Optimal Sparsity Criteria for Network Inference

Andreas Tjärnberg

Goals

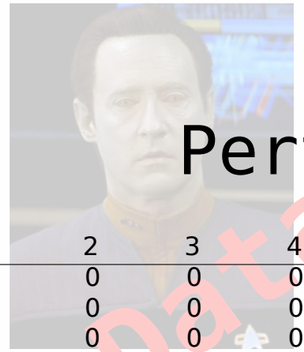
1. Find structure of a network

2. Estimate influence of each parameter (link) in the network

3. Be able to estimate the *correct* network and parameters from data.

$$\mathbf{G} : \mathbf{P} \mapsto \mathbf{Y}$$


Data



Perturbation

	1	2	3	4	5	6	7	8	9	10
A	0	0	0	0	0.3300	0	0	0	-0.2660	0
B	0	0	0	0	0	1.4770	0	0	0	0
C	-3.5550	0	0	0	0	-5.6840	0	0	0	0
D	0	-7.7710	-6.1450	-1.1230	0	0	-1.2540	0	0	0
E	0	0	0	0	0	0.8610	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	0
H	0	2.9740	2.3550	0	0	0	0	-0.0080	0	0
I	0	0	0	0	0	0	0	0	0	-0.3960
J	0	0	0	0	0	0	0	0	0	0

	11	12	13	14	15	16	17	18	19	20
A	0	0	0	0	0	0	0	0	-0.7150	0
B	0	0	0	0	0	0	0	0	0	0
C	-3.4010	0	0	0	0	0	0	0	-2.7680	0
D	0	0	0	0	0	-4.9700	0	0	0	0
E	0.9960	0	0	0	0	0	0	0	0	-0.5980
F	0	0	0	-0.8440	-0.0030	0	0	0	0	0
G	0	-0.0690	-0.0490	0.1010	0	0	0	0	0	0
H	0	0	0	0	0	1.9040	0	0	2.0940	0
I	0	0	0	0	0	0	0	0.4040	0	0
J	0	0	0	0	0	0	0.7810	0	0	0

Expression

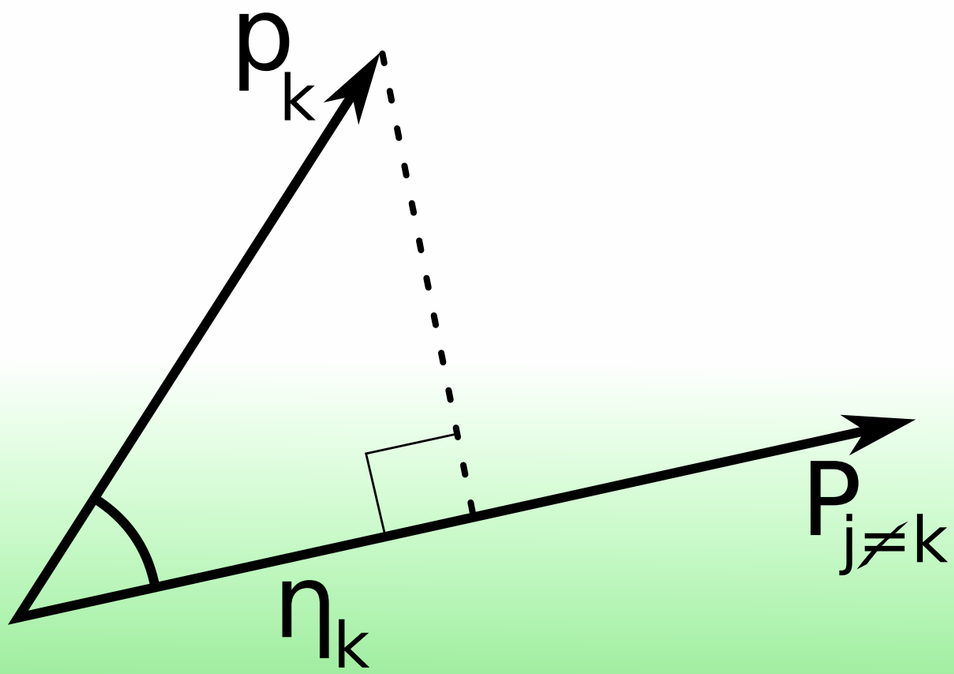
	1	2	3	4	5	6	7	8	9	10
A	0.0812	0.0499	0.0414	-0.2448	0.6609	-0.1038	-0.2734	-0.0047	-0.5327	-0.0698
B	-0.5492	0.0183	0.0151	-0.0895	0.0111	0.6028	-0.1000	-0.0017	-0.0089	-0.0255
C	-0.4476	0.0149	0.0123	-0.0730	0.0090	-0.7853	-0.0815	-0.0014	-0.0073	-0.0208
D	-0.0036	-0.6914	-0.5467	-0.1059	0.0136	-0.0020	-0.1183	-0.0001	-0.0109	-0.0302
E	-0.4914	0.0163	0.0136	-0.0801	0.0099	0.0075	-0.0895	-0.0015	-0.0080	-0.0228
F	-0.0096	-0.0059	-0.0049	0.0291	-0.0785	0.0123	0.0325	0.0006	0.0633	0.0083
G	-0.0113	-0.0070	-0.0058	0.0342	-0.0923	0.0145	0.0382	0.0007	0.0744	0.0097
H	0.0615	0.0378	0.0313	-0.1853	0.0076	-0.0785	-0.2069	-0.0036	-0.0061	-0.0528
I	-0.1236	0.0047	0.0039	-0.0232	0.0104	0.0007	-0.0259	-0.0004	-0.0084	-0.6847
J	0.0182	0.0112	0.0093	-0.0548	0.1480	-0.0232	-0.0612	-0.0011	-0.1193	-0.0156

	11	12	13	14	15	16	17	18	19	20
A	0.1039	-0.0021	-0.0015	0.0000	-0.0000	0.0331	0.0680	0.0712	-0.1407	-0.0158
B	-0.5158	0.0653	0.0464	-0.0007	0.0003	0.0121	0.0249	0.0260	-0.0024	-0.0058
C	-0.4204	0.0532	0.0378	-0.0006	0.0003	0.0099	0.0203	0.0212	-0.0020	-0.0047
D	0.0079	0.0035	0.0025	-0.0000	0.0000	-0.4422	0.0873	0.0308	-0.0029	-0.0068
E	0.5445	0.0585	0.0415	-0.0006	0.0003	0.0108	0.0222	0.0233	-0.0022	-0.6092
F	-0.0123	0.0003	0.0002	-1.0024	-0.0036	-0.0039	-0.0081	-0.0085	0.0167	0.0019
G	-0.0145	-0.8115	-0.5763	0.0090	-0.0042	-0.0046	-0.0095	-0.0099	0.0197	0.0022
H	0.0786	-0.0016	-0.0011	0.0000	-0.0000	0.0250	0.0514	0.0539	0.9607	-0.0119
I	0.1392	0.0843	0.0599	-0.0009	0.0004	0.0031	0.0065	0.6985	-0.0022	-0.1546
J	0.0233	-0.0005	-0.0003	0.0000	-0.0000	0.0074	0.9818	0.0159	-0.0315	-0.0035

Data

Informative

$$\eta_{p_k} = ||\mathbf{P}_{j \neq k}^T \mathbf{p}_k||_1 \qquad \eta_{y_k} = ||\mathbf{Y}_{j \neq k}^T \mathbf{y}_k||_1$$



$$\eta_{p_k} = ||\mathbf{P}_{j \neq k}^T \mathbf{p}_k||_1 \qquad \eta_{y_k} = ||\mathbf{Y}_{j \neq k}^T \mathbf{y}_k||_1$$

Choosing an eta-limit to indentify informative samples

:: eta should be larger than the smallest estimated signal in the system

:: smallest singular value of the P and Y matrix

Perturbation

	1	2	3	4	5	6	7	8	9	10
A	0	0	0	0	0.3300	0	0	0	-0.2660	0
B	0	0	0	0	0	1.4770	0	0	0	0
C	-3.5550	0	0	0	0	-5.6840	0	0	0	0
D	0	-7.7710	-6.1450	-1.1230	0	0	-1.2540	0	0	0
E	0	0	0	0	0	0.8610	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	0
H	0	2.9740	2.3550	0	0	0	0	-0.0080	0	0
I	0	0	0	0	0	0	0	0	0	-0.3960
J	0	0	0	0	0	0	0	0	0	0

	11	12	13	14	15	16	17	18	19	20
A	0	0	0	0	0	0	0	0	-0.7150	0
B	0	0	0	0	0	0	0	0	0	0
C	-3.4010	0	0	0	0	0	0	0	-2.7680	0
D	0	0	0	0	0	-4.9700	0	0	0	0
E	0.9960	0	0	0	0	0	0	0	0	-0.5980
F	0	0	0	-0.8440	-0.0030	0	0	0	0	0
G	0	-0.0690	-0.0490	0.1010	0	0	0	0	0	0
H	0	0	0	0	0	1.9040	0	0	2.0940	0
I	0	0	0	0	0	0	0	0.4040	0	0
J	0	0	0	0	0	0	0.7810	0	0	0

Expression

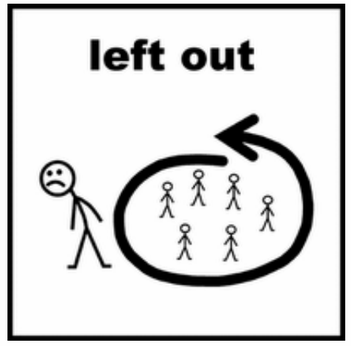
	1	2	3	4	5	6	7	8	9	10
A	0.0812	0.0499	0.0414	-0.2448	0.6609	-0.1038	-0.2734	-0.0047	-0.5327	-0.0698
B	-0.5492	0.0183	0.0151	-0.0895	0.0111	0.6028	-0.1000	-0.0017	-0.0089	-0.0255
C	-0.4476	0.0149	0.0123	-0.0730	0.0090	-0.7853	-0.0815	-0.0014	-0.0073	-0.0208
D	-0.0036	-0.6914	-0.5467	-0.1059	0.0136	-0.0020	-0.1183	-0.0001	-0.0109	-0.0302
E	-0.4914	0.0163	0.0136	-0.0801	0.0099	0.0075	-0.0895	-0.0015	-0.0080	-0.0228
F	-0.0096	-0.0059	-0.0049	0.0291	-0.0785	0.0123	0.0325	0.0006	0.0633	0.0083
G	-0.0113	-0.0070	-0.0058	0.0342	-0.0923	0.0145	0.0382	0.0007	0.0744	0.0097
H	0.0615	0.0378	0.0313	-0.1853	0.0076	-0.0785	-0.2069	-0.0036	-0.0061	-0.0528
I	-0.1236	0.0047	0.0039	-0.0232	0.0104	0.0007	-0.0259	-0.0004	-0.0084	-0.6847
J	0.0182	0.0112	0.0093	-0.0548	0.1480	-0.0232	-0.0612	-0.0011	-0.1193	-0.0156

	11	12	13	14	15	16	17	18	19	20
A	0.1039	-0.0021	-0.0015	0.0000	-0.0000	0.0331	0.0680	0.0712	-0.1407	-0.0158
B	-0.5158	0.0653	0.0464	-0.0007	0.0003	0.0121	0.0249	0.0260	-0.0024	-0.0058
C	-0.4204	0.0532	0.0378	-0.0006	0.0003	0.0099	0.0203	0.0212	-0.0020	-0.0047
D	0.0079	0.0035	0.0025	-0.0000	0.0000	-0.4422	0.0873	0.0308	-0.0029	-0.0068
E	0.5445	0.0585	0.0415	-0.0006	0.0003	0.0108	0.0222	0.0233	-0.0022	-0.6092
F	-0.0123	0.0003	0.0002	-1.0024	-0.0036	-0.0039	-0.0081	-0.0085	0.0167	0.0019
G	-0.0145	-0.8115	-0.5763	0.0090	-0.0042	-0.0046	-0.0095	-0.0099	0.0197	0.0022
H	0.0786	-0.0016	-0.0011	0.0000	-0.0000	0.0250	0.0514	0.0539	0.9607	-0.0119
I	0.1392	0.0843	0.0599	-0.0009	0.0004	0.0031	0.0065	0.6985	-0.0022	-0.1546
J	0.0233	-0.0005	-0.0003	0.0000	-0.0000	0.0074	0.9818	0.0159	-0.0315	-0.0035

Data

Informative

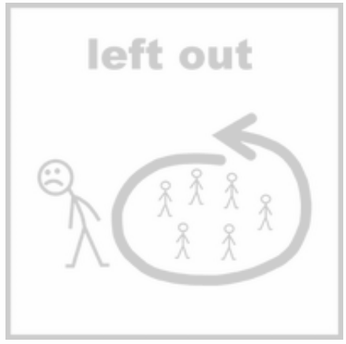
LOOCO



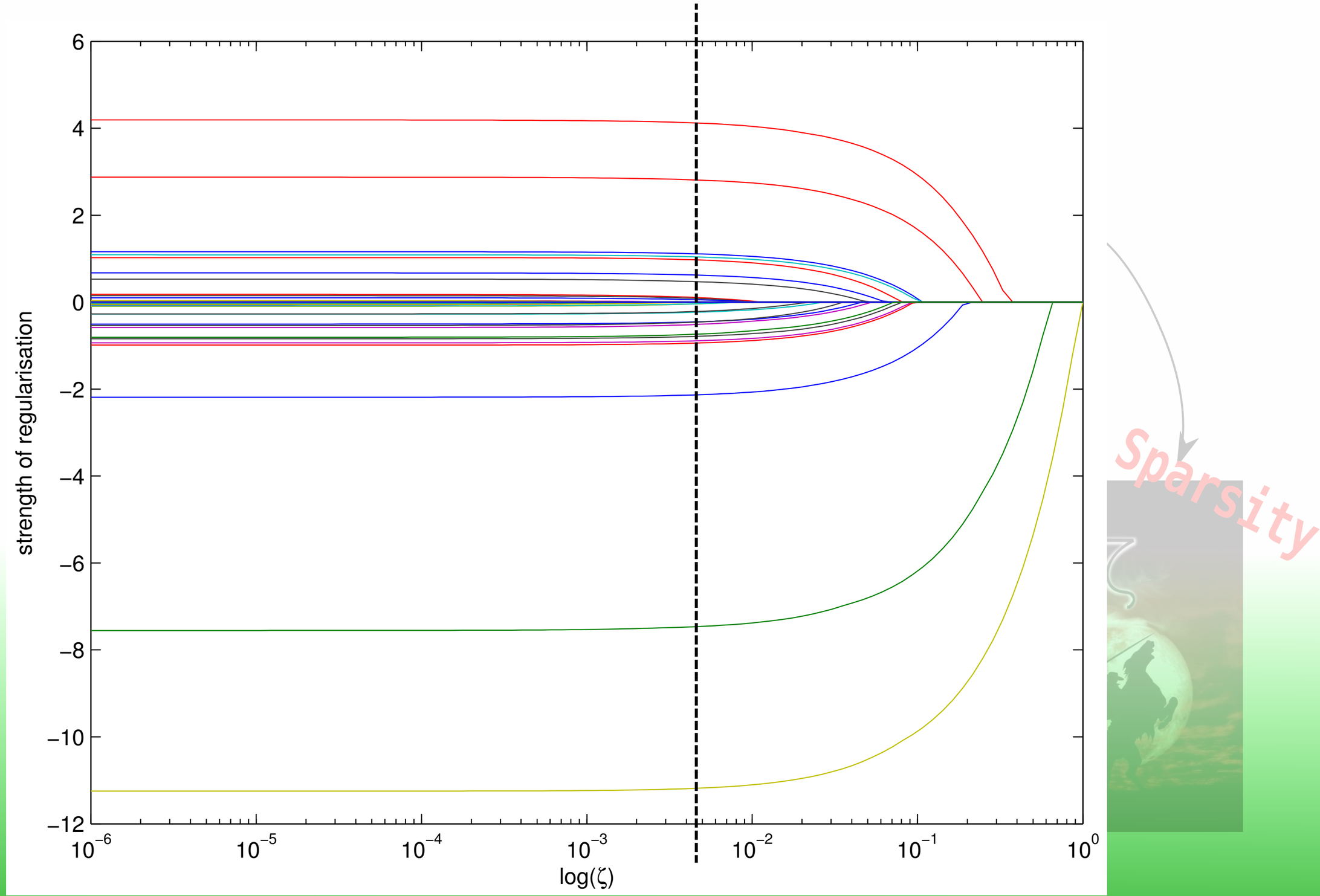
Data

Informative

LOOCO



Sparsity

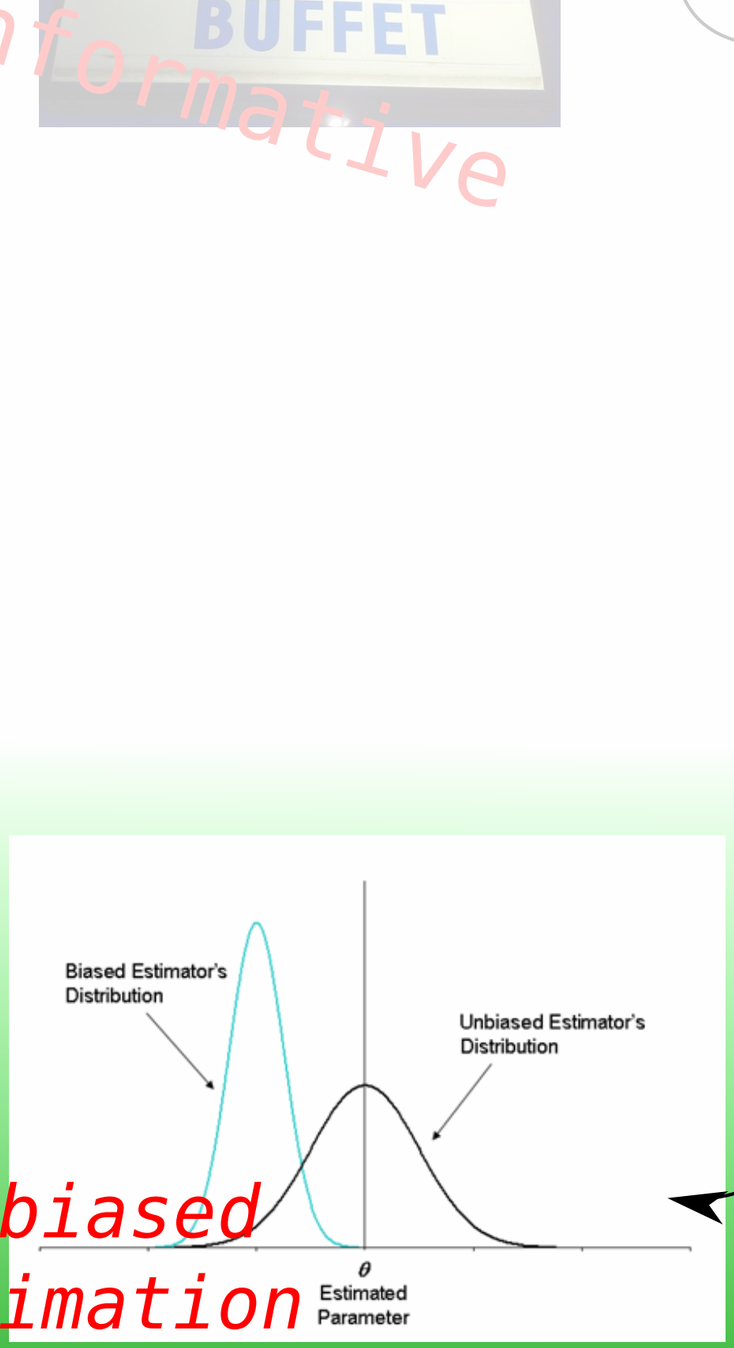

6
4
2
0
−2
−4
−6
−8
−10
−12
strength of regularisation
10^{-6}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}
log(ζ)
Sparsity

Data

Informative

LOOCO

Sparsity



Unbiased Estimation

Unbiased Estimation

LASSO: $\min_{\mathbf{A}} ||\mathbf{AY} + \mathbf{P}||_{l_2} + \zeta||\mathbf{A}||_{l_1}$

Unbiased Estimation

LASSO: $\min_{\mathbf{A}} ||\mathbf{AY} + \mathbf{P}||_{l_2} + \zeta||\mathbf{A}||_{l_1}$

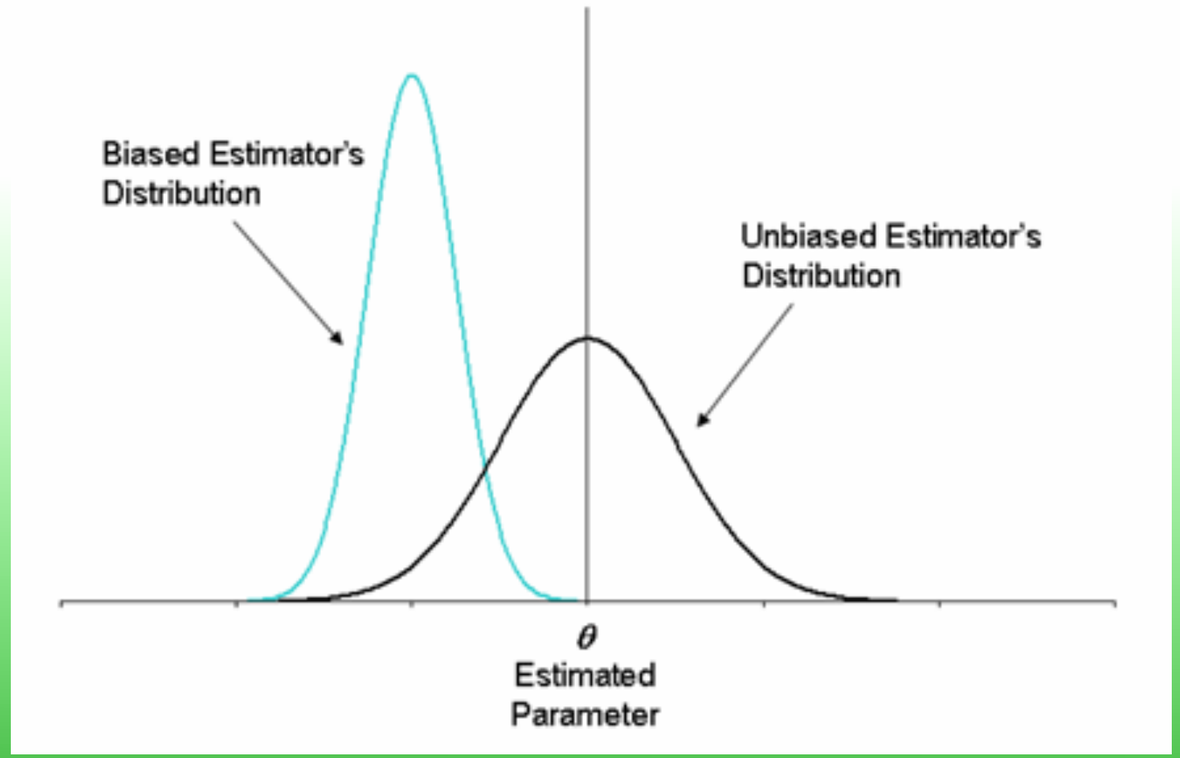
Problematic

Unbiased Estimation

LASSO: $\min_{\mathbf{A}} ||\mathbf{AY} + \mathbf{P}||_{l_2} + \zeta||\mathbf{A}||_{l_1}$

MLE: $\min_{\mathbf{A}_\zeta} ||\mathbf{A}_\zeta\mathbf{Y} + \mathbf{P}||_{l_2}$

Problematic

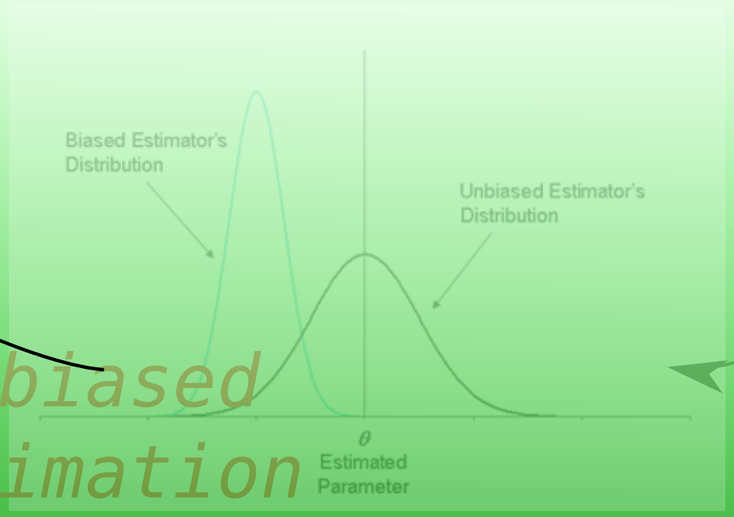


Data

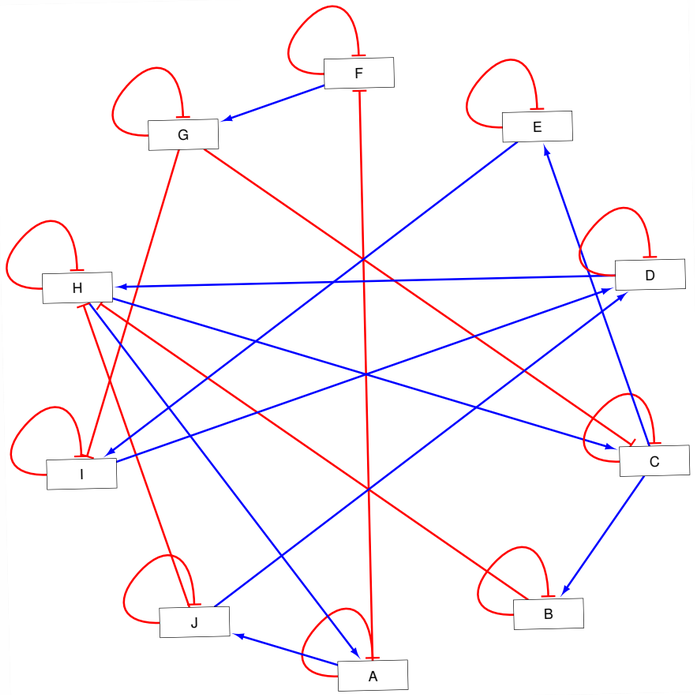
Informative

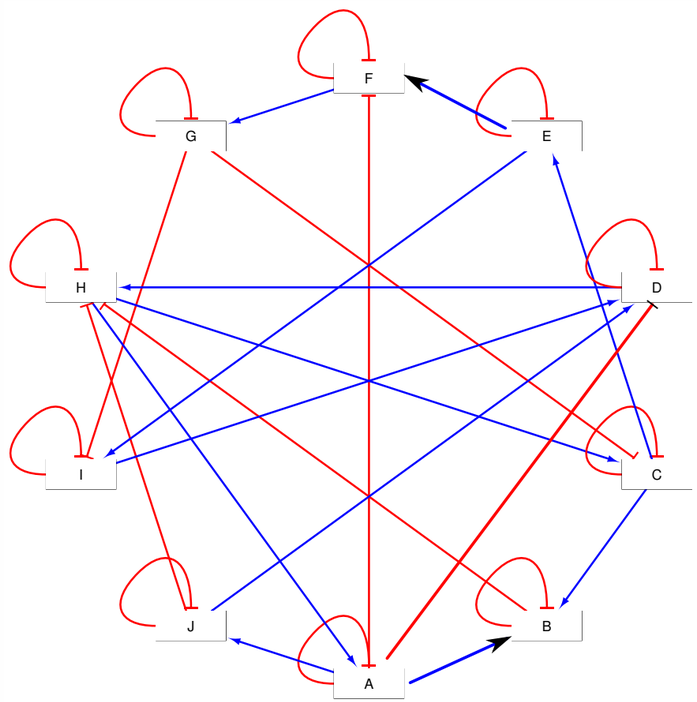
LOOCO

Sparsity

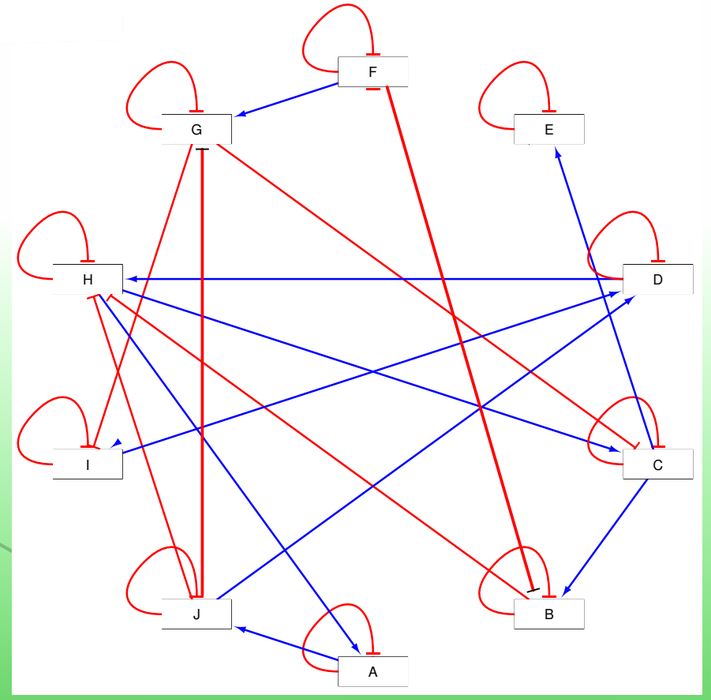


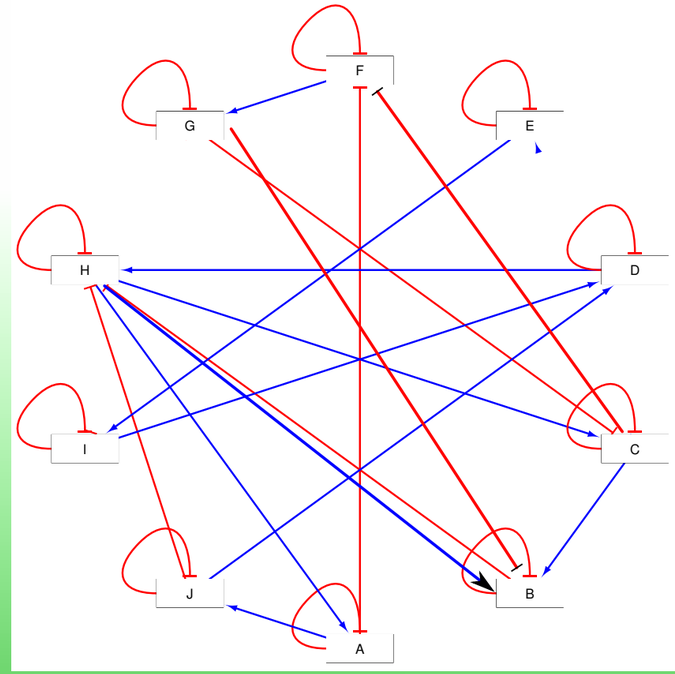
Unbiased Estimation


F
G
E
H
D
I
C
J
B
A


F
G
E
H
D
I
C
J
B
A


Models

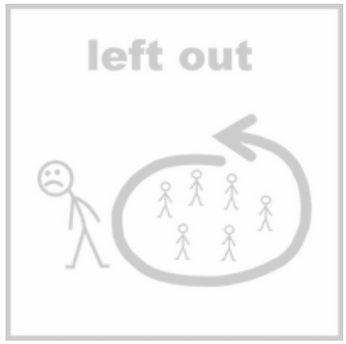

F
G
E
H
D
I
C
J
B
A


F
G
E
H
D
I
C
J
B
A

Data

Informative

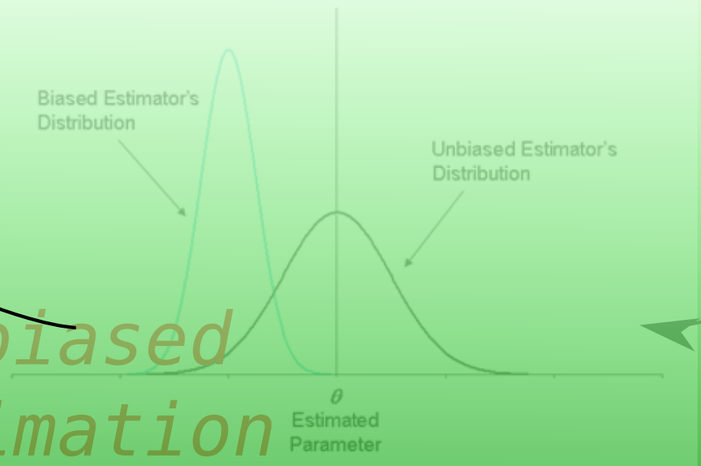
LOOCO



Predicting Data

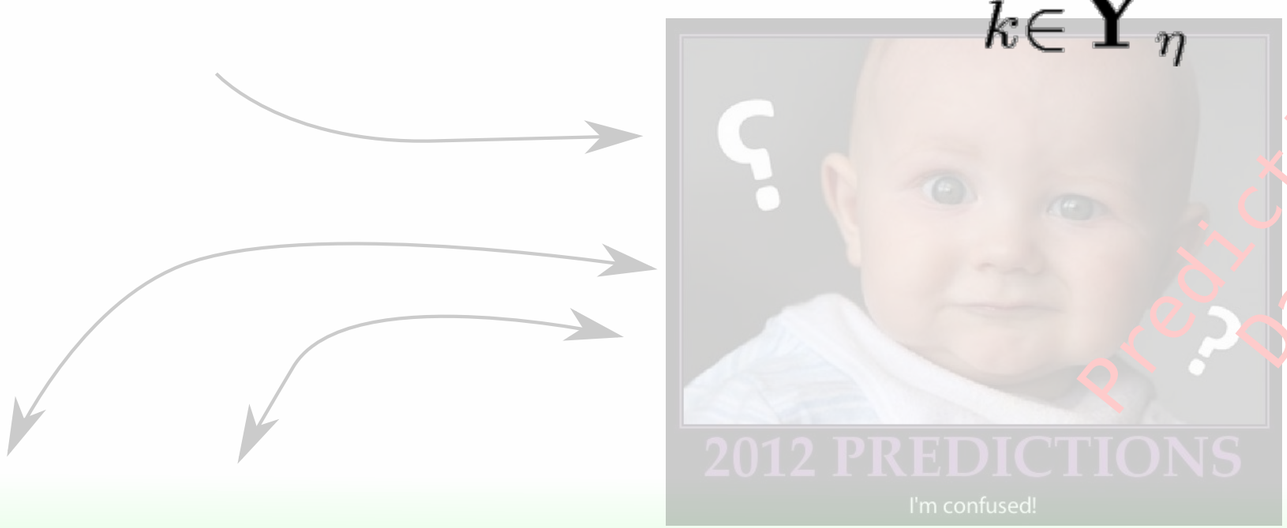
Sparsity

Models

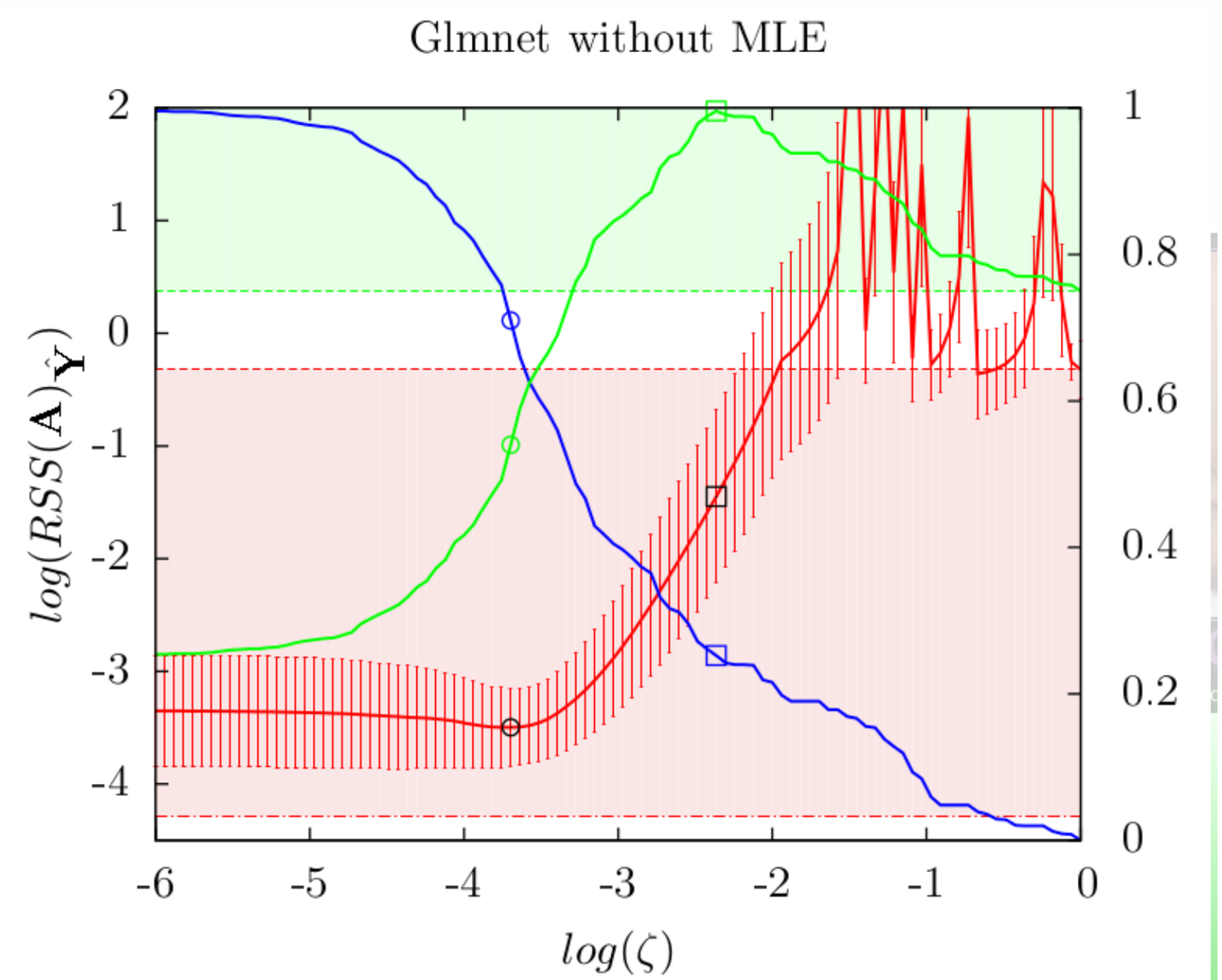


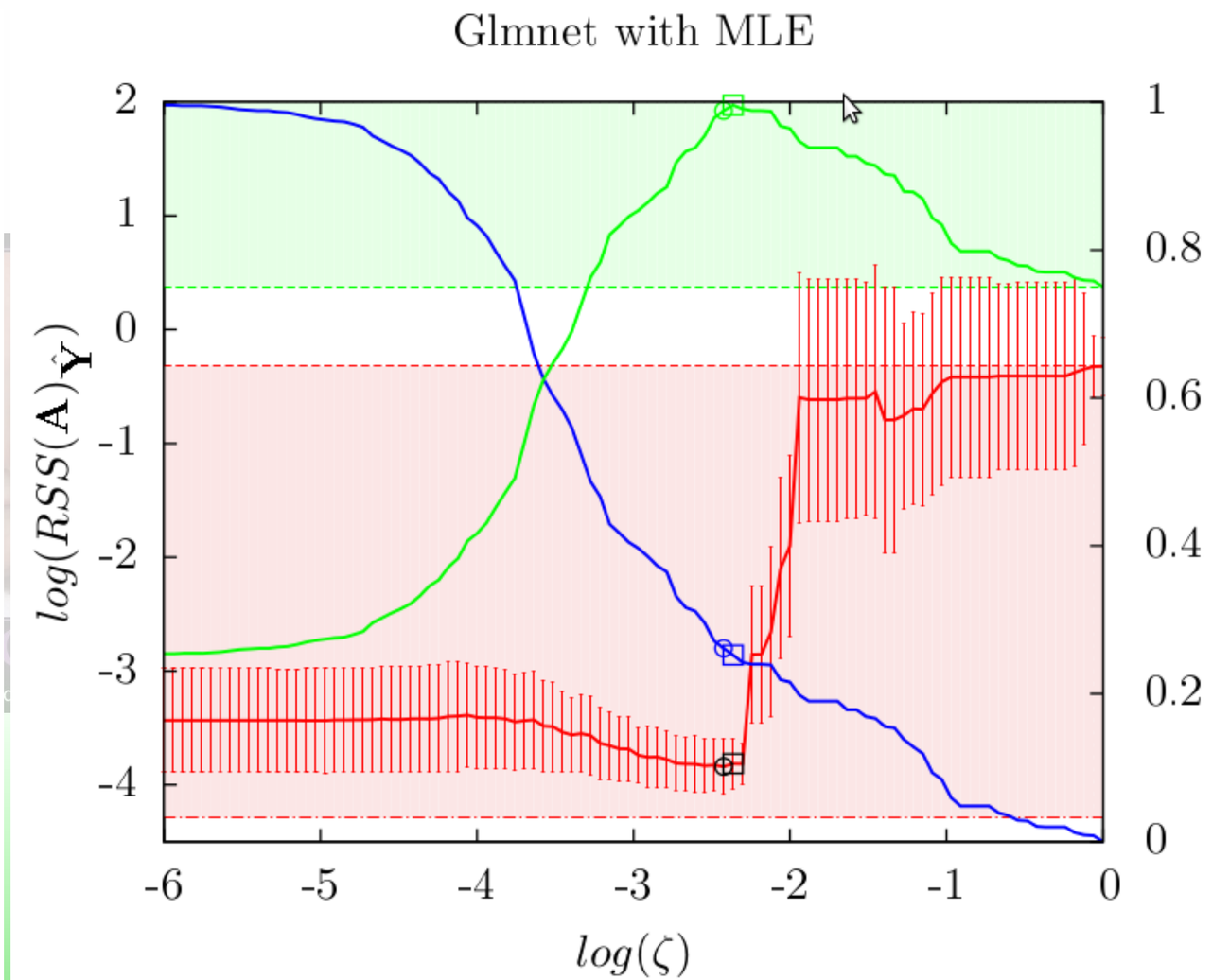
Unbiased Estimation

$$RSS = |\mathbf{Y}_\eta|^{-1} \sum_{k \in \mathbf{Y}_\eta} ||\hat{\mathbf{y}}_k - \mathbf{y}_k||^2$$



Glmnet predictions

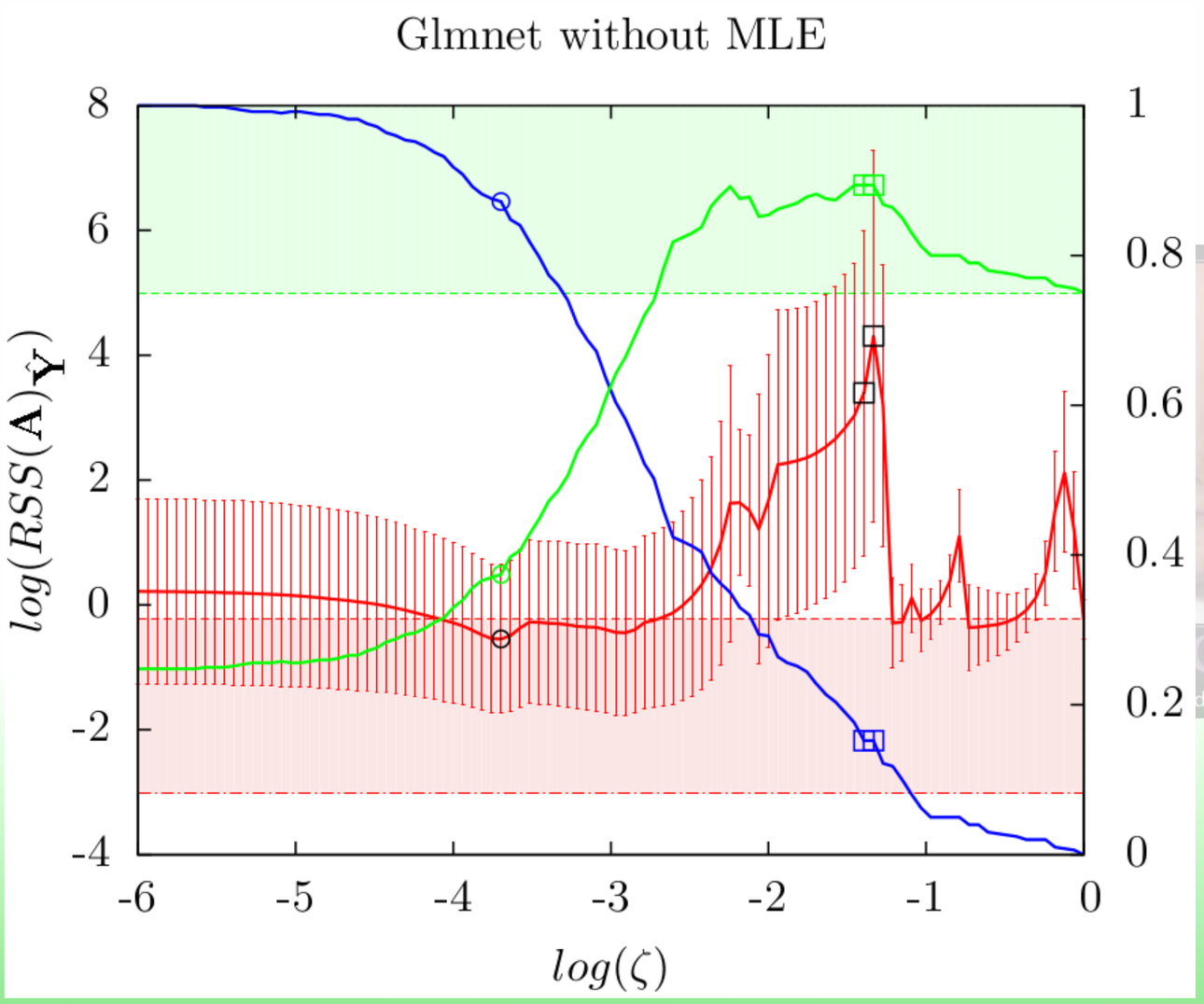


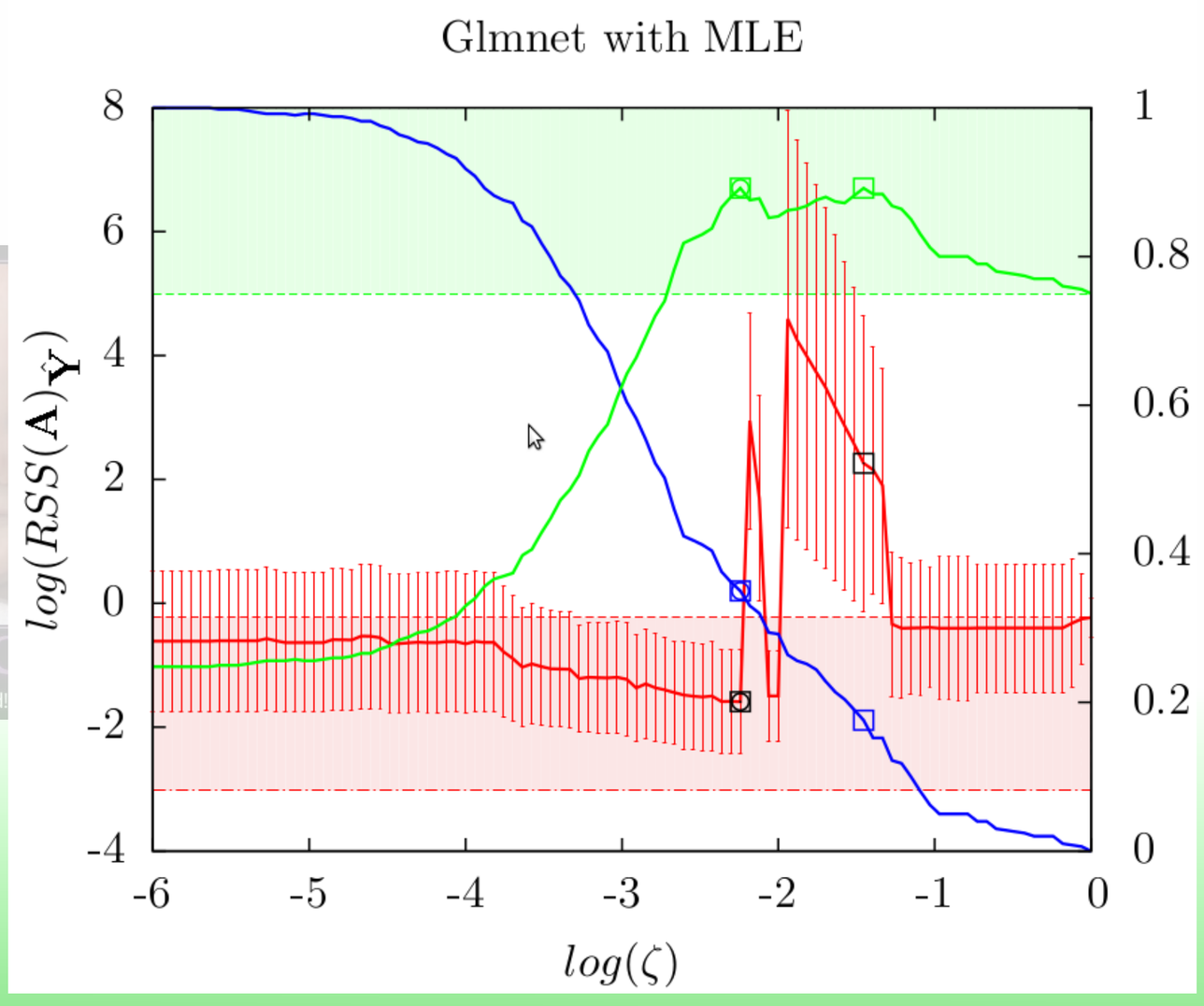


prediction error **similarity of signed topology** **fraction of links**

optimal similarity **minimum of prediction error**

Glmnet predictions

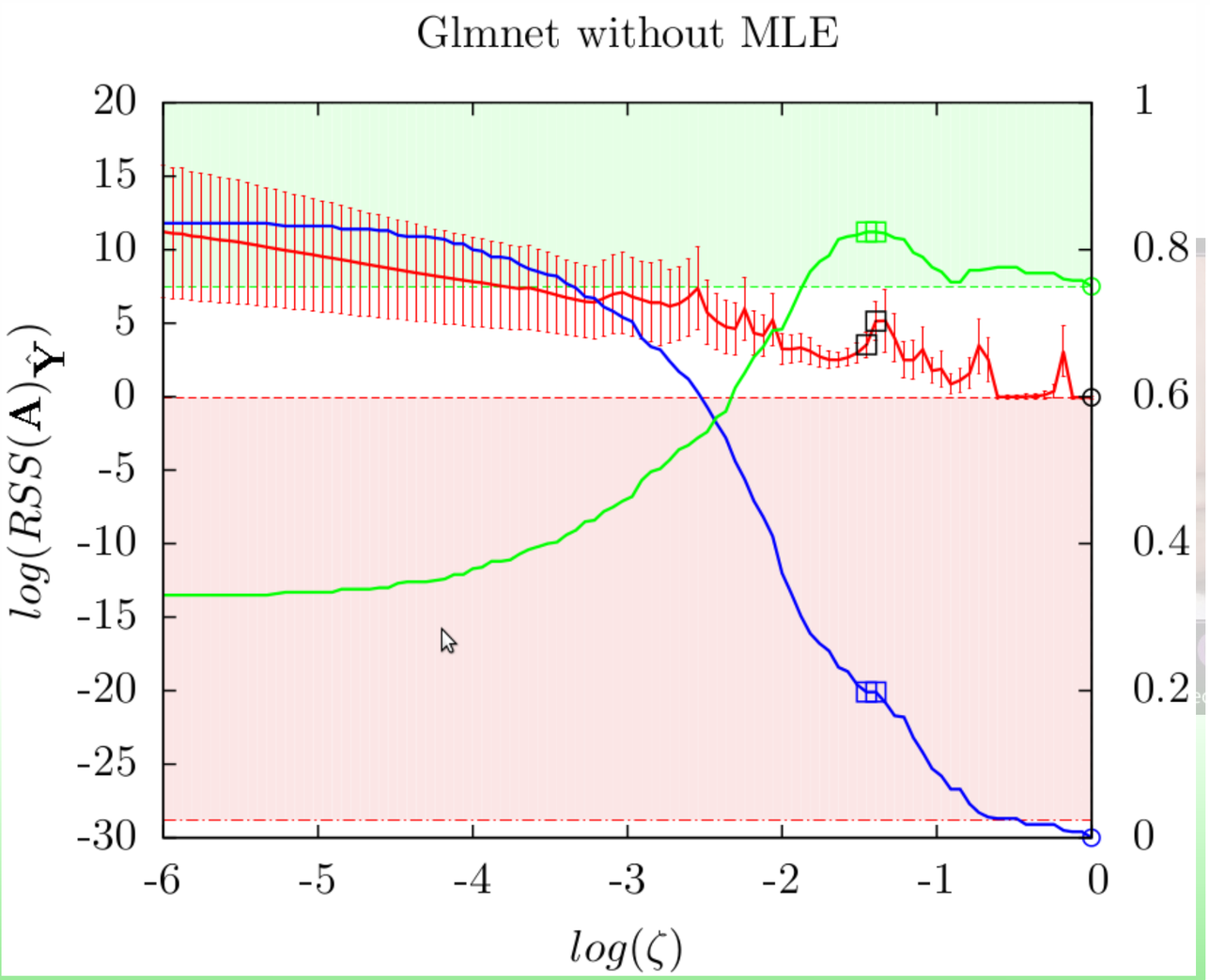


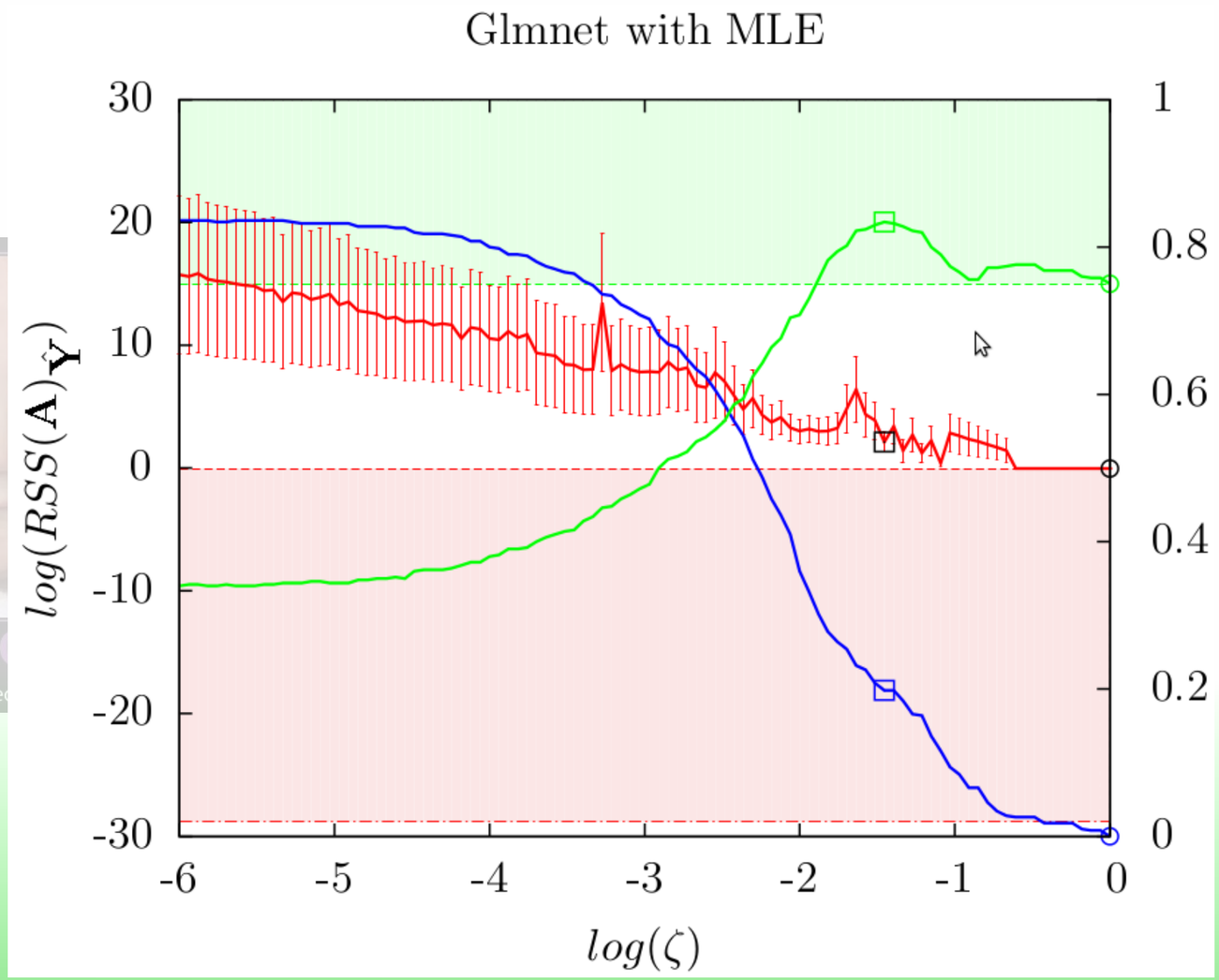


prediction error **similarity of signed topology** **fraction of links**

optimal similarity **minimum of prediction error**

Glmnet predictions





prediction error **similarity of signed topology** **fraction of links**

optimal similarity **minimum of prediction error**

Thanks for your attention