0.1 Additional Experiment results

The following sections gather all of our experiment results we collected to compare our sampling-based robust tensor power method to a sketching-based robust tensor power method. Our algorithm has the following parameters to set:

1. T: Number of power iterations
2. L: The number of starting vectors of the robust tensor power method
3. B: The number of sketches used in sketching, or the number of repetitions of sampling
4. b: The size of the sketch, or the number of indices sampled

In all of our experiments, we fix $T = 30$, $L = 50$, and change B and b for each input tensor. We observe the squared residual Frobenius norm to evaluate the performance of each algorithm. We only compute the first eigenvalue and eigenvector (rank-1 recovery) of each tensor. This enables us to run a large collection of tensors within a reasonable time.

0.1.1 Synthetic tensors with inverse decaying eigenvalues

The tables shown in this section use synthetic tensors generated with eigenvalues decaying as $\lambda_i = \frac{1}{i}$. All tensors have rank $k = 100$ and added noise with $\sigma = 0.01$.

Sketching based robust power method: $n = \mathbf{200}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	0.5981	0.4352	0.4277		0.2365	0.7794	1.395		0.2519	0.2383	0.3068
2^{12}	0.4343	0.4062	0.4047		1.08	3.055	5.078		0.3667	0.4322	0.495
2^{14}	0.4051	0.4006	0.4002		4.585	14.13	24.54		0.9685	1.074	1.149
2^{16}	0.4002	0.3989	0.3988		22.76	67.67	111.7		3.557	3.539	3.65

Table 1: Sketching based robust power method: $n = \mathbf{200}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{200}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.4068	0.4039	0.4003		0.3984	1.186	2		0.0	0.0	0.0
$10n$	0.4048	0.4	0.4003		0.6185	1.824	3.014		0.0	0.0	0.0
$20n$	0.4017	0.3994	0.3995		1.118	3.206	5.499		0.0	0.0	0.0
$30n$	0.4019	0.3992	0.3992		1.557	4.64	7.516		0.0	0.0	0.0
$40n$	0.4007	0.3991	0.3989		2.035	5.786	9.626		0.0	0.0	0.0

Table 2: Importance sampling based robust power method, without prescanning: $n = \mathbf{200}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{200}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.417	0.4023	0.3997		0.4084	1.203	2.062		0.01061	0.01067	0.01061
$10n$	0.4079	0.4001	0.3999		0.6833	2.107	3.287		0.01104	0.01068	0.01101
$20n$	0.4016	0.3997	0.3991		1.246	3.376	5.582		0.0105	0.01102	0.0104
$30n$	0.4012	0.3993	0.3988		1.674	4.724	7.866		0.01101	0.01105	0.01042
$40n$	0.4	0.399	0.3987		2.11	6.131	10.06		0.01101	0.011	0.01045

Table 3: Importance sampling based robust power method with prescanning: $n = \mathbf{200}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method: $n = \mathbf{400}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	0.9478	0.5163	0.455		0.3032	1.124	1.992		0.3742	0.84	1.307
2^{12}	0.5154	0.4225	0.4104		1.034	3.246	5.679		0.645	1.068	1.557
2^{14}	0.4149	0.4037	0.402		4.651	14.49	25.12		1.255	1.921	2.581
2^{16}	0.4023	0.3998	0.3991		21.89	67.64	113.6		3.812	4.633	5.398

Table 4: Sketching based robust power method: $n = \mathbf{400}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{400}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.4532	0.4032	0.4008		0.8437	2.724	4.627		0.0	0.0	0.0
$10n$	0.4018	0.4002	0.3995		1.461	4.504	7.593		0.0	0.0	0.0
$20n$	0.4058	0.3998	0.399		2.782	8.111	13.67		0.0	0.0	0.0
$30n$	0.4034	0.3997	0.3995		3.885	11.81	19.74		0.0	0.0	0.0
$40n$	0.3997	0.399	0.3987		5.108	15.42	25.75		0.0	0.0	0.0

Table 5: Importance sampling based robust power method, without prescanning: $n = \mathbf{400}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{400}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.4106	0.4036	0.4007		0.9649	3.052	5.219		0.08252	0.08253	0.08249
$10n$	0.4017	0.3998	0.3995		1.771	5.194	8.786		0.08351	0.08258	0.08574
$20n$	0.401	0.4006	0.3992		3.102	9.688	15.77		0.08351	0.0836	0.0868
$30n$	0.4019	0.3992	0.3994		4.507	13.63	22.81		0.08261	0.08248	0.08282
$40n$	0.3996	0.3992	0.3998		5.898	17.89	30.19		0.08253	0.0826	0.0836

Table 6: Importance sampling based robust power method with prescanning: $n = \mathbf{400}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method: $n = \mathbf{600}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	1.059	0.5516	0.5023		0.474	1.433	2.576		0.8617	2.295	3.745
2^{12}	0.4969	0.4366	0.4199		1.124	3.592	6.215		1.093	2.629	4.166
2^{14}	0.4291	0.4074	0.4037		4.592	14.61	25.64		2	4.058	6.103
2^{16}	0.4057	0.4003	0.3997		22.28	68.13	113.8		4.71	7.3	9.991

Table 7: Sketching based robust power method: $n = \mathbf{600}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{600}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.4071	0.4076	0.4007		1.24	3.978	6.828		0.0	0.0	0.0
$10n$	0.4055	0.401	0.4		2.186	6.755	11.43		0.0	0.0	0.0
$20n$	0.4042	0.3999	0.4001		3.996	12.18	20.57		0.0	0.0	0.0
$30n$	0.4027	0.3996	0.3991		5.807	17.69	29.73		0.0	0.0	0.0
$40n$	0.4011	0.3996	0.3989		7.654	23.21	39.02		0.0	0.0	0.0

Table 8: Importance sampling based robust power method, without prescanning: $n = \mathbf{600}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{600}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.4154	0.4005	0.4001		1.36	4.34	7.342		0.2784	0.2781	0.2781
$10n$	0.4092	0.4004	0.4031		2.333	7.219	12.17		0.279	0.2789	0.2789
$20n$	0.406	0.3993	0.399		4.272	13.04	21.9		0.2787	0.2794	0.2787
$30n$	0.3999	0.3996	0.399		6.21	18.83	31.57		0.2813	0.279	0.2787
$40n$	0.4005	0.3996	0.3994		8.147	24.64	41.19		0.292	0.2805	0.2798

Table 9: Importance sampling based robust power method with prescanning: $n = \mathbf{600}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method: $n = \mathbf{800}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	1.071	0.691	0.5281		0.4544	1.758	3.174		1.766	5.019	8.287
2^{12}	0.5351	0.4489	0.4378		1.192	3.962	6.994		2.075	5.541	9.039
2^{14}	0.439	0.4084	0.4053		4.77	14.79	26.82		3.313	8.165	12.96
2^{16}	0.4054	0.4014	0.4		21.98	71.85	112.9		6.473	12.64	18.83

Table 10: Sketching based robust power method: $n = \mathbf{800}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{800}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.4128	0.4018	0.3994		1.701	5.436	9.271		0.0	0.0	0.0
$10n$	0.4023	0.4013	0.4003		2.918	9.075	15.35		0.0	0.0	0.0
$20n$	0.4032	0.3994	0.399		5.336	16.39	27.48		0.0	0.0	0.0
$30n$	0.4009	0.399	0.3992		7.718	23.57	39.46		0.0	0.0	0.0
$40n$	0.3997	0.4002	0.3987		10.09	30.74	51.53		0.0	0.0	0.0

Table 11: Importance sampling based robust power method, without prescanning: $n = \mathbf{800}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{800}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.4132	0.4039	0.4009		1.871	6.031	10.18		0.6598	0.6598	0.6602
$10n$	0.4054	0.4	0.4006		3.209	10.11	16.87		0.6603	0.6598	0.6604
$20n$	0.4072	0.3998	0.3991		5.843	18.06	30.14		0.6611	0.6599	0.6605
$30n$	0.4019	0.3999	0.3992		8.506	26.16	43.3		0.6607	0.6597	0.6599
$40n$	0.4019	0.3992	0.3987		11.1	34.13	56.49		0.661	0.6597	0.66

Table 12: Importance sampling based robust power method with prescanning: $n = \mathbf{800}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method: $n = \mathbf{1000}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	1.01	0.8482	0.5804		0.5285	2.152	3.771		3.244	9.45	15.72
2^{12}	0.7134	0.451	0.4344		1.271	4.239	7.448		3.687	10.37	17.05
2^{14}	0.4374	0.4124	0.4066		4.819	15.13	27.11		5.607	14.83	24.19
2^{16}	0.4072	0.4019	0.4008		22.25	68.96	115.4		9.469	21.59	33.78

Table 13: Sketching based robust power method: $n = \mathbf{1000}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{1000}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.4061	0.402	0.4003		2.151	6.891	12.8		0.0	0.0	0.0
$10n$	0.4043	0.4006	0.4003		3.675	11.42	21.38		0.0	0.0	0.0
$20n$	0.4026	0.4	0.3993		6.659	20.49	38.17		0.0	0.0	0.0
$30n$	0.4014	0.3991	0.3989		9.679	29.39	54.82		0.0	0.0	0.0
$40n$	0.4003	0.3998	0.3988		12.6	38.44	71.96		0.0	0.0	0.0

Table 14: Importance sampling based robust power method, without prescanning: $n = \mathbf{1000}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{1000}$, **inverse** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.4336	0.4017	0.4009		2.963	9.396	16.35		1.293	1.293	1.293
$10n$	0.4029	0.4035	0.3993		5.063	15.95	27.87		1.288	1.288	1.288
$20n$	0.4084	0.4006	0.3992		9.33	29.15	50.35		1.287	1.288	1.288
$30n$	0.4005	0.3995	0.3989		13.58	42.17	72.84		1.288	1.288	1.288
$40n$	0.3995	0.4002	0.3993		17.81	55.11	95.41		1.288	1.288	1.288

Table 15: Importance sampling based robust power method with prescanning: $n = \mathbf{1000}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	1.011	0.9263	0.6358		0.6034	2.407	4.382		5.386	15.96	26.47
2^{12}	1.01	0.4912	0.4451		1.342	4.603	8.012		5.972	17.34	28.72
2^{14}	0.4394	0.4161	0.4085		4.921	15.64	27.93		8.877	24.81	40.87
2^{16}	0.4089	0.4029	0.4006		22.48	69.67	115.2		13.79	34.86	55.65

Table 16: Sketching based robust power method: $n = \mathbf{1200}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning)										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.4075	0.4021	0.4		2.611	8.325	16.34		0.0	0.0	0.0
$10n$	0.4031	0.4007	0.3997		4.764	13.79	27.56		0.0	0.0	0.0
$20n$	0.4046	0.3997	0.3995		8.519	24.61	49.8		0.0	0.0	0.0
$30n$	0.3999	0.3995	0.399		12.32	35.52	71.59		0.0	0.0	0.0
$40n$	0.4009	0.3994	0.3989		16.15	46.25	94.76		0.0	0.0	0.0

Table 17: Importance sampling based robust power method, without prescanning: $n = \mathbf{1200}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning)										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.408	0.4025	0.4005		3.621	11.9	20.49		2.234	2.235	2.241
$10n$	0.4051	0.4003	0.3997		6.302	20.56	35.24		2.225	2.223	2.226
$20n$	0.4057	0.3991	0.3993		11.67	37.89	64.52		2.225	2.226	2.225
$30n$	0.4027	0.3993	0.399		16.99	54.83	93.14		2.225	2.225	2.226
$40n$	0.4024	0.399	0.3988		22.23	72.13	122.2		2.223	2.225	2.225

Table 18: Importance sampling based robust power method with prescanning: $n = \mathbf{1200}$, **inverse** decay, $\|\mathbf{T}\|_F^2$ =1.01

0.1.2 Synthetic tensors with inverse square decaying eigenvalues

The tables shown in this section use synthetic tensors generated with eigenvalues decaying as $\lambda_i = \frac{1}{i^2}$. All tensors have rank $k = 100$ and added noise with $\sigma = 0.01$.

Sketching based robust power method: $n = \mathbf{200}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	0.2519	0.1437	0.1235		0.2295	0.8136	1.372		0.1725	0.2676	0.302
2^{12}	0.1306	0.1019	0.09501		0.9692	2.936	4.843		0.3599	0.4267	0.4962
2^{14}	0.09876	0.08909	0.08836		4.655	13.81	24.36		0.9991	1.053	1.133
2^{16}	0.08781	0.08666	0.08641		21.96	67.03	112.9		3.455	3.539	3.795

Table 19: Sketching based robust power method: $n = \mathbf{200}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{200}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.08857	0.08643	0.08635		0.3984	1.181	2.015		0.0	0.0	0.0
$10n$	0.08684	0.08637	0.08613		0.76	1.849	3.185		0.0	0.0	0.0
$20n$	0.08894	0.08645	0.08621		1.13	3.266	5.445		0.0	0.0	0.0
$30n$	0.08727	0.08675	0.08625		1.747	4.669	7.681		0.0	0.0	0.0
$40n$	0.08632	0.08623	0.08611		2.026	5.921	9.841		0.0	0.0	0.0

Table 20: Importance sampling based robust power method, without prescanning: $n = \mathbf{200}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{200}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.08851	0.08779	0.08628		0.4237	1.176	1.974		0.01051	0.01045	0.01055
$10n$	0.08744	0.08642	0.08639		0.7775	1.897	3.121		0.01049	0.01104	0.01067
$20n$	0.08639	0.08627	0.08612		1.149	3.214	5.329		0.01106	0.01105	0.0104
$30n$	0.08712	0.08627	0.08633		1.623	4.549	7.413		0.01056	0.01061	0.011
$40n$	0.08688	0.08614	0.08612		2.01	5.701	9.498		0.01104	0.011	0.01104

Table 21: Importance sampling based robust power method with prescanning: $n = \mathbf{200}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method: $n = \mathbf{400}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	0.7762	0.2374	0.1531		0.3067	1.106	1.985		0.3803	0.8405	1.307
2^{12}	0.2208	0.1187	0.1039		1.059	3.283	5.669		0.5859	1.06	1.556
2^{14}	0.1082	0.09359	0.09064		4.729	14.22	25.27		1.498	1.915	2.616
2^{16}	0.09005	0.08781	0.08703		22.7	68.36	112.6		3.781	4.68	5.388

Table 22: Sketching based robust power method: $n = \mathbf{400}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{400}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.08747	0.08649	0.0867		0.8425	2.697	4.604		0.0	0.0	0.0
$10n$	0.0869	0.08677	0.08649		1.461	4.5	7.561		0.0	0.0	0.0
$20n$	0.0866	0.08624	0.08615		2.658	8.091	13.64		0.0	0.0	0.0
$30n$	0.08674	0.08621	0.08646		3.878	11.74	19.72		0.0	0.0	0.0
$40n$	0.08653	0.08621	0.08629		5.08	15.35	25.68		0.0	0.0	0.0

Table 23: Importance sampling based robust power method, without prescanning: $n = \mathbf{400}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{400}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.08823	0.08701	0.08656		0.9648	3.031	5.128		0.08254	0.08263	0.08248
$10n$	0.08661	0.08642	0.08618		1.737	5.278	8.939		0.08257	0.0875	0.08293
$20n$	0.08649	0.08621	0.08617		3.17	10.01	16.19		0.08257	0.08908	0.08364
$30n$	0.08651	0.0863	0.08615		4.598	13.88	23.24		0.08649	0.08264	0.08705
$40n$	0.08657	0.08615	0.08621		6.034	18.12	30.21		0.08398	0.08753	0.08257

Table 24: Importance sampling based robust power method with prescanning: $n = \mathbf{400}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method: $n = \mathbf{600}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	1.012	0.3249	0.2027		0.3788	1.45	2.611		0.8633	2.3	3.737
2^{12}	0.3143	0.1376	0.1184		1.11	3.652	6.203		1.09	2.629	4.159
2^{14}	0.1187	0.09597	0.09389		4.676	14.74	26.18		1.959	4.097	6.12
2^{16}	0.09312	0.08877	0.08775		21.94	68.15	114.6		4.709	7.308	9.941

Table 25: Sketching based robust power method: $n = \mathbf{600}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{600}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.08792	0.08656	0.08701		1.236	3.959	6.795		0.0	0.0	0.0
$10n$	0.08675	0.08646	0.08625		2.176	6.754	11.41		0.0	0.0	0.0
$20n$	0.08653	0.08678	0.08619		4.009	12.18	20.63		0.0	0.0	0.0
$30n$	0.08636	0.08628	0.08615		5.832	17.68	29.78		0.0	0.0	0.0
$40n$	0.08647	0.08621	0.08628		7.658	23.1	39.05		0.0	0.0	0.0

Table 26: Importance sampling based robust power method, without prescanning: $n = \mathbf{600}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{600}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.08863	0.08806	0.087		1.381	4.343	7.402		0.2786	0.2781	0.2782
$10n$	0.0934	0.0864	0.08645		2.427	7.477	12.62		0.2787	0.2788	0.2793
$20n$	0.08631	0.08673	0.08619		4.469	13.61	22.91		0.2787	0.2843	0.2872
$30n$	0.08652	0.08623	0.08634		6.528	19.76	33.11		0.2812	0.279	0.2789
$40n$	0.08636	0.0862	0.08618		8.533	25.84	43.17		0.2786	0.2788	0.2787

Table 27: Importance sampling based robust power method with prescanning: $n = \mathbf{600}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method: $n = \mathbf{800}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	1.015	0.6313	0.2574		0.458	1.759	3.191		1.788	5.091	8.291
2^{12}	0.2933	0.1546	0.1283		1.216	3.979	6.805		2.09	5.604	9.063
2^{14}	0.1386	0.1005	0.09644		4.744	15.04	26.33		3.347	8.17	12.84
2^{16}	0.09721	0.08996	0.08854		21.98	69.06	113.7		6.497	12.65	18.82

Table 28: Sketching based robust power method: $n = \mathbf{800}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{800}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.0886	0.0866	0.08621		1.692	5.421	9.288		0.0	0.0	0.0
$10n$	0.08745	0.08647	0.08723		2.903	9.078	15.34		0.0	0.0	0.0
$20n$	0.08682	0.08619	0.08615		5.319	16.28	27.49		0.0	0.0	0.0
$30n$	0.08663	0.08623	0.08613		7.719	23.52	39.52		0.0	0.0	0.0
$40n$	0.0864	0.08619	0.08612		10.09	30.7	51.4		0.0	0.0	0.0

Table 29: Importance sampling based robust power method, without prescanning: $n = \mathbf{800}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{800}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.0867	0.08766	0.08665		1.916	5.962	10.3		0.663	0.6635	0.6625
$10n$	0.08677	0.08645	0.08622		3.298	10.12	17.1		0.6596	0.66	0.6602
$20n$	0.08733	0.08622	0.08624		6.069	18.33	30.89		0.6598	0.662	0.6612
$30n$	0.08744	0.08649	0.08635		8.794	26.48	44.21		0.6596	0.6601	0.6609
$40n$	0.0864	0.08618	0.08618		11.49	34.59	57.64		0.6594	0.6614	0.6607

Table 30: Importance sampling based robust power method with prescanning: $n = \mathbf{800}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method: $n = \mathbf{1000}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	1.01	0.682	0.2942		0.5317	2.076	3.777		3.262	9.702	15.63
2^{12}	1.01	0.1816	0.1407		1.267	4.303	7.406		3.691	10.35	16.96
2^{14}	0.1308	0.1079	0.09943		4.792	15.22	27.12		5.583	14.77	23.9
2^{16}	0.101	0.09097	0.0893		22.26	69.24	115.4		9.425	21.25	33.53

Table 31: Sketching based robust power method: $n = \mathbf{1000}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{1000}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.09244	0.08656	0.08647		2.143	7.278	11.91		0.0	0.0	0.0
$10n$	0.08697	0.08628	0.0862		3.672	11.39	19.35		0.0	0.0	0.0
$20n$	0.08783	0.08624	0.08614		6.656	20.41	35.8		0.0	0.0	0.0
$30n$	0.08712	0.08615	0.08614		9.633	29.51	49.43		0.0	0.0	0.0
$40n$	0.08647	0.08613	0.08614		12.63	38.4	64.38		0.0	0.0	0.0

Table 32: Importance sampling based robust power method, without prescanning: $n = \mathbf{1000}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{1000}$, **inverse-square** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.08776	0.08709	0.08647		2.711	8.991	14.72		1.294	1.293	1.294
$10n$	0.08714	0.08631	0.08622		4.029	12.5	25.2		1.288	1.289	1.288
$20n$	0.08645	0.08622	0.0864		7.323	22.32	45.43		1.29	1.288	1.288
$30n$	0.08831	0.08618	0.08615		10.63	32.27	67.92		1.29	1.288	1.288
$40n$	0.08749	0.08629	0.08622		13.89	42.25	88.09		1.288	1.289	1.289

Table 33: Importance sampling based robust power method with prescanning: $n = \mathbf{1000}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	1.01	1.014	0.5437		0.6114	2.423	4.374		5.361	15.85	26.08
2^{12}	1.02	0.2271	0.1549		1.344	4.563	8.022		5.978	17.23	28.31
2^{14}	0.1513	0.1097	0.1003		4.928	15.51	27.87		8.788	24.72	40.4
2^{16}	0.1065	0.09242	0.08936		22.28	69.7	116.3		13.76	34.74	55.34

Table 34: Sketching based robust power method: $n = \mathbf{1200}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning)										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.08684	0.08637	0.08639		2.595	8.3	15.46		0.0	0.0	0.0
$10n$	0.08784	0.08671	0.08627		4.42	13.68	25.84		0.0	0.0	0.0
$20n$	0.08704	0.087	0.08618		8.02	24.51	46.37		0.0	0.0	0.0
$30n$	0.08697	0.08645	0.08625		11.63	35.35	66.71		0.0	0.0	0.0
$40n$	0.08653	0.08664	0.08611		15.19	46.12	87.24		0.0	0.0	0.0

Table 35: Importance sampling based robust power method, without prescanning: $n = \mathbf{1200}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning)										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	0.08657	0.08684	0.08636		3.1	10.47	18		2.234	2.236	2.234
$10n$	0.08741	0.08677	0.08668		5.427	17.43	30.26		2.232	2.233	2.233
$20n$	0.08648	0.08624	0.08634		9.843	31.42	54.49		2.226	2.226	2.226
$30n$	0.08635	0.08634	0.08615		14.33	45.4	63.85		2.226	2.224	2.227
$40n$	0.08622	0.08652	0.08619		18.68	59.32	82.83		2.225	2.225	2.225

Table 36: Importance sampling based robust power method with prescanning: $n = \mathbf{1200}$, **inverse-square** decay, $\|\mathbf{T}\|_F^2$ =1.01

0.1.3 Synthetic tensors with linearly decaying eigenvalues

The tables shown in this section use synthetic tensors generated with eigenvalues decaying as $\lambda_i = 1 - \frac{i-1}{k}$. All tensors have rank $k = 100$ and added noise $\sigma = 0.01$. Due to the slowly decaying eigenvalues, we expect to see large residual values of rank-1 decomposition, and both algorithms have to work harder to get a better recovery. Our sampling based method works better especially when the dimension is large.

Sketching based robust power method: $n = \mathbf{200}$, **linear** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	1.013	0.9935	0.9874		0.2514	0.7949	1.4		0.1729	0.2378	0.3013
2^{12}	0.9939	0.9834	0.9826		0.966	2.986	4.99		0.3609	0.4264	0.4963
2^{14}	0.9823	0.981	0.9814		4.659	14.06	24.37		0.9906	1.061	1.144
2^{16}	0.9815	0.9806	0.9806		21.75	67.43	112.6		3.435	3.524	3.652

Table 37: Sketching based robust power method: $n = \mathbf{200}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{200}$, **linear** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	1.01	0.9917	0.988		0.4123	1.236	2.069		0.0	0.0	0.0
$10n$	1.007	0.9872	0.9845		0.6577	1.993	3.252		0.0	0.0	0.0
$20n$	0.992	0.9855	0.9838		1.433	3.325	5.687		0.0	0.0	0.0
$30n$	0.9892	0.9836	0.983		1.625	4.627	7.719		0.0	0.0	0.0
$40n$	0.9864	0.9825	0.9817		2.087	5.998	9.874		0.0	0.0	0.0

Table 38: Importance sampling based robust power method, without prescanning: $n = \mathbf{200}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{200}$, **linear** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	1.01	0.9901	0.9871		0.4061	1.309	2.09		0.01056	0.0108	0.01036
$10n$	1.002	0.9855	0.987		0.676	1.973	3.236		0.01044	0.01042	0.01047
$20n$	0.9919	0.984	0.9829		1.199	3.484	5.545		0.0104	0.0106	0.01065
$30n$	0.9896	0.9844	0.9825		1.66	4.705	7.734		0.0104	0.01039	0.01085
$40n$	0.9868	0.9833	0.9818		2.109	6.063	9.984		0.01107	0.01104	0.0104

Table 39: Importance sampling based robust power method with prescanning: $n = \mathbf{200}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method: $n = \mathbf{400}$, **linear** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	1.01	1.01	1.01		0.3087	1.114	2.011		0.3731	0.8392	1.305
2^{12}	1.01	0.9875	0.9844		1.031	3.209	5.73		0.5768	1.066	1.57
2^{14}	0.9845	0.9823	0.9818		4.527	14.29	25.19		1.241	1.901	2.536
2^{16}	0.981	0.9807	0.9806		21.91	68.05	112		3.768	4.572	5.349

Table 40: Sketching based robust power method: $n = \mathbf{400}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{400}$, **linear** decay											
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**		
b \ B	10	30	50		10	30	50		10	30	50
$5n$	1.01	0.9922	0.9876		0.856	2.724	4.642		0.0	0.0	0.0
$10n$	1.01	0.9872	0.9855		1.48	4.541	7.653		0.0	0.0	0.0
$20n$	0.9971	0.9843	0.983		2.702	8.135	13.67		0.0	0.0	0.0
$30n$	0.9916	0.9835	0.9823		3.911	11.82	19.78		0.0	0.0	0.0
$40n$	0.9877	0.9832	0.9825		5.134	15.47	25.95		0.0	0.0	0.0

Table 41: Importance sampling based robust power method, without prescanning: $n = \mathbf{400}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{400}$, **linear** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	1.01	0.9896	0.9844		0.8837	2.814	4.826		0.0827	0.08262	0.08259
$10n$	1.01	0.988	0.9851		1.521	4.699	7.95		0.08738	0.08383	0.08749
$20n$	0.9911	0.9846	0.9835		2.753	8.312	13.94		0.08276	0.08252	0.0825
$30n$	0.9919	0.9853	0.9824		4.051	12.03	20.11		0.08303	0.08284	0.08727
$40n$	0.9882	0.9839	0.9833		5.226	15.71	26.26		0.08256	0.08254	0.08267

Table 42: Importance sampling based robust power method with prescanning: $n = \mathbf{400}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method: $n = \mathbf{600}$, **linear** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	1.01	1.01	1.01		0.3862	1.441	2.578		0.8606	2.295	3.738
2^{12}	1.01	1.01	0.9869		1.113	3.615	6.222		1.104	2.618	4.167
2^{14}	1.01	0.9827	0.9815		4.683	14.71	25.74		1.959	4.043	6.1
2^{16}	0.9816	0.9815	0.9813		22.26	67.67	113.1		4.697	7.281	9.91

Table 43: Sketching based robust power method: $n = \mathbf{600}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{600}$, **linear** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	1.01	0.9936	0.9879		1.248	4	6.833		0.0	0.0	0.0
$10n$	1.006	0.9868	0.9847		2.189	6.75	11.45		0.0	0.0	0.0
$20n$	0.9927	0.9851	0.9836		4.025	12.24	20.7		0.0	0.0	0.0
$30n$	0.9908	0.9838	0.9825		5.843	17.68	29.84		0.0	0.0	0.0
$40n$	0.9898	0.9838	0.9825		7.668	23.21	38.87		0.0	0.0	0.0

Table 44: Importance sampling based robust power method, without prescanning: $n = \mathbf{600}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{600}$, **linear** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	1.01	0.9913	0.9869		1.307	4.167	7.102		0.2782	0.2783	0.2781
$10n$	1.009	0.9895	0.9845		2.249	6.936	11.74		0.2888	0.2803	0.2788
$20n$	0.9939	0.9856	0.9838		4.107	12.53	21.05		0.2789	0.2815	0.2787
$30n$	0.9874	0.9828	0.982		5.931	18.03	30.2		0.279	0.2872	0.2786
$40n$	0.9885	0.9837	0.9818		7.778	23.56	39.45		0.295	0.2788	0.2796

Table 45: Importance sampling based robust power method with prescanning: $n = \mathbf{600}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method: $n = \mathbf{800}$, **linear** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	1.01	1.01	1.01		0.4554	1.761	3.186		1.777	5.046	8.331
2^{12}	1.01	1.01	1.01		1.199	3.94	6.819		2.098	5.598	9.069
2^{14}	1.01	0.9844	0.9826		4.795	14.91	26.39		3.336	8.102	12.88
2^{16}	0.9824	0.9811	0.9809		22.02	68.19	113.5		6.466	12.66	18.85

Table 46: Sketching based robust power method: $n = \mathbf{800}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{800}$, **linear** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	1.01	0.9916	0.9868		1.708	5.435	9.608		0.0	0.0	0.0
$10n$	1.01	0.9892	0.9849		2.94	9.08	15.82		0.0	0.0	0.0
$20n$	0.992	0.9853	0.9835		5.331	16.4	28.25		0.0	0.0	0.0
$30n$	0.9894	0.984	0.9836		7.747	23.61	40.45		0.0	0.0	0.0
$40n$	0.9881	0.9826	0.9819		10.15	30.74	52.89		0.0	0.0	0.0

Table 47: Importance sampling based robust power method, without prescanning: $n = \mathbf{800}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{800}$, **linear** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	1.01	0.9902	0.986		1.859	6.067	10.67		0.663	0.6598	0.6613
$10n$	1.01	0.9882	0.9856		3.164	10.06	17.79		0.6595	0.6594	0.6596
$20n$	0.9897	0.9846	0.9833		5.725	18.18	31.41		0.6597	0.6596	0.6599
$30n$	0.9915	0.9855	0.9825		8.265	25.93	44.79		0.6596	0.6594	0.6595
$40n$	0.9885	0.9831	0.9824		10.82	33.58	58.05		0.6598	0.6594	0.6594

Table 48: Importance sampling based robust power method with prescanning: $n = \mathbf{800}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method: $n = \mathbf{1000}$, **linear** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	1.01	1.01	1.01		0.5316	2.096	3.822		3.249	9.442	15.59
2^{12}	1.01	1.01	1.01		1.327	4.31	7.386		3.664	10.29	16.94
2^{14}	1.01	1.01	0.9829		4.807	15.37	27.51		5.616	14.77	23.96
2^{16}	0.9828	0.9817	0.9809		22.36	68.58	115.4		9.395	21.4	33.48

Table 49: Sketching based robust power method: $n = \mathbf{1000}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning): $n = \mathbf{1000}$, **linear** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	1.01	0.9905	0.9883		2.161	6.938	13.24		0.0	0.0	0.0
$10n$	1.01	0.9886	0.9844		3.694	11.48	22.15		0.0	0.0	0.0
$20n$	0.9923	0.9843	0.9838		6.674	20.45	39.15		0.0	0.0	0.0
$30n$	0.9931	0.9843	0.9823		9.683	29.55	56.41		0.0	0.0	0.0
$40n$	0.9878	0.9824	0.9819		12.63	38.43	73.41		0.0	0.0	0.0

Table 50: Importance sampling based robust power method, without prescanning: $n = \mathbf{1000}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning): $n = \mathbf{1000}$, **linear** decay										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	1.01	0.9929	0.9879		2.942	8.791	15.01		1.294	1.294	1.294
$10n$	1.005	0.9881	0.9863		4.504	14.88	25.29		1.288	1.288	1.288
$20n$	0.9936	0.9842	0.9832		8.171	26.67	45.35		1.288	1.288	1.288
$30n$	0.9918	0.9836	0.9829		11.89	38.5	65.5		1.288	1.288	1.289
$40n$	0.9874	0.9832	0.9823		15.39	50.23	85.27		1.288	1.289	1.288

Table 51: Importance sampling based robust power method with prescanning: $n = \mathbf{1000}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Sketching based robust power method										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
2^{10}	1.01	1.01	1.01		0.6153	2.457	4.381		5.364	15.78	26.25
2^{12}	1.01	1.01	1.01		1.356	4.704	7.994		5.975	17.24	28.48
2^{14}	1.01	1.01	0.9829		4.942	15.84	27.82		8.851	24.62	40.5
2^{16}	0.9835	0.9813	0.9811		22.66	72.01	115.1		13.75	35.13	55.59

Table 52: Sketching based robust power method: $n = \mathbf{1200}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (without prescanning)										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	1.01	0.9884	0.9879		2.627	8.699	17.61		0.0	0.0	0.0
$10n$	1.005	0.9881	0.9865		4.419	14.43	29.98		0.0	0.0	0.0
$20n$	0.9924	0.9853	0.9847		8.05	25.88	53.99		0.0	0.0	0.0
$30n$	0.9899	0.984	0.982		11.67	37.32	77.92		0.0	0.0	0.0
$40n$	0.9878	0.9831	0.9822		15.26	48.55	102		0.0	0.0	0.0

Table 53: Importance sampling based robust power method, without prescanning: $n = \mathbf{1200}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

Importance sampling based robust power method (with prescanning)										
	Squared residual norm				**Running time (s)**				**Preprocessing time (s)**	
b \ B	10	30	50		10	30	50		10	30	50
$5n$	1.01	0.991	0.9882		3.668	11.64	19.48		2.234	2.234	2.234
$10n$	1.009	0.9883	0.9851		6.512	20.14	33.45		2.225	2.225	2.226
$20n$	0.9918	0.9852	0.9828		11.9	36.43	61.01		2.225	2.225	2.226
$30n$	0.9873	0.9849	0.9826		17.14	52.89	87.36		2.226	2.226	2.226
$40n$	0.9882	0.9829	0.9822		22.38	69.01	114.6		2.226	2.227	2.226

Table 54: Importance sampling based robust power method with prescanning: $n = \mathbf{1200}$, **linear** decay, $\|\mathbf{T}\|_F^2$ =1.01

0.1.4 Results from real-life datasets

Dataset	# Documents	# Vocabulary	# Words	Short Description
Wiki	114274	10000	58508288	Wikipedia articles, preprocessed by authors of [WTSA15]
Enron [1]	186501	10000	16971591	Enron Email Dataset, preprocessed by authors of [WTSA15]
AP [2]	2246	10473	435838	Associated Press [BNJ03]
NIPS [3]	1500	12419	1900000	NIPS full papers, preprocessed by UCI
KOS [4]	3430	6906	467714	KOS blog, preprocessed by UCI
NSF [5]	49078	22170	4406190	NSF research award abstracts 1990-2003

Table 55: List of six real-life datasets

We used the two same datasets as the previous work [WTSA15]: Wiki and Enron, as well as four additional real-life datasets, shown in Table 55. We use the program from the authors of [TWZS15] to build a $K \times K \times K$ tensor $\mathbf{T}_{\text{LDA}}$ for each dataset. Due to limited time our experiments only use $K = 200$. We expect to gain more over the sketching-based method on larger tensors.

Since these tensors comes from real-world applications and are not normalized, their Frobenius norm is usually very large.

Sketching based robust power method: dataset **Enron**							
	Squared residual norm			**Running time (s)**		**Preprocessing time (s)**	
b \ B	10	30		10	30	10	30
2^{10}	4.733e+07	4.608e+07		0.3667	0.7942	0.2429	0.2378
2^{11}	4.677e+07	4.58e+07		0.4434	1.486	0.2468	0.3238
2^{12}	4.581e+07	4.559e+07		0.995	3.243	0.3647	0.6283
2^{13}	4.568e+07	4.546e+07		2.023	6.017	0.5745	0.6458
2^{14}	4.556e+07	4.542e+07		4.561	13.97	0.9698	1.061

Table 56: Sketching based robust power method: dataset **Enron**, $\|\mathbf{T}\|_F^2$ =4.96e+07

Importance sampling based robust power method (without prescanning): dataset **Enron**							
	Squared residual norm			**Running time (s)**		**Preprocessing time (s)**	
b \ B	10	30		10	30	10	30
$5n$	4.579e+07	4.557e+07		0.3784	1.139	0.0	0.0
$10n$	4.563e+07	4.559e+07		0.5689	1.603	0.0	0.0
$20n$	4.56e+07	4.545e+07		1.003	2.706	0.0	0.0
$30n$	4.555e+07	4.543e+07		1.362	3.717	0.0	0.0
$40n$	4.564e+07	4.544e+07		1.693	4.753	0.0	0.0

Table 57: Importance sampling based robust power method, without prescanning: dataset **Enron**, $\|\mathbf{T}\|_F^2$ =4.96e+07

Importance sampling based robust power method (with prescanning): dataset **Enron**							
	Squared residual norm			**Running time (s)**		**Preprocessing time (s)**	
b \ B	10	30		10	30	10	30
$5n$	4.605e+07	4.545e+07		0.4014	1.176	0.0105	0.01043
$10n$	4.573e+07	4.55e+07		0.6255	1.773	0.01036	0.01044
$20n$	4.561e+07	4.545e+07		1.318	2.923	0.01041	0.01039
$30n$	4.563e+07	4.546e+07		1.459	4.041	0.01101	0.01104
$40n$	4.56e+07	4.544e+07		1.841	5.092	0.01039	0.01102

Table 58: Importance sampling based robust power method with prescanning: dataset **Enron**, $\|\mathbf{T}\|_F^2$ =4.96e+07

[1] `http://www.cs.cmu.edu/~enron/`
[2] `http://www.cs.princeton.edu/~blei/lda-c/`
[3] `https://archive.ics.uci.edu/ml/machine-learning-databases/bag-of-words`
[4] `https://archive.ics.uci.edu/ml/machine-learning-databases/bag-of-words`
[5] `http://archive.ics.uci.edu/ml/datasets/NSF+Research+Award+Abstracts+1990-2003`

Sketching based robust power method: dataset **AP**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
2^{10}	7.065e+06	6.703e+06		0.2316	0.8128		0.172	0.2472
2^{11}	6.674e+06	6.613e+06		0.4404	1.417		0.2419	0.3072
2^{12}	6.656e+06	6.57e+06		0.9689	2.936		0.3586	0.4257
2^{13}	6.585e+06	6.573e+06		2.058	6.172		0.5826	0.6592
2^{14}	6.57e+06	6.555e+06		4.558	13.99		0.9601	1.059

Table 59: Sketching based robust power method: dataset **AP**, $\|\mathbf{T}\|_F^2$ =6.905e+06

Importance sampling based robust power method (without prescanning): dataset **AP**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
$5n$	6.859e+06	6.805e+06		0.3762	1.548		0.0	0.0
$10n$	6.631e+06	6.904e+06		0.5469	1.577		0.0	0.0
$20n$	6.572e+06	6.72e+06		0.9535	2.587		0.0	0.0
$30n$	6.64e+06	6.557e+06		1.263	3.53		0.0	0.0
$40n$	6.565e+06	6.564e+06		1.613	4.478		0.0	0.0

Table 60: Importance sampling based robust power method, without prescanning: dataset **AP**, $\|\mathbf{T}\|_F^2$ =6.905e+06

Importance sampling based robust power method (with prescanning): dataset **AP**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
$5n$	6.58e+06	6.893e+06		0.4007	1.233		0.01045	0.01103
$10n$	6.604e+06	6.567e+06		0.7035	2.011		0.01061	0.01051
$20n$	6.803e+06	6.563e+06		1.166	3.246		0.01042	0.01061
$30n$	6.58e+06	6.573e+06		1.588	4.815		0.01103	0.01122
$40n$	6.66e+06	6.555e+06		1.989	5.601		0.01084	0.01103

Table 61: Importance sampling based robust power method with prescanning: dataset **AP**, $\|\mathbf{T}\|_F^2$ =6.905e+06

Sketching based robust power method: dataset **wiki**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
2^{10}	2.091e+07	1.951e+07		0.2346	0.8749		0.1727	0.2535
2^{11}	1.971e+07	1.938e+07		0.4354	1.439		0.2408	0.3167
2^{12}	1.947e+07	1.93e+07		1.035	2.912		0.4226	0.4275
2^{13}	1.931e+07	1.927e+07		2.04	5.94		0.5783	0.6493
2^{14}	1.928e+07	1.926e+07		4.577	13.93		1.045	1.121

Table 62: Sketching based robust power method: dataset **wiki**, $\|\mathbf{T}\|_F^2$ =2.135e+07

Importance sampling based robust power method (without prescanning): dataset **wiki**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
$5n$	1.931e+07	1.928e+07		0.3698	1.146		0.0	0.0
$10n$	1.931e+07	1.929e+07		0.5623	1.623		0.0	0.0
$20n$	1.935e+07	1.926e+07		0.9767	2.729		0.0	0.0
$30n$	1.929e+07	1.926e+07		1.286	3.699		0.0	0.0
$40n$	1.928e+07	1.925e+07		1.692	4.552		0.0	0.0

Table 63: Importance sampling based robust power method, without prescanning: dataset **wiki**, $\|\mathbf{T}\|_F^2$ =2.135e+07

Importance sampling based robust power method (with prescanning): dataset **wiki**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
$5n$	1.931e+07	1.93e+07		0.4376	1.168		0.01038	0.01103
$10n$	1.928e+07	1.93e+07		0.6357	1.8		0.0104	0.01044
$20n$	1.931e+07	1.927e+07		1.083	2.962		0.01102	0.01042
$30n$	1.929e+07	1.925e+07		1.457	4.049		0.01102	0.01043
$40n$	1.929e+07	1.925e+07		1.905	5.246		0.01105	0.01105

Table 64: Importance sampling based robust power method with prescanning: dataset **wiki**, $\|\mathbf{T}\|_F^2$ =2.135e+07

Sketching based robust power method: dataset **NSF**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
2^{10}	4.342e+06	3.983e+06		0.2329	1.118		0.1731	0.2938
2^{11}	4.183e+06	3.906e+06		0.4415	1.434		0.2439	0.3061
2^{12}	3.929e+06	3.869e+06		0.9758	2.952		0.3554	0.4268
2^{13}	3.885e+06	3.86e+06		2.045	6.015		0.6591	0.6489
2^{14}	3.865e+06	3.854e+06		4.591	13.89		1.081	1.047

Table 65: Sketching based robust power method: dataset **NSF**, $\|\mathbf{T}\|_F^2$ =4.765e+06

Importance sampling based robust power method (without prescanning): dataset **NSF**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
$5n$	3.899e+06	3.885e+06		0.3701	1.166		0.0	0.0
$10n$	3.891e+06	3.862e+06		0.5678	1.698		0.0	0.0
$20n$	3.896e+06	3.861e+06		1.029	2.927		0.0	0.0
$30n$	3.889e+06	3.856e+06		1.415	3.941		0.0	0.0
$40n$	3.867e+06	3.857e+06		1.823	5.287		0.0	0.0

Table 66: Importance sampling based robust power method, without prescanning: dataset **NSF**, $\|\mathbf{T}\|_F^2$ =4.765e+06

Importance sampling based robust power method (with prescanning): dataset **NSF**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
$5n$	3.938e+06	3.902e+06		0.5219	1.566		0.01039	0.01108
$10n$	3.898e+06	3.867e+06		0.8981	2.568		0.01104	0.01043
$20n$	3.947e+06	3.864e+06		1.561	4.546		0.01097	0.01057
$30n$	3.889e+06	3.852e+06		2.207	6.477		0.011	0.01107
$40n$	3.891e+06	3.853e+06		2.827	8.446		0.01103	0.01065

Table 67: Importance sampling based robust power method with prescanning: dataset **NSF**, $\|\mathbf{T}\|_F^2$ =4.765e+06

Sketching based robust power method: dataset **NIPS**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
2^{10}	5.649e+09	5.432e+09		0.2383	0.8686		0.1724	0.2625
2^{11}	5.522e+09	5.383e+09		0.45	1.52		0.2417	0.3247
2^{12}	5.394e+09	5.366e+09		1.264	3.304		0.5082	0.6358
2^{13}	5.375e+09	5.353e+09		2.07	6.068		0.5766	0.7604
2^{14}	5.36e+09	5.35e+09		4.591	13.84		0.9817	1.044

Table 68: Sketching based robust power method: dataset **NIPS**, $\|\mathbf{T}\|_F^2$ =6.058e+09

Importance sampling based robust power method (without prescanning): dataset **NIPS**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
$5n$	5.648e+09	5.403e+09		0.3205	0.9938		0.0	0.0
$10n$	5.417e+09	5.347e+09		0.4739	1.753		0.0	0.0
$20n$	5.352e+09	5.349e+09		0.8345	2.464		0.0	0.0
$30n$	5.361e+09	5.35e+09		1.19	3.317		0.0	0.0
$40n$	5.352e+09	5.349e+09		1.532	4.153		0.0	0.0

Table 69: Importance sampling based robust power method, without prescanning: dataset **NIPS**, $\|\mathbf{T}\|_F^2$ =6.058e+09

Importance sampling based robust power method (with prescanning): dataset **NIPS**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
$5n$	6.019e+09	5.943e+09		0.4818	1.401		0.01039	0.01057
$10n$	5.352e+09	5.389e+09		0.797	2.197		0.01063	0.01105
$20n$	5.387e+09	5.349e+09		1.335	3.849		0.01042	0.01039
$30n$	5.363e+09	5.346e+09		1.895	5.491		0.0104	0.01042
$40n$	5.352e+09	5.382e+09		2.523	7.019		0.01064	0.011

Table 70: Importance sampling based robust power method with prescanning: dataset **NIPS**, $\|\mathbf{T}\|_F^2$ =6.058e+09

Sketching based robust power method: dataset **KOS**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
2^{10}	5.015e+04	4.807e+04		0.2356	0.8581		0.1781	0.2507
2^{11}	4.858e+04	4.787e+04		0.4482	1.417		0.2411	0.3072
2^{12}	4.797e+04	4.773e+04		0.9652	2.947		0.3556	0.4265
2^{13}	4.776e+04	4.764e+04		2.046	5.985		0.5722	0.6517
2^{14}	4.771e+04	4.761e+04		4.453	13.87		0.968	1.052

Table 71: Sketching based robust power method: dataset **KOS**, $\|\mathbf{T}\|_F^2$ =5.016e+04

Importance sampling based robust power method (without prescanning): dataset **KOS**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
$5n$	4.783e+04	4.766e+04		0.3968	1.181		0.0	0.0
$10n$	4.817e+04	4.767e+04		0.6259	1.704		0.0	0.0
$20n$	4.775e+04	4.766e+04		1.039	2.79		0.0	0.0
$30n$	4.771e+04	4.762e+04		1.357	3.57		0.0	0.0
$40n$	4.767e+04	4.762e+04		1.76	4.892		0.0	0.0

Table 72: Importance sampling based robust power method, without prescanning: dataset **KOS**, $\|\mathbf{T}\|_F^2$ =5.016e+04

Importance sampling based robust power method (with prescanning): dataset **KOS**								
	Squared residual norm			**Running time (s)**			**Preprocessing time (s)**	
b \ B	10	30		10	30		10	30
$5n$	4.773e+04	4.766e+04		0.4808	1.35		0.01061	0.01099
$10n$	4.792e+04	4.765e+04		0.7363	2.024		0.01108	0.01044
$20n$	4.795e+04	4.766e+04		1.336	3.308		0.01056	0.01104
$30n$	4.785e+04	4.765e+04		1.728	4.631		0.01039	0.0104
$40n$	4.767e+04	4.764e+04		2.295	5.773		0.01063	0.01102

Table 73: Importance sampling based robust power method with prescanning: dataset **KOS**, $\|\mathbf{T}\|_F^2$ =5.016e+04

Bibliography

[BNJ03] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *JMLR*, 3:993–1022, 2003.

[TWZS15] Hsiao-Yu Fish Tung, Chao-Yuan Wu, Manzil Zaheer, and Alexander J Smola. Spectral methods for the hierarchical dirichlet process. 2015.

[WTSA15] Yining Wang, Hsiao-Yu Tung, Alex J Smola, and Anima Anandkumar. Fast and guaranteed tensor decomposition via sketching. In *NIPS*, pages 991–999, 2015.