Journal of Machine Learning Research XX (2024) 1-14 Submitted 4/24; Revised X/XX; Published X/XX

Online Appendix for 'Optimal Distributed Prediction in Massive Data'

Jun Liao JUNLIAO@RUC.EDU.CN
Center for Applied Statistics and School of Statistics
Renmin University of China
Beijing, 100872, China

Rui Liu RAYLIU@MAIL.USTC.EDU.CN
School of Management
University of Science and Technology of China
Hefei, 230026, China

Xinyu Zhang* XINYU@AMSS.AC.CN
Academy of Mathematics and Systems Science
Chinese Academy of Sciences
Beijing, 100190, China
* *Corresponding author*

Liping Zhu ZHU.LIPING@RUC.EDU.CN
Center for Applied Statistics and Institute of Statistics and Big Data
Renmin University of China
Beijing, 100872, China

Editor: My editor

Online Appendix A. More Simulation Results

This part includes Tables S1-S4 and some additional tables and figures:

- We add the case of $n_m = 10000$ and $M = 150$ for the least squares estimation situation and the simulation results are shown in Tables S5-S7.
- Tables S8-S10 provide a richer result about the SPEs of mehtod CSL corresponding to the coefficient estimators $\hat{\boldsymbol{\beta}}^{(a)}_{CSL,1}$, $\hat{\boldsymbol{\beta}}^{(a)}_{CSL,2}$, $\hat{\boldsymbol{\beta}}^{(a)}_{CSL,3}$, $\hat{\boldsymbol{\beta}}^{(c)}_{CSL,1}$, $\hat{\boldsymbol{\beta}}^{(c)}_{CSL,2}$ and $\hat{\boldsymbol{\beta}}^{(c)}_{CSL,3}$, which are the complements to the CSL columns listed in Tables 1, S1 and S2 (Overall, it appears that the CSL estimator using the average initial value $\bar{\boldsymbol{\beta}}$ performs better than those with the central site's initial value $\hat{\boldsymbol{\beta}}^{local}_{\lfloor M/2 \rfloor}$. Also, as the number of iterations increases, the SPE decreases only when the sample size of local sites is large enough (e.g., $n_m = 1000$). This phenomenon is more pronounced for the estimators with the central site's initial value. And it seems that iterations do not help with decreasing SPE if $\bar{\boldsymbol{\beta}}$ is set as the initial value, which is most likely attributable to the variance reduction effect of averaging that leads $\bar{\boldsymbol{\beta}}$ a good initial value that is closer to the true parameter relative to $\hat{\boldsymbol{\beta}}^{local}_{\lfloor M/2 \rfloor}$).
- Tables S11 and S12 present the computing time of five methods per replication in the situations of OLS and lasso estimation with $\alpha = 0.5$. When $\alpha = 1$ and $\alpha = 1.5$, the results are similar so we do not show them for brevity.
- Figures S1-S5 are the complements of Figure 2, which display the situations with $(q, \alpha) = (15, 0.5)$, $(q, \alpha) = (30, 0.5)$, $(q, \alpha) = (15, 1)$, $(q, \alpha) = (15, 1.5)$ and $(q, \alpha) = (30, 1.5)$, respectively.

σ_m^2	n_m	M	PD	DC_{ew}	$Full_l$	$Full_A$	$CSL_1^{(a)}$	DC_{opt}	PD	DC_{ew}	$Full_l$	$Full_A$	$CSL_1^{(a)}$	DC_{opt}
			$q = 15$						$q = 30$					
0.5	50	10	27.606	7.798	**6.182**	*6.285*	7.022	7.008	50.074	7.233	**4.580**	*4.957*	95.990	6.013
	50	100	25.570	5.781	**4.622**	**4.622**	4.723	*4.623*	52.968	2.774	**1.596**	*1.597*	9.454	1.600
	100	10	14.684	6.000	**5.286**	*5.298*	5.322	5.590	18.850	4.199	**2.830**	*2.989*	3.599	3.885
	100	100	15.570	5.103	**4.553**	**4.553**	*4.556*	**4.553**	17.886	2.803	**1.453**	**1.453**	1.526	*1.506*
	1000	10	5.336	4.542	**4.538**	**4.538**	**4.538**	*4.539*	3.042	1.488	**1.451**	**1.451**	**1.451**	*1.471*
	1000	100	*5.309*	**4.479**	**4.479**	**4.479**	**4.479**	**4.479**	2.991	*1.339*	**1.314**	**1.314**	**1.314**	**1.314**
1	50	10	42.341	10.414	**7.754**	*8.176*	9.342	9.216	79.706	11.390	**7.781**	*7.865*	189.387	9.027
	50	100	37.149	7.078	**4.764**	**4.764**	4.959	*4.777*	85.899	4.118	**1.891**	*1.903*	17.453	1.914
	100	10	22.215	7.662	**6.021**	6.151	*6.092*	6.871	29.563	6.397	**4.316**	*4.661*	5.826	5.792
	100	100	23.820	6.051	**4.633**	**4.633**	*4.639*	**4.633**	26.580	4.221	**1.604**	*1.605*	1.746	1.885
	1000	10	6.174	4.639	**4.615**	**4.615**	**4.615**	*4.624*	4.676	1.777	**1.601**	1.603	*1.602*	1.700
	1000	100	6.189	*4.496*	**4.486**	**4.486**	**4.486**	**4.486**	4.754	*1.478*	**1.329**	**1.329**	**1.329**	**1.329**
$0.5+\frac{m}{M}$	50	10	53.598	10.426	**7.783**	*8.120*	9.640	9.115	104.455	11.658	*8.112*	**8.020**	196.043	9.013
	50	100	45.448	6.977	*4.760*	**4.758**	4.963	4.774	112.989	3.972	*1.908*	1.910	18.255	**1.904**
	100	10	28.559	7.653	**6.031**	6.172	*6.162*	6.792	38.548	6.386	**4.354**	*4.694*	6.041	5.610
	100	100	30.890	5.986	**4.624**	**4.624**	4.639	*4.631*	32.556	4.118	**1.590**	*1.592*	1.745	1.838
	1000	10	7.129	4.653	**4.610**	**4.610**	4.623	*4.622*	6.261	1.785	**1.591**	*1.593*	1.617	1.672
	1000	100	7.073	4.497	**4.485**	**4.485**	*4.486*	*4.486*	6.245	1.471	**1.326**	**1.326**	1.329	*1.328*

Table S1: SPEs of various methods based on the OLS estimation with $\alpha = 0.5$ ($\times 10^{-2}$). The SPEs for the best-performing and the second-best methods are highlighted in bold and italics.

σ_m^2	n_m	M	PD	DC_{ew}	$Full_l$	$Full_A$	$CSL_1^{(a)}$	DC_{opt}	PD	DC_{ew}	$Full_l$	$Full_A$	$CSL_1^{(a)}$	DC_{opt}
			$q = 15$						$q = 30$					
0.5	50	10	11.174	1.455	1.591	*1.162*	2.323	**1.140**	21.081	2.747	3.204	**1.209**	94.209	*1.361*
	50	100	9.126	0.456	*0.170*	0.184	0.261	**0.160**	29.074	0.498	0.299	**0.199**	8.052	*0.225*
	100	10	4.894	0.786	0.756	*0.692*	0.792	**0.670**	6.554	0.851	1.488	**0.699**	2.224	*0.707*
	100	100	5.009	0.321	**0.106**	0.113	*0.109*	*0.109*	5.471	0.309	0.154	*0.119*	0.223	**0.114**
	1000	10	0.684	0.139	**0.105**	*0.116*	**0.105**	0.128	0.697	0.136	0.153	**0.123**	0.153	*0.127*
	1000	100	0.682	0.080	**0.034**	**0.034**	**0.034**	*0.035*	0.695	0.073	**0.018**	*0.019*	**0.018**	0.026
1	50	10	20.676	2.468	3.154	*2.107*	4.617	**1.951**	40.794	4.988	6.406	**2.107**	188.468	*2.356*
	50	100	15.643	0.726	0.314	*0.313*	0.496	**0.278**	54.278	0.886	0.595	**0.342**	16.104	*0.418*
	100	10	8.247	1.293	1.485	**1.148**	1.556	*1.152*	11.812	1.439	2.974	**1.168**	4.444	*1.215*
	100	100	8.967	0.504	*0.185*	0.194	0.192	**0.174**	9.820	0.493	0.305	*0.208*	0.442	**0.193**
	1000	10	1.206	0.223	**0.184**	*0.197*	**0.184**	0.207	1.206	0.222	0.304	*0.210*	0.304	**0.209**
	1000	100	1.181	0.125	**0.042**	**0.042**	**0.042**	*0.047*	1.181	0.118	**0.033**	*0.034*	**0.033**	0.041
$0.5+\frac{m}{M}$	50	10	29.117	2.574	3.195	*2.139*	4.914	**1.964**	57.463	5.067	6.691	**2.306**	194.763	*2.373*
	50	100	22.778	0.714	*0.304*	*0.304*	0.502	**0.280**	75.233	0.884	0.608	**0.348**	16.876	*0.418*
	100	10	11.666	1.335	1.491	*1.164*	1.631	**1.154**	17.243	1.487	3.006	*1.232*	4.661	**1.220**
	100	100	12.597	0.494	*0.175*	0.182	0.192	**0.173**	13.353	0.485	0.290	*0.197*	0.442	**0.192**
	1000	10	1.634	0.227	**0.179**	*0.192*	*0.192*	0.204	1.634	0.226	0.294	**0.204**	0.320	*0.206*
	1000	100	1.622	0.122	**0.040**	*0.041*	0.042	0.046	1.622	0.115	**0.030**	*0.031*	0.033	0.039

Table S2: SPEs of various methods based on the OLS estimation with $\alpha = 1.5$ ($\times 10^{-2}$). The SPEs for the best-performing and the second-best methods are highlighted in bold and italics.

σ_m^2	n_m	M	PD	DC_{ew}	$Full$	$CSL_{1,cv}^{(c)}$	$CSL_{1,o}^{(c)}$	DC_{opt}	PD	DC_{ew}	$Full$	$CSL_{1,cv}^{(c)}$	$CSL_{1,o}^{(c)}$	DC_{opt}
			$q = 15$						$q = 30$					
0.5	50	10	24.062	9.627	**6.176**	20.685	20.058	*7.510*	32.530	13.933	**4.579**	25.373	31.362	*9.562*
	50	100	25.330	8.002	*4.631*	20.905	21.680	**4.627**	34.221	11.772	**1.615**	24.443	28.837	*2.512*
	100	10	14.595	6.059	**5.329**	9.323	9.272	*5.597*	18.608	6.178	**2.946**	11.991	11.592	*4.658*
	100	100	14.311	5.218	*4.550*	8.725	8.401	**4.544**	17.208	5.159	**1.459**	11.802	10.775	*1.708*
	1000	10	5.276	4.551	*4.549*	4.579	4.555	**4.547**	2.916	1.471	**1.458**	1.531	1.509	*1.466*
	1000	100	5.290	4.484	*4.483*	4.511	4.492	**4.476**	2.855	1.332	*1.321*	1.399	1.378	**1.314**
1	50	10	36.137	15.350	**7.727**	25.535	24.559	*10.635*	49.607	23.785	**7.389**	30.925	41.762	*15.565*
	50	100	40.047	12.741	**4.786**	26.469	25.919	*4.796*	52.907	20.124	**1.914**	27.366	33.495	*3.669*
	100	10	22.830	8.261	**6.105**	12.413	11.674	*6.874*	29.793	11.394	**4.508**	15.644	14.314	*7.987*
	100	100	21.785	6.924	*4.625*	11.430	10.593	**4.620**	27.304	9.721	**1.610**	14.388	12.894	*2.357*
	1000	10	6.026	*4.630*	**4.624**	4.689	4.641	**4.624**	4.444	1.692	**1.608**	1.759	1.716	*1.664*
	1000	100	6.031	4.494	*4.491*	4.552	4.513	**4.484**	4.347	1.430	*1.336*	1.493	1.453	**1.330**
$0.5{+}\frac{m}{M}$	50	10	48.760	15.735	**7.872**	25.540	24.520	*10.425*	61.789	24.092	**7.628**	31.147	41.901	*15.253*
	50	100	50.887	12.496	**4.788**	26.487	25.829	*4.793*	68.258	19.572	**1.917**	27.456	33.365	*3.516*
	100	10	30.679	8.428	**6.193**	12.520	11.793	*6.855*	38.824	11.553	**4.676**	15.695	14.494	*7.735*
	100	100	28.197	6.863	*4.627*	11.425	10.604	**4.619**	35.367	9.474	**1.614**	14.302	12.921	*2.221*
	1000	10	6.794	4.638	*4.630*	4.696	4.648	**4.626**	5.915	1.719	**1.624**	1.777	1.734	*1.660*
	1000	100	6.759	4.495	*4.491*	4.552	4.513	**4.484**	5.743	1.433	*1.336*	1.494	1.453	**1.329**

Table S3: SPEs of various methods based on the lasso estimation with $\alpha = 0.5$ ($\times 10^{-2}$). The SPEs for the best-performing and the second-best methods are highlighted in bold and italics.

σ_m^2	n_m	M	PD	DC_{ew}	$Full$	$CSL_{1,cv}^{(c)}$	$CSL_{1,o}^{(c)}$	DC_{opt}	PD	DC_{ew}	$Full$	$CSL_{1,cv}^{(c)}$	$CSL_{1,o}^{(c)}$	DC_{opt}
			$q = 15$						$q = 30$					
0.5	50	10	11.464	4.416	**1.379**	4.425	4.149	*2.293*	16.169	7.066	**1.903**	5.662	10.174	*4.061*
	50	100	12.355	3.459	**0.182**	4.111	5.437	*0.193*	16.757	6.002	**0.245**	4.540	6.364	*0.673*
	100	10	6.387	1.980	**0.742**	1.896	1.555	*1.122*	8.417	3.546	**1.042**	2.296	*2.105*	2.136
	100	100	5.893	1.628	*0.110*	1.569	1.274	**0.105**	7.888	3.093	**0.132**	1.785	1.606	*0.397*
	1000	10	0.697	0.171	**0.111**	0.149	*0.113*	0.128	0.992	0.370	**0.135**	0.193	*0.155*	0.260
	1000	100	0.681	0.116	0.049	0.072	*0.046*	**0.034**	0.958	0.306	**0.029**	0.079	*0.045*	0.048
1	50	10	20.353	8.347	**2.550**	7.160	6.610	*4.034*	27.765	12.708	**3.360**	9.116	18.299	*6.968*
	50	100	21.804	6.714	**0.324**	6.197	9.207	*0.349*	29.246	10.907	**0.444**	6.550	8.682	*1.144*
	100	10	11.247	3.883	**1.348**	3.163	2.504	*2.080*	14.347	6.487	**1.830**	3.663	*3.395*	3.767
	100	100	10.539	3.270	**0.179**	2.454	1.999	*0.190*	13.713	5.714	**0.248**	2.794	2.363	*0.686*
	1000	10	1.313	0.349	**0.180**	0.267	*0.194*	0.239	1.788	0.725	**0.251**	0.345	*0.279*	0.490
	1000	100	1.271	0.251	*0.056*	0.113	0.064	**0.042**	1.744	0.618	**0.039**	0.142	*0.076*	0.089
$0.5{+}\frac{m}{M}$	50	10	27.184	8.658	**2.645**	7.258	6.689	*3.992*	38.536	12.925	**3.542**	9.152	18.474	*6.656*
	50	100	29.879	6.589	**0.326**	6.214	9.155	*0.351*	40.294	10.681	**0.446**	6.604	8.632	*1.104*
	100	10	15.611	3.889	**1.406**	3.212	2.563	*2.030*	19.710	6.552	**1.931**	3.728	*3.470*	3.683
	100	100	14.414	3.191	**0.180**	2.452	1.996	*0.186*	19.152	5.595	**0.248**	2.784	2.360	*0.651*
	1000	10	1.873	0.362	**0.185**	0.268	*0.200*	0.237	2.537	0.745	**0.262**	0.364	*0.289*	0.461
	1000	100	1.860	0.247	*0.057*	0.113	0.064	**0.042**	2.457	0.604	**0.039**	0.143	*0.076*	0.083

Table S4: SPEs of various methods based on the lasso estimation with $\alpha = 1.5$ ($\times 10^{-2}$). The SPEs for the best-performing and the second-best methods are highlighted in bold and italics.

σ_m^2	n_m	M	PD	DC_{ew}	$Full_l$	$Full_A$	$CSL_1^{(a)}$	DC_{opt}	PD	DC_{ew}	$Full_l$	$Full_A$	$CSL_1^{(a)}$	DC_{opt}
			$q = 15$						$q = 30$					
0.5	50	150	27.842	5.740	**4.574**	**4.574**	*4.644*	**4.574**	57.743	2.598	*1.499*	*1.499*	5.845	**1.498**
	100	150	14.637	5.060	**4.517**	**4.517**	*4.519*	**4.517**	18.278	2.732	**1.399**	**1.399**	1.452	*1.409*
	1000	150	5.335	*4.475*	**4.474**	**4.474**	**4.474**	**4.474**	3.048	*1.334*	**1.309**	**1.309**	**1.309**	**1.309**
	10000	10	*4.552*	**4.480**	**4.480**	**4.480**	**4.480**	**4.480**	*1.452*	**1.313**	**1.313**	**1.313**	**1.313**	**1.313**
	10000	100	*4.547*	**4.469**	**4.469**	**4.469**	**4.469**	**4.469**	*1.450*	**1.300**	**1.300**	**1.300**	**1.300**	**1.300**
	10000	150	*4.554*	**4.470**	**4.470**	**4.470**	**4.470**	**4.470**	*1.457*	**1.299**	**1.299**	**1.299**	**1.299**	**1.299**
1	50	150	42.536	7.014	**4.672**	**4.672**	*4.806*	**4.672**	90.166	3.845	*1.697*	1.699	10.344	**1.694**
	100	150	22.452	5.972	**4.568**	**4.568**	*4.572*	**4.568**	29.441	4.129	**1.500**	**1.500**	*1.604*	1.606
	1000	150	6.222	*4.488*	**4.479**	**4.479**	**4.479**	**4.479**	4.778	*1.467*	**1.319**	**1.319**	**1.319**	**1.319**
	10000	10	*4.628*	**4.488**	**4.488**	**4.488**	**4.488**	**4.488**	*1.603*	**1.329**	**1.329**	**1.329**	**1.329**	**1.329**
	10000	100	*4.618*	**4.470**	**4.470**	**4.470**	**4.470**	**4.470**	*1.600*	**1.301**	**1.301**	**1.301**	**1.301**	**1.301**
	10000	150	*4.631*	**4.471**	**4.471**	**4.471**	**4.471**	**4.471**	*1.612*	**1.300**	**1.300**	**1.300**	**1.300**	**1.300**
$0.5+\frac{m}{M}$	50	150	54.663	6.896	*4.667*	**4.666**	4.809	4.677	121.008	3.710	1.705	*1.700*	10.473	**1.693**
	100	150	27.738	5.873	**4.564**	**4.564**	4.570	*4.566*	36.389	3.994	**1.490**	**1.490**	1.604	*1.587*
	1000	150	7.107	*4.488*	**4.479**	**4.479**	**4.479**	**4.479**	6.393	1.460	**1.317**	**1.317**	*1.319*	*1.319*
	10000	10	4.703	4.489	**4.487**	**4.487**	4.489	*4.488*	1.760	1.331	**1.327**	**1.327**	1.331	*1.329*
	10000	100	*4.689*	**4.470**	**4.470**	**4.470**	**4.470**	**4.470**	*1.758*	**1.301**	**1.301**	**1.301**	**1.301**	**1.301**
	10000	150	4.709	*4.471*	**4.470**	**4.470**	*4.471*	*4.471*	*1.774*	**1.300**	**1.300**	**1.300**	**1.300**	**1.300**

Table S5: SPEs of various methods based on the OLS estimation with $\alpha = 0.5$ ($\times 10^{-2}$). The SPEs for the best-performing and the second-best methods are highlighted in bold and italics.

σ_m^2	n_m	M	PD	DC_{ew}	$Full_l$	$Full_A$	$CSL_1^{(a)}$	DC_{opt}	PD	DC_{ew}	$Full_l$	$Full_A$	$CSL_1^{(a)}$	DC_{opt}
			$q = 15$						$q = 30$					
0.5	50	150	15.997	1.176	**0.475**	0.478	0.541	*0.476*	37.592	0.804	*0.267*	0.284	4.599	**0.256**
	100	150	7.228	0.932	**0.428**	**0.428**	*0.429*	*0.429*	9.245	0.778	**0.169**	*0.178*	0.220	0.214
	1000	150	1.269	*0.413*	**0.382**	**0.382**	**0.382**	**0.382**	1.381	0.220	**0.079**	**0.079**	**0.079**	*0.088*
	10000	10	*0.454*	**0.385**	**0.385**	**0.385**	**0.385**	**0.385**	0.230	0.088	**0.084**	**0.084**	**0.084**	*0.087*
	10000	100	*0.449*	**0.377**	**0.377**	**0.377**	**0.377**	**0.377**	0.227	*0.074*	**0.070**	**0.070**	**0.070**	**0.070**
	10000	150	*0.456*	**0.377**	**0.377**	**0.377**	**0.377**	**0.377**	0.238	*0.073*	**0.069**	**0.069**	**0.069**	**0.069**
1	50	150	27.089	1.711	**0.573**	0.586	0.704	*0.574*	61.314	1.295	*0.465*	0.477	9.138	**0.435**
	100	150	11.700	1.335	**0.479**	*0.482*	0.483	0.489	15.296	1.192	**0.270**	*0.289*	0.372	0.335
	1000	150	1.999	*0.486*	**0.387**	**0.387**	**0.387**	**0.387**	2.155	0.346	**0.089**	**0.089**	**0.089**	*0.123*
	10000	10	0.536	*0.394*	**0.393**	**0.393**	**0.393**	**0.393**	0.394	0.118	**0.099**	*0.100*	**0.099**	0.112
	10000	100	*0.525*	**0.378**	**0.378**	**0.378**	**0.378**	**0.378**	0.387	0.089	**0.071**	**0.071**	**0.071**	*0.072*
	10000	150	*0.540*	**0.378**	**0.378**	**0.378**	**0.378**	**0.378**	0.400	0.088	**0.070**	**0.070**	**0.070**	*0.071*
$0.5+\frac{m}{M}$	50	150	37.847	1.683	**0.567**	*0.571*	0.707	0.576	89.969	1.280	*0.471*	0.473	9.266	**0.432**
	100	150	15.636	1.306	**0.474**	*0.475*	0.481	0.486	20.759	1.159	**0.259**	*0.273*	0.373	0.325
	1000	150	2.778	0.482	**0.386**	**0.386**	*0.387*	*0.387*	2.935	0.335	**0.087**	**0.087**	*0.089*	0.115
	10000	10	0.619	0.395	**0.392**	**0.392**	*0.394*	*0.394*	0.534	0.118	**0.098**	**0.098**	*0.101*	0.110
	10000	100	*0.608*	**0.378**	**0.378**	**0.378**	**0.378**	**0.378**	0.527	0.088	**0.071**	**0.071**	**0.071**	*0.072*
	10000	150	*0.628*	**0.378**	**0.378**	**0.378**	**0.378**	**0.378**	0.546	*0.087*	**0.070**	**0.070**	**0.070**	**0.070**

Table S6: SPEs of various methods based on the OLS estimation with $\alpha = 1$ ($\times 10^{-2}$). The SPEs for the best-performing and the second-best methods are highlighted in bold and italics.

σ_m^2	n_m	M	PD	DC_{ew}	$Full_l$	$Full_A$	$CSL_1^{(a)}$	DC_{opt}	PD	DC_{ew}	$Full_l$	$Full_A$	$CSL_1^{(a)}$	DC_{opt}
			$q=15$						$q=30$					
0.5	50	150	11.757	0.425	*0.124*	0.135	0.190	**0.119**	29.092	0.407	0.201	**0.144**	4.547	*0.167*
	100	150	4.715	0.301	**0.077**	0.083	*0.079*	*0.079*	6.542	0.285	0.104	*0.085*	0.155	**0.082**
	1000	150	0.693	*0.078*	**0.031**	**0.031**	**0.031**	**0.031**	0.725	0.071	**0.013**	*0.014*	**0.013**	0.019
	10000	10	0.109	0.037	**0.034**	**0.034**	**0.034**	*0.036*	0.120	0.024	**0.018**	*0.019*	**0.018**	0.023
	10000	100	0.108	*0.029*	**0.027**	**0.027**	**0.027**	**0.027**	0.120	0.015	**0.004**	**0.004**	**0.004**	*0.007*
	10000	150	0.112	*0.029*	**0.027**	**0.027**	**0.027**	**0.027**	0.126	0.014	**0.004**	**0.004**	**0.004**	*0.005*
1	50	150	20.623	0.683	*0.222*	0.230	0.353	**0.207**	55.283	0.720	0.399	**0.244**	9.094	*0.316*
	100	150	7.390	0.478	*0.129*	0.137	0.132	**0.125**	10.119	0.460	0.205	*0.149*	0.306	**0.140**
	1000	150	1.201	0.121	**0.036**	**0.036**	**0.036**	*0.039*	1.231	0.114	**0.023**	*0.025*	**0.023**	0.030
	10000	10	0.189	0.049	**0.042**	*0.043*	**0.042**	0.047	0.210	0.040	**0.034**	**0.034**	**0.034**	*0.038*
	10000	100	0.185	*0.035*	**0.028**	**0.028**	**0.028**	**0.028**	0.206	0.024	**0.006**	**0.006**	**0.006**	*0.010*
	10000	150	0.197	*0.035*	**0.027**	**0.027**	**0.027**	**0.027**	0.217	0.023	**0.005**	**0.005**	**0.005**	*0.008*
$0.5+\frac{m}{M}$	50	150	30.160	0.677	*0.215*	0.219	0.356	**0.209**	81.279	0.720	0.405	**0.247**	9.222	*0.316*
	100	150	10.184	0.465	**0.123**	0.130	0.131	*0.125*	13.601	0.448	*0.194*	**0.139**	0.307	**0.139**
	1000	150	1.711	0.118	**0.035**	**0.035**	*0.036*	0.038	1.737	0.110	**0.022**	*0.023*	0.024	0.029
	10000	10	0.268	0.050	**0.042**	**0.042**	*0.043*	0.047	0.293	0.041	**0.032**	*0.033*	0.036	0.038
	10000	100	0.249	0.035	**0.027**	**0.027**	*0.028*	*0.028*	0.272	0.023	**0.006**	**0.006**	**0.006**	*0.010*
	10000	150	0.269	*0.034*	**0.027**	**0.027**	**0.027**	**0.027**	0.292	0.023	**0.005**	**0.005**	**0.005**	*0.008*

Table S7: SPEs of various methods based on the OLS estimation with $\alpha = 1.5$ ($\times 10^{-2}$). The SPEs for the best-performing and the second-best methods are highlighted in bold and italics.

σ_m^2	n_m	M	$CSL_1^{(a)}$	$CSL_2^{(a)}$	$CSL_3^{(a)}$	$CSL_1^{(c)}$	$CSL_2^{(c)}$	$CSL_3^{(c)}$	$CSL_1^{(a)}$	$CSL_2^{(a)}$	$CSL_3^{(a)}$	$CSL_1^{(c)}$	$CSL_2^{(c)}$	$CSL_3^{(c)}$
			$q=15$						$q=30$					
0.5	50	10	7.022	10.480	48.710	51.045	291.749	-	95.990	-	-	-	-	-
	50	100	4.723	5.422	17.592	85.581	-	-	9.454	-	-	-	-	-
	100	10	5.322	5.331	5.342	8.703	9.898	13.171	3.599	7.762	50.831	49.094	342.989	-
	100	100	4.556	4.558	4.560	9.081	11.059	18.210	1.526	1.871	5.344	53.151	381.293	-
	1000	10	4.538	4.538	4.538	4.549	4.538	4.538	1.451	1.451	1.451	1.506	1.457	1.451
	1000	100	4.479	4.479	4.479	4.492	4.479	4.479	1.314	1.314	1.314	1.374	1.319	1.314
1	50	10	9.342	15.724	84.548	94.131	557.990	-	189.387	-	-	-	-	-
	50	100	4.959	6.258	28.624	156.542	-	-	17.453	-	-	-	-	-
	100	10	6.092	6.106	6.129	12.593	14.751	21.012	5.826	13.927	97.251	94.671	663.669	-
	100	100	4.639	4.643	4.646	13.280	17.211	31.663	1.746	2.431	9.304	103.877	754.645	-
	1000	10	4.615	4.615	4.615	4.637	4.616	4.615	1.602	1.601	1.601	1.710	1.613	1.602
	1000	100	4.486	4.486	4.486	4.513	4.488	4.486	1.329	1.329	1.329	1.447	1.340	1.330
$0.5+\frac{m}{M}$	50	10	9.640	16.020	90.718	95.421	562.580	-	196.043	-	-	-	-	-
	50	100	4.963	6.258	28.112	156.614	-	-	18.255	-	-	-	-	-
	100	10	6.162	6.163	6.198	12.718	14.765	21.035	6.041	14.430	101.587	95.294	664.646	-
	100	100	4.639	4.642	4.645	13.280	17.224	31.721	1.745	2.412	9.217	103.834	754.126	-
	1000	10	4.623	4.623	4.623	4.647	4.624	4.623	1.617	1.616	1.616	1.728	1.628	1.617
	1000	100	4.486	4.486	4.486	4.514	4.488	4.486	1.329	1.329	1.329	1.448	1.340	1.331

Table S8: SPEs of CSL based on the OLS estimation with $\alpha = 0.5$ ($\times 10^{-2}$), where the superscripts (a) and (c) denote the initial value of the simple average of all local estimators and the central site's estimator, and the subscripts 1, 2, and 3 denote the number of iterations. For neatness, if the value in the table is greater than 1000, we will not present it.

σ_m^2	n_m	M	$CSL_1^{(a)}$	$CSL_2^{(a)}$	$CSL_3^{(a)}$	$CSL_1^{(c)}$	$CSL_2^{(c)}$	$CSL_3^{(c)}$	$CSL_1^{(a)}$	$CSL_2^{(a)}$	$CSL_3^{(a)}$	$CSL_1^{(c)}$	$CSL_2^{(c)}$	$CSL_3^{(c)}$
			$q = 15$						$q = 30$					
0.5	50	10	2.703	5.598	35.529	43.970	271.971	-	94.005	-	-	-	-	-
	50	100	0.614	1.213	11.334	71.458	824.061	-	8.104	-	-	-	-	-
	100	10	1.152	1.155	1.169	4.334	5.279	8.350	2.295	6.204	46.019	45.389	317.363	-
	100	100	0.460	0.462	0.463	4.600	6.613	14.189	0.289	0.631	4.046	50.952	375.601	-
	1000	10	0.454	0.454	0.454	0.465	0.455	0.454	0.219	0.219	0.219	0.273	0.224	0.219
	1000	100	0.384	0.384	0.384	0.398	0.385	0.384	0.083	0.083	0.083	0.142	0.089	0.084
1	50	10	5.002	10.745	69.664	87.669	546.374	-	188.099	-	-	-	-	-
	50	100	0.849	2.032	21.904	141.268	-	-	16.146	-	-	-	-	-
	100	10	1.918	1.923	1.953	8.282	10.149	16.302	4.517	12.305	91.513	90.401	631.241	-
	100	100	0.543	0.546	0.550	8.779	12.850	28.230	0.509	1.193	8.024	101.825	751.651	-
	1000	10	0.533	0.533	0.533	0.555	0.534	0.533	0.370	0.369	0.369	0.478	0.381	0.370
	1000	100	0.392	0.392	0.392	0.419	0.393	0.392	0.099	0.099	0.099	0.215	0.109	0.100
$0.5+\frac{m}{M}$	50	10	5.300	11.005	75.096	88.996	550.825	-	194.460	-	-	-	-	-
	50	100	0.854	2.037	21.577	141.321	-	-	16.925	-	-	-	-	-
	100	10	1.992	1.984	2.024	8.413	10.164	16.324	4.734	12.793	95.695	91.014	632.257	-
	100	100	0.543	0.546	0.550	8.775	12.862	28.282	0.508	1.173	7.912	101.803	751.266	-
	1000	10	0.541	0.541	0.541	0.564	0.542	0.541	0.385	0.385	0.385	0.496	0.395	0.386
	1000	100	0.392	0.392	0.392	0.419	0.393	0.392	0.099	0.099	0.099	0.216	0.109	0.100

Table S9: SPEs of CSL based on the OLS estimation with $\alpha = 1$ ($\times 10^{-2}$), where the superscripts (a) and (c) denote the initial value of the simple average of all local estimators and the central site's estimator, and the subscripts 1, 2, and 3 denote the number of iterations. For neatness, if the value in the table is greater than 1000, we will not present it.

σ_m^2	n_m	M	$CSL_1^{(a)}$	$CSL_2^{(a)}$	$CSL_3^{(a)}$	$CSL_1^{(c)}$	$CSL_2^{(c)}$	$CSL_3^{(c)}$	$CSL_1^{(a)}$	$CSL_2^{(a)}$	$CSL_3^{(a)}$	$CSL_1^{(c)}$	$CSL_2^{(c)}$	$CSL_3^{(c)}$
			$q = 15$						$q = 30$					
0.5	50	10	2.323	5.160	33.908	43.844	275.842	-	94.209	-	-	-	-	-
	50	100	0.261	0.842	10.533	69.712	787.147	-	8.052	-	-	-	-	-
	100	10	0.792	0.793	0.809	3.986	4.901	8.001	2.224	6.094	45.357	44.930	312.838	-
	100	100	0.109	0.111	0.112	4.204	6.281	14.167	0.223	0.566	3.984	50.901	376.494	-
	1000	10	0.105	0.105	0.105	0.116	0.105	0.105	0.153	0.153	0.153	0.208	0.159	0.154
	1000	100	0.034	0.034	0.034	0.047	0.034	0.034	0.018	0.018	0.018	0.076	0.023	0.019
1	50	10	4.617	10.288	67.673	87.786	553.160	-	188.468	-	-	-	-	-
	50	100	0.496	1.654	20.955	139.204	-	-	16.104	-	-	-	-	-
	100	10	1.556	1.559	1.591	7.949	9.780	15.988	4.444	12.180	90.633	89.803	625.046	-
	100	100	0.192	0.195	0.199	8.377	12.547	28.394	0.442	1.129	7.967	101.808	753.185	-
	1000	10	0.184	0.184	0.184	0.206	0.185	0.184	0.304	0.304	0.304	0.413	0.315	0.305
	1000	100	0.042	0.042	0.042	0.068	0.043	0.042	0.033	0.033	0.033	0.149	0.044	0.035
$0.5+\frac{m}{M}$	50	10	4.914	10.542	72.963	89.130	557.610	-	194.763	-	-	-	-	-
	50	100	0.502	1.662	20.678	139.251	-	-	16.876	-	-	-	-	-
	100	10	1.631	1.622	1.663	8.082	9.796	16.010	4.661	12.664	94.781	90.414	626.071	-
	100	100	0.192	0.195	0.199	8.371	12.559	28.443	0.442	1.108	7.848	101.792	752.831	-
	1000	10	0.192	0.192	0.192	0.215	0.193	0.192	0.320	0.319	0.319	0.431	0.329	0.320
	1000	100	0.042	0.042	0.042	0.069	0.043	0.042	0.033	0.033	0.033	0.150	0.044	0.035

Table S10: SPEs of CSL based on the OLS estimation with $\alpha = 1.5$ ($\times 10^{-2}$), where the superscripts (a) and (c) denote the initial value of the simple average of all local estimators and the central site's estimator, and the subscripts 1, 2, and 3 denote the number of iterations. For neatness, if the value in the table is greater than 1000, we will not present it.

σ_m^2	n_m	M	PD	DC_{ew}	$Full_A$	DC_{opt}	$CSL_1^{(a)}$	$CSL_2^{(a)}$	$CSL_3^{(a)}$	PD	DC_{ew}	$Full_A$	DC_{opt}	$CSL_1^{(a)}$	$CSL_2^{(a)}$	$CSL_3^{(a)}$
						$q = 15$							$q = 30$			
0.5	50	10	0.0011	0.0019	0.0019	0.0030	0.0010	0.0012	0.0016	0.0023	0.0040	0.0070	0.0052	0.0015	0.0020	0.0026
	50	100	0.0012	0.0055	0.0185	0.0239	0.0026	0.0036	0.0045	0.0026	0.0079	0.0555	0.0128	0.0029	0.0042	0.0056
	100	10	0.0014	0.0050	0.0026	0.0064	0.0011	0.0015	0.0019	0.0029	0.0040	0.0084	0.0054	0.0018	0.0026	0.0035
	100	100	0.0013	0.0049	0.0299	0.0187	0.0028	0.0038	0.0051	0.0029	0.0133	0.1154	0.0188	0.0033	0.0049	0.0061
	1000	10	0.0026	0.0065	0.0162	0.0083	0.0026	0.0039	0.0051	0.0083	0.0101	0.0726	0.0120	0.0034	0.0052	0.0071
	1000	100	0.0055	0.0250	0.5909	0.0521	0.0088	0.0101	0.0115	0.0134	0.0363	2.2449	0.0614	0.0106	0.0127	0.0150
1	50	10	0.0011	0.0012	0.0018	0.0023	0.0008	0.0011	0.0014	0.0024	0.0039	0.0051	0.0051	0.0013	0.0018	0.0024
	50	100	0.0010	0.0051	0.0130	0.0104	0.0027	0.0038	0.0049	0.0026	0.0126	0.0565	0.0183	0.0029	0.0042	0.0054
	100	10	0.0011	0.0017	0.0025	0.0028	0.0011	0.0014	0.0018	0.0029	0.0044	0.0082	0.0059	0.0019	0.0027	0.0034
	100	100	0.0011	0.0057	0.0243	0.0109	0.0026	0.0039	0.0050	0.0029	0.0125	0.1172	0.0187	0.0030	0.0046	0.0059
	1000	10	0.0025	0.0047	0.0145	0.0066	0.0023	0.0038	0.0049	0.0081	0.0100	0.0711	0.0119	0.0035	0.0054	0.0075
	1000	100	0.0055	0.0091	0.4592	0.0288	0.0055	0.0069	0.0081	0.0133	0.0273	2.2971	0.0496	0.0131	0.0151	0.0174
$0.5+\frac{m}{M}$	50	10	0.0010	0.0016	0.0122	0.0026	0.0009	0.0012	0.0014	0.0023	0.0032	0.0283	0.0044	0.0013	0.0018	0.0023
	50	100	0.0011	0.0059	0.1154	0.0115	0.0025	0.0036	0.0046	0.0025	0.0153	0.2973	0.0197	0.0028	0.0041	0.0056
	100	10	0.0011	0.0016	0.0136	0.0027	0.0010	0.0013	0.0017	0.0027	0.0035	0.0356	0.0049	0.0016	0.0025	0.0033
	100	100	0.0013	0.0060	0.1316	0.0110	0.0029	0.0053	0.0065	0.0029	0.0103	0.3782	0.0159	0.0030	0.0044	0.0057
	1000	10	0.0024	0.0033	0.0371	0.0053	0.0024	0.0036	0.0049	0.0084	0.0096	0.1504	0.0115	0.0036	0.0054	0.0074
	1000	100	0.0053	0.0110	0.6508	0.0315	0.0057	0.0074	0.0092	0.0134	0.0283	2.9363	0.0526	0.0091	0.0118	0.0139

† The simulations are conducted in R version 4.1.3 by a server with four Intel(R) Xeon(R) Gold 6230R Processors (2.10 GHz) with 128 GB memory.

‡ The transmission time of the statistic between sites is not included. When multiple sites are assigned the same computation task simultaneously and do the job in a parallel manner, we take the computation time for all sites to complete the task.

Table S11: Computing time measured in seconds of the five methods based on OLS with $\alpha = 0.5$.

σ_m^2	n_m	M	PD	DC_{ew}	$Full$	DC_{opt}	$CSL_{1,cv}^{(c)}$	$CSL_{1,o}^{(c)}$	PD	DC_{ew}	$Full$	DC_{opt}	$CSL_{1,cv}^{(c)}$	$CSL_{1,o}^{(c)}$
			$q=15$						$q=30$					
0.5	50	10	0.0518	0.0604	0.0555	0.0619	196.6028	19.5516	0.0611	0.0726	0.0639	0.0744	198.0919	19.6973
	50	100	0.0512	0.1044	0.1158	0.1194	197.9122	19.6043	0.0600	0.0938	0.1507	0.1077	199.3290	19.7644
	100	10	0.0528	0.0669	0.0639	0.0687	199.0085	19.6990	0.0603	0.0768	0.0680	0.0787	199.9214	19.8555
	100	100	0.0538	0.0962	0.2153	0.1202	198.5428	19.7727	0.0574	0.1055	0.2526	0.1236	201.1087	19.9214
	1000	10	0.0638	0.0781	0.1566	0.0820	224.8075	23.2449	0.0690	0.0989	0.2019	0.1076	239.5410	24.4848
	1000	100	0.0659	0.1657	1.5325	0.2324	224.3786	23.2077	0.0750	0.2162	2.7024	0.3128	241.2756	25.1631
1	50	10	0.0546	0.0696	0.0568	0.0713	197.8334	19.6502	0.0610	0.0767	0.0632	0.0786	200.0140	19.8427
	50	100	0.0516	0.0998	0.1030	0.1127	199.1997	19.7185	0.0620	0.1156	0.1490	0.1340	198.7041	19.8290
	100	10	0.0528	0.0627	0.0614	0.0646	199.8655	19.7229	0.0582	0.0716	0.0685	0.0736	200.2241	19.9260
	100	100	0.0509	0.0982	0.1553	0.1132	197.6680	19.7618	0.0578	0.1209	0.2441	0.1410	200.0689	19.9421
	1000	10	0.0623	0.0836	0.1489	0.0870	223.3077	22.5400	0.0730	0.0897	0.2081	0.0956	240.7182	24.5418
	1000	100	0.0670	0.1806	1.5821	0.2348	223.4397	23.0230	0.0750	0.2238	2.7407	0.3239	241.4459	25.0862
$0.5+\frac{m}{M}$	50	10	0.0542	0.0643	0.0571	0.0660	197.3043	19.6344	0.0659	0.0817	0.0619	0.0835	199.2562	19.9143
	50	100	0.0528	0.1056	0.1001	0.1177	197.2048	19.6186	0.0638	0.1178	0.1561	0.1323	200.2775	19.8498
	100	10	0.0537	0.0630	0.0610	0.0649	199.0610	19.7045	0.0607	0.0734	0.0695	0.0753	201.9974	20.0709
	100	100	0.0527	0.0929	0.1511	0.1089	200.0594	19.8445	0.0600	0.1120	0.2470	0.1365	201.5327	20.2067
	1000	10	0.0617	0.0775	0.1450	0.0816	223.3021	22.5739	0.0703	0.0892	0.1975	0.0959	240.0915	24.5373
	1000	100	0.0641	0.1833	1.5731	0.2460	222.9046	23.0348	0.0752	0.2282	2.7547	0.3183	240.5886	25.1712

† The simulations are conducted in R version 4.1.3 by a server with four Intel(R) Xeon(R) Gold 6230R Processors (2.10 GHz) with 128 GB memory.

‡ The transmission time of the statistic between sites is not included. When multiple sites are assigned the same computation task simultaneously and do the job in a parallel manner, we take the computation time for all sites to complete the task.

§ Since the global likelihood function is replaced with the surrogate loss function, $\tilde{\boldsymbol{\beta}}_{CSL,1,cv}^{(c)}$ and $\tilde{\boldsymbol{\beta}}_{CSL,1,o}^{(c)}$ can not be directly obtained by R-package *glmnet*, thus R-package $CVXR$ is used to minimize the surrogate likelihood function plus the lasso penalty. However, the solving speed of $CVXR$ is slower than *glmnet*.

§§ $CSL_{1,cv}^{(c)}$ calls the $CVXR$ function 1001 times, in which 1000 times for tuning, and the last to calculate the final parameter estimator. $CSL_{1,o}^{(c)}$ calls the $CVXR$ function 100 times for different hyperparameters.

Table S12: Computing time measured in seconds of the five methods based on the lasso with $\alpha = 0.5$.

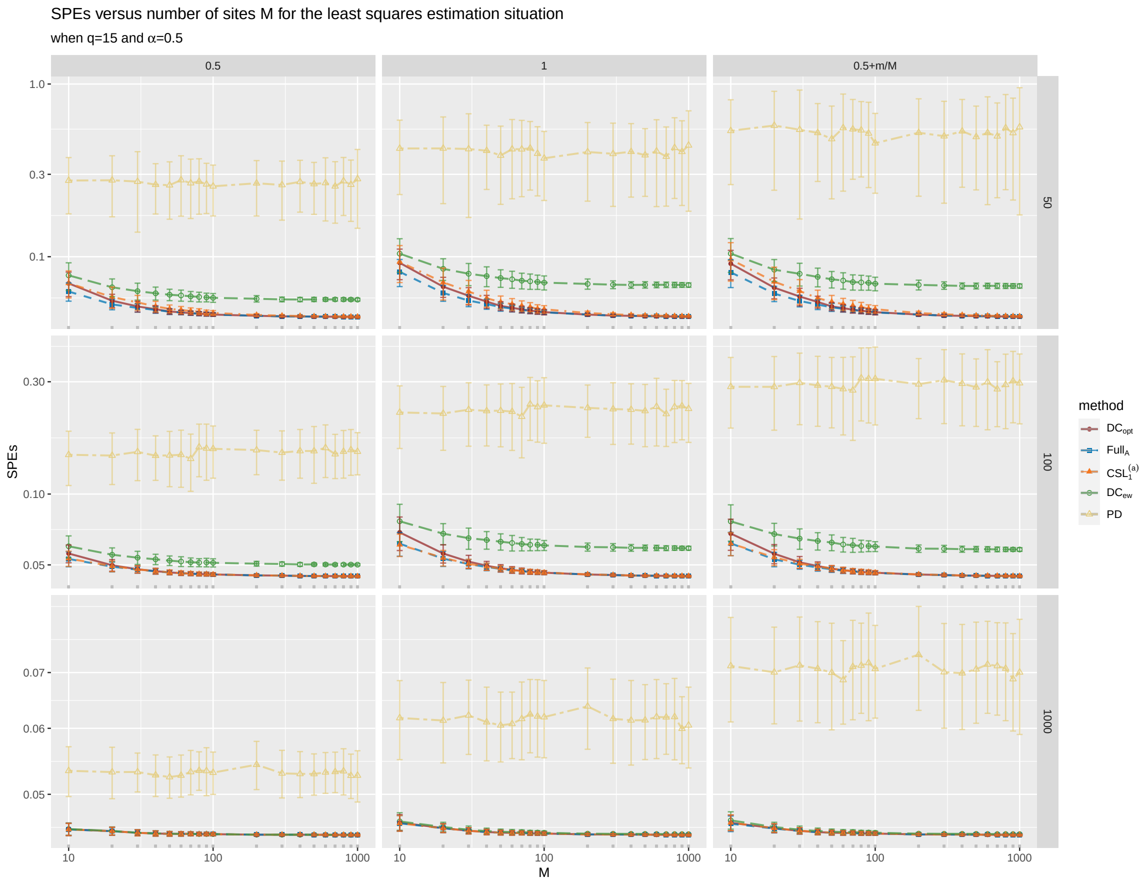

Figure S1: SPEs versus number of sites M for the least squares estimation situation when $q = 15$ and $\alpha = 0.5$. The three rows of the whole panel correspond to different local sample sizes n of 50, 100, and 1000 from top to bottom respectively, while the three columns of it correspond to different local error variances σ_m^2 of 0.5, 1, and $0.5 + m/M$ from left to right respectively. The log-log scale is applied to each panel. In all cases, each point corresponds to the average of 100 trials, with standard errors also shown in the form of error bar. Note that there exist some error bars that are only displayed the top half at some points for PD and $CSL_1^{(a)}$ in some panels at the first and second rows because the standard error is larger than the mean for these points.

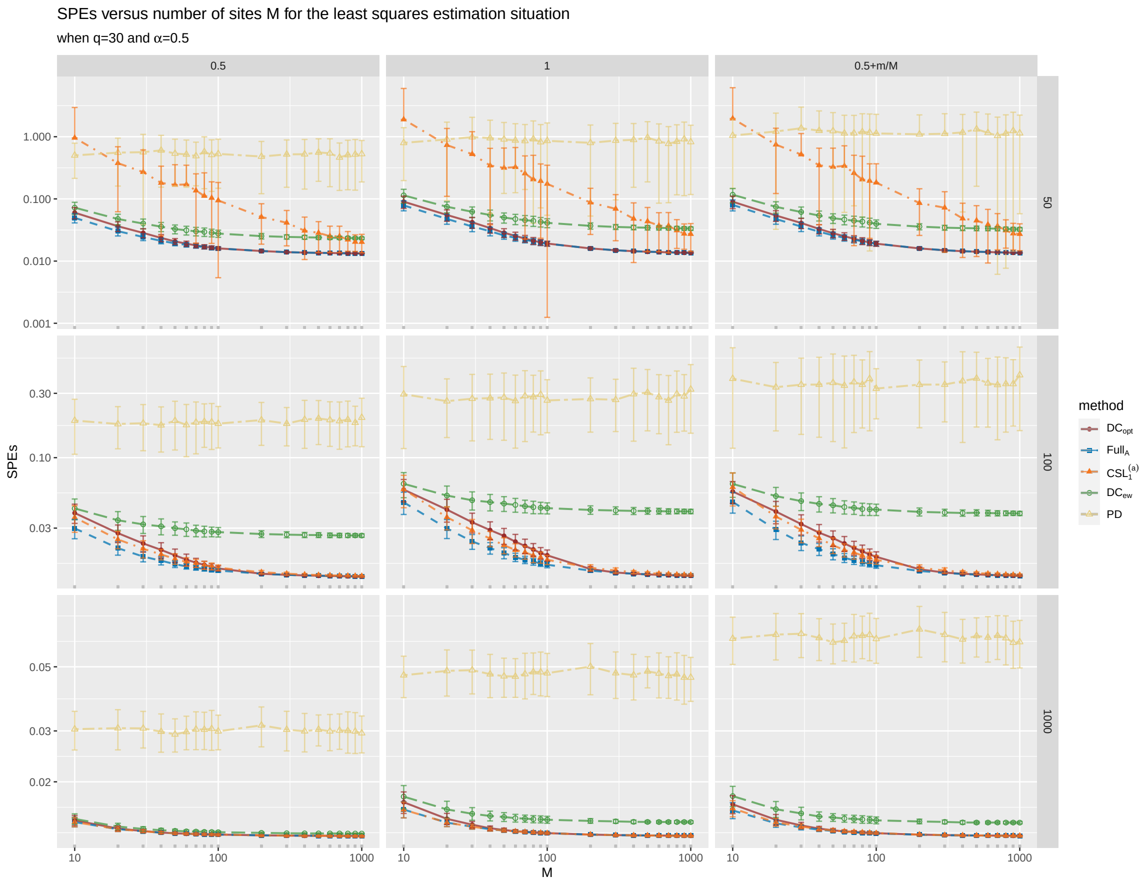

Figure S2: SPEs versus number of sites M for the least squares estimation situation when $q = 30$ and $\alpha = 0.5$. The three rows of the whole panel correspond to different local sample sizes n of 50, 100, and 1000 from top to bottom respectively, while the three columns of it correspond to different local error variances σ_m^2 of 0.5, 1, and $0.5 + m/M$ from left to right respectively. The log-log scale is applied to each panel. In all cases, each point corresponds to the average of 100 trials, with standard errors also shown in the form of error bar. Note that there exist some error bars that are only displayed the top half at some points for PD and $CSL_1^{(a)}$ in some panels at the first and second rows because the standard error is larger than the mean for these points.

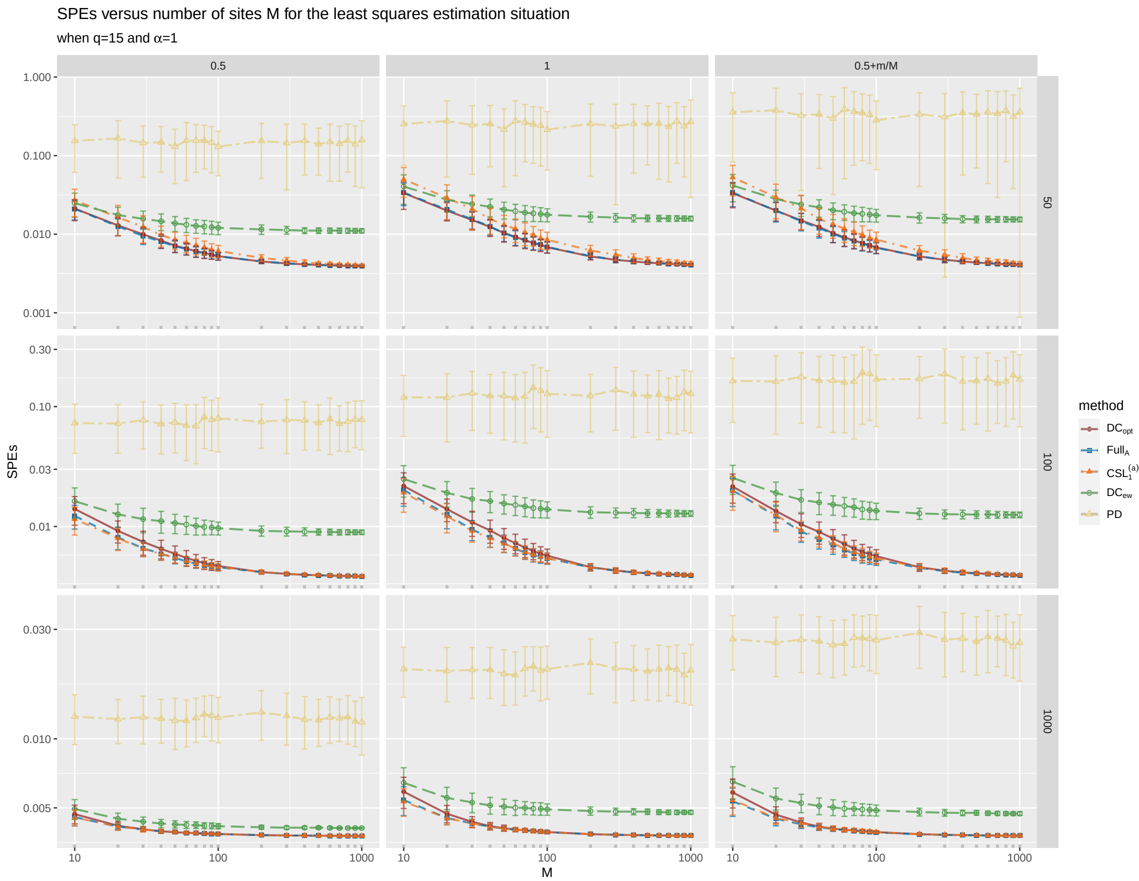

Figure S3: SPEs versus number of sites M for the least squares estimation situation when $q = 15$ and $\alpha = 1$. The three rows of the whole panel correspond to different local sample sizes n of 50, 100, and 1000 from top to bottom respectively, while the three columns of it correspond to different local error variances σ_m^2 of 0.5, 1, and $0.5 + m/M$ from left to right respectively. The log-log scale is applied to each panel. In all cases, each point corresponds to the average of 100 trials, with standard errors also shown in the form of error bar. Note that there exist some error bars that are only displayed the top half at some points for PD and $CSL_1^{(a)}$ in some panels at the first and second rows because the standard error is larger than the mean for these points.

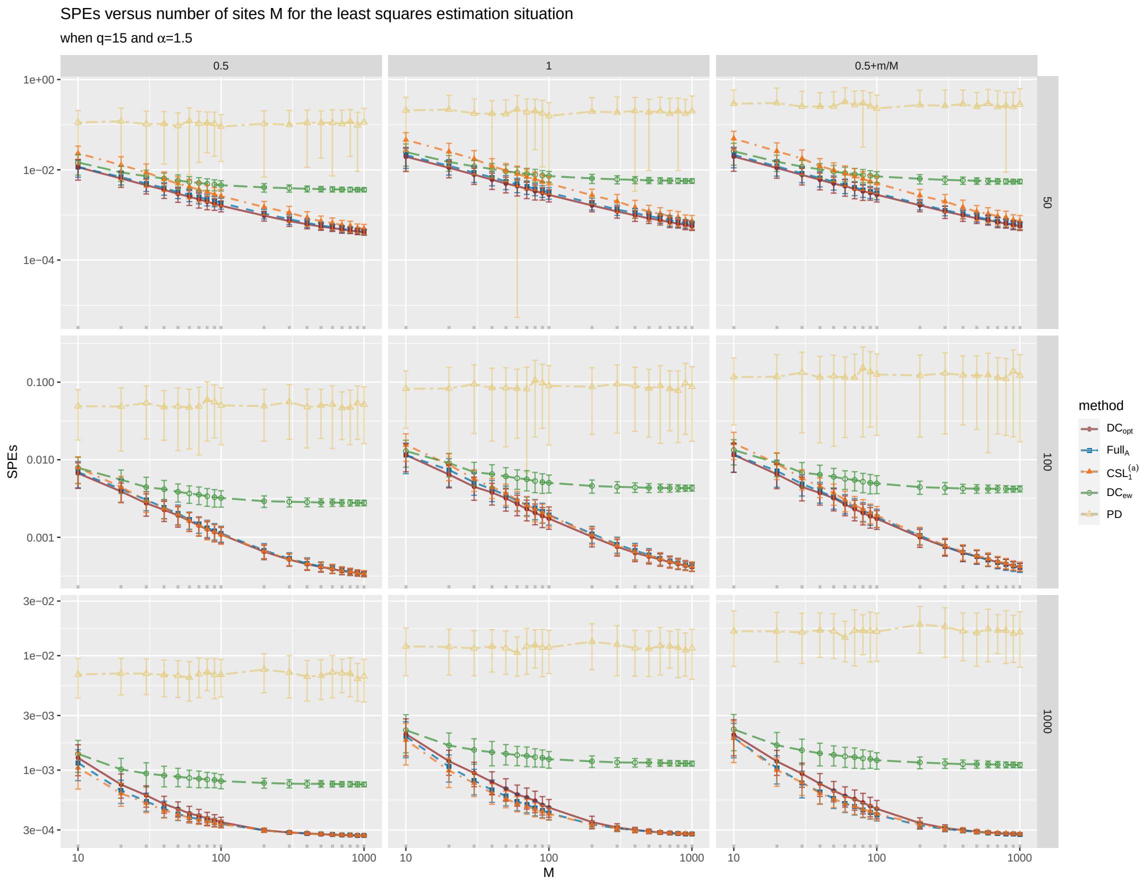

Figure S4: SPEs versus number of sites M for the least squares estimation situation when $q = 15$ and $\alpha = 1.5$. The three rows of the whole panel correspond to different local sample sizes n of 50, 100, and 1000 from top to bottom respectively, while the three columns of it correspond to different local error variances σ_m^2 of 0.5, 1, and $0.5 + m/M$ from left to right respectively. The log-log scale is applied to each panel. In all cases, each point corresponds to the average of 100 trials, with standard errors also shown in the form of error bar. Note that there exist some error bars that are only displayed the top half at some points for PD and $CSL_1^{(a)}$ in some panels at the first and second rows because the standard error is larger than the mean for these points.

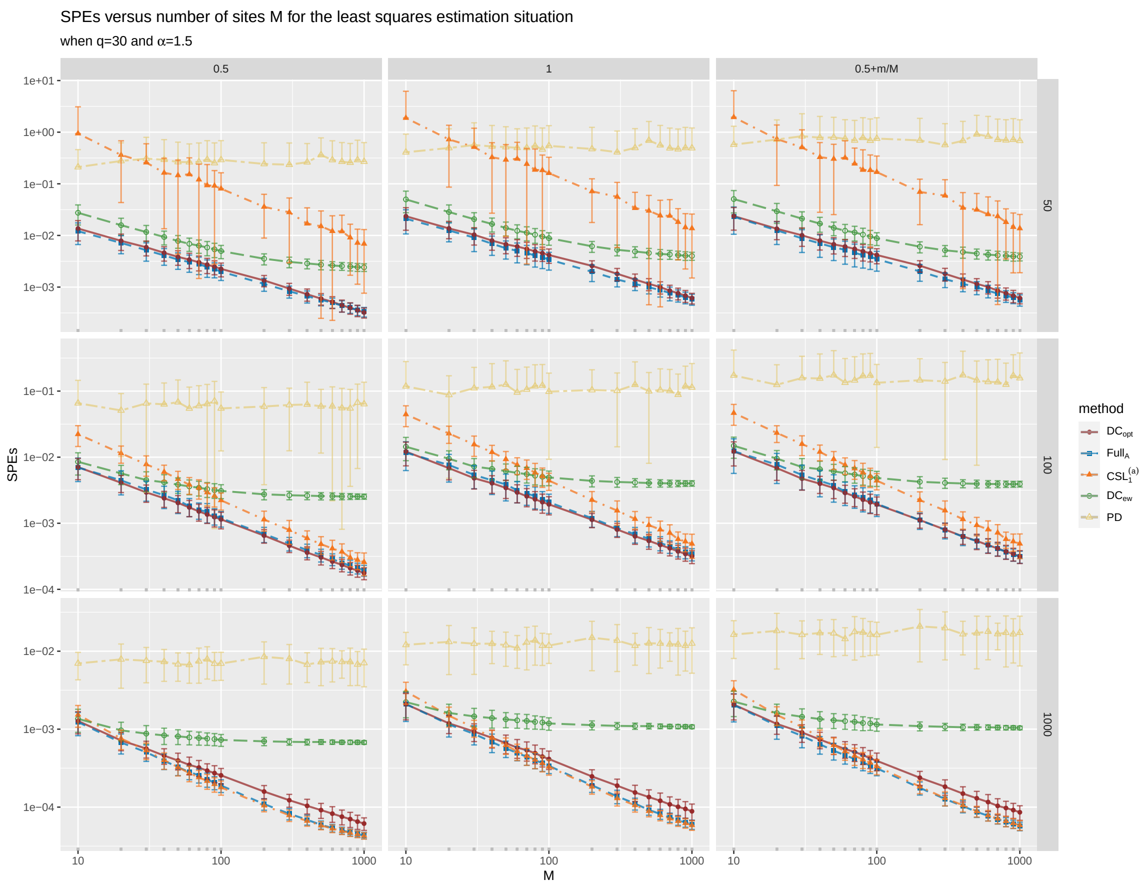

Figure S5: SPEs versus number of sites M for the least squares estimation situation when $q = 30$ and $\alpha = 1.5$. The three rows of the whole panel correspond to different local sample sizes n of 50, 100, and 1000 from top to bottom respectively, while the three columns of it correspond to different local error variances σ_m^2 of 0.5, 1, and $0.5 + m/M$ from left to right respectively. The log-log scale is applied to each panel. In all cases, each point corresponds to the average of 100 trials, with standard errors also shown in the form of error bar. Note that there exist some error bars that are only displayed the top half at some points for PD and $CSL_1^{(a)}$ in some panels at the first and second rows because the standard error is larger than the mean for these points.