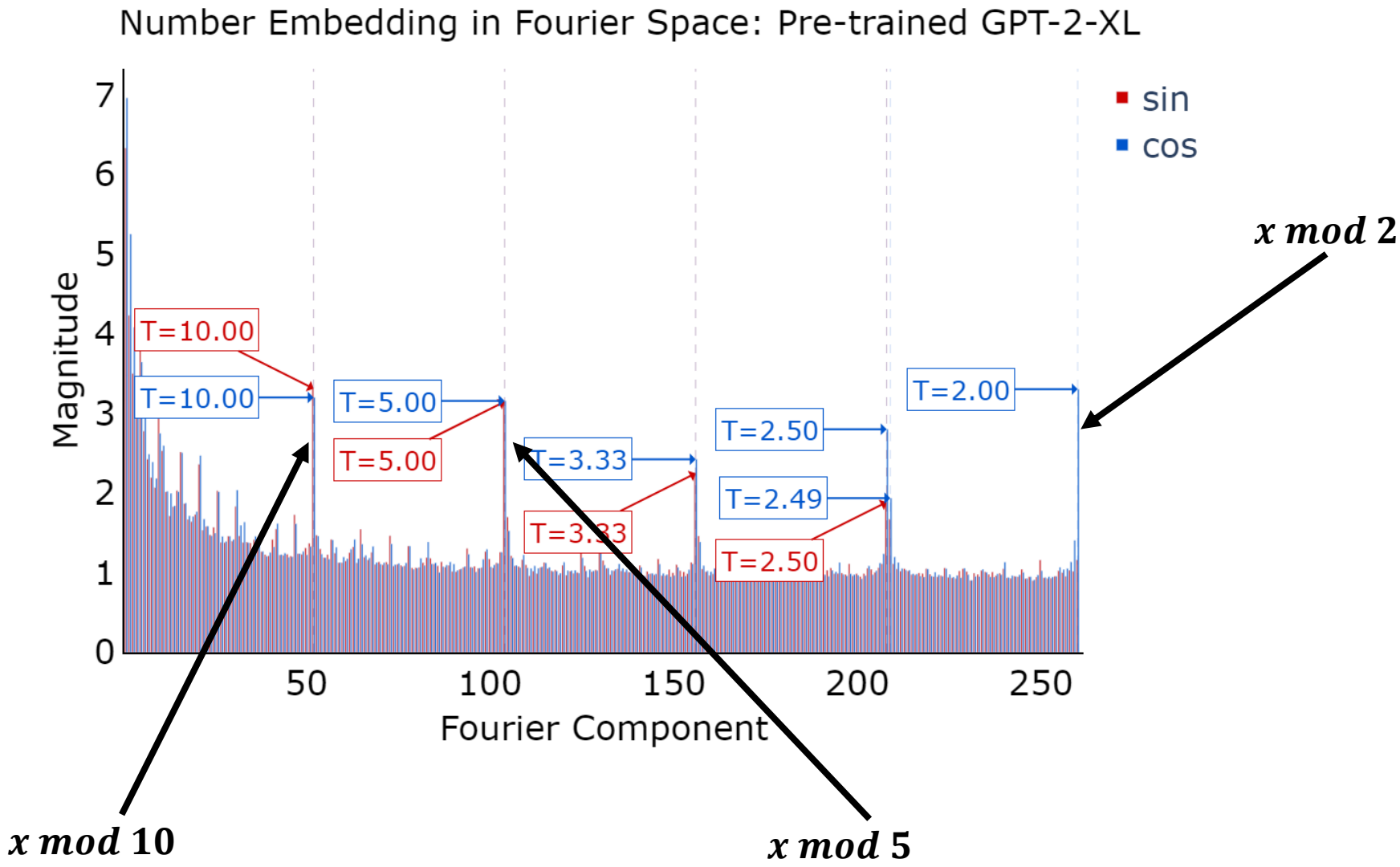
⋮


Number Embeddings
Tokens
18 0.31 0.67 1 0.65 0.33 0.01 0.68 0.72 0.81 0.93 0.62 -0.81 0.49 0.31 0.62 0.67 0.36 0.03 0.44 0.05
19 0.81 0.35 -1 0.34 0.34 0.29 0.23 0.64 0.73 0.66 0.9 0.31 0.97 0.49 0.38 0.35 0.27 0.1 0.86 0.13
20 1 0.88 1 0.08 0.22 0.22 0.54 0.91 0.06 0.51 0.94 1 0.4 0.9 0.23 0.88 0.49 0.44 0.04 0.49
21 0.81 0.56 -1 0.68 0.95 0.71 0.98 0.14 0.54 0.75 0.91 0.31 0.14 0.54 0.28 0.56 0.11 0.37 0.54 0.18
22 0.31 0.6 1 0.78 0.54 0.24 0.16 0.52 0.4 0.15 0.09 -0.81 0.66 0.49 0.59 0.6 0.85 0.74 0.49 0.31
23 -0.31 0.98 -1 0.32 0.22 0.17 0.79 0.23 0.82 0.32 0.76 -0.81 0.34 0.6 0.54 0.98 0.03 0.42 0.71 0.54
24 -0.81 0.82 1 0.83 0.3 0.32 0.48 0.24 0.02 0.5 0.57 0.31 0.24 0.19 0.59 0.82 0.55 0.67 0.25 0.44
25 -1 0.59 -1 0.96 0.62 0.42 0.49 1 0 0.08 0.93 1 0.77 0.35 0.82 0.59 0.32 0.55 0.92 0.16
26 -0.81 0.23 1 0.57 0.14 0.65 0.3 0.68 0.38 0.37 0.16 0.31 0.67 0.66 0.24 0.23 0.19 0.07 0.16 0.05
27 -0.31 0.21 -1 0.77 0.02 0.54 0.41 0.73 0.1 0.69 0.19 -0.81 0.15 0.67 0.83 0.21 0.52 0.14 0.5 0.04
28 0.31 0.12 1 0.45 0.35 0.4 0.98 0.93 0.15 0.03 0.51 -0.81 0.15 0.23 0.77 0.12 0.68 0.7 0.11 0.02
29 0.81 0.44 -1 0.06 0.06 0.64 0.29 0.19 0.6 0.55 0.89 0.31 0.59 0.45 0.76 0.44 0.33 0.28 0.25 0.93
...
⋮

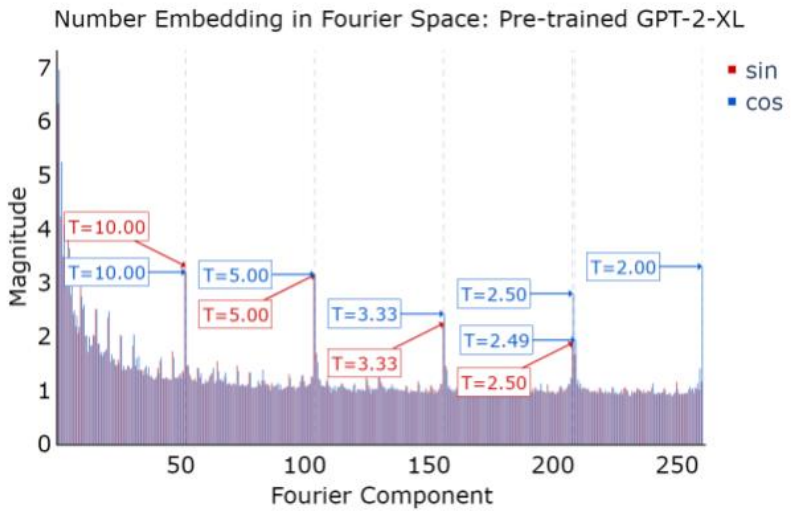

Period/T = 2
Number Embeddings
Tokens
18 0.31 0.67 1 0.65 0.33 0.01 0.68 0.72 0.81 0.93 0.62 -0.81 0.49 0.31 0.62 0.67 0.36 0.03 0.44 0.05
19 0.81 0.35 -1 0.34 0.34 0.29 0.23 0.64 0.73 0.66 0.9 0.31 0.97 0.49 0.38 0.35 0.27 0.1 0.86 0.13
20 1 0.88 1 0.08 0.22 0.22 0.54 0.91 0.06 0.51 0.94 1 0.4 0.9 0.23 0.88 0.49 0.44 0.04 0.49
21 0.81 0.56 -1 0.68 0.95 0.71 0.98 0.14 0.54 0.75 0.91 0.31 0.14 0.54 0.28 0.56 0.11 0.37 0.54 0.18
22 0.31 0.6 1 0.78 0.54 0.24 0.16 0.52 0.4 0.15 0.09 -0.81 0.66 0.49 0.59 0.6 0.85 0.74 0.49 0.31
23 -0.31 0.98 -1 0.32 0.22 0.17 0.79 0.23 0.82 0.32 0.76 -0.81 0.34 0.6 0.54 0.98 0.03 0.42 0.71 0.54
24 -0.81 0.82 1 0.83 0.3 0.32 0.48 0.24 0.02 0.5 0.57 0.31 0.24 0.19 0.59 0.82 0.55 0.67 0.25 0.44
25 -1 0.59 -1 0.96 0.62 0.42 0.49 1 0 0.08 0.93 1 0.77 0.35 0.82 0.59 0.32 0.55 0.92 0.16
26 -0.81 0.23 1 0.57 0.14 0.65 0.3 0.68 0.38 0.37 0.16 0.31 0.67 0.66 0.24 0.23 0.19 0.07 0.16 0.05
27 -0.31 0.21 -1 0.77 0.02 0.54 0.41 0.73 0.1 0.69 0.19 -0.81 0.15 0.67 0.83 0.21 0.52 0.14 0.5 0.04
28 0.31 0.12 1 0.45 0.35 0.4 0.98 0.93 0.15 0.03 0.51 -0.81 0.15 0.23 0.77 0.12 0.68 0.7 0.11 0.02
29 0.81 0.44 -1 0.06 0.06 0.64 0.29 0.19 0.6 0.55 0.89 0.31 0.59 0.45 0.76 0.44 0.33 0.28 0.25 0.93
...
Period/T = 10
Period/T = 5


Parity of number
Number Embeddings
Tokens
18 0.31 0.67 1 0.65 0.33 0.01 0.68 0.72 0.81 0.93 0.62 -0.81 0.49 0.31 0.62 0.67 0.36 0.03 0.44 0.05
19 0.81 0.35 -1 0.34 0.34 0.29 0.23 0.64 0.73 0.66 0.9 0.31 0.97 0.49 0.38 0.35 0.27 0.1 0.86 0.13
20 1 0.88 1 0.08 0.22 0.22 0.54 0.91 0.06 0.51 0.94 1 0.4 0.9 0.23 0.88 0.49 0.44 0.04 0.49
21 0.81 0.56 -1 0.68 0.95 0.71 0.98 0.14 0.54 0.75 0.91 0.31 0.14 0.54 0.28 0.56 0.11 0.37 0.54 0.18
22 0.31 0.6 1 0.78 0.54 0.24 0.16 0.52 0.4 0.15 0.09 -0.81 0.66 0.49 0.59 0.6 0.85 0.74 0.49 0.31
23 -0.31 0.98 -1 0.32 0.22 0.17 0.79 0.23 0.82 0.32 0.76 -0.81 0.34 0.6 0.54 0.98 0.03 0.42 0.71 0.54
24 -0.81 0.82 1 0.83 0.3 0.32 0.48 0.24 0.02 0.5 0.57 0.31 0.24 0.19 0.59 0.82 0.55 0.67 0.25 0.44
25 -1 0.59 -1 0.96 0.62 0.42 0.49 1 0 0.08 0.93 1 0.77 0.35 0.82 0.59 0.32 0.55 0.92 0.16
26 -0.81 0.23 1 0.57 0.14 0.65 0.3 0.68 0.38 0.37 0.16 0.31 0.67 0.66 0.24 0.23 0.19 0.07 0.16 0.05
27 -0.31 0.21 -1 0.77 0.02 0.54 0.41 0.73 0.1 0.69 0.19 -0.81 0.15 0.67 0.83 0.21 0.52 0.14 0.5 0.04
28 0.31 0.12 1 0.45 0.35 0.4 0.98 0.93 0.15 0.03 0.51 -0.81 0.15 0.23 0.77 0.12 0.68 0.7 0.11 0.02
29 0.81 0.44 -1 0.06 0.06 0.64 0.29 0.19 0.6 0.55 0.89 0.31 0.59 0.45 0.76 0.44 0.33 0.28 0.25 0.93
...
x mod 10
x mod 5

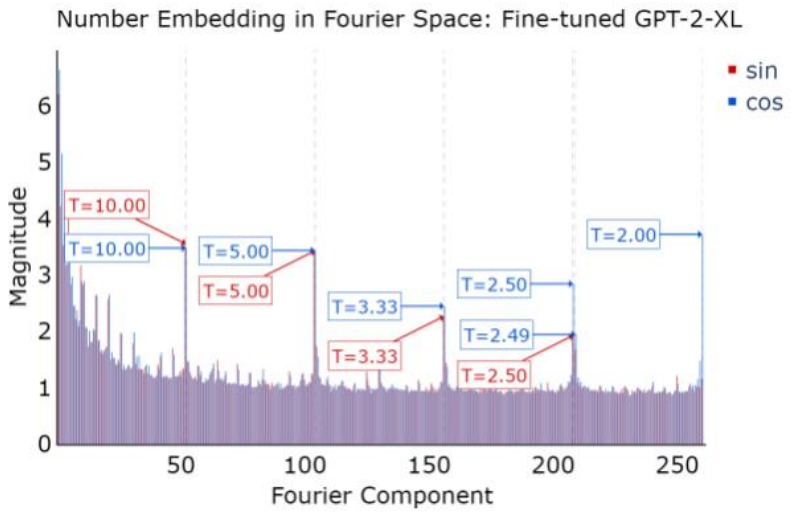
Transform to Fourier Space for better analysis



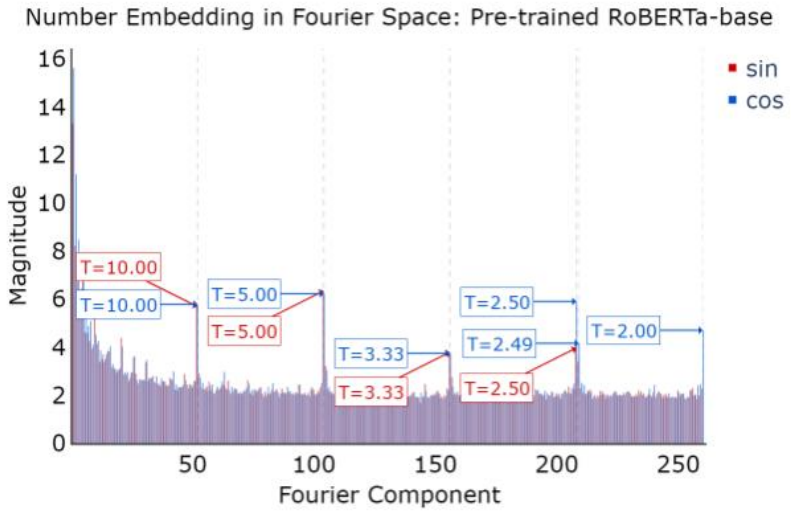
All these pretrained LLMs have similar Fourier features!!



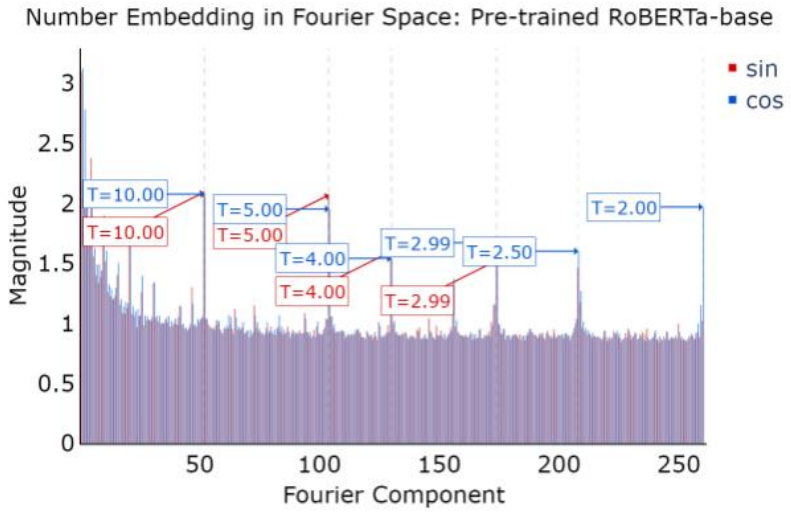
(a) pre-trained GPT-2-XL



(b) fine-tuned GPT-2-XL



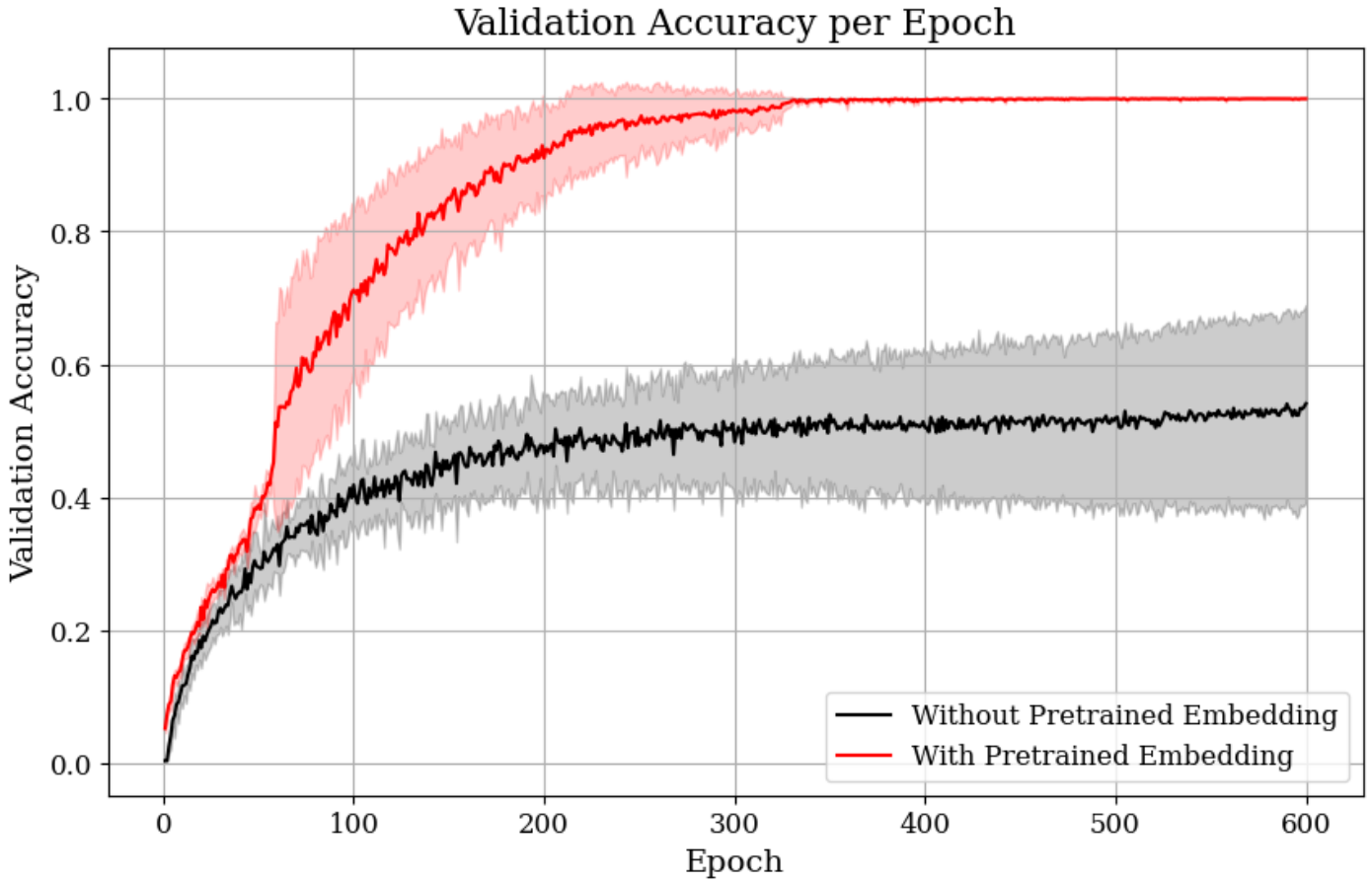
(c) pre-trained RoBERTa



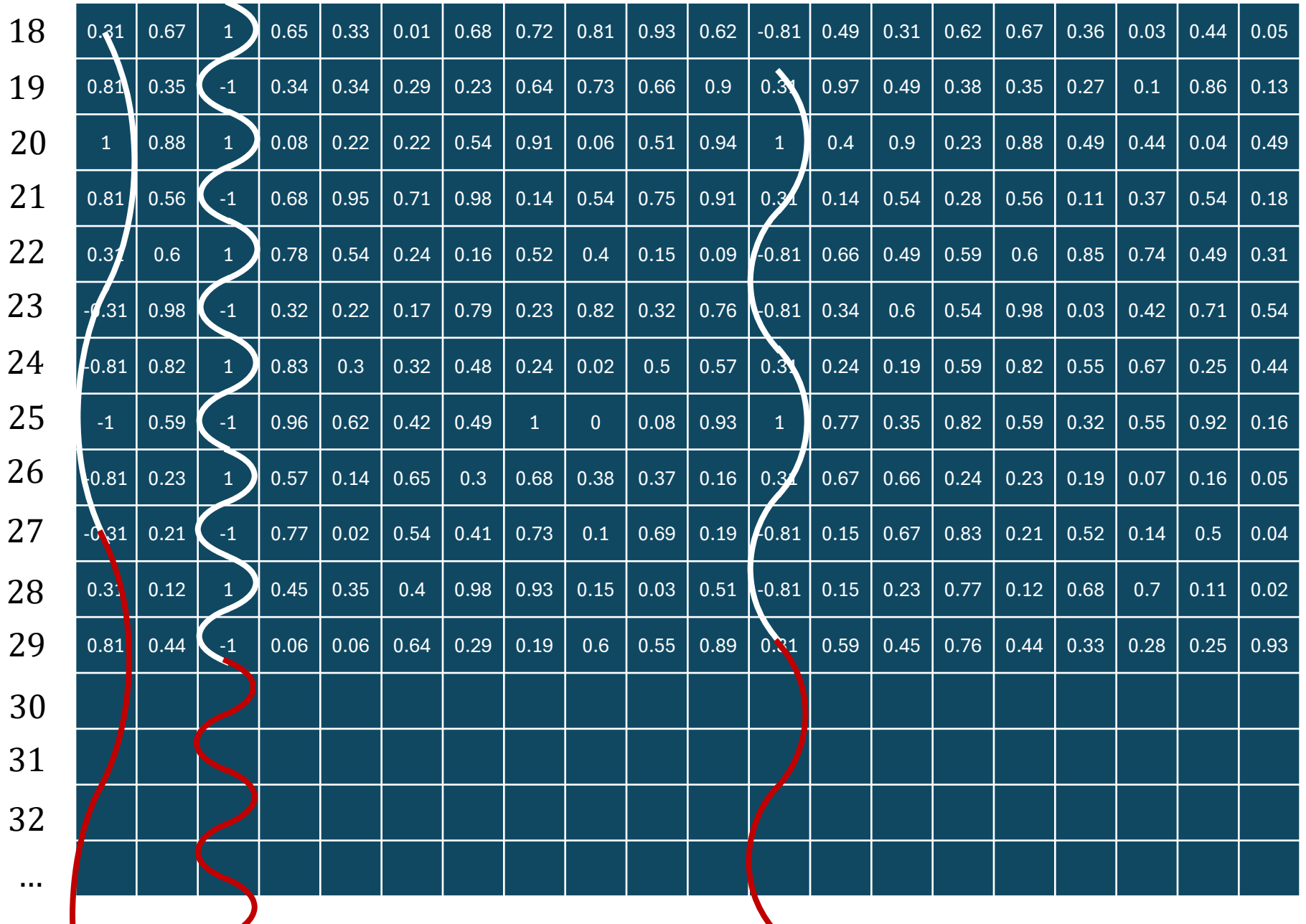
(d) pre-trained Phi2

Figure 15: Number embedding in Fourier space for different pre-trained models.

With solely these Fourier Features, models trained from scratch can achieve 100% accuracy!!



We extend these waves and add waves with a larger period to obtain the embedding for all numbers.



18	0.31	0.67	1	0.65	0.33	0.01	0.68	0.72	0.81	0.93	0.62	-0.81	0.49	0.31	0.62	0.67	0.36	0.03	0.44	0.05
19	0.81	0.35	-1	0.34	0.34	0.29	0.23	0.64	0.73	0.66	0.9	0.31	0.97	0.49	0.38	0.35	0.27	0.1	0.86	0.13
20	1	0.88	1	0.08	0.22	0.22	0.54	0.91	0.06	0.51	0.94	1	0.4	0.9	0.23	0.88	0.49	0.44	0.04	0.49
21	0.81	0.56	-1	0.68	0.95	0.71	0.98	0.14	0.54	0.75	0.91	0.31	0.14	0.54	0.28	0.56	0.11	0.37	0.54	0.18
22	0.31	0.6	1	0.78	0.54	0.24	0.16	0.52	0.4	0.15	0.09	-0.81	0.66	0.49	0.59	0.6	0.85	0.74	0.49	0.31
23	-0.31	0.98	-1	0.32	0.22	0.17	0.79	0.23	0.82	0.32	0.76	-0.81	0.34	0.6	0.54	0.98	0.03	0.42	0.71	0.54
24	-0.81	0.82	1	0.83	0.3	0.32	0.48	0.24	0.02	0.5	0.57	0.31	0.24	0.19	0.59	0.82	0.55	0.67	0.25	0.44
25	-1	0.59	-1	0.96	0.62	0.42	0.49	1	0	0.08	0.93	1	0.77	0.35	0.82	0.59	0.32	0.55	0.92	0.16
26	-0.81	0.23	1	0.57	0.14	0.65	0.3	0.68	0.38	0.37	0.16	0.31	0.67	0.66	0.24	0.23	0.19	0.07	0.16	0.05
27	-0.31	0.21	-1	0.77	0.02	0.54	0.41	0.73	0.1	0.69	0.19	-0.81	0.15	0.67	0.83	0.21	0.52	0.14	0.5	0.04
28	0.31	0.12	1	0.45	0.35	0.4	0.98	0.93	0.15	0.03	0.51	-0.81	0.15	0.23	0.77	0.12	0.68	0.7	0.11	0.02
29	0.81	0.44	-1	0.06	0.06	0.64	0.29	0.19	0.6	0.55	0.89	0.81	0.59	0.45	0.76	0.44	0.33	0.28	0.25	0.93
30																				
31																				
32																				
...																				