Bias–Variance decomposition of the error in bagged SVM ensembles: results and graphics

Giorgio Valentini
DSI - Dipartimento di Scienze dell' Informazione
Università degli Studi di Milano
e-mail: valentini@dsi.unimi.it

Contents

1 Experimental setup

In order to estimate the decomposition of the error in bias, unbiased and biased variance with bagged ensembles of SVMs we drew from each data set 100 sets of samples using a uniform probability distribution. From each extracted data set we drew 60 samples by bootstrap; each sample is composed by 100 patterns.

For each SVM model (specified by a choice of a kernel and its associated parameters) we trained 100 bagged ensembles using each time a different set of samples. To save room, each time a set of 60 samples is bootstrapped and a pseudorandom generator is initialized with the same seed before the generation of the first set of samples. In this way we can generate the same 100 sets of samples for each SVM model, storing only 1 set of samples (60 learning sets composed each one by 100 patterns) at a time.

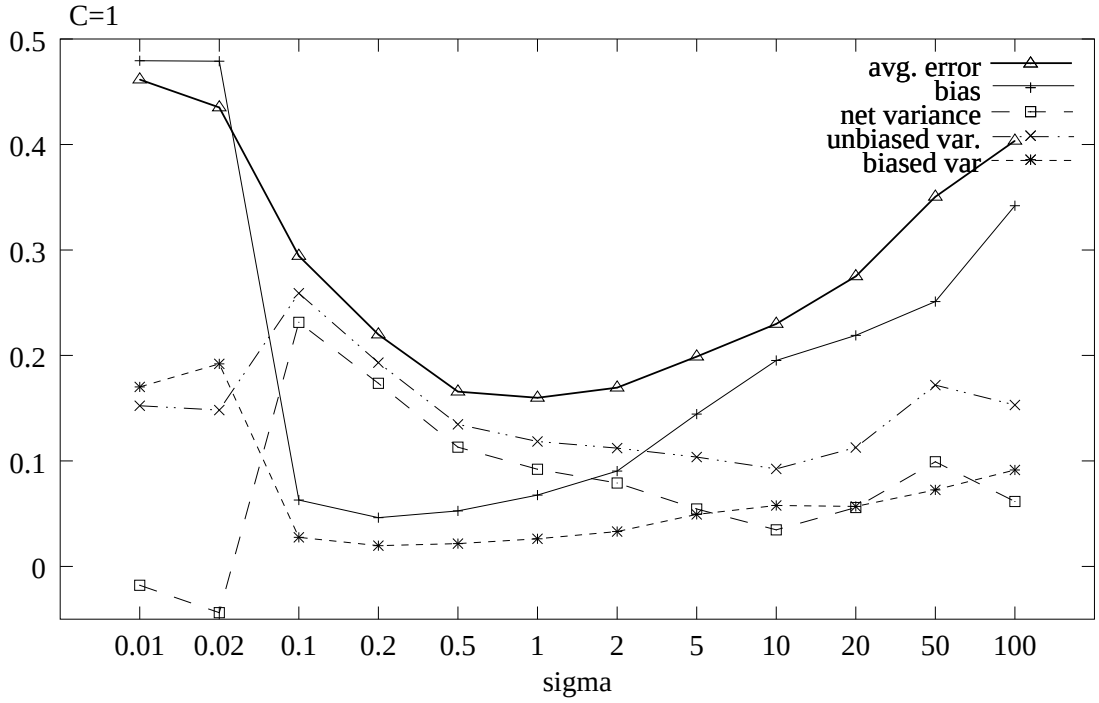
The bias-variance decomposition of the error is computed with respect to a separate test set significantly larger than the learning sets.

2 Bias–variance decomposition of the error in bagged SVM ensembles

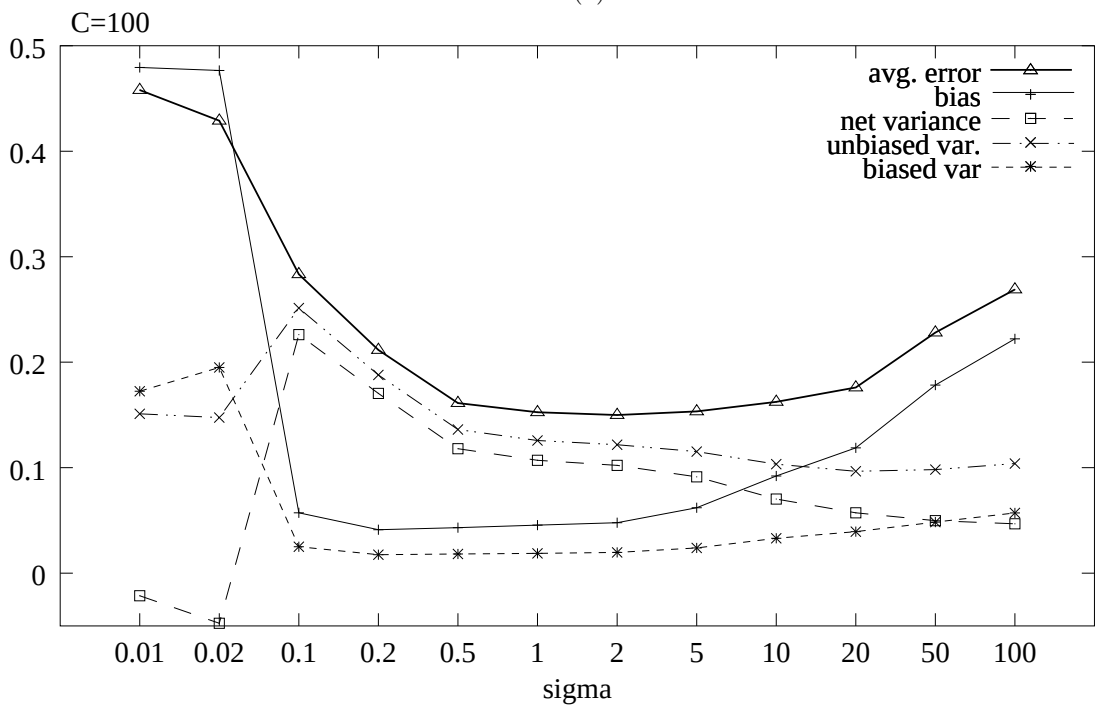
We evaluated about 110 different SVM models considering different combinations of the type of the kernel (gaussian, polynomial and dot-product) and its parameters for each data set. We trained and tested about 80000 different bagged SVM ensembles and a total of about 5 millions of single SVMs.

2.1 Decomposition in bagged RBF-SVMs

The decomposition of the error is represented with respect to different values of σ and for fixed values of C.

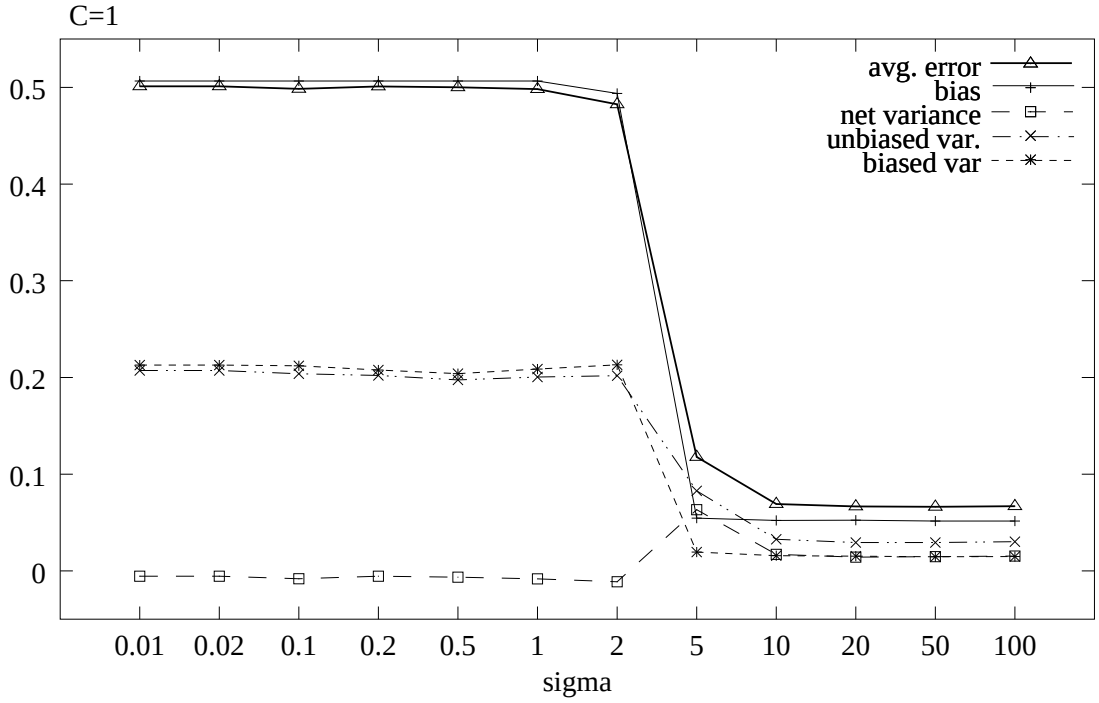


(a)

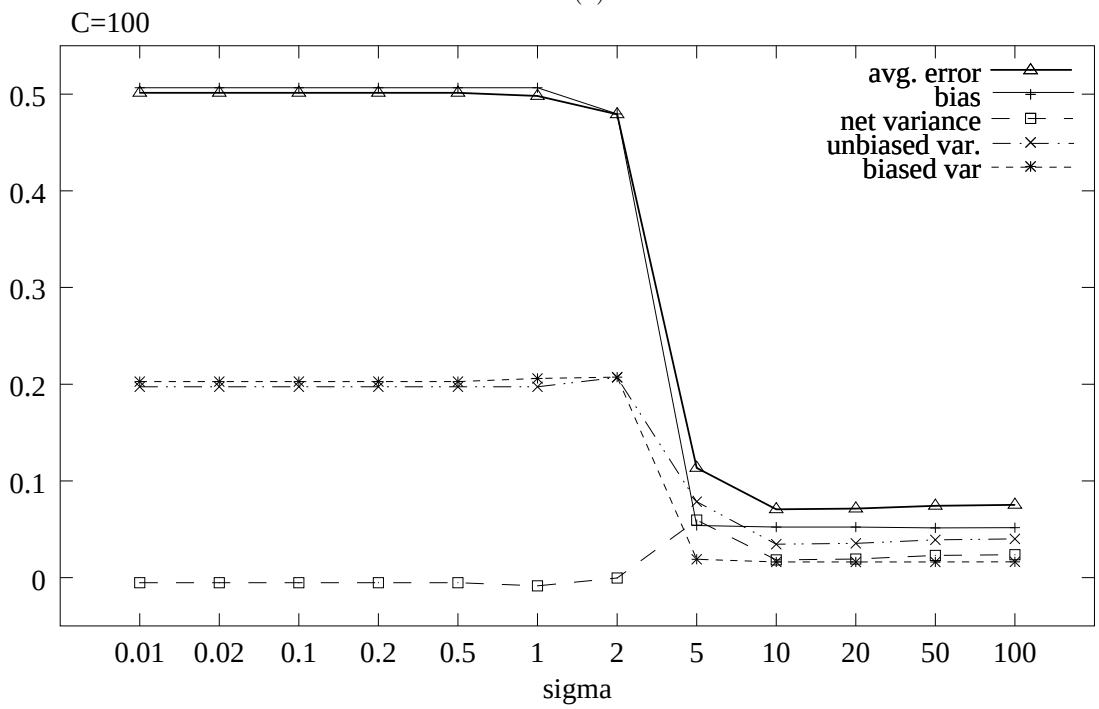


(b)

Figure 1: P2 data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF, while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

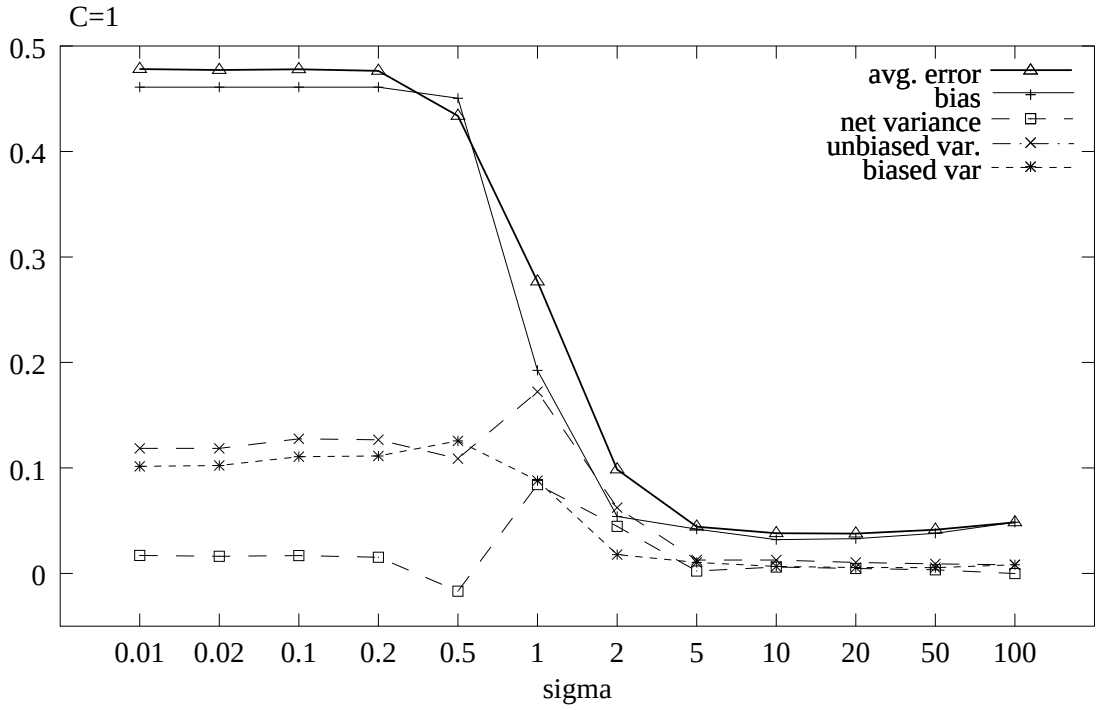


(a)

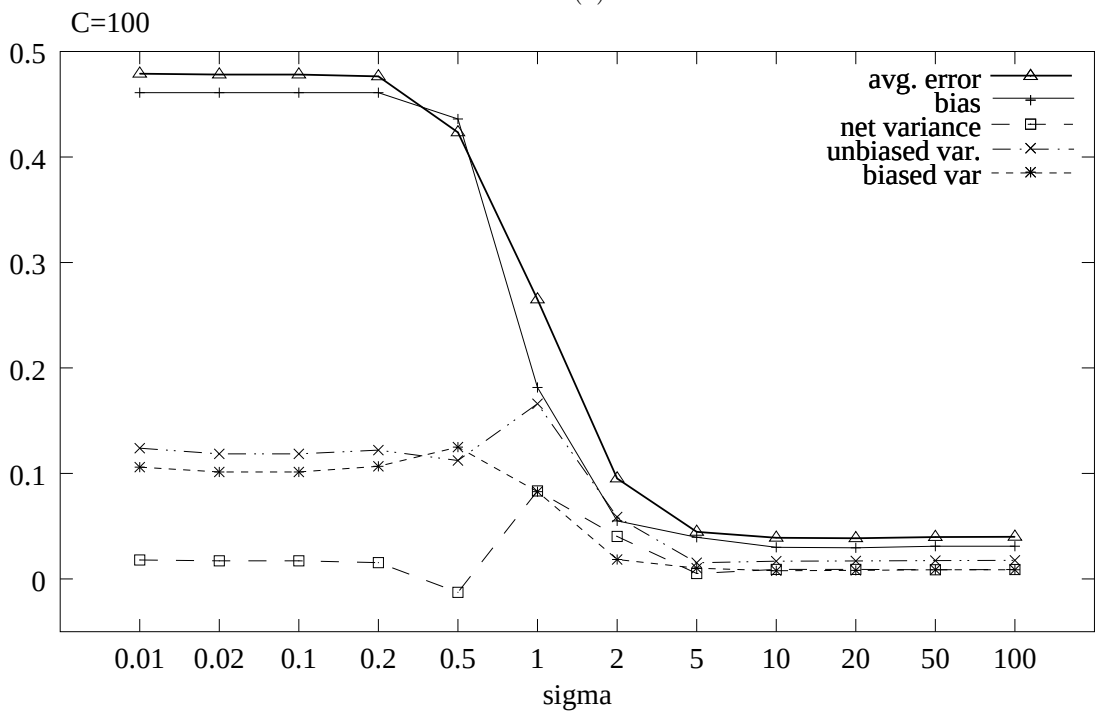


(b)

Figure 2: Waveform data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF, while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

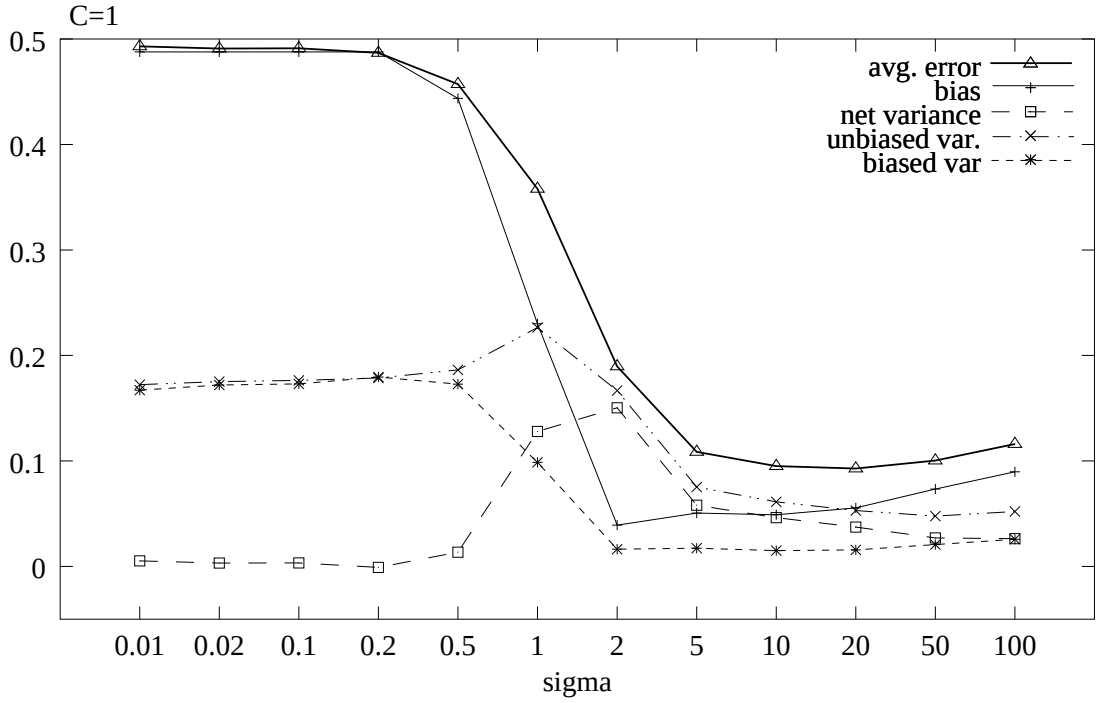


(a)

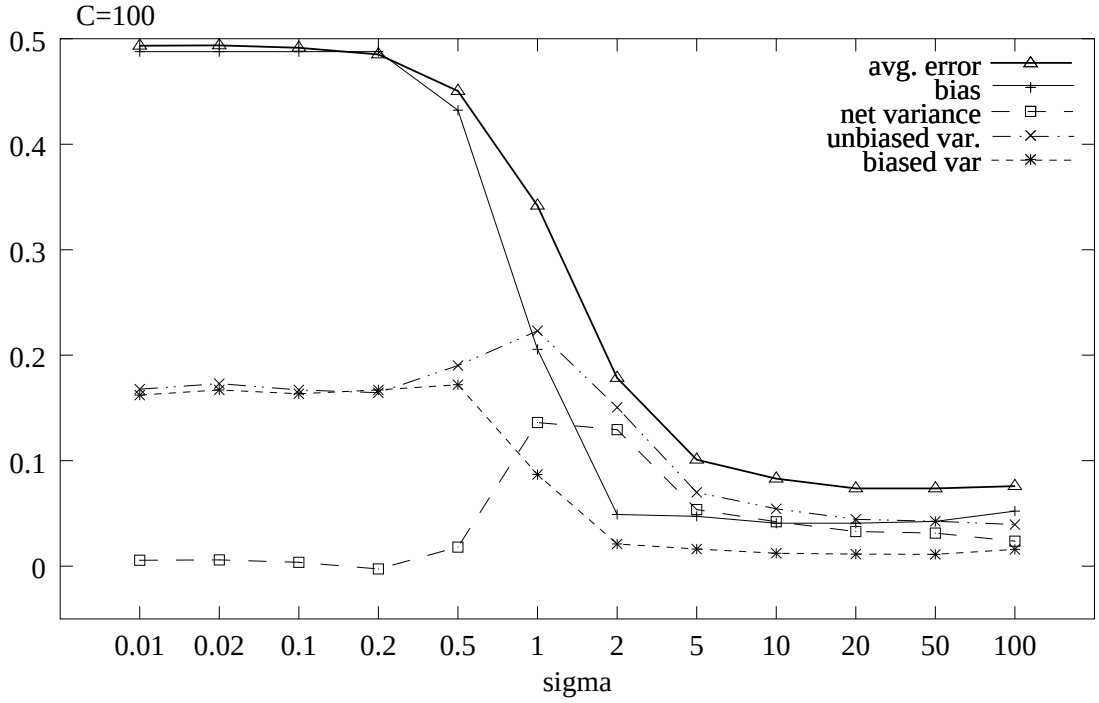


(b)

Figure 3: Grey-Landsat data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF, while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

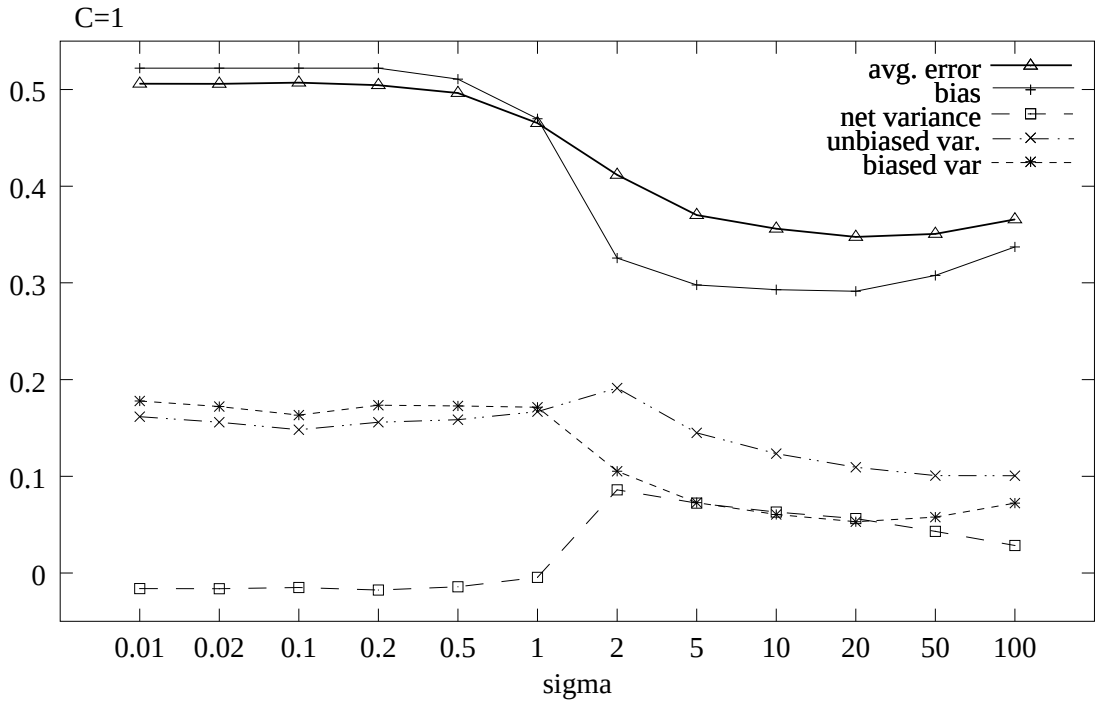


(a)

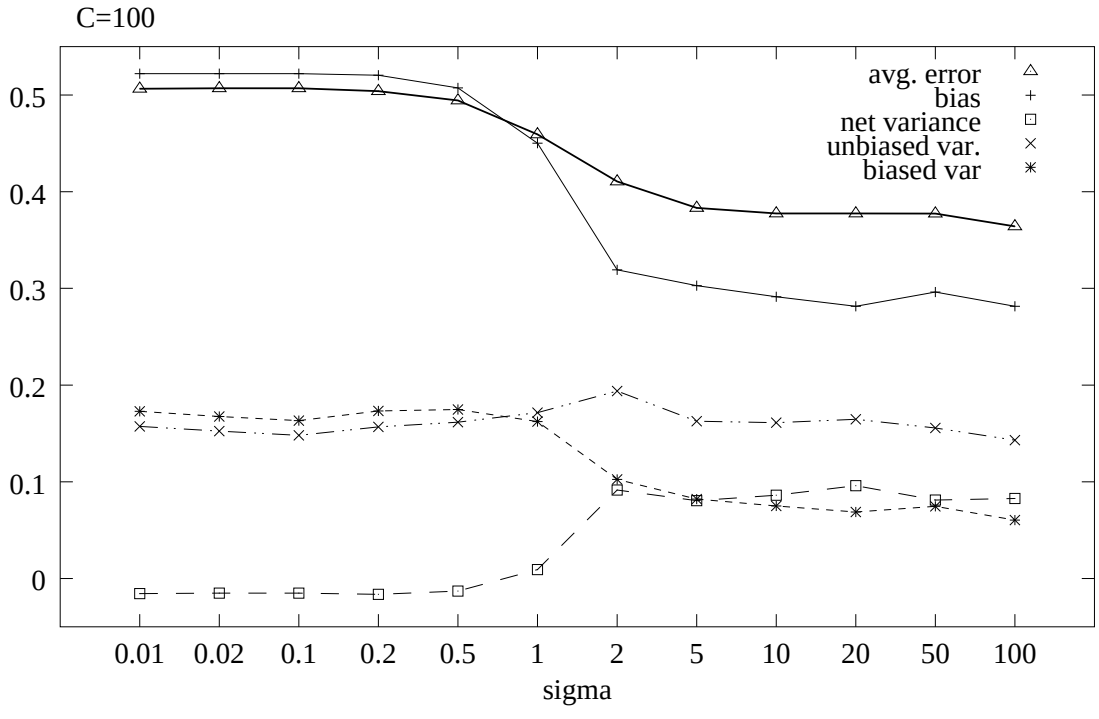


(b)

Figure 4: Letter-Two data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF, while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

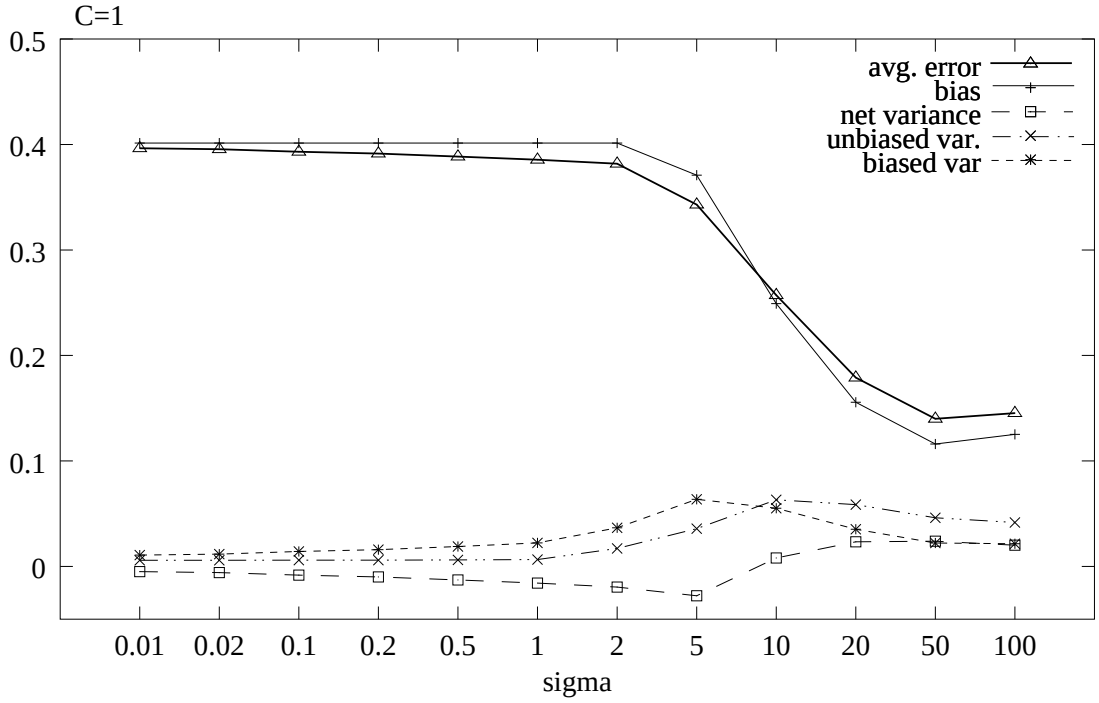


(a)

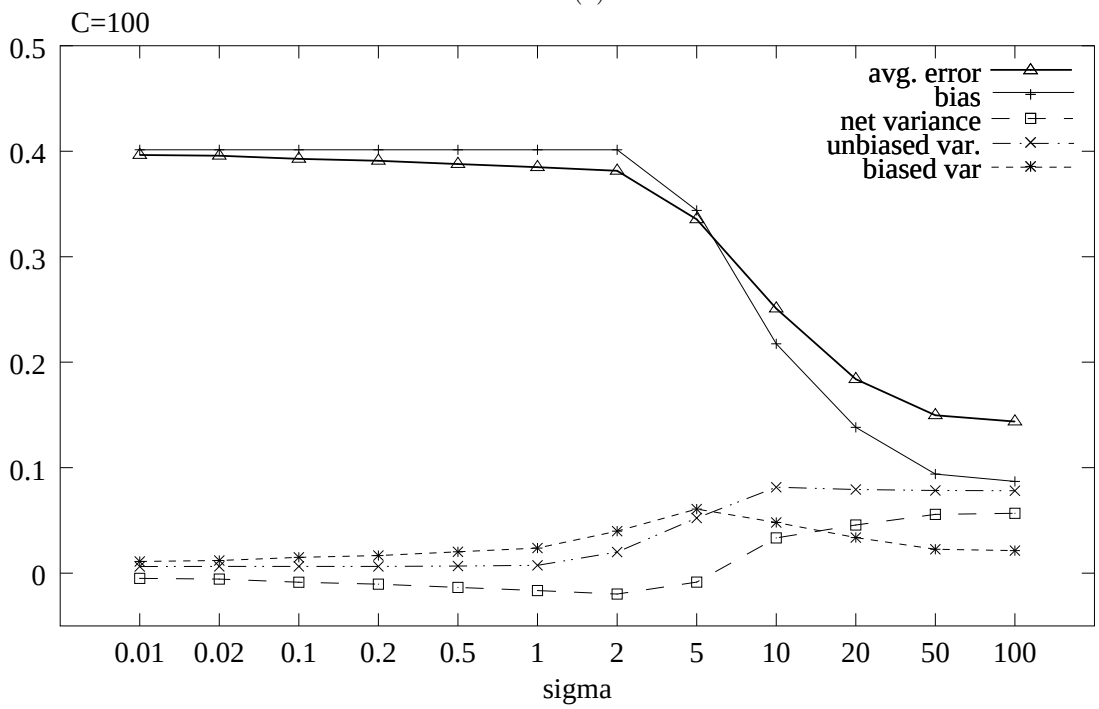


(b)

Figure 5: Letter-Two with noise data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF, while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

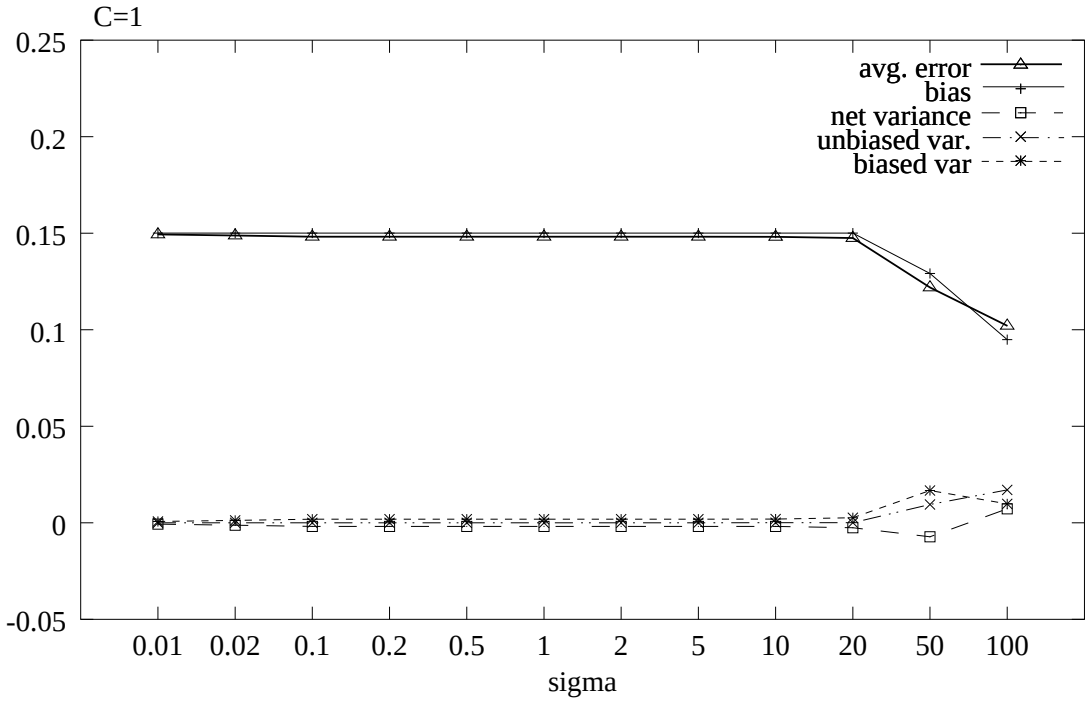


(a)

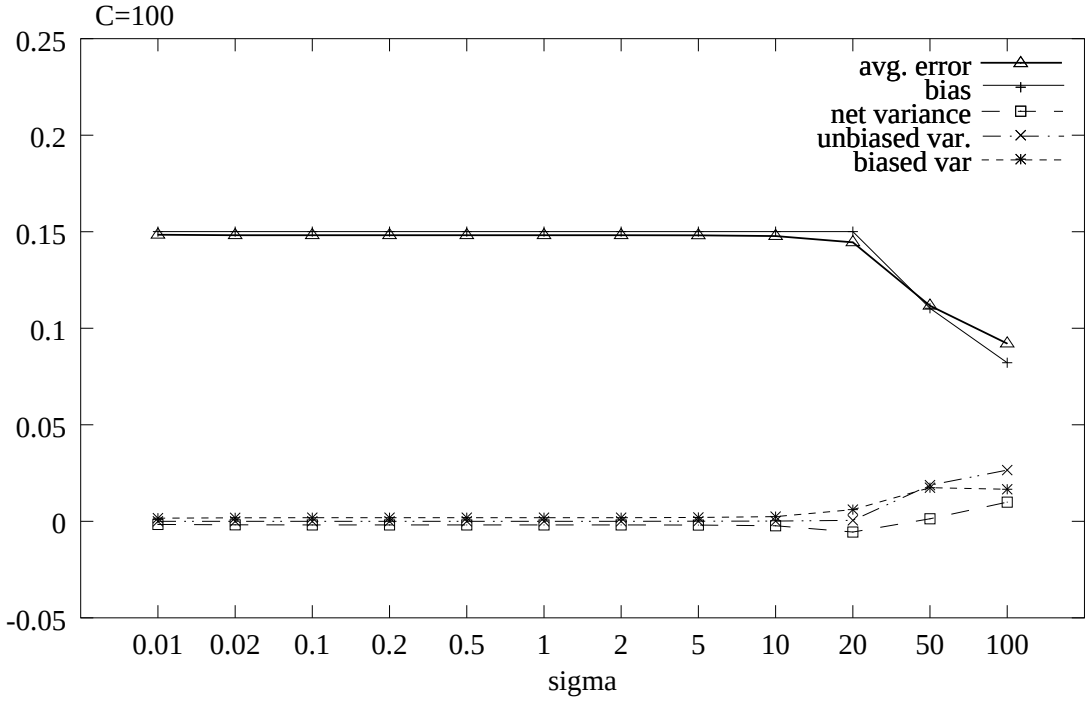


(b)

Figure 6: Spam data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF, while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.



(a)

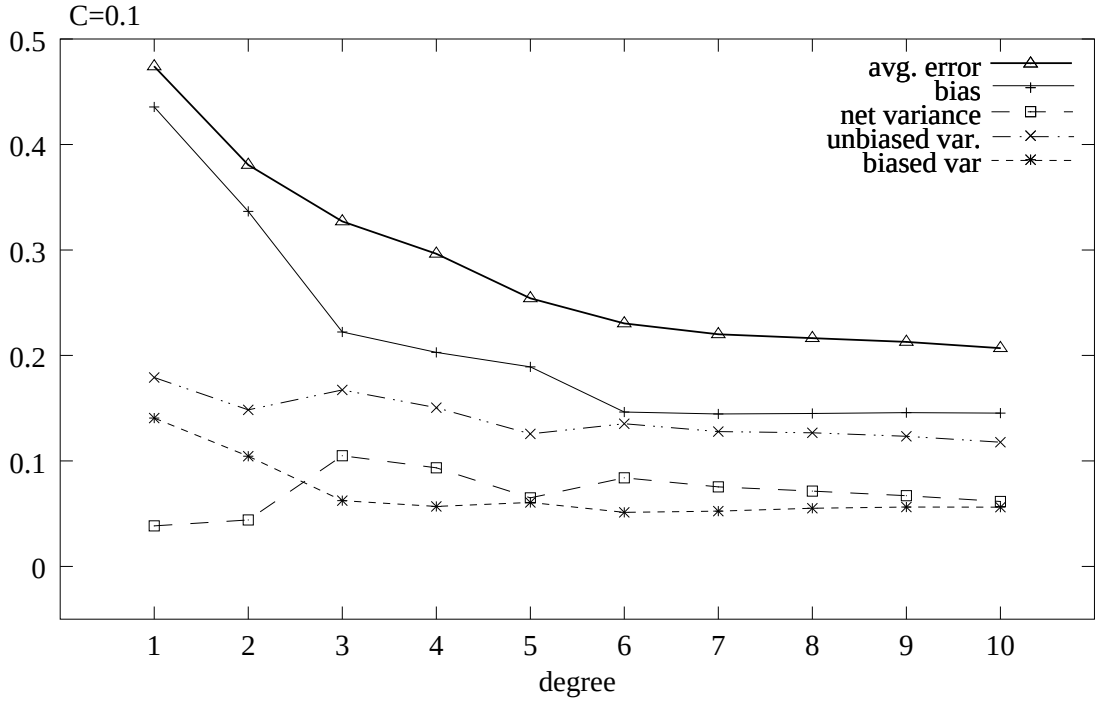


(b)

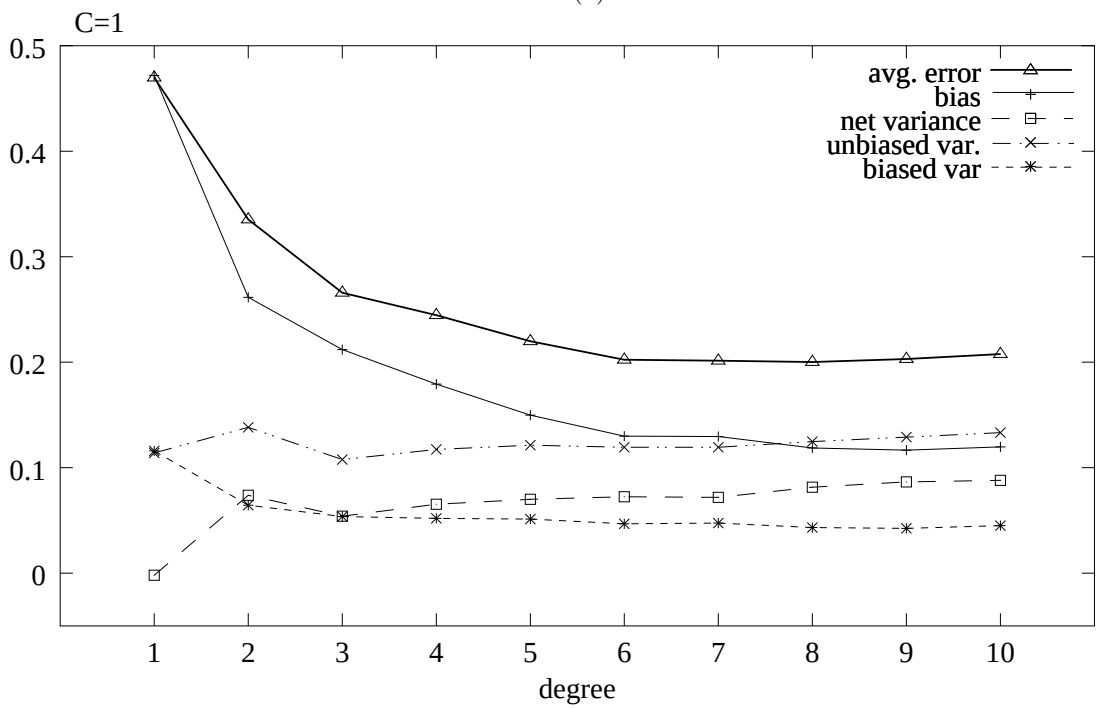
Figure 7: Musk data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF, while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

2.2 Decomposition in bagged Polynomial SVMs

The decomposition of the error is represented with respect to different values of the polynomial degree and for fixed values of C.

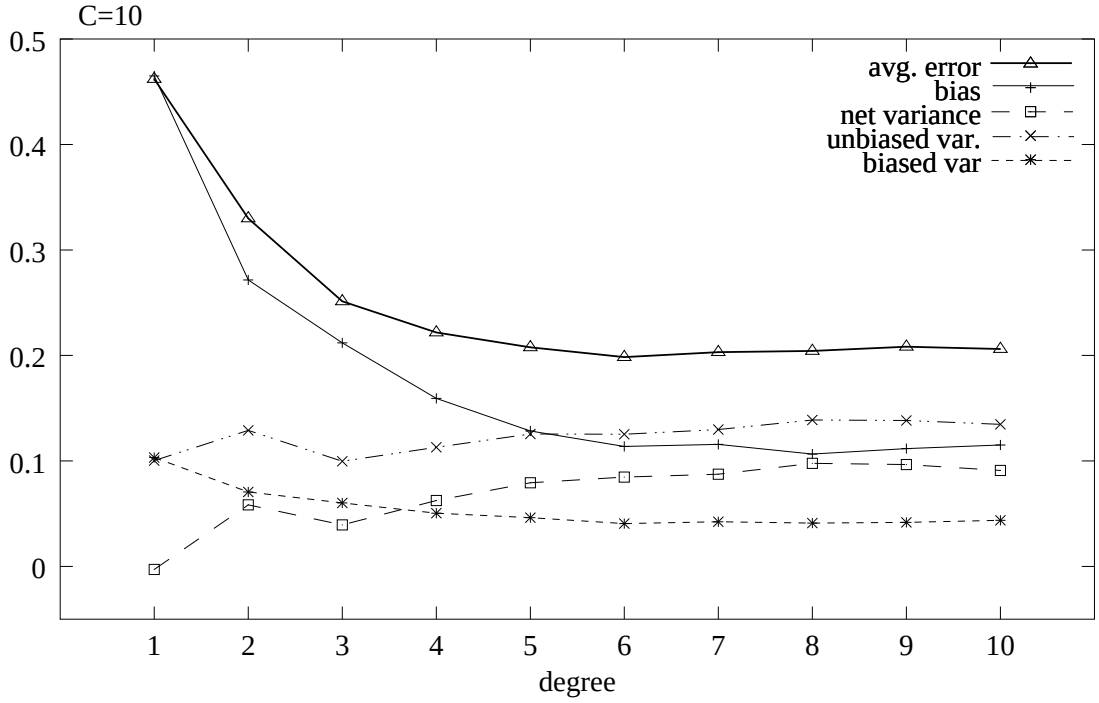


(a)

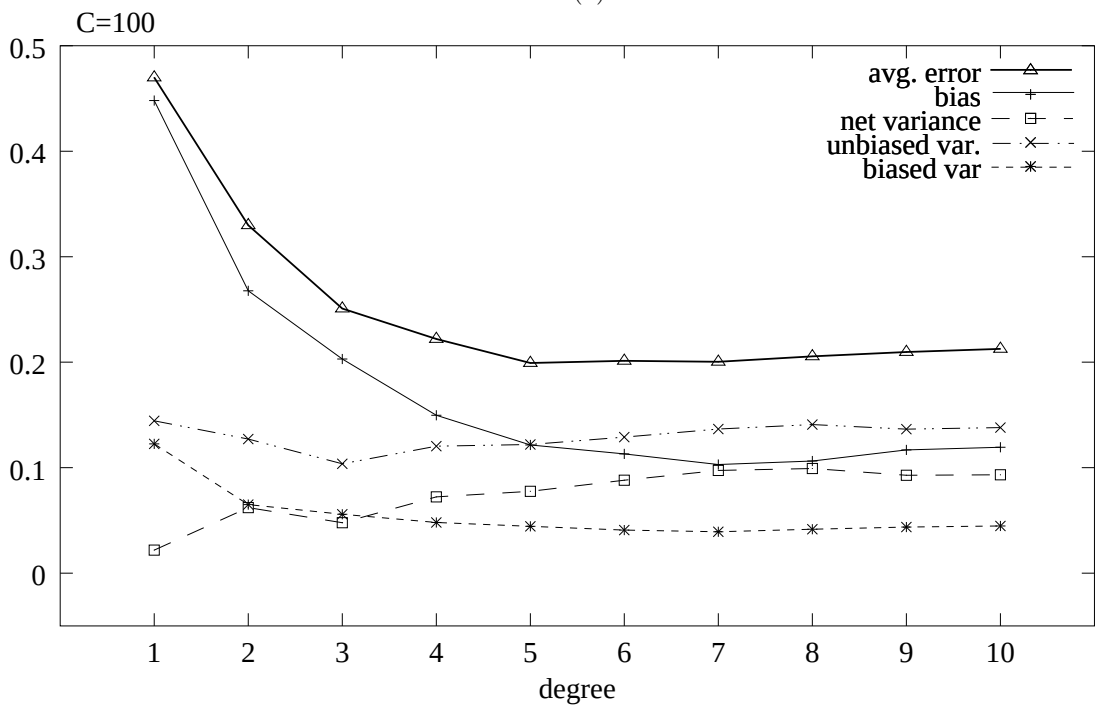


(b)

Figure 8: P2 data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 1$.

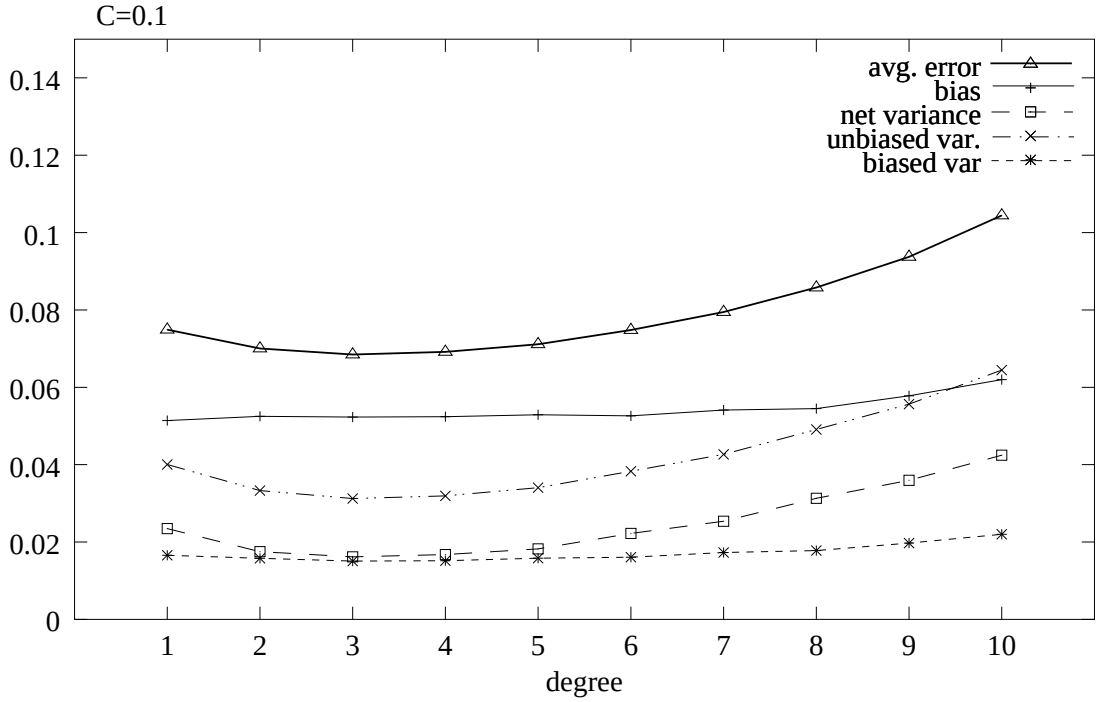


(a)

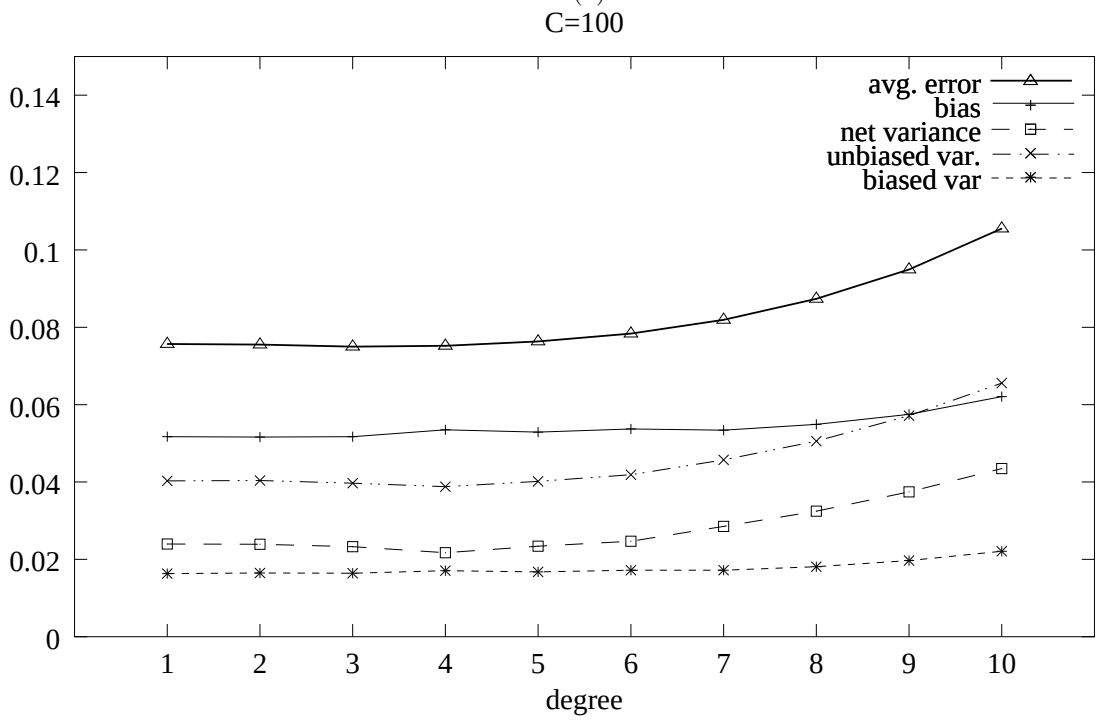


(b)

Figure 9: P2 data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 10$, (b) $C = 100$.

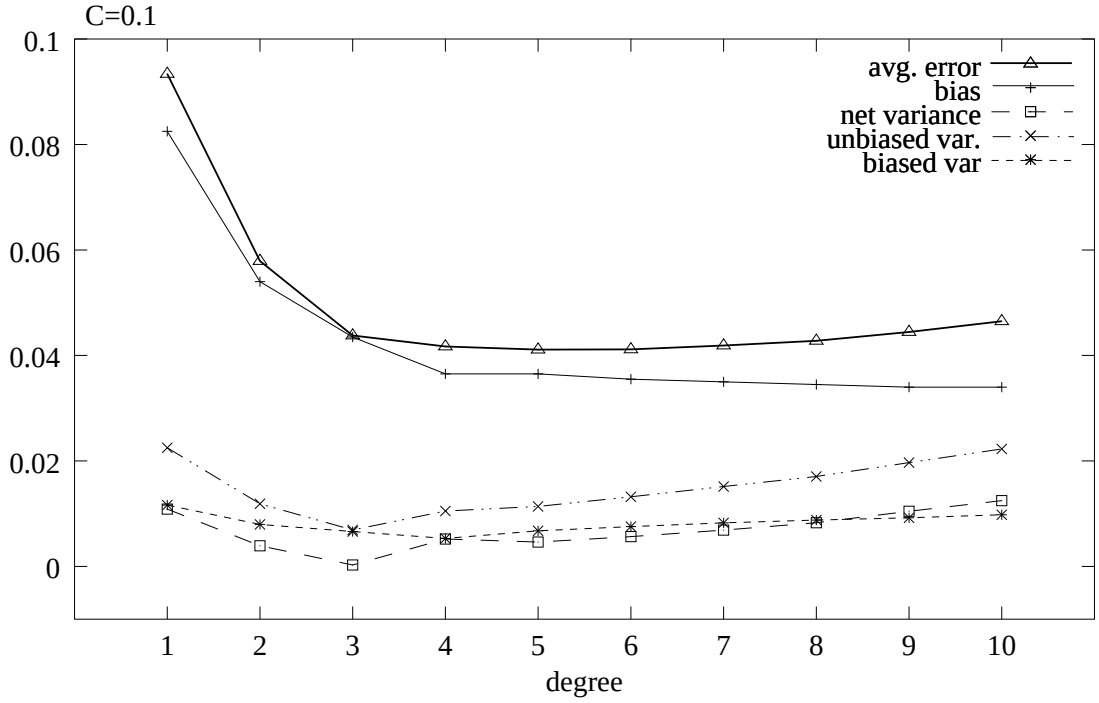


(a)

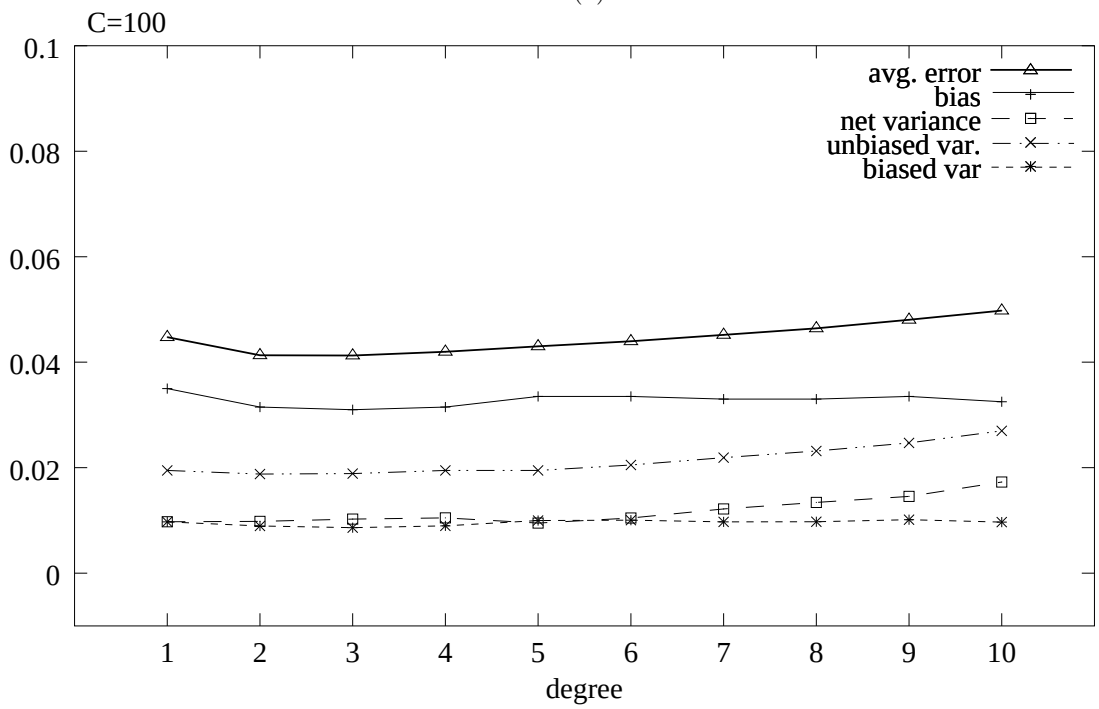


(b)

Figure 10: Waveform data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 100$.

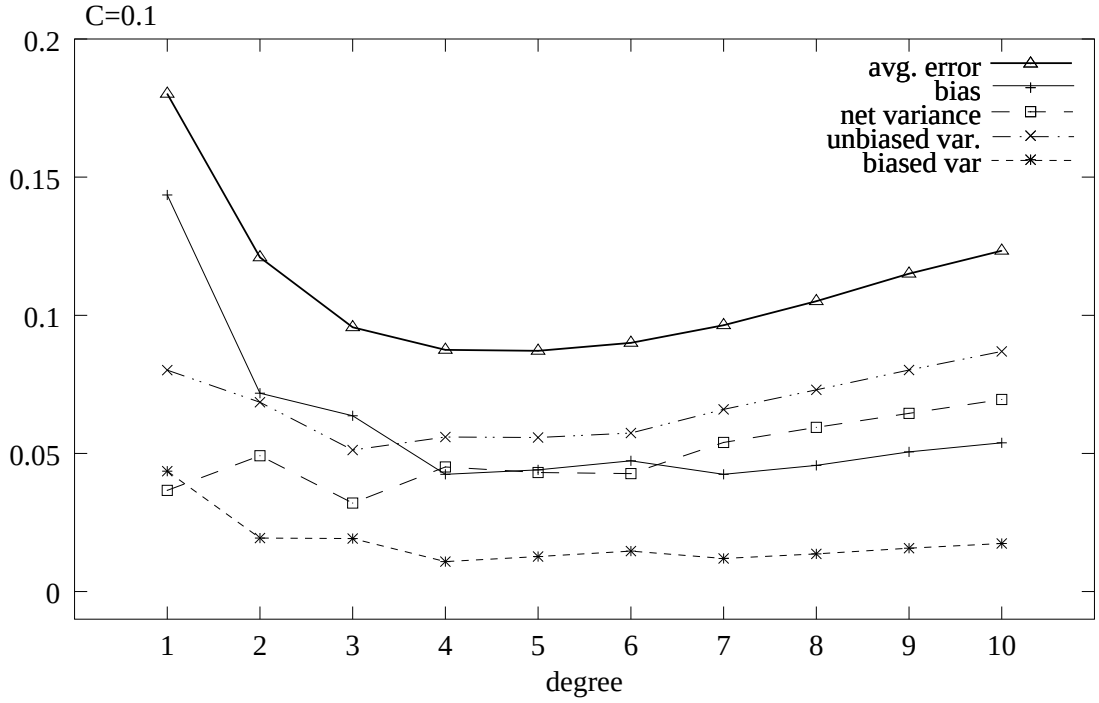


(a)

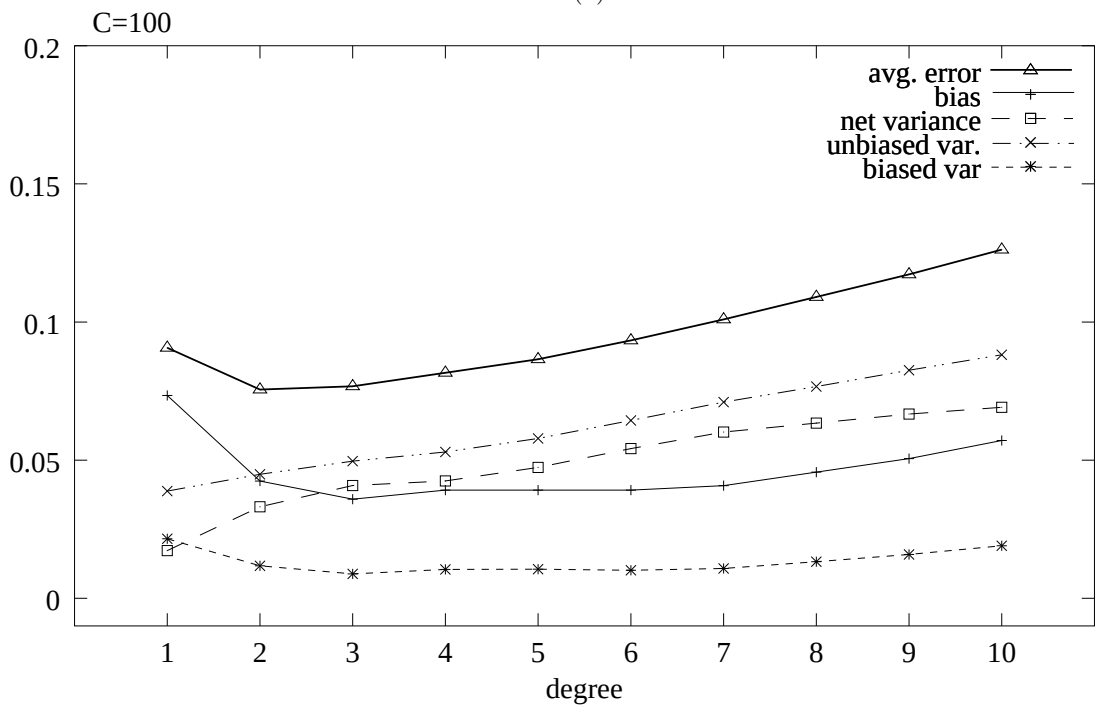


(b)

Figure 11: Grey-Landsat data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 100$.

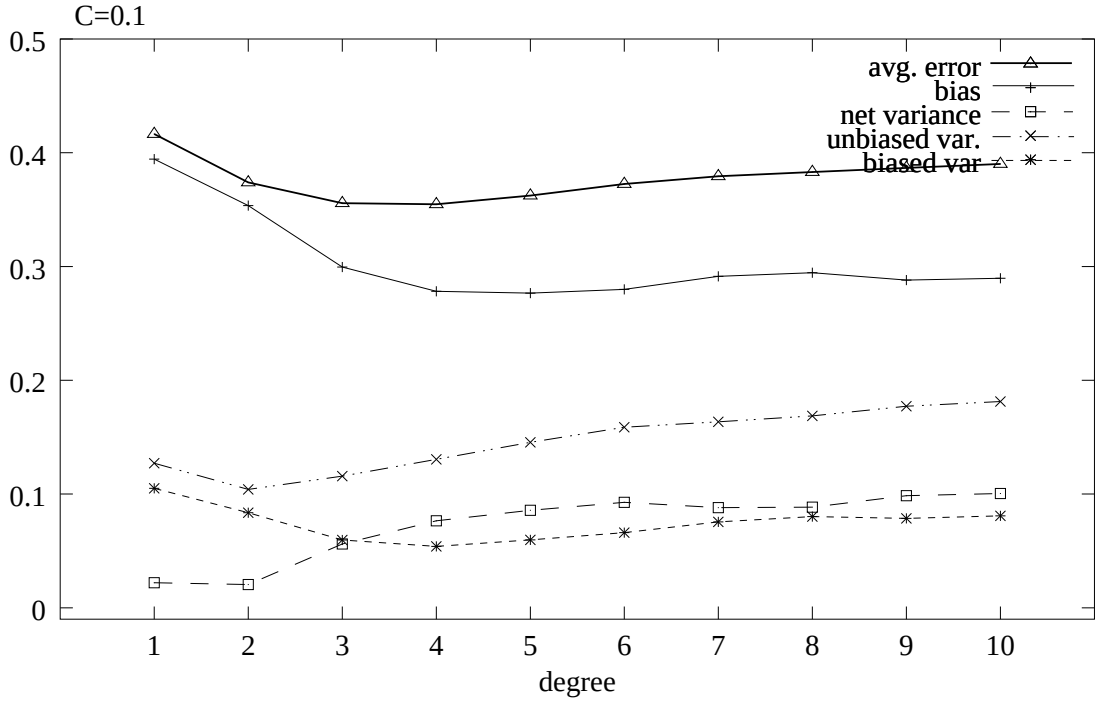


(a)

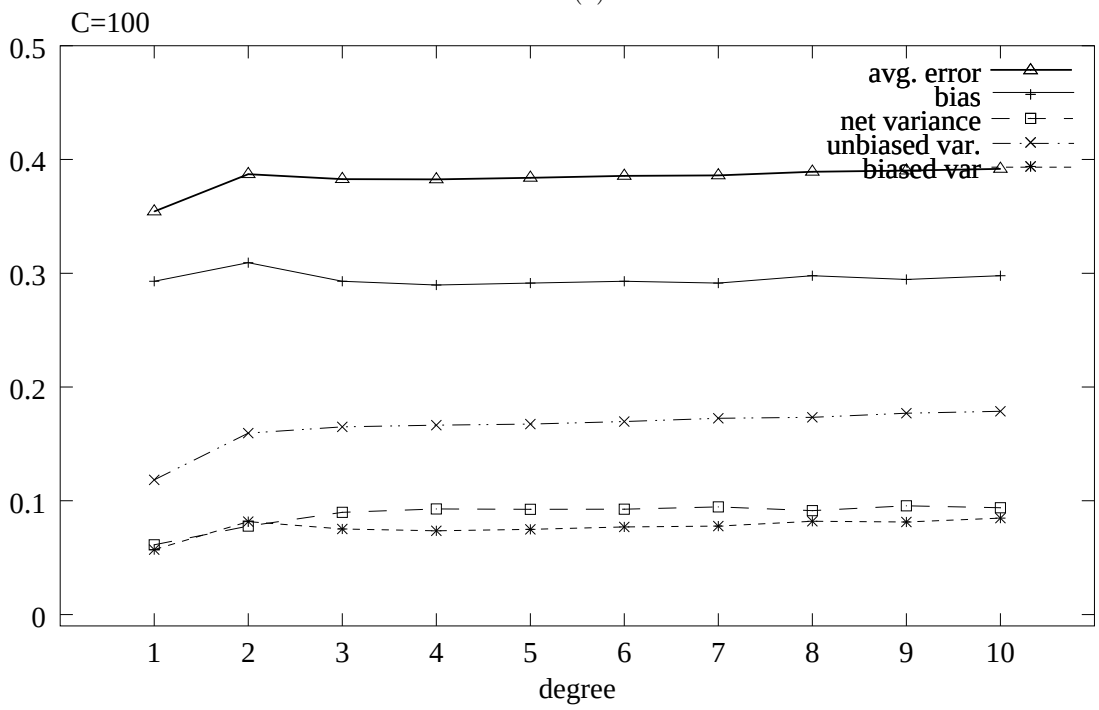


(b)

Figure 12: Letter-Two data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 100$.

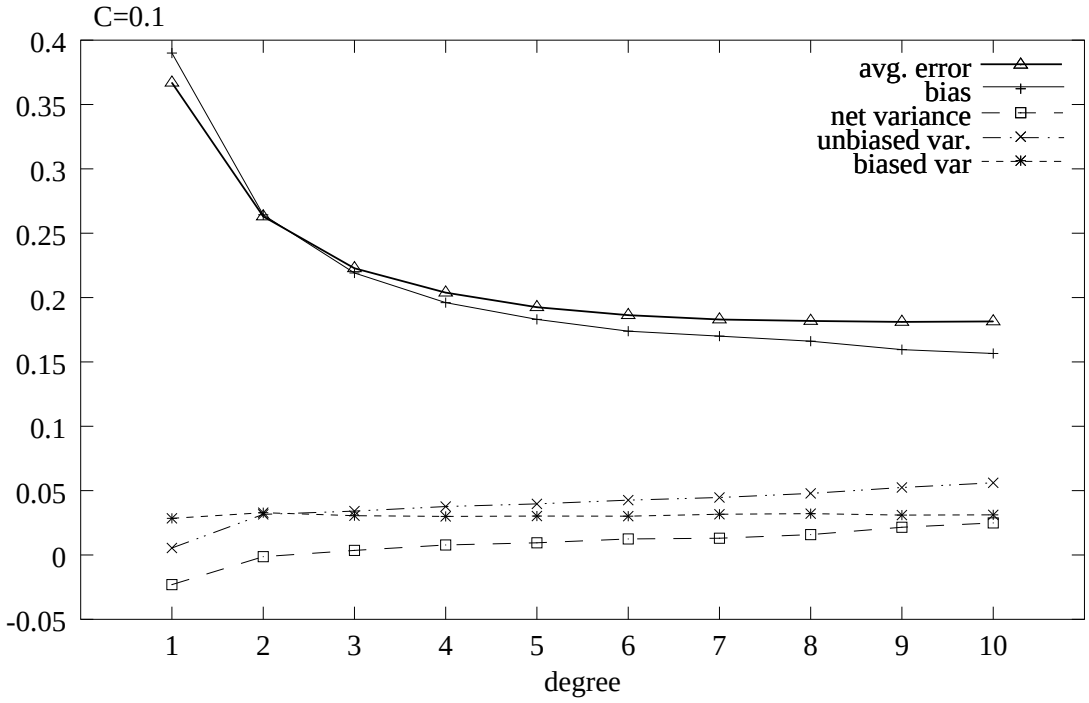


(a)

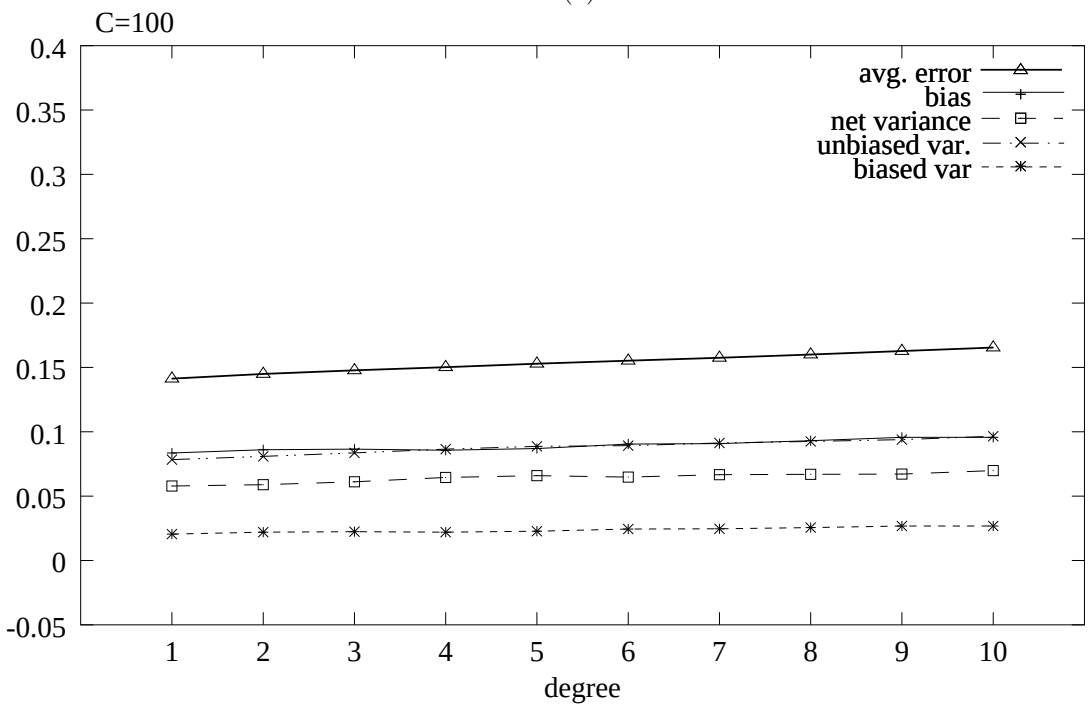


(b)

Figure 13: Letter-Two with noise data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 100$.

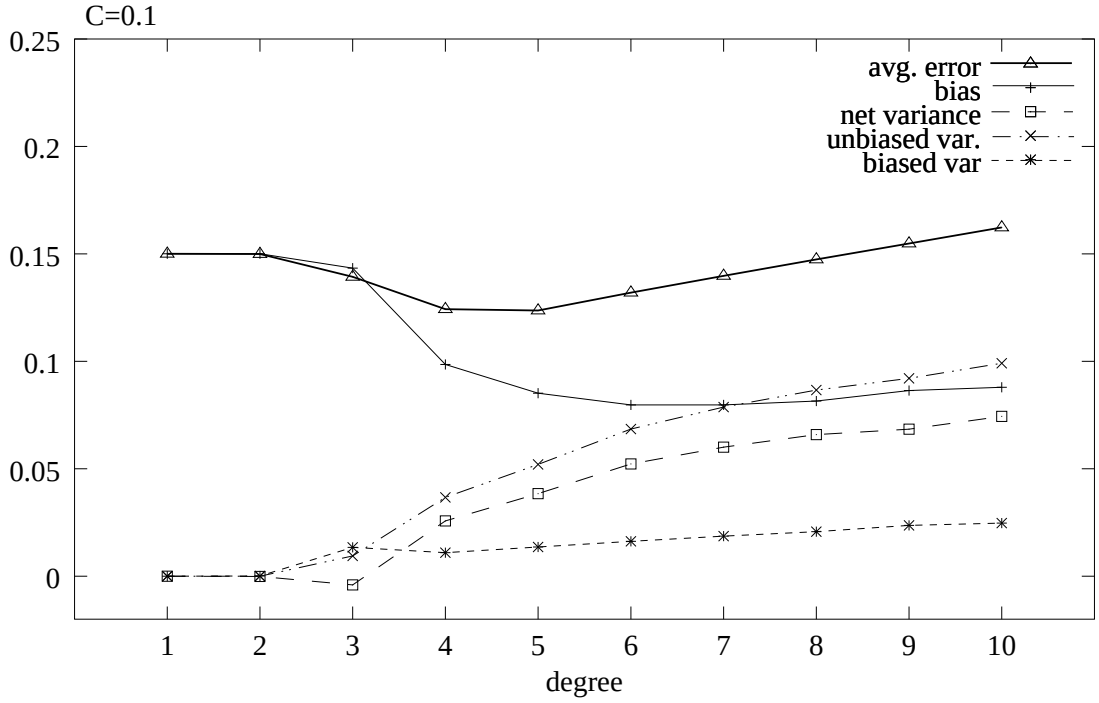


(a)

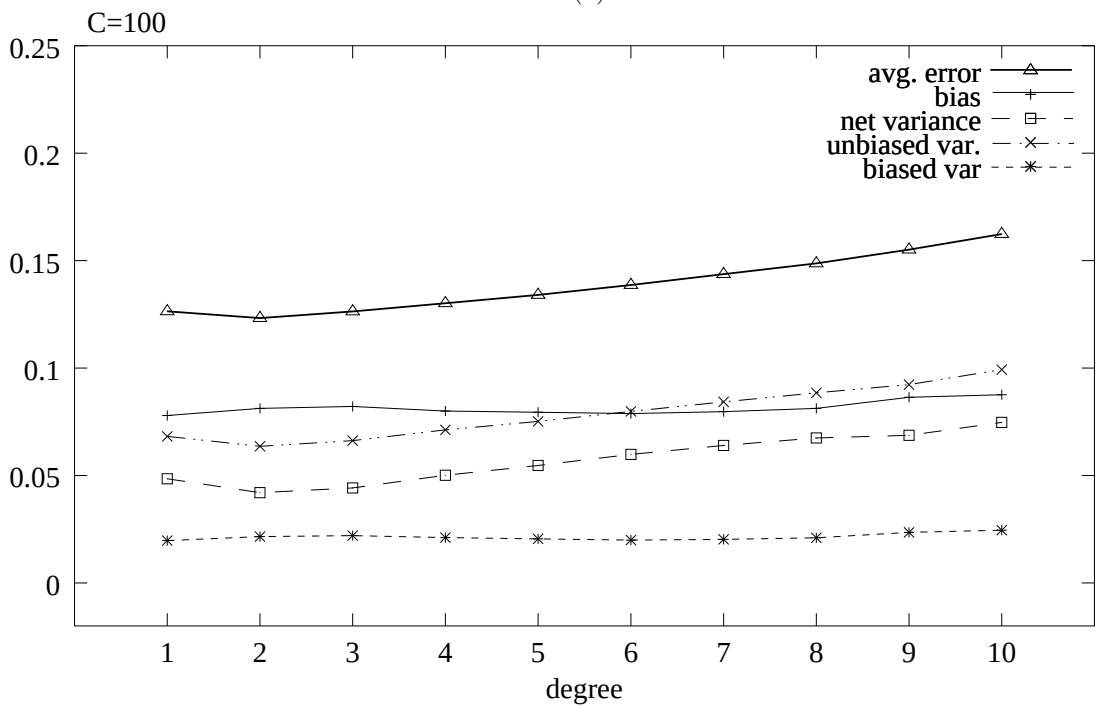


(b)

Figure 14: Spam data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 100$.



(a)

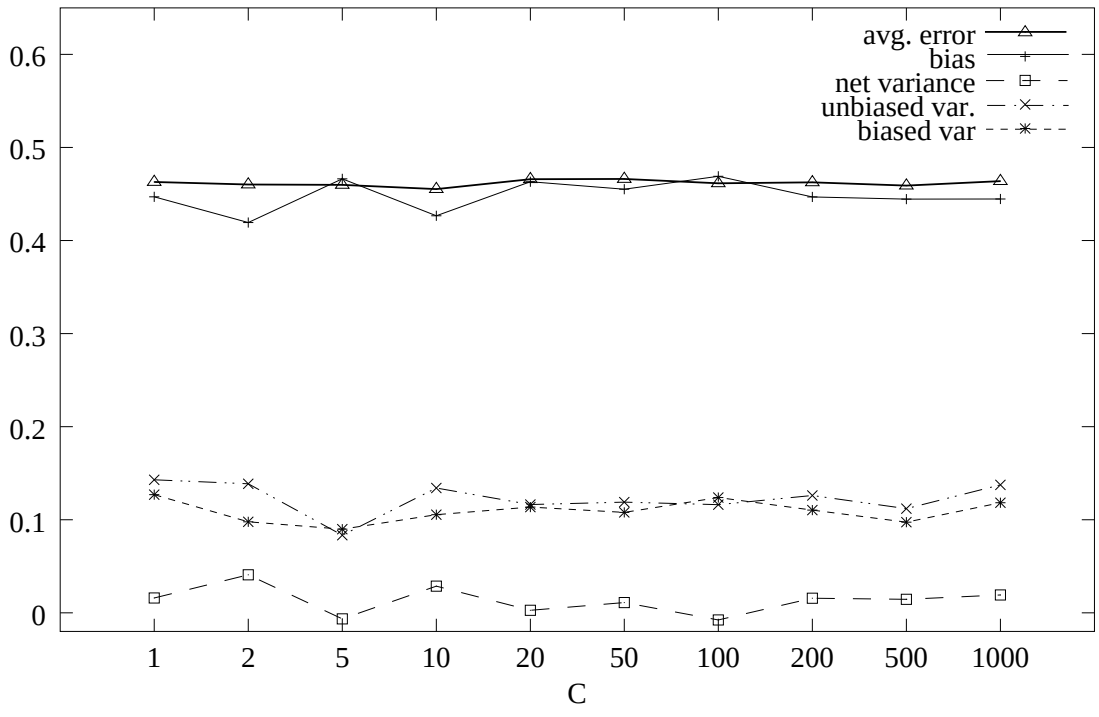


(b)

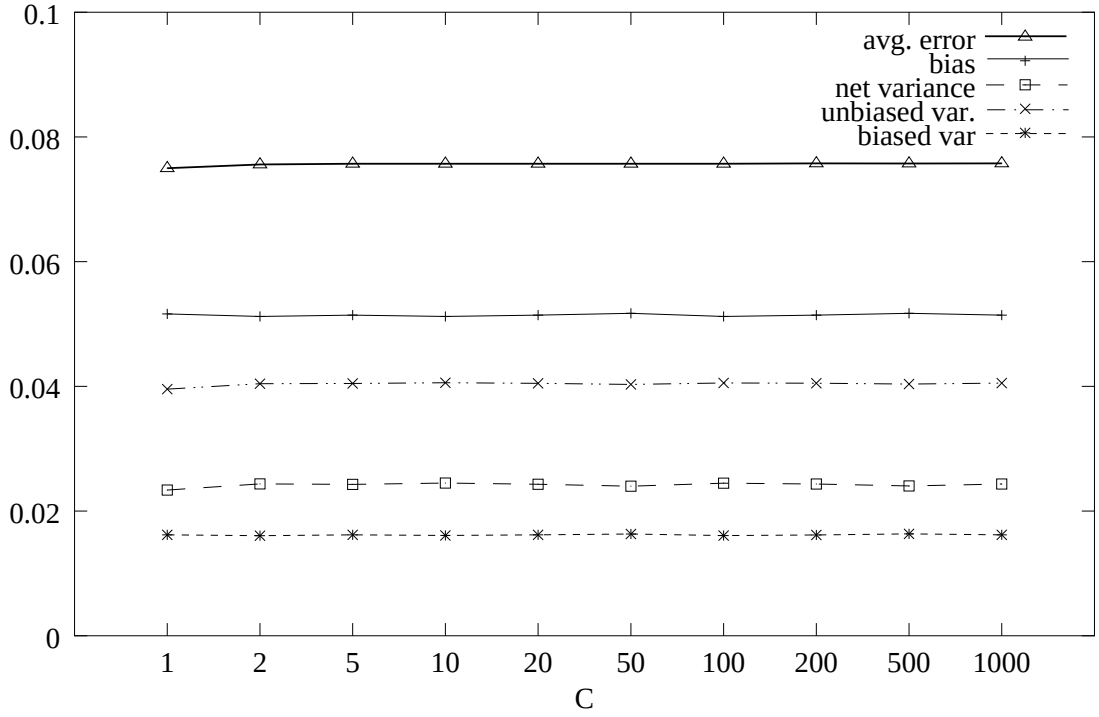
Figure 15: Musk data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 100$.

2.3 Decomposition in bagged Dot–product SVMs

The decomposition of the error is represented with respect to different values of C.

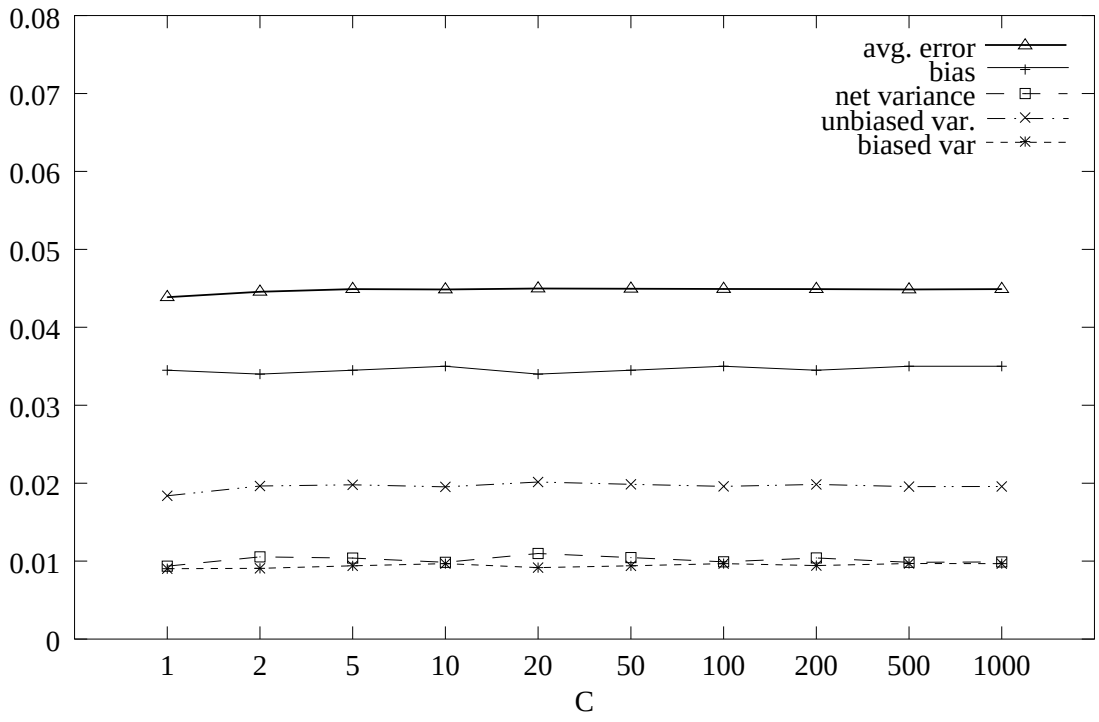


(a)

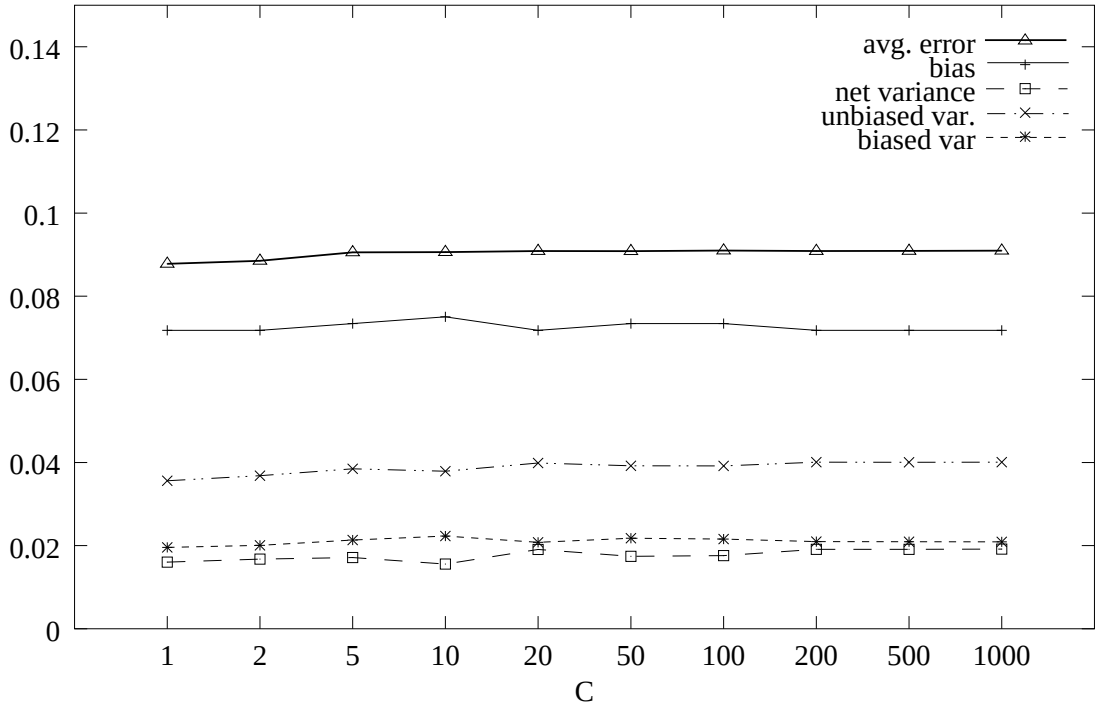


(b)

Figure 16: Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged dot-product SVM, while varying C: (a) P2, (b) Waveform.

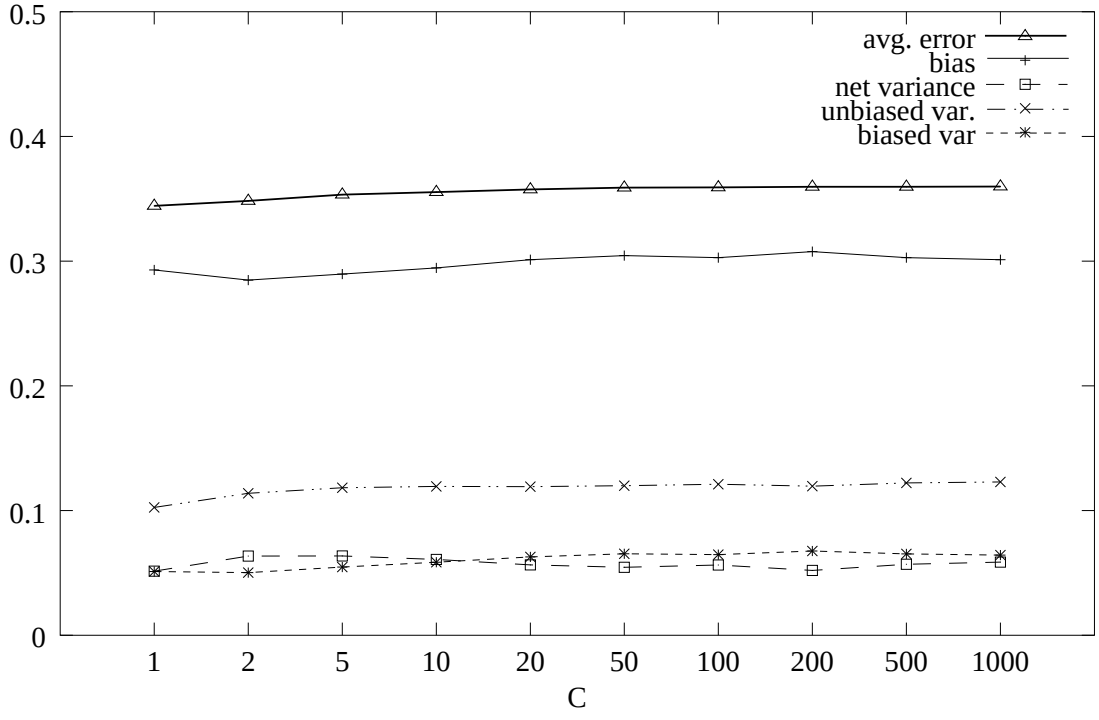


(a)

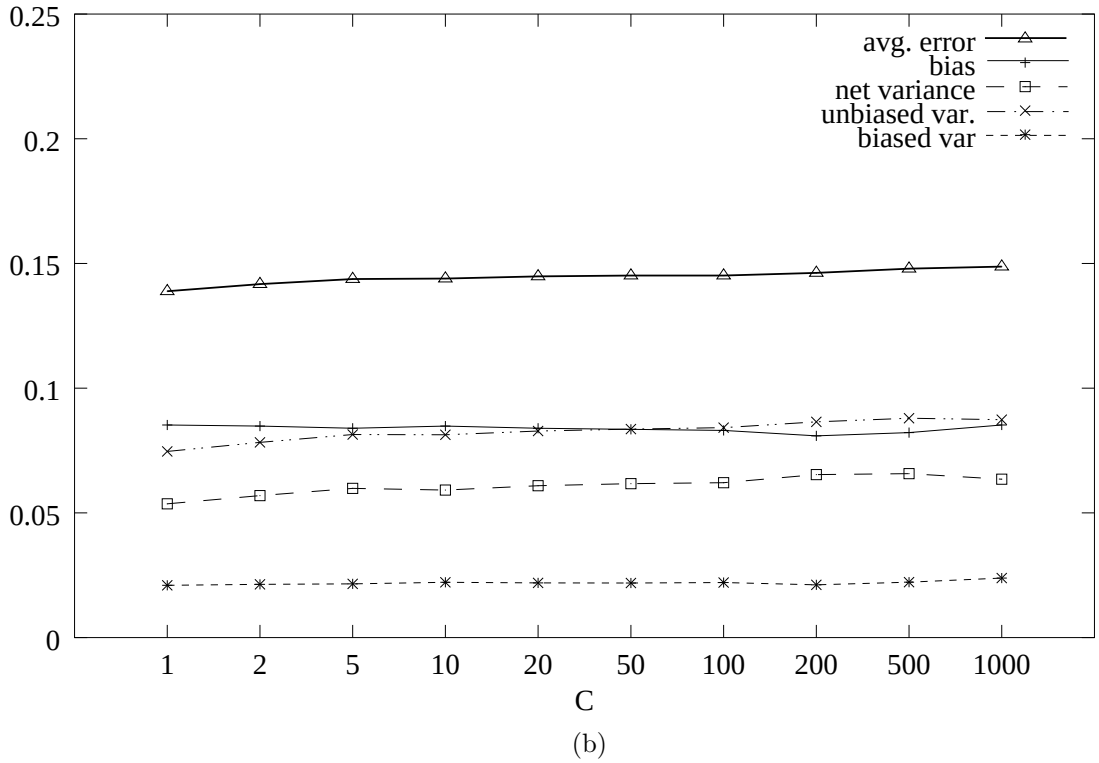


(b)

Figure 17: Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged dot-product SVM, while varying C: (a) Grey-Landsat, (b) Letter-Two.



(a)



(b)

Figure 18: Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged dot-product SVM, while varying C: (a) Letter-Two with noise, (b) Spam

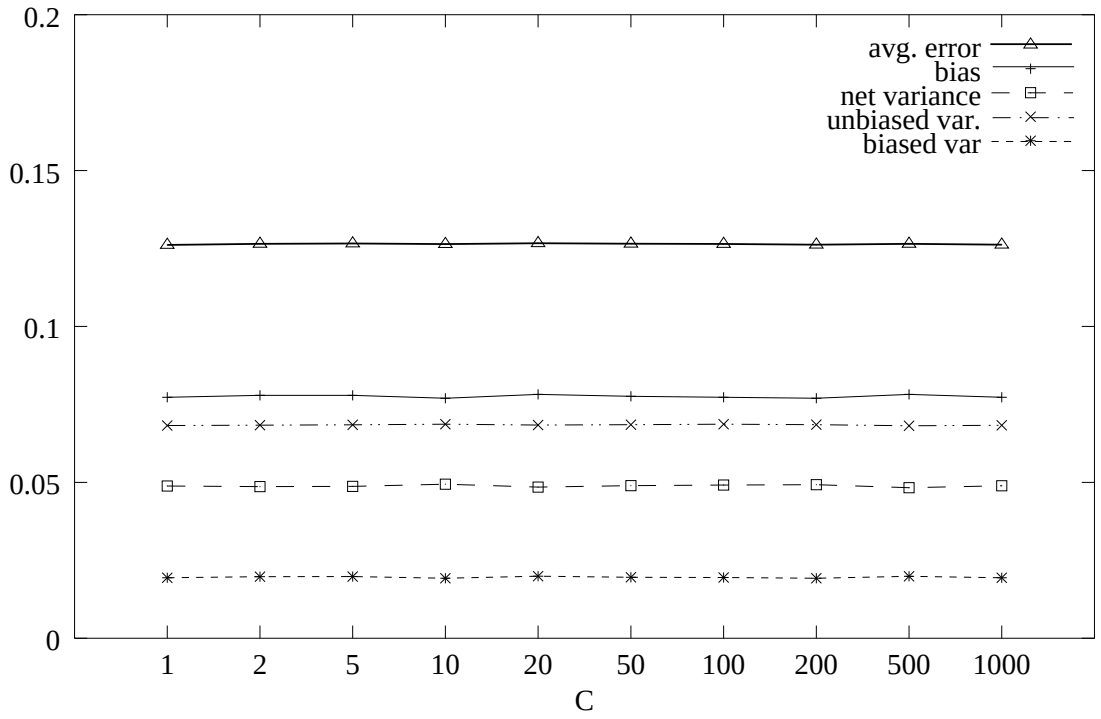


Figure 19: Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged dot-product SVM, while varying C with Musk data set

3 Decomposition of bias–variance with respect to the number of base learners

This section reports data and graphs about the decomposition of bias–variance in bagged SVMs with respect to the number of iterations of bagging, that is the number of base learners used.

3.1 Decomposition with respect to the number of base learners in RBF-SVM bagged ensembles

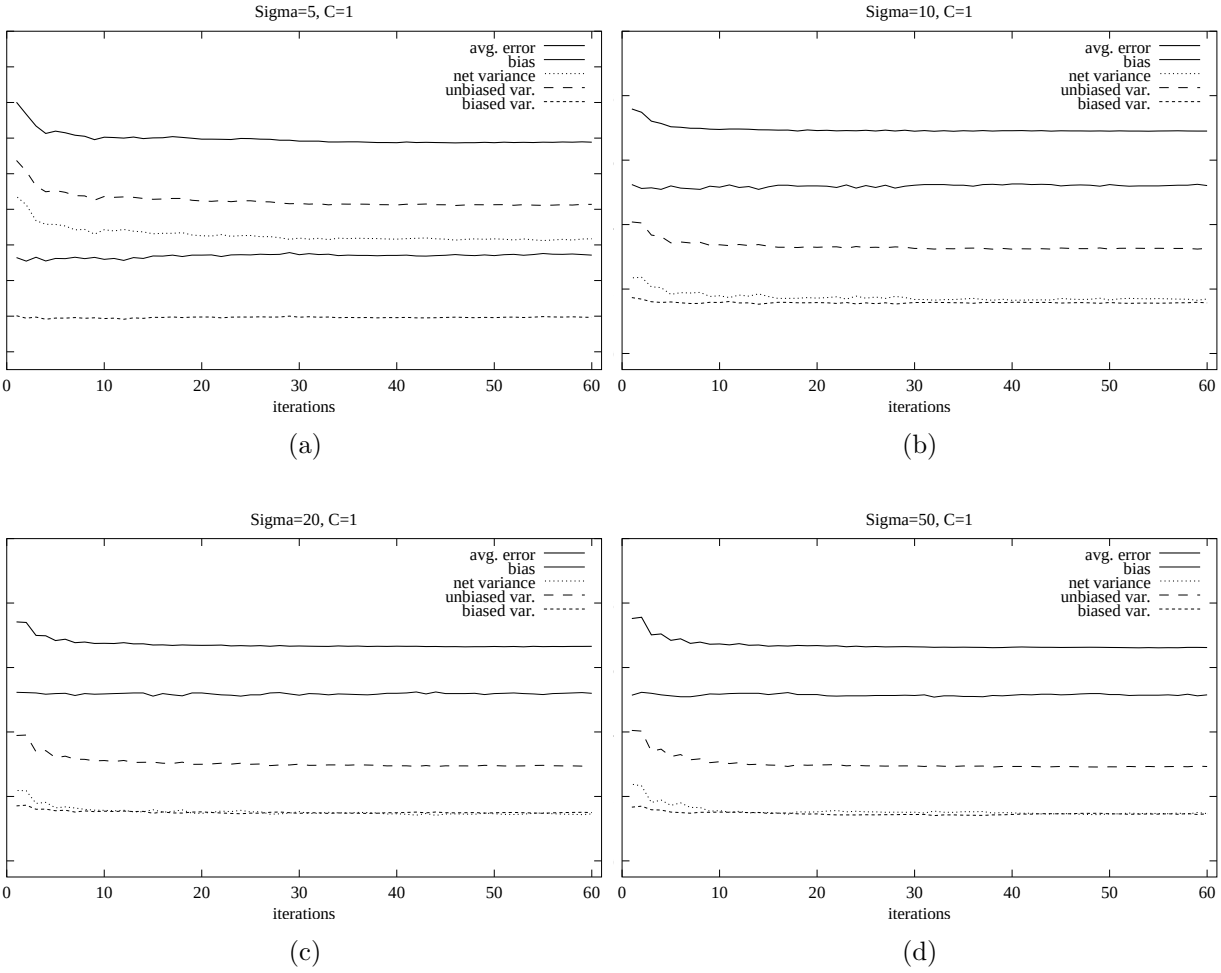


Figure 20: Waveform data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF with respect to the number of iterations for some values of σ and $C = 1$: (a) $\sigma = 5$, (b) $\sigma = 10$, (c) $\sigma = 20$, (d) $\sigma = 50$.

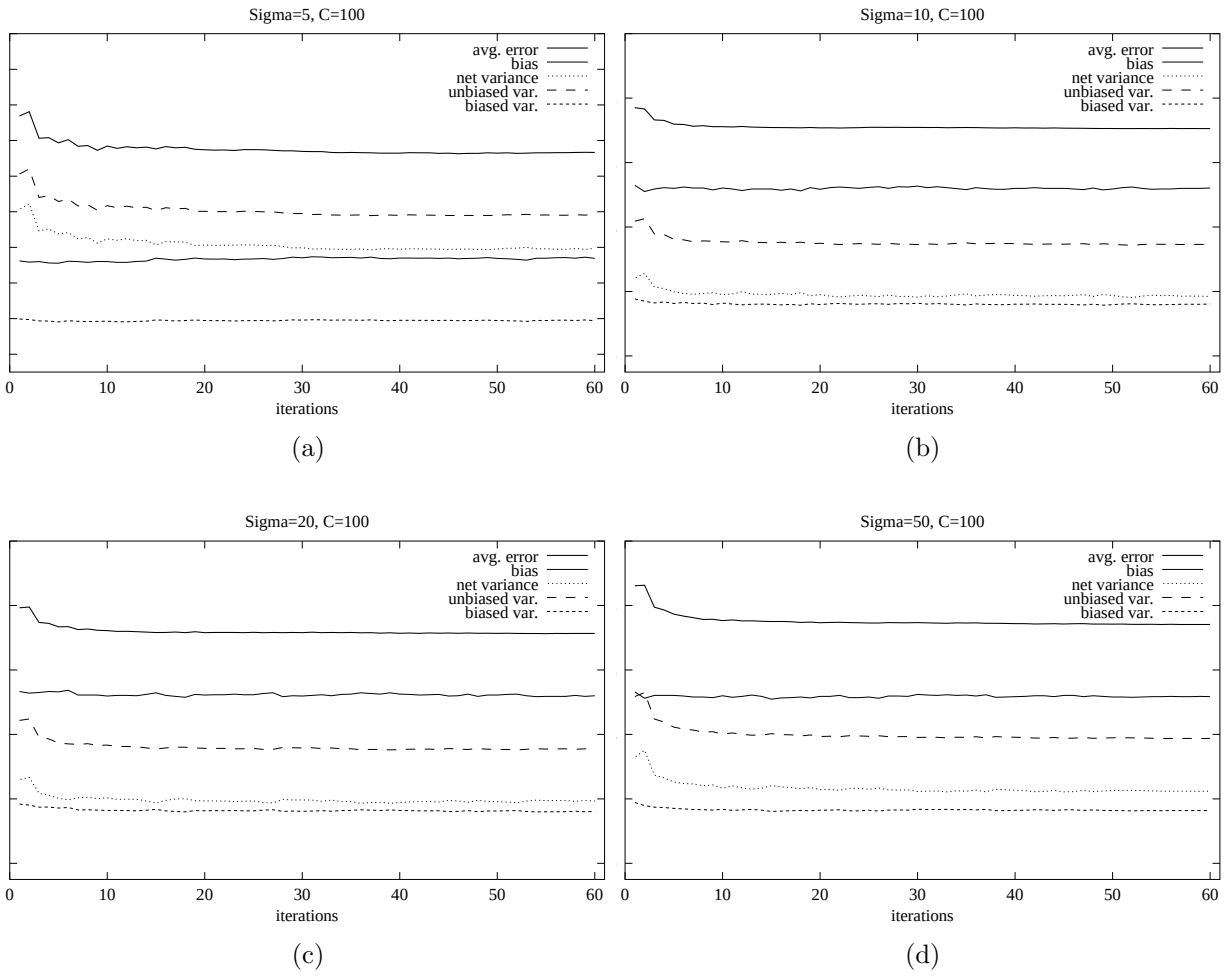


Figure 21: Waveform data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF with respect to the number of iterations for some values of σ and $C = 100$: (a) $\sigma = 5$, (b) $\sigma = 10$, (c) $\sigma = 20$, (d) $\sigma = 50$.

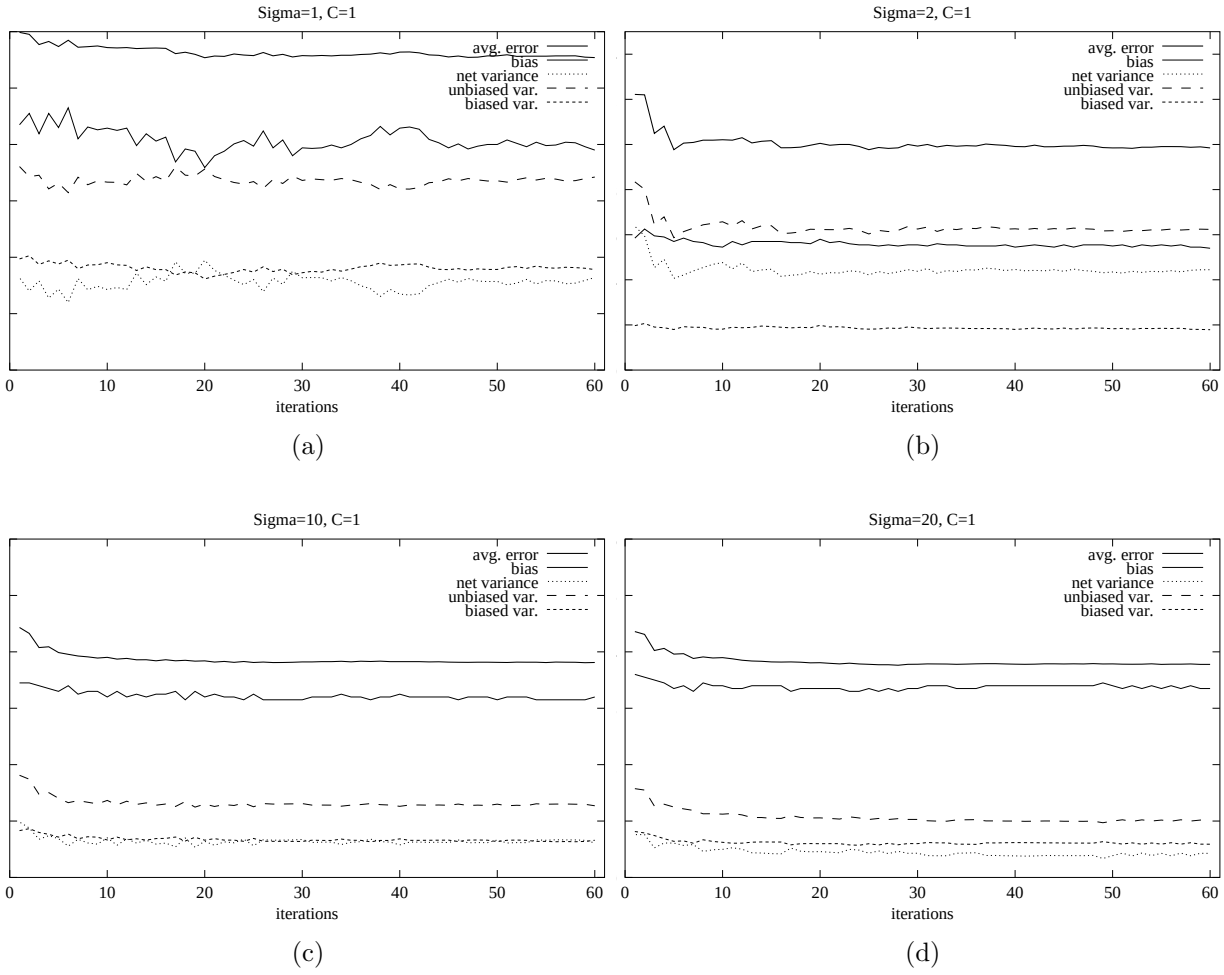


Figure 22: Grey-Landsat data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF with respect to the number of iterations for some values of σ and $C = 1$: (a) $\sigma = 1$, (b) $\sigma = 2$, (c) $\sigma = 10$, (d) $\sigma = 20$.

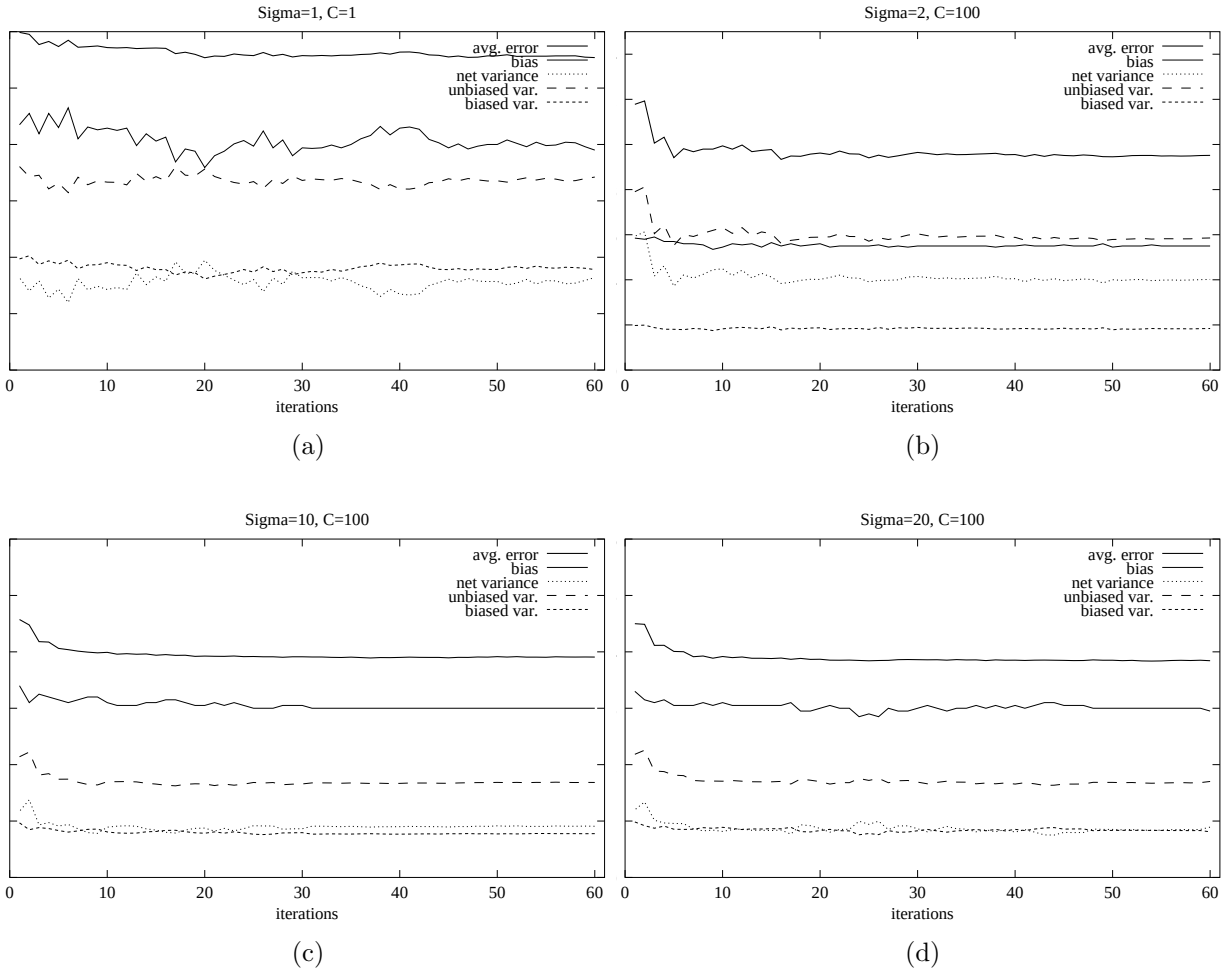


Figure 23: Grey-Landsat data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF with respect to the number of iterations for some values of σ and $C = 100$: (a) $\sigma = 1$, (b) $\sigma = 2$, (c) $\sigma = 10$, (d) $\sigma = 20$.

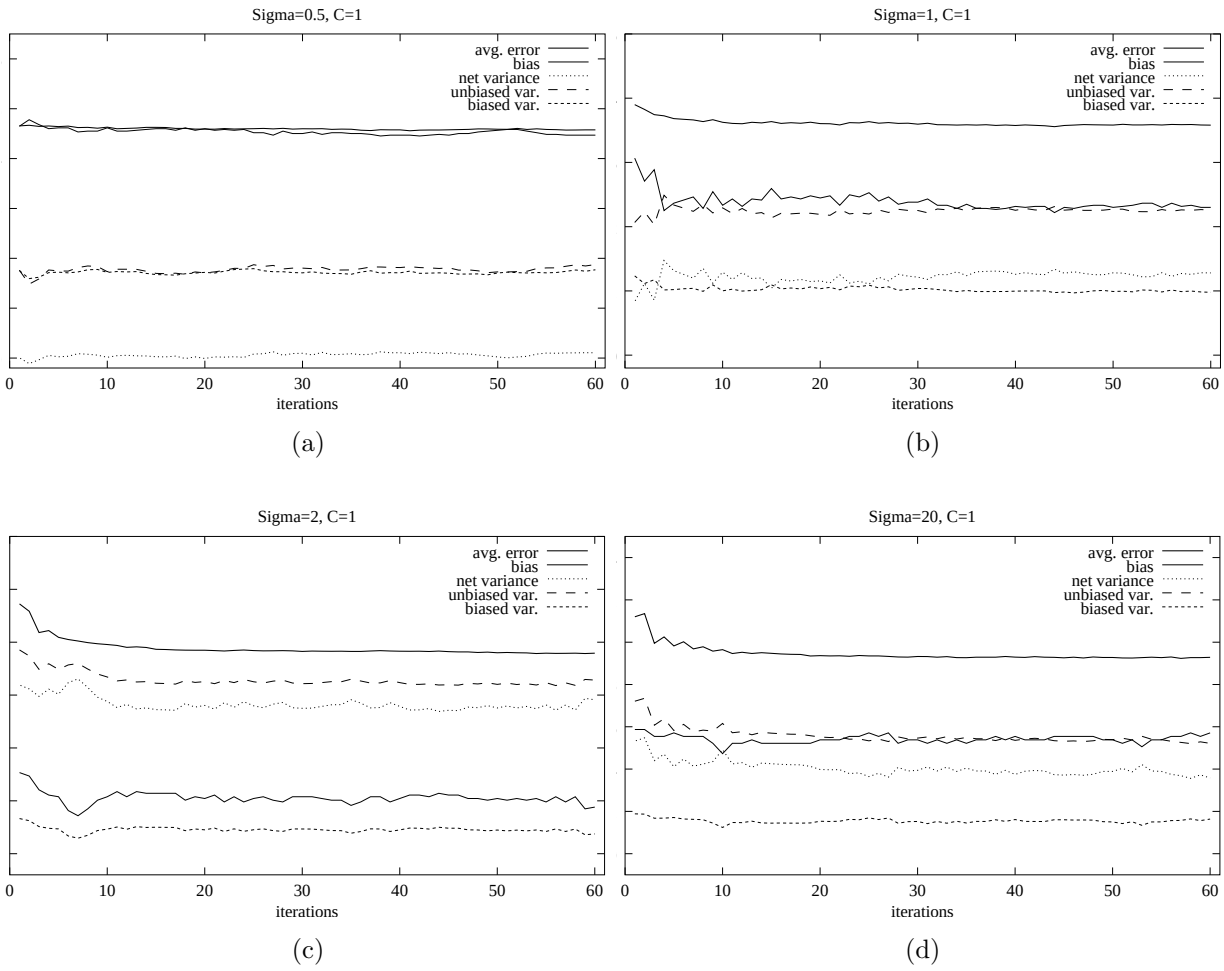


Figure 24: Letter-two data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF with respect to the number of iterations for some values of σ and $C = 1$: (a) $\sigma = 0.5$, (b) $\sigma = 1$, (c) $\sigma = 2$, (d) $\sigma = 20$.

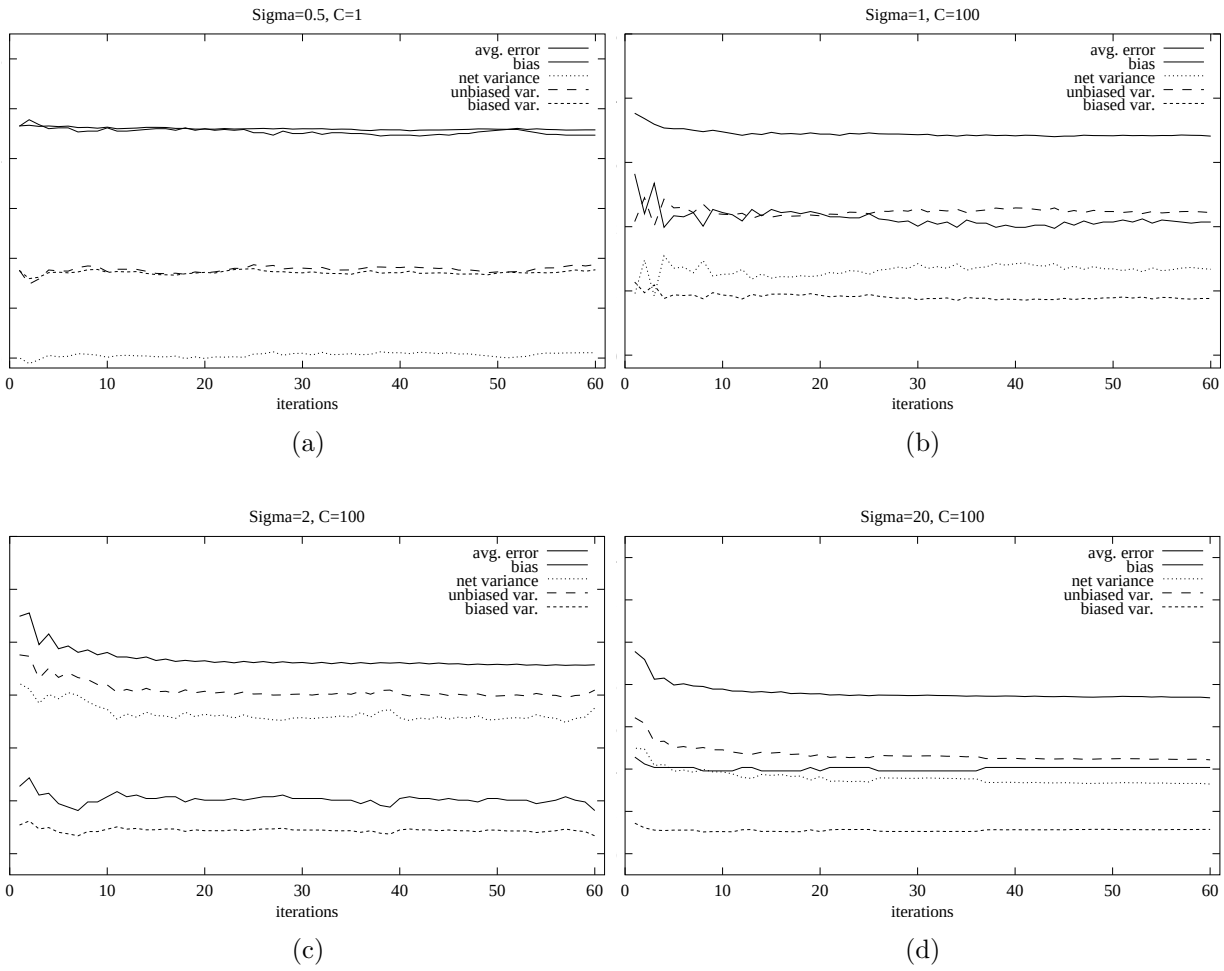


Figure 25: Letter-two data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF with respect to the number of iterations for some values of σ and $C = 100$: (a) $\sigma = 0.5$, (b) $\sigma = 1$, (c) $\sigma = 2$, (d) $\sigma = 20$.

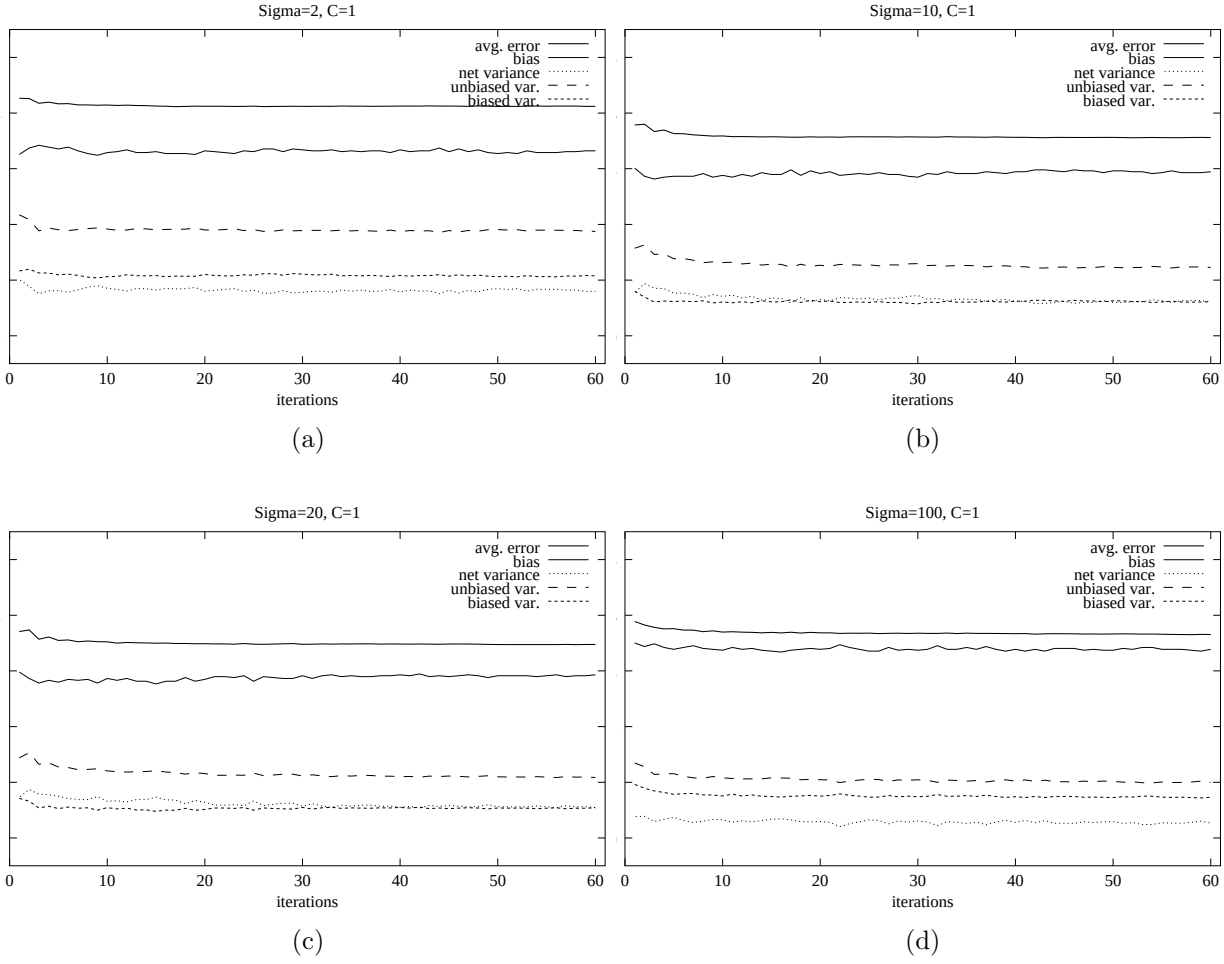


Figure 26: Letter-two with noise data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF with respect to the number of iterations for some values of σ and $C = 1$: (a) $\sigma = 2$, (b) $\sigma = 10$, (c) $\sigma = 20$, (d) $\sigma = 100$.

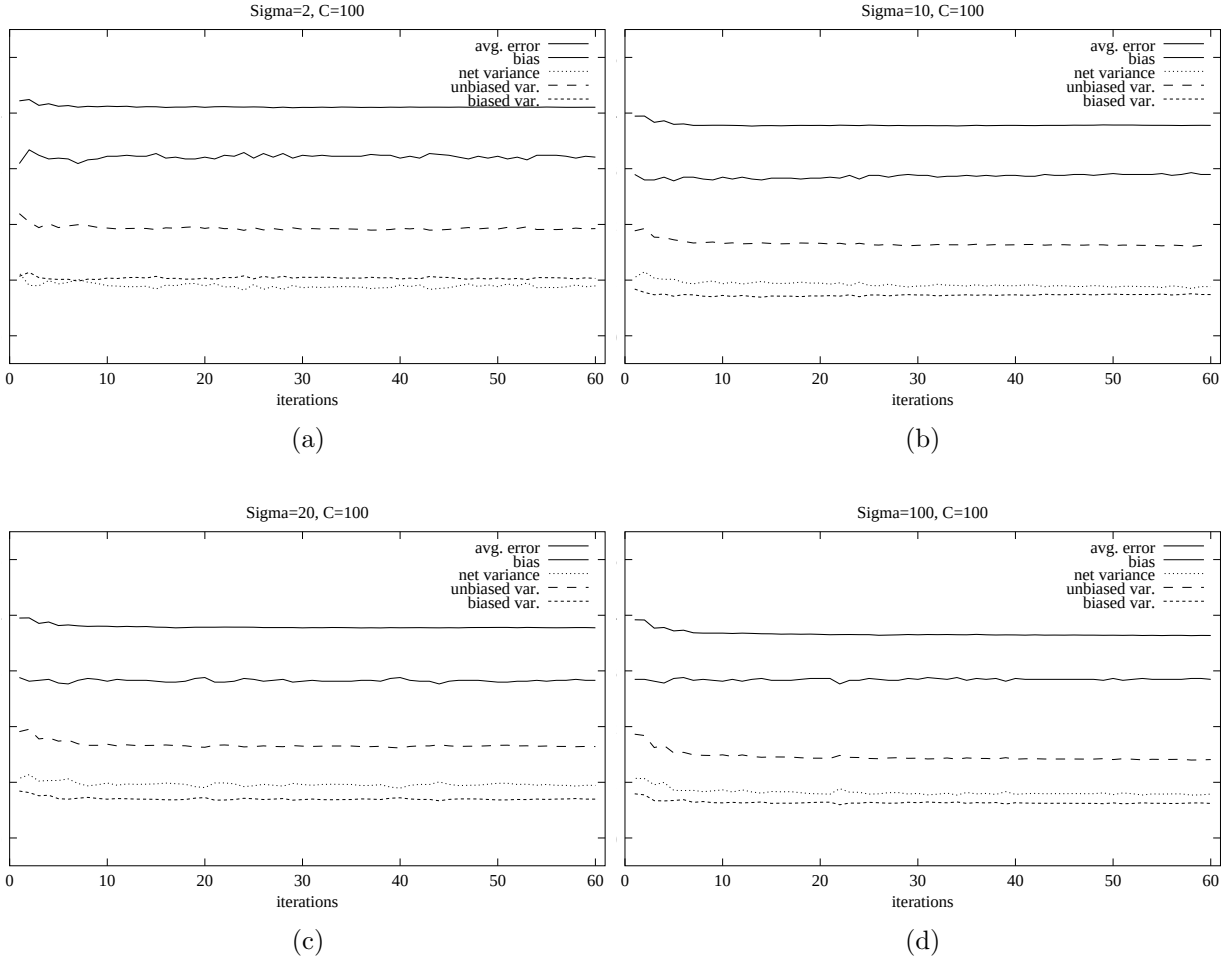


Figure 27: Letter-two with noise data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF with respect to the number of iterations for some values of σ and $C = 100$: (a) $\sigma = 2$, (b) $\sigma = 10$, (c) $\sigma = 20$, (d) $\sigma = 100$.

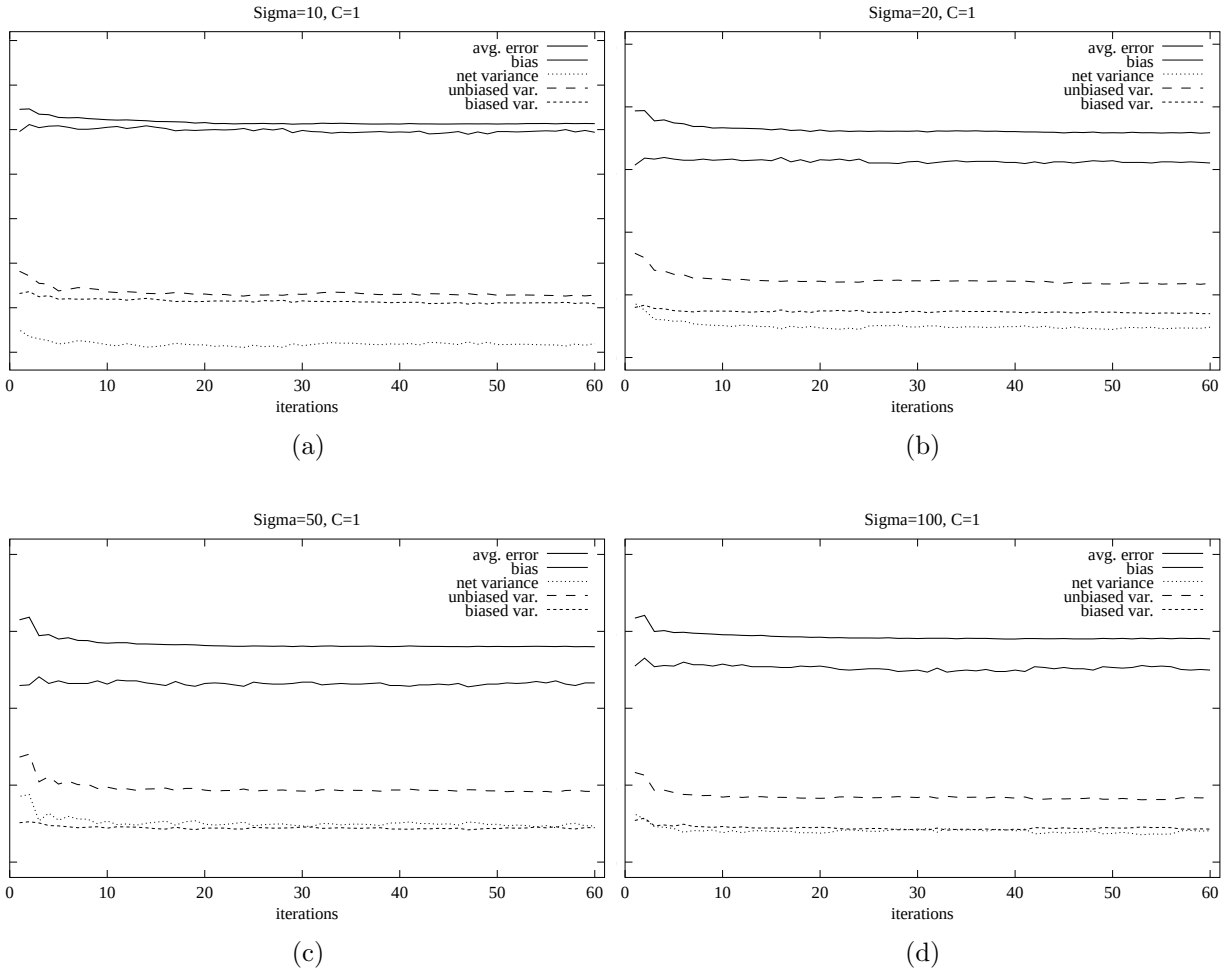


Figure 28: Spam data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF with respect to the number of iterations for some values of σ and $C = 1$: (a) $\sigma = 10$, (b) $\sigma = 20$, (c) $\sigma = 50$, (d) $\sigma = 100$.

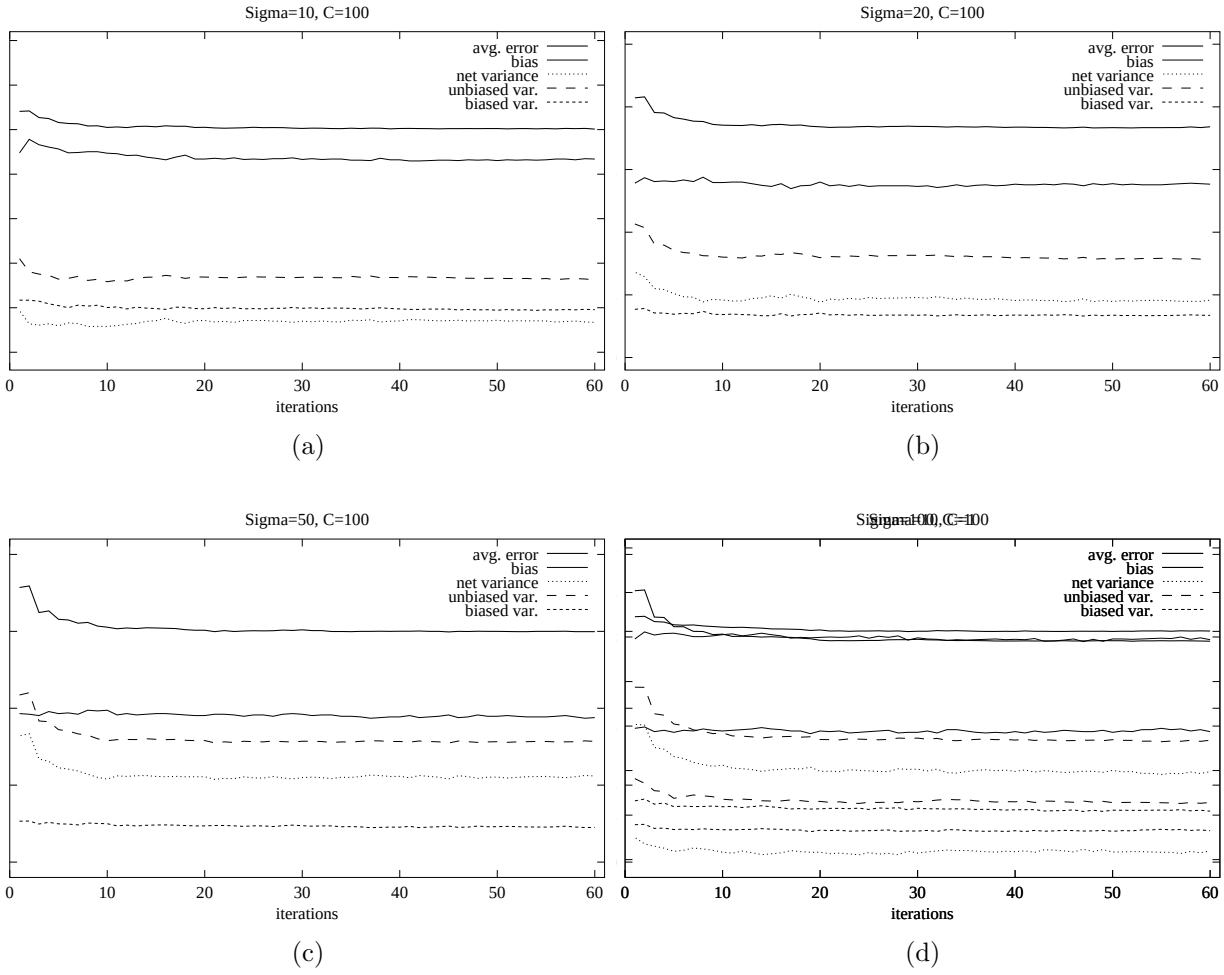


Figure 29: Spam data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF with respect to the number of iterations for some values of σ and $C = 100$: (a) $\sigma = 10$, (b) $\sigma = 20$, (c) $\sigma = 50$, (d) $\sigma = 100$.

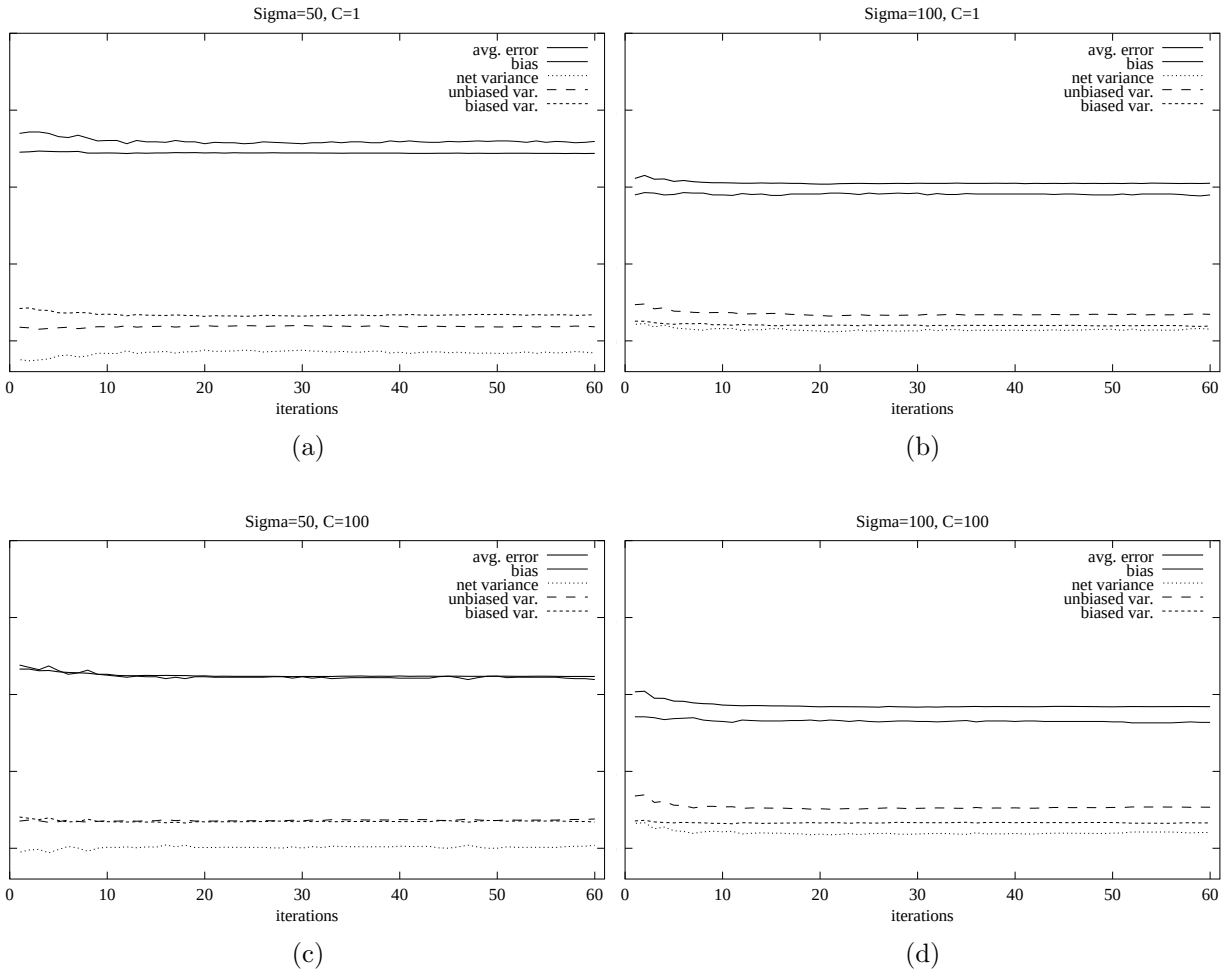


Figure 30: Musk data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in bagged SVM RBF with respect to the number of iterations for some values of σ and C: (a) $\sigma = 50$ and $C = 1$, (b) $\sigma = 100$ and $C = 1$, (c) $\sigma = 50$ and $C = 100$, (d) $\sigma = 100$ and $C = 100$.

3.2 Decomposition with respect to the number of base learners in Polynomial SVM bagged ensembles

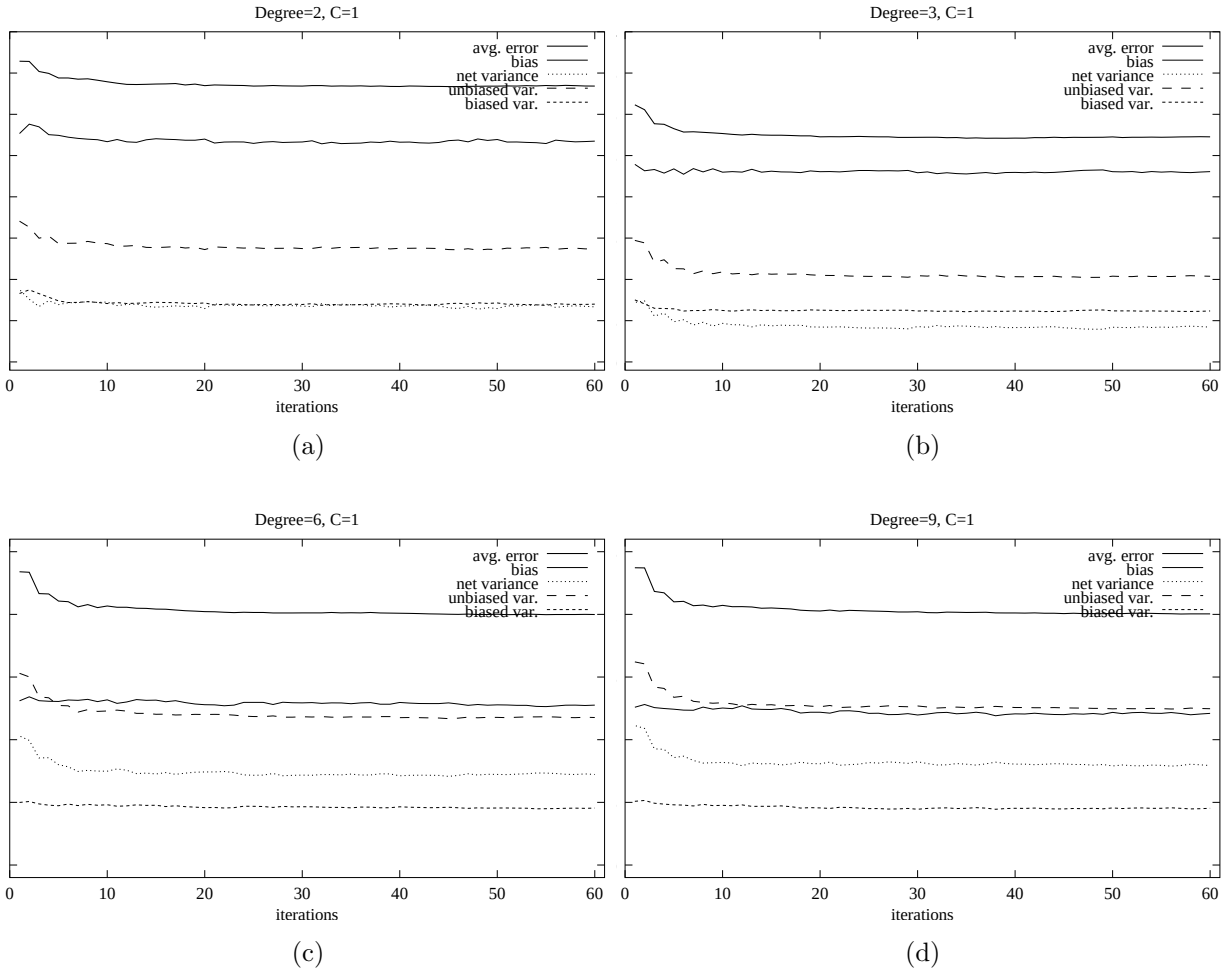


Figure 31: P2 data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 1$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

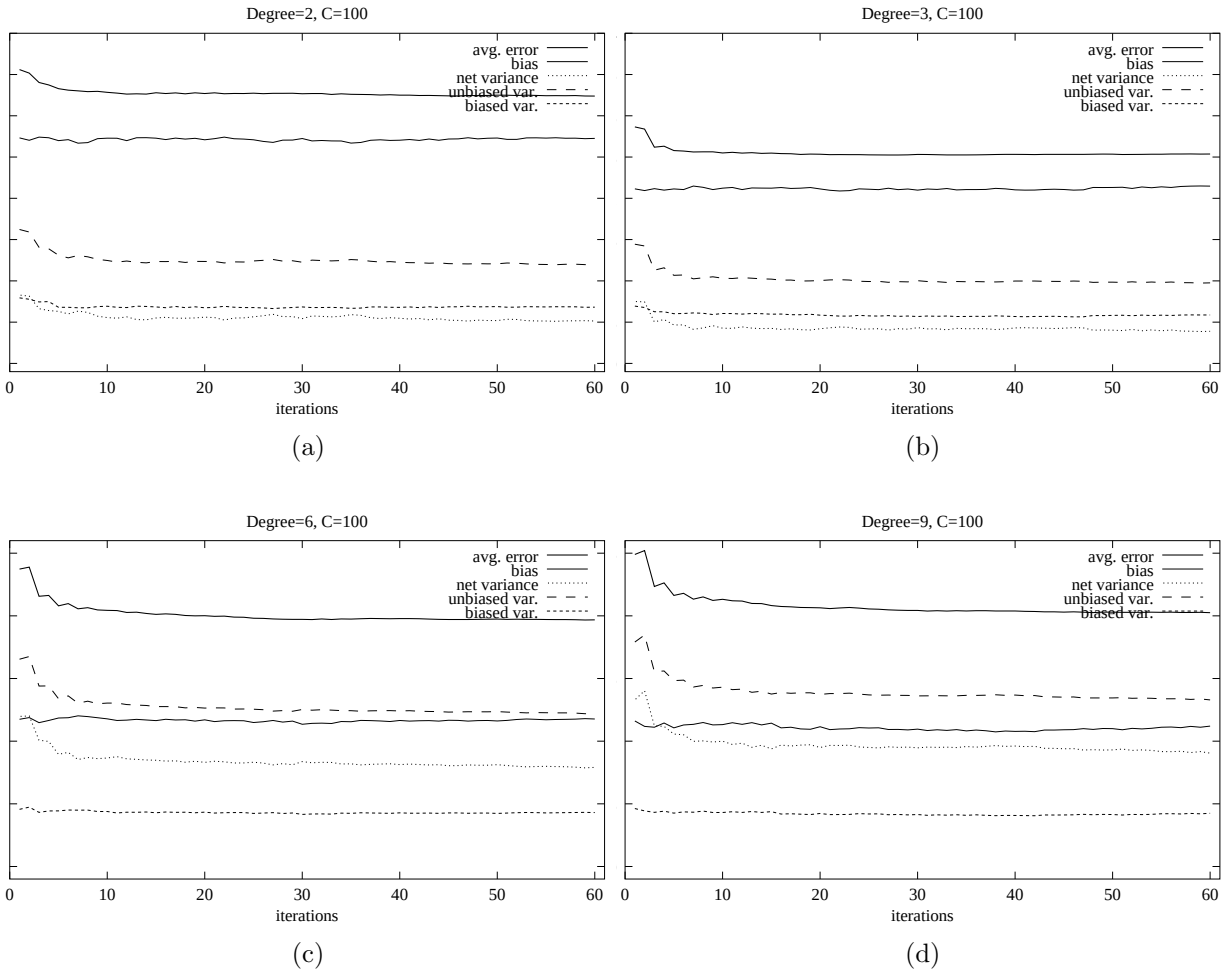


Figure 32: P2 data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 100$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

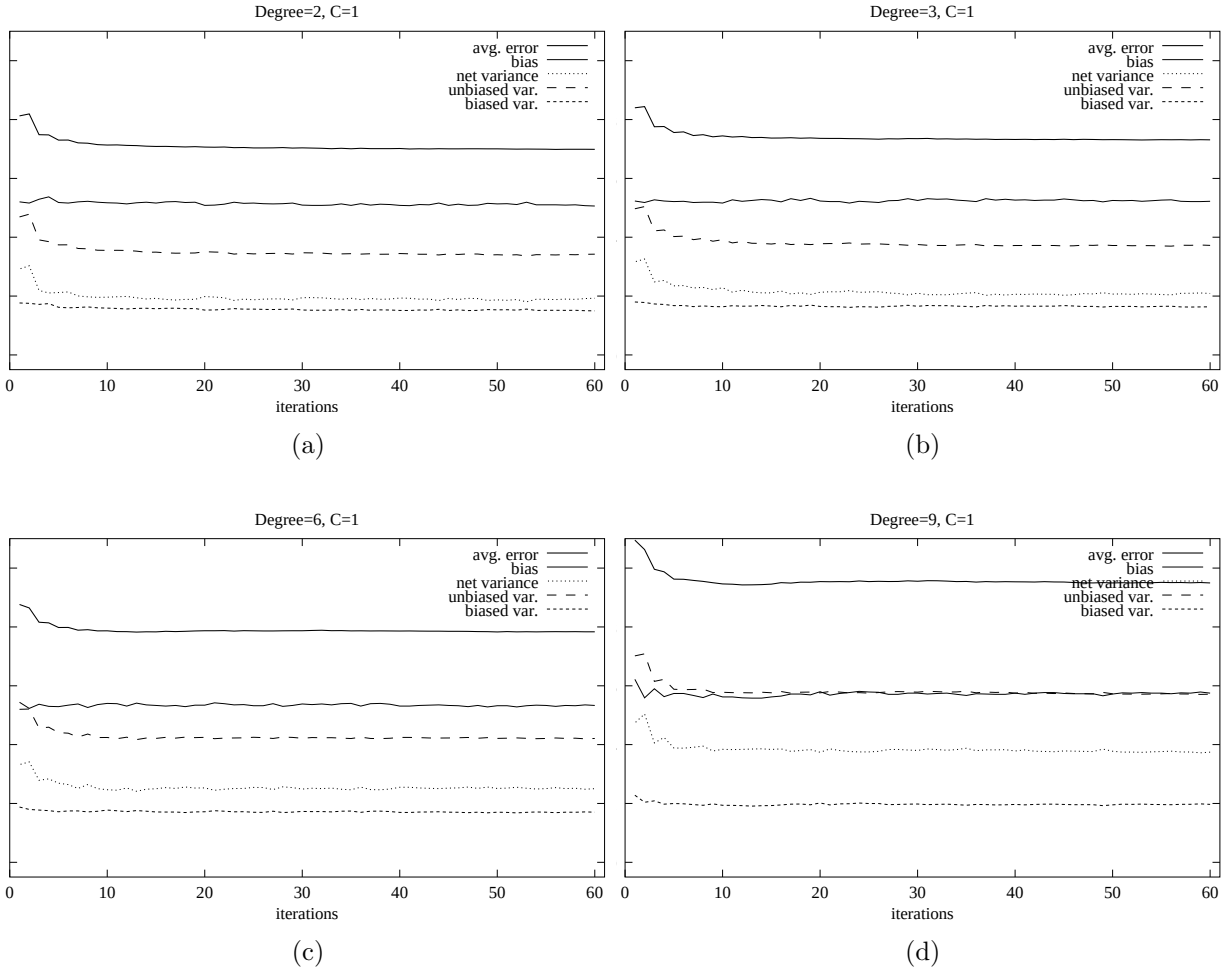


Figure 33: Waveform data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 1$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

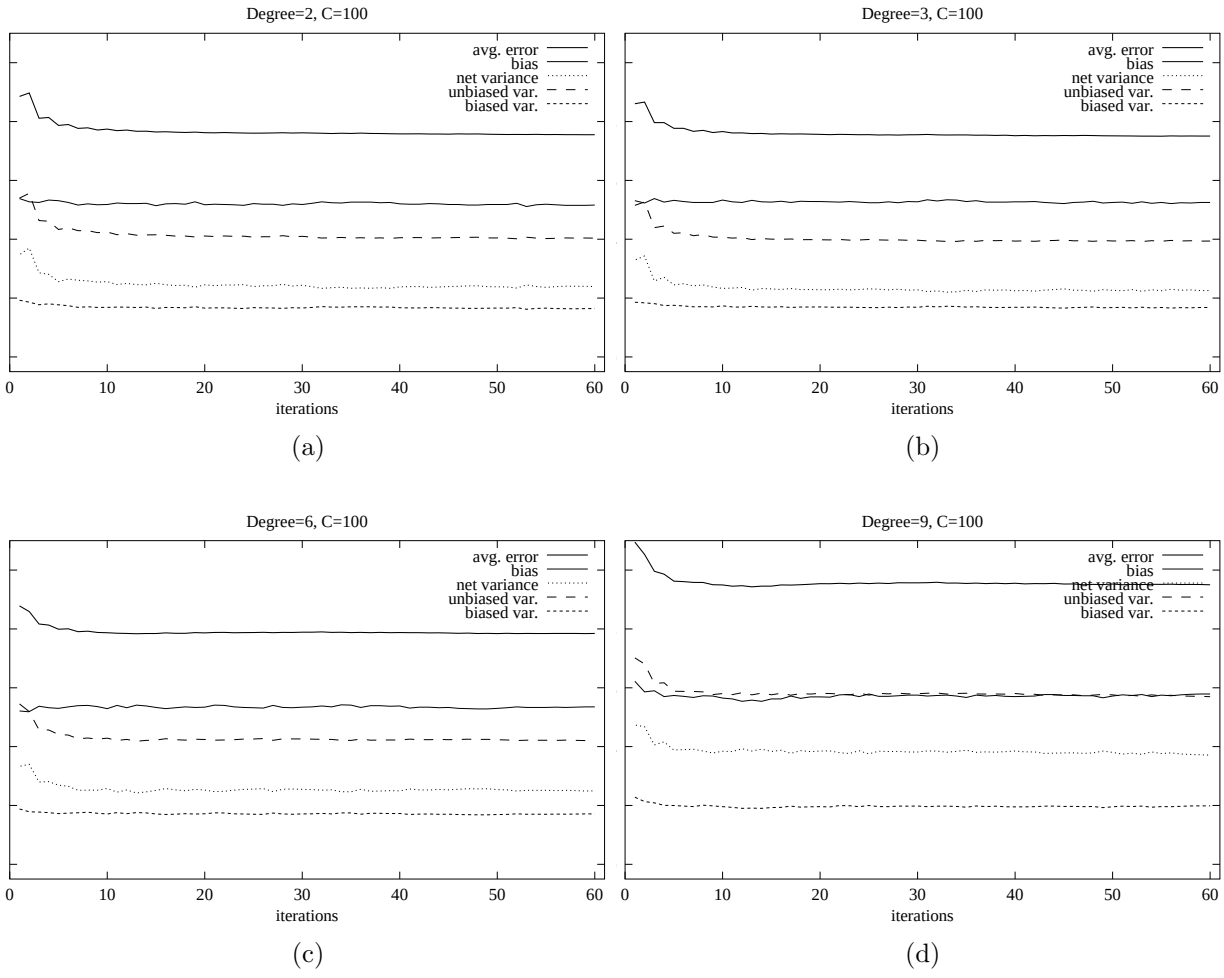


Figure 34: Waveform data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 100$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

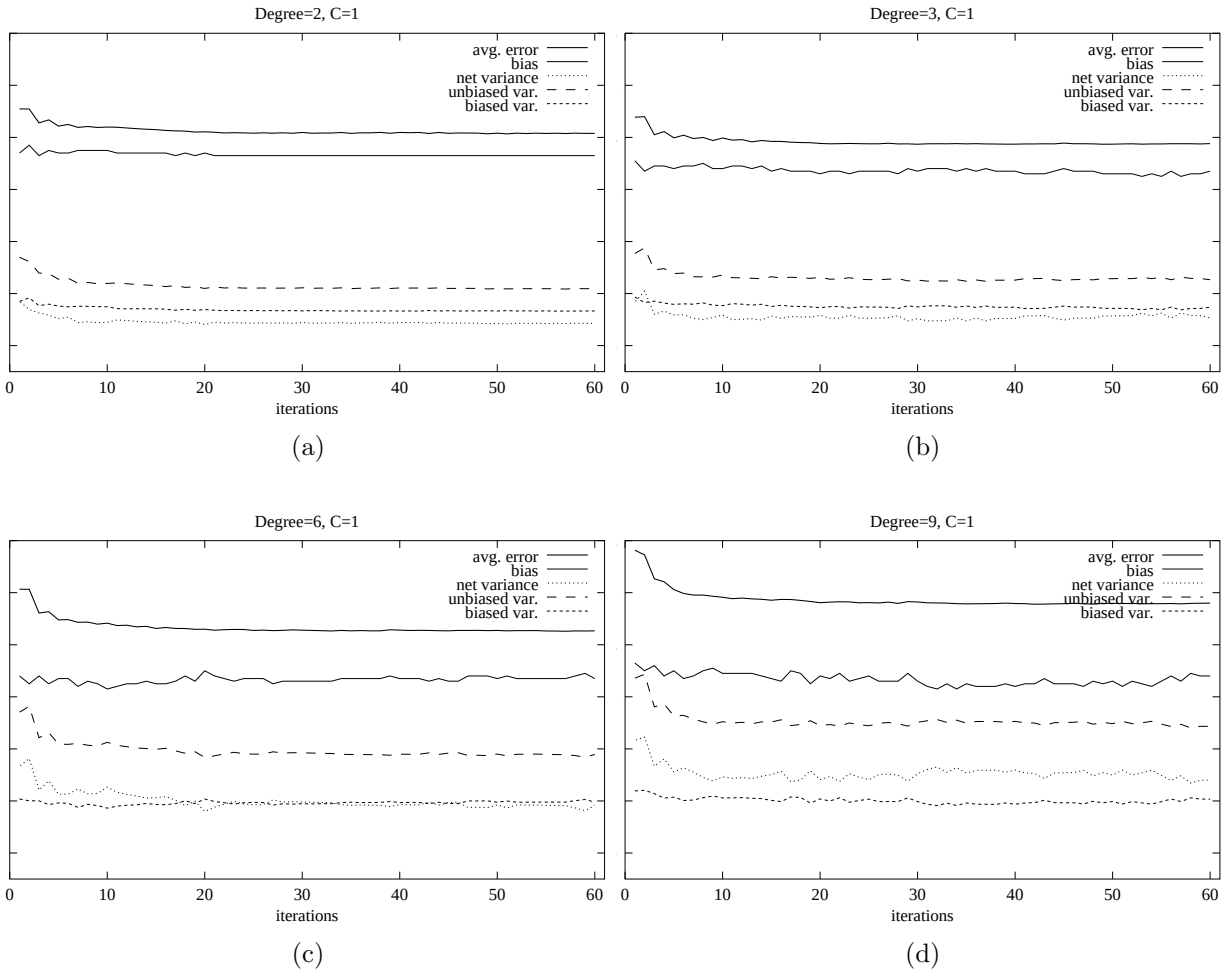


Figure 35: Grey-Landsat data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 1$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

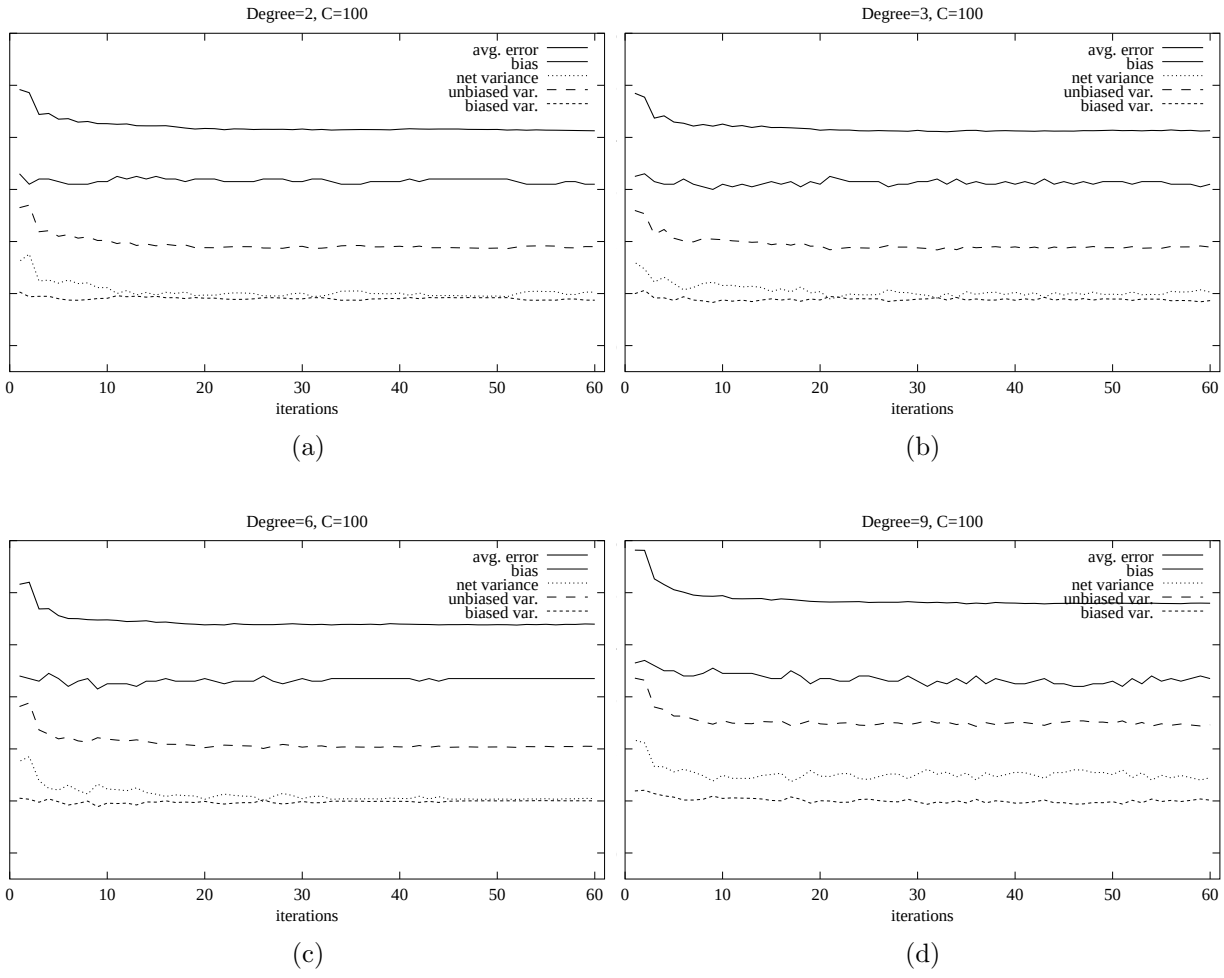


Figure 36: Grey-Landsat data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 100$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

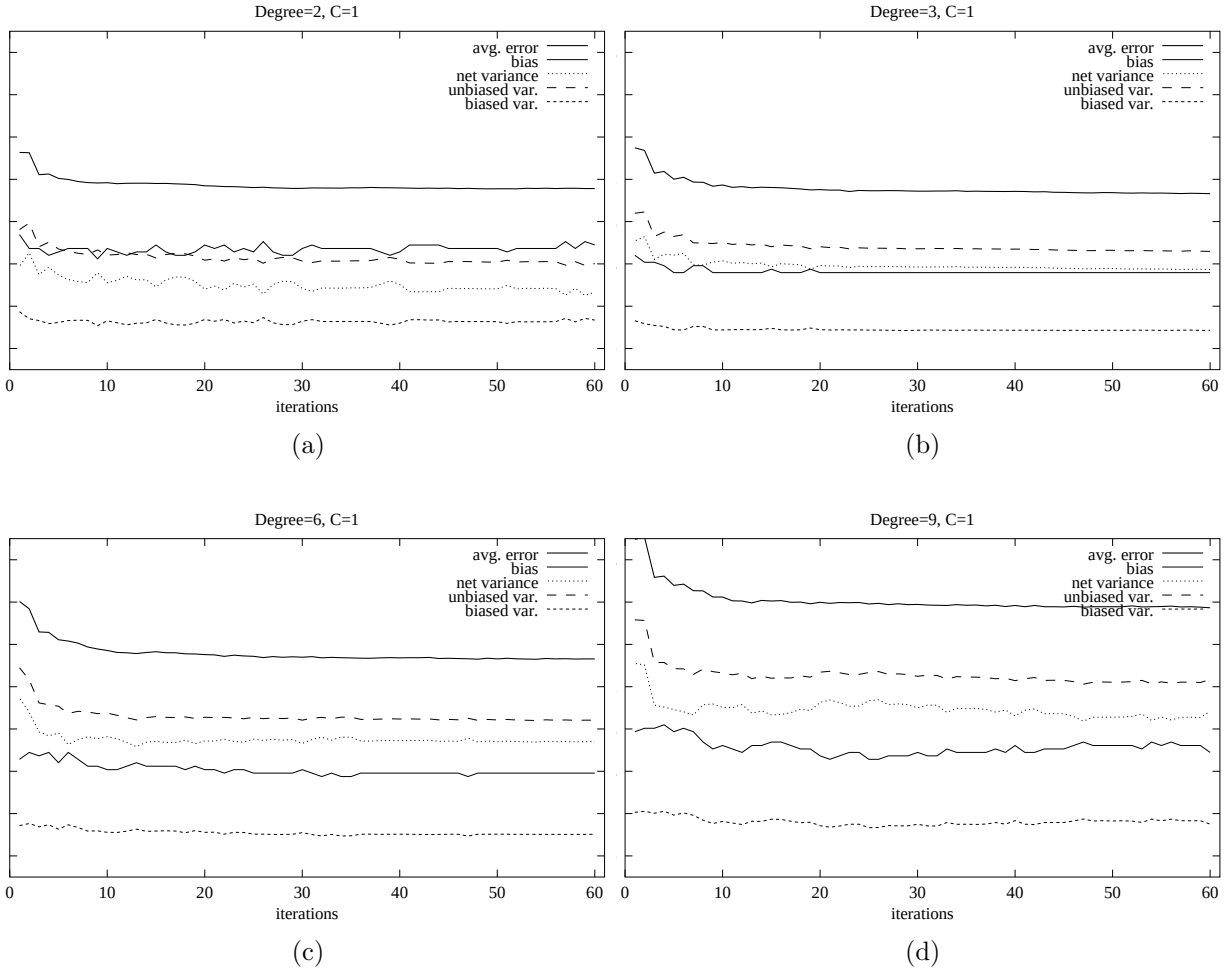


Figure 37: Letter-Two data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 1$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

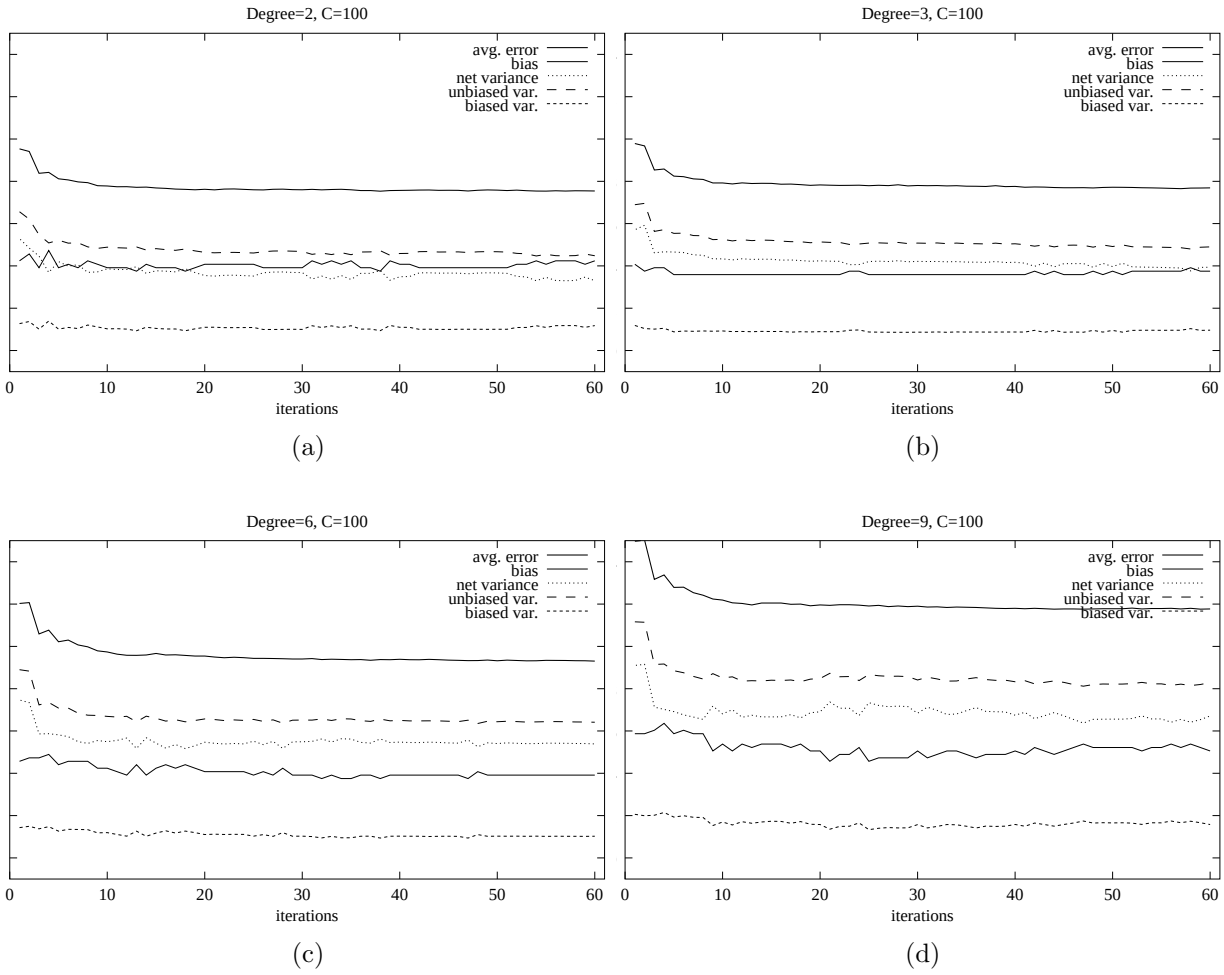


Figure 38: Letter-Two data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 100$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

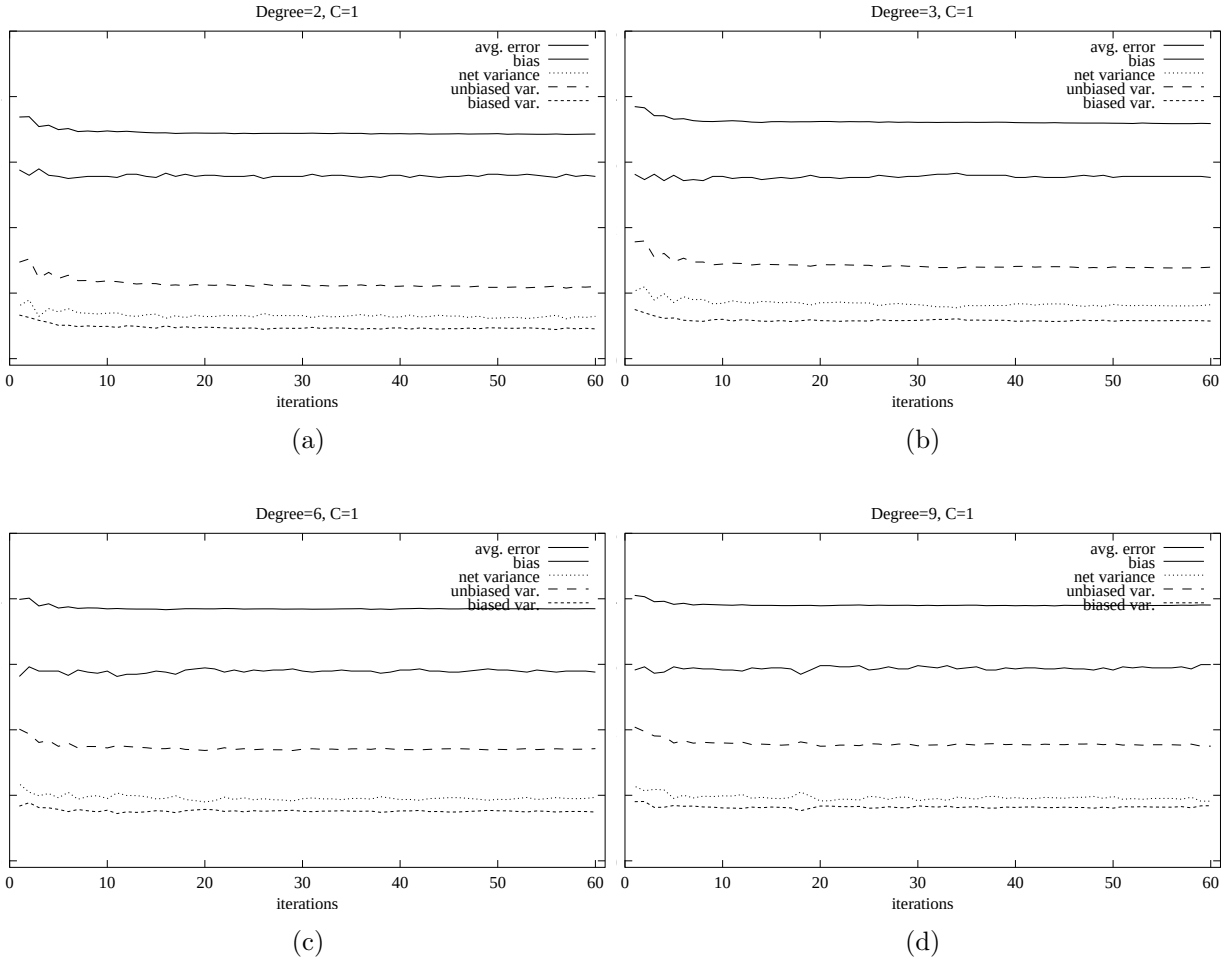


Figure 39: Letter-Two with noise data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 1$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

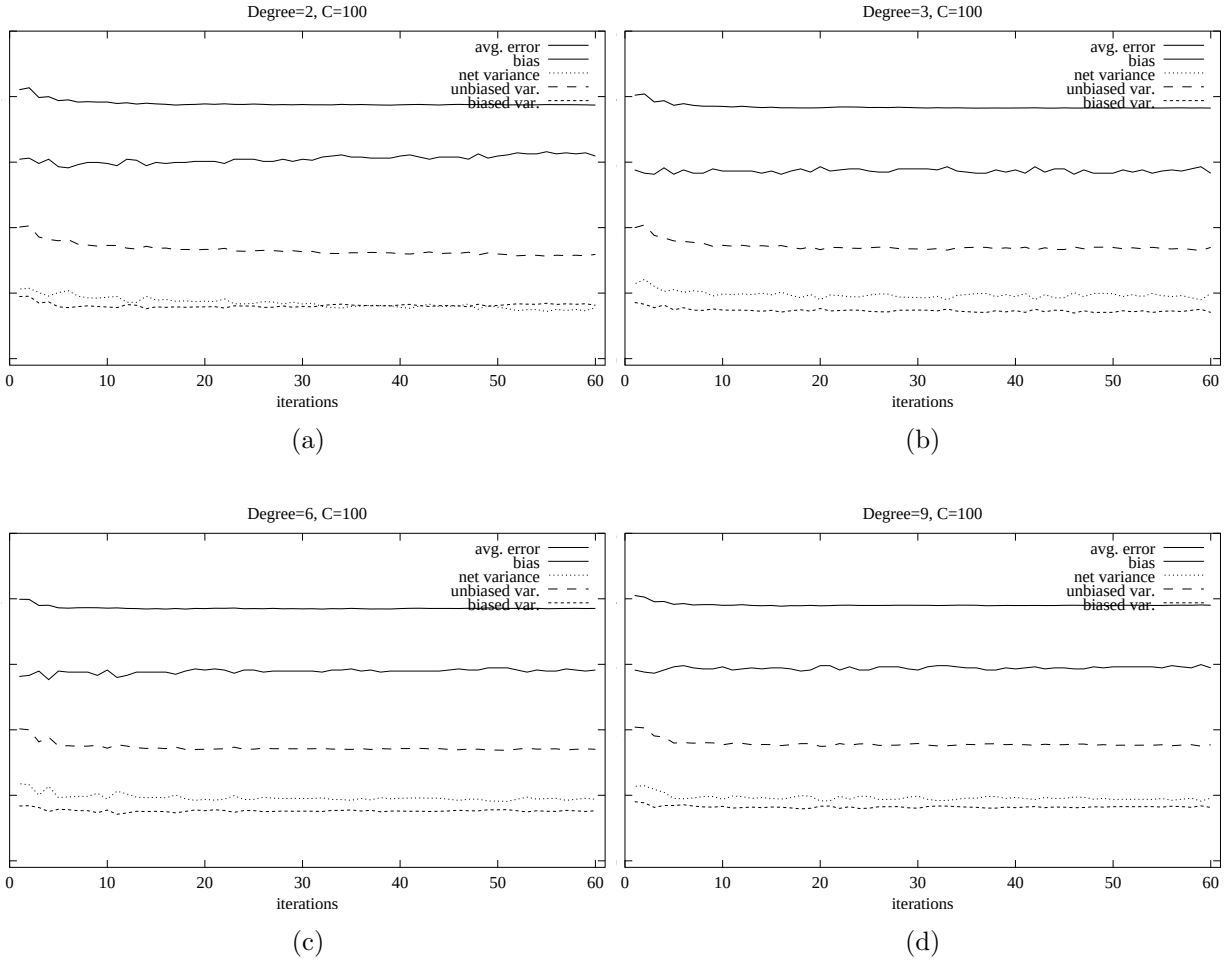


Figure 40: Letter-Two with noise data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 100$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

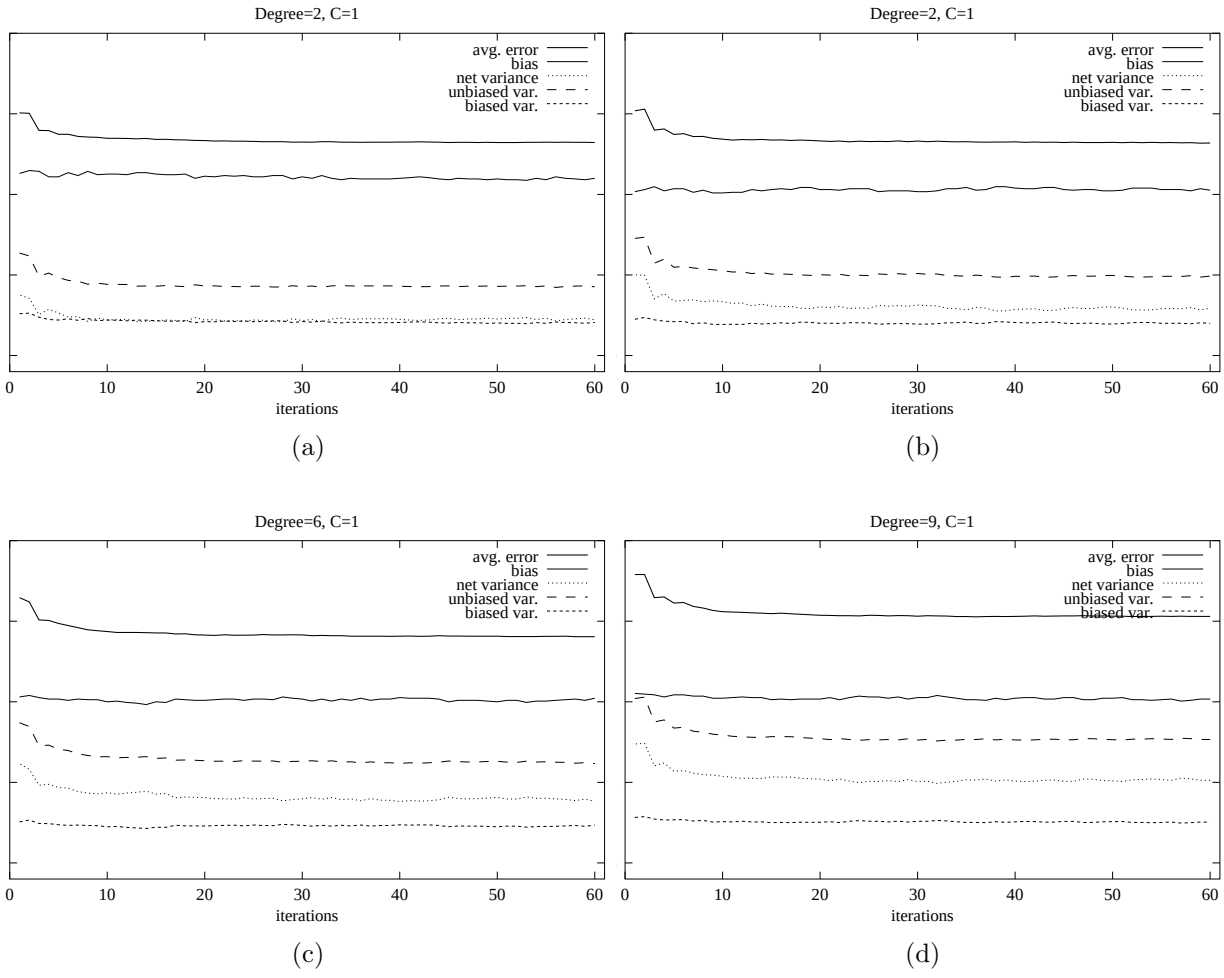


Figure 41: Spam data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 1$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

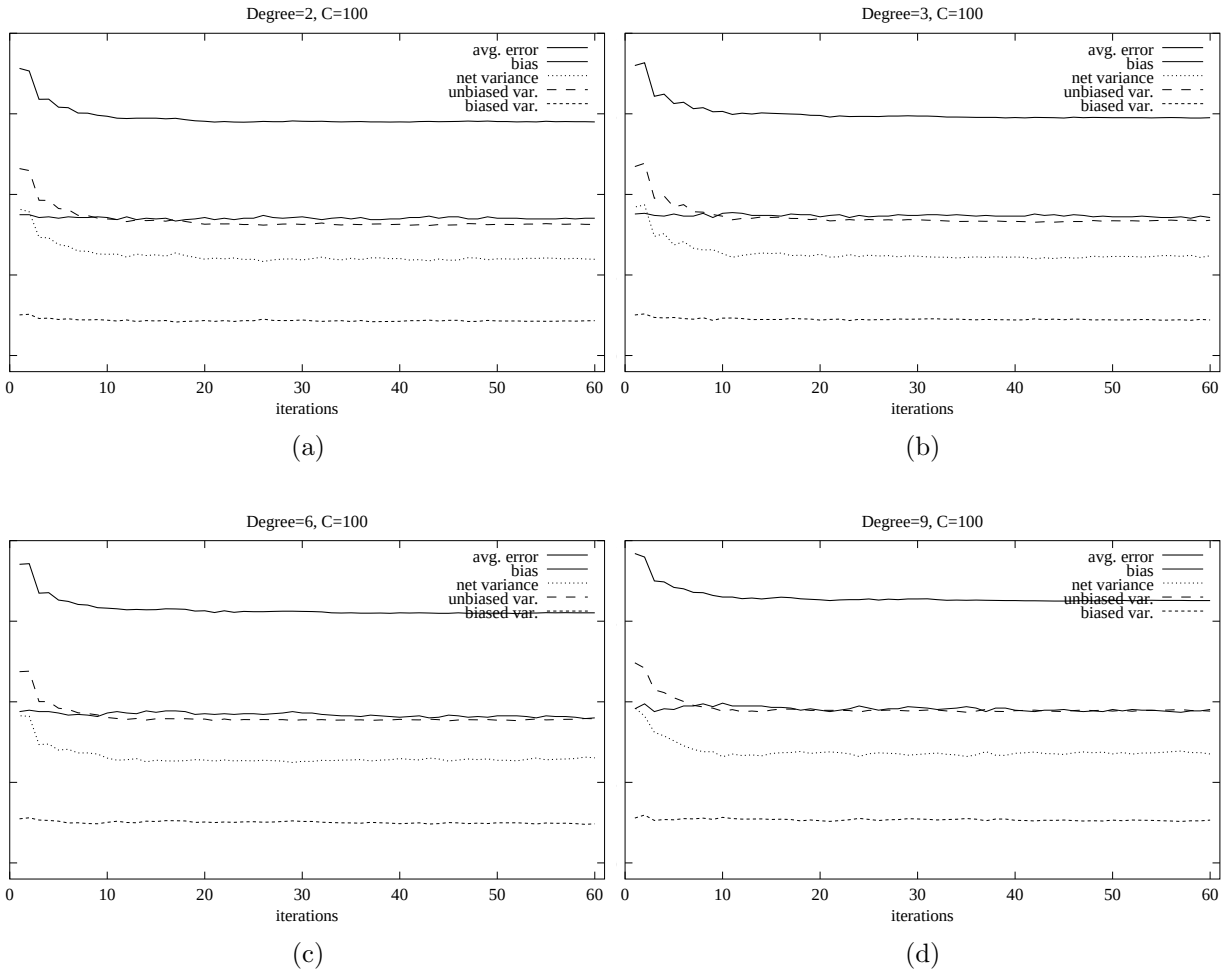


Figure 42: Spam data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 100$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

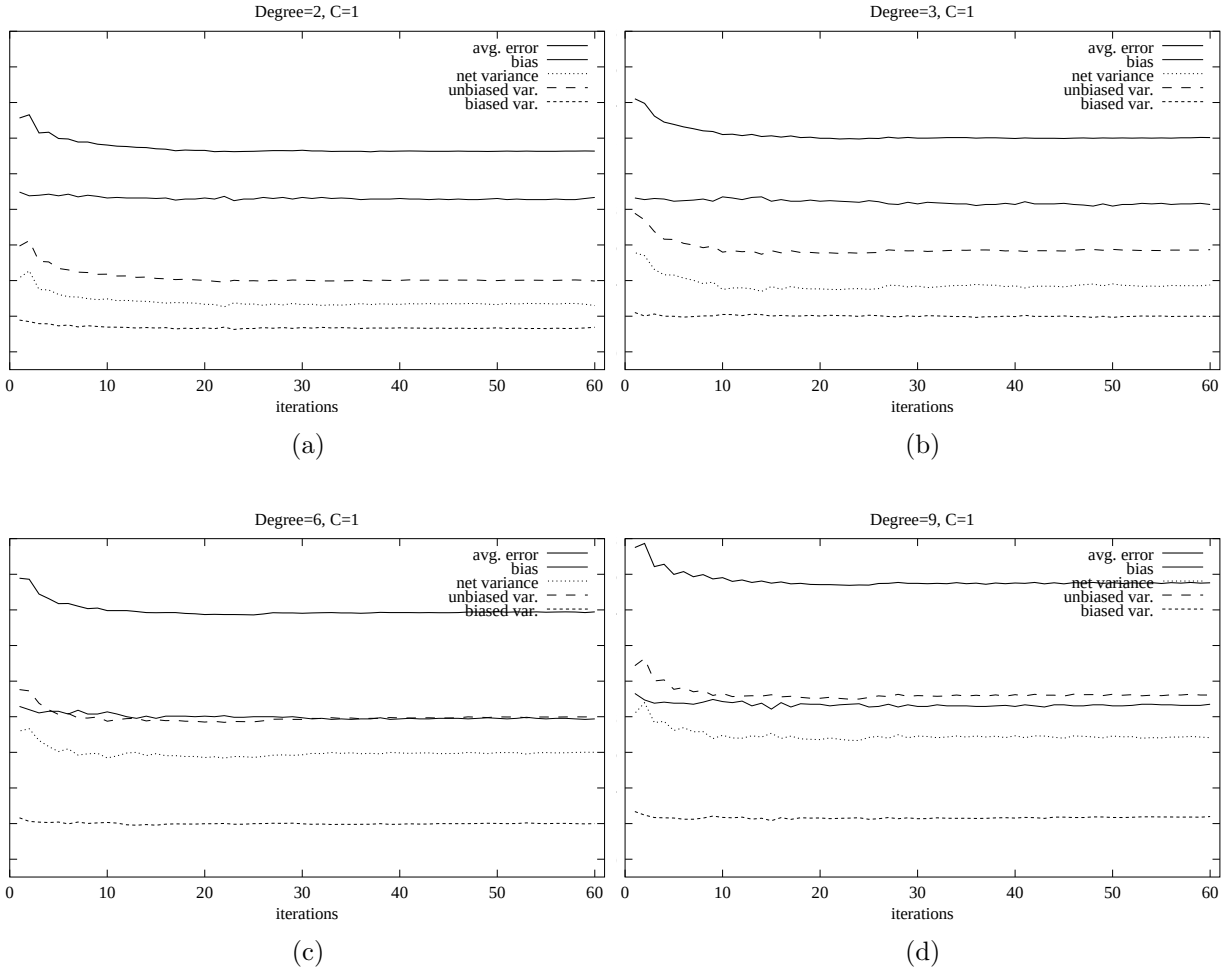


Figure 43: Musk data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 1$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

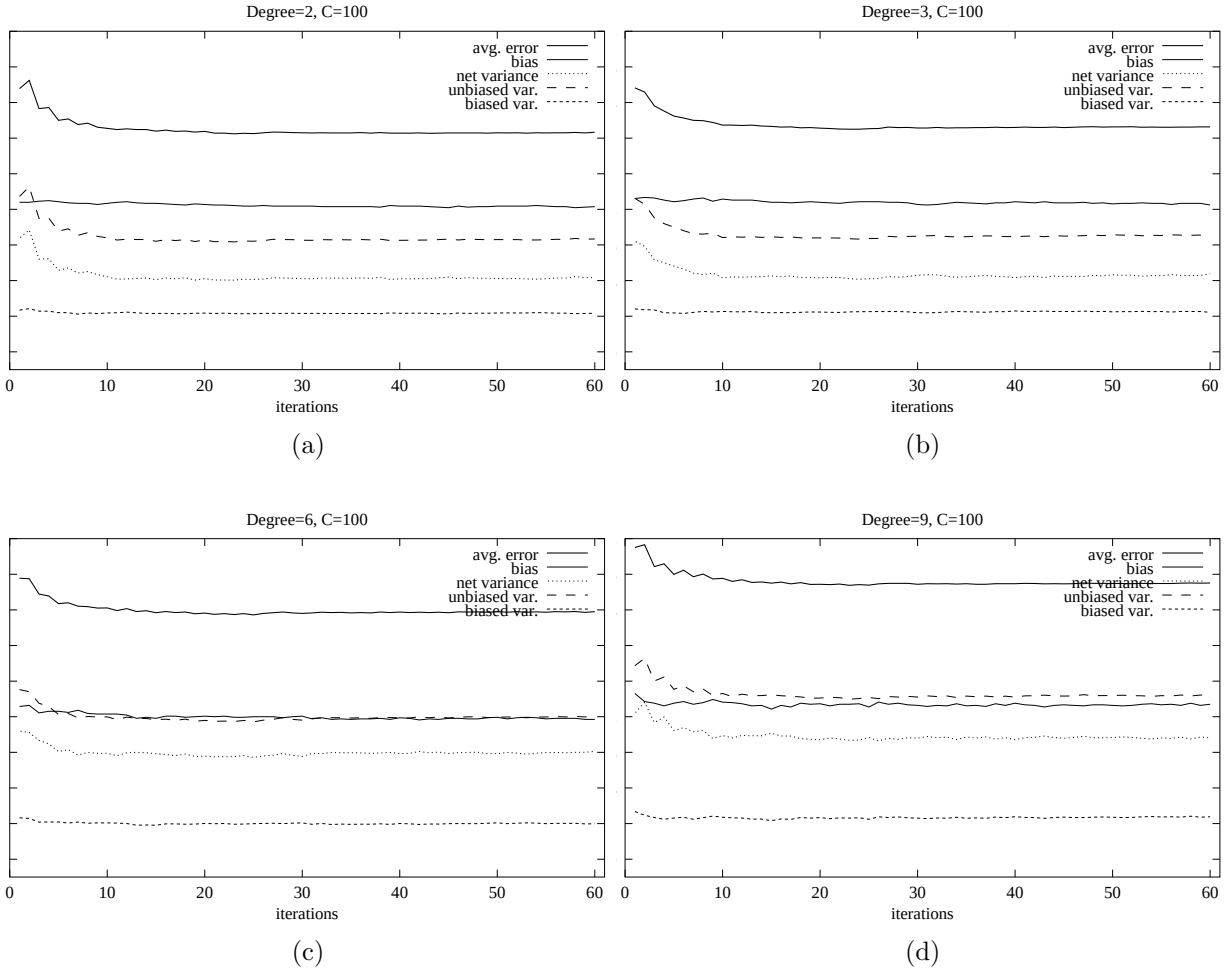


Figure 44: Musk data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 100$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

3.3 Decomposition with respect to the number of base learners in Dot-product SVM bagged ensembles

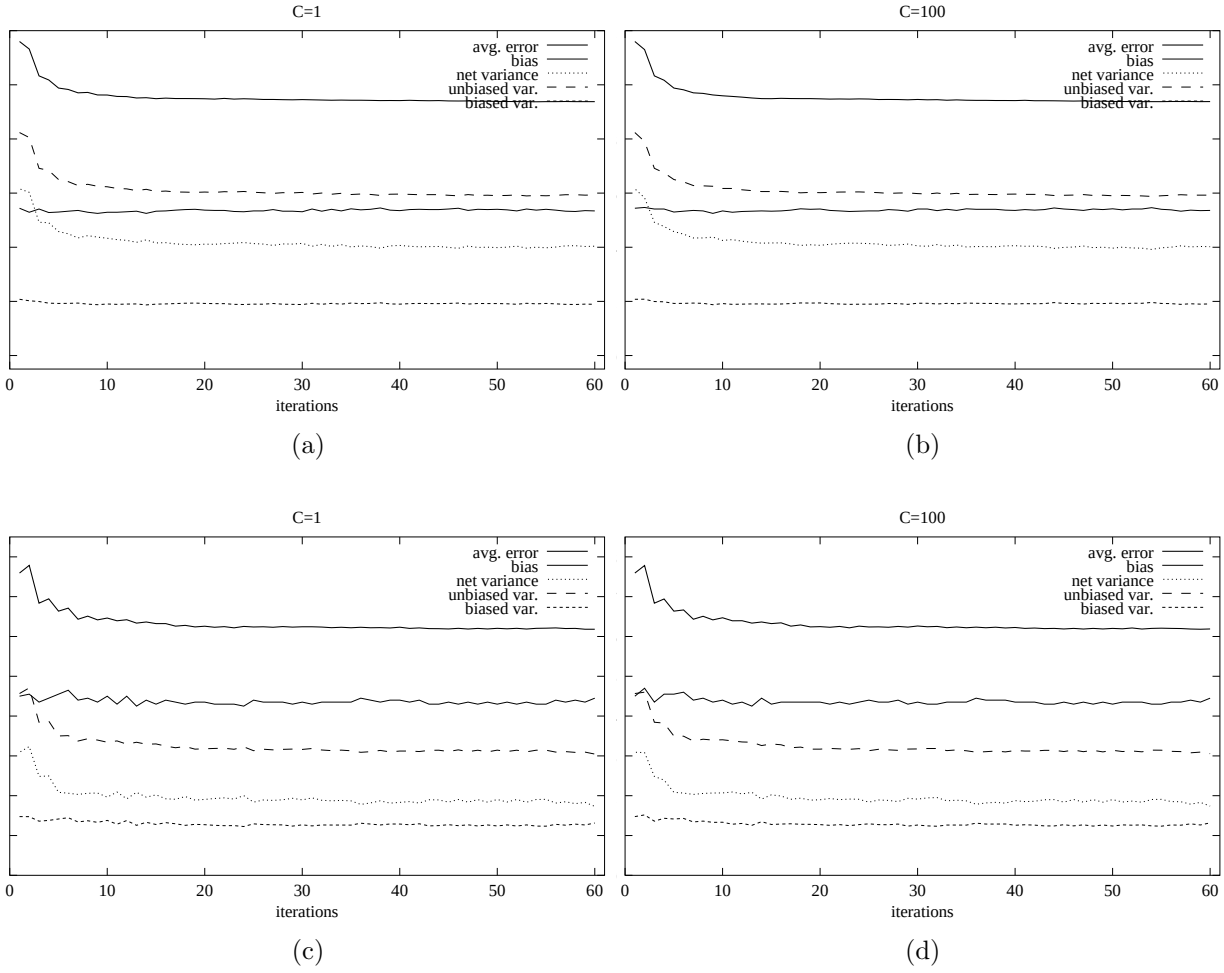


Figure 45: Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged dot-product SVM with respect to the number of iterations: (a) Waveform, $C = 1$, (b) Waveform, $C = 100$, (c) Grey-Landsat, $C = 1$, (d) Grey-Landsat, $C = 100$.

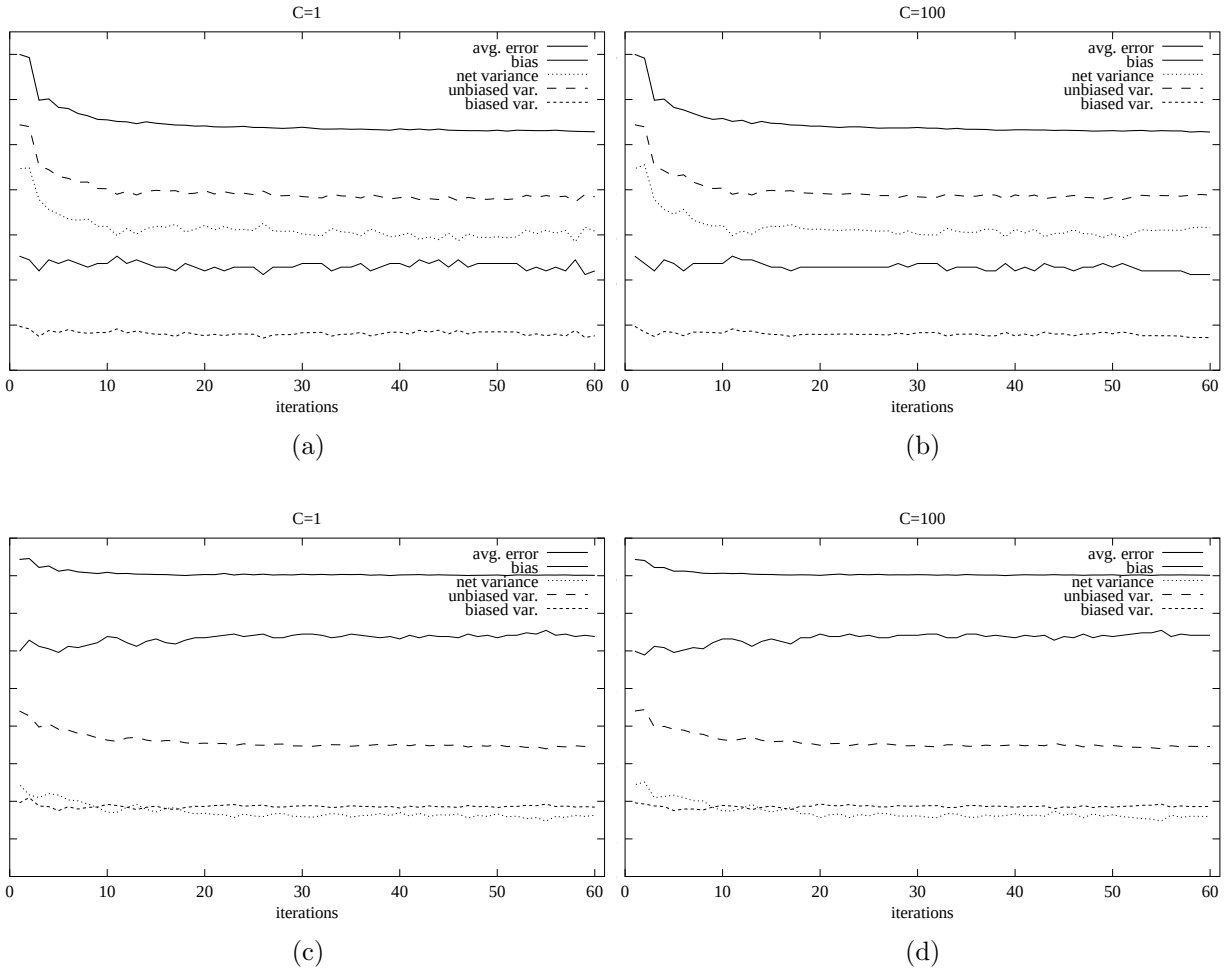


Figure 46: Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged dot-product SVM with respect to the number of iterations: (a) Letter-Two, $C = 1$, (b) Letter-Two, $C = 100$, (c) Letter-Two woth noise, $C = 1$, (d) Letter-Two woth noise, $C = 100$.

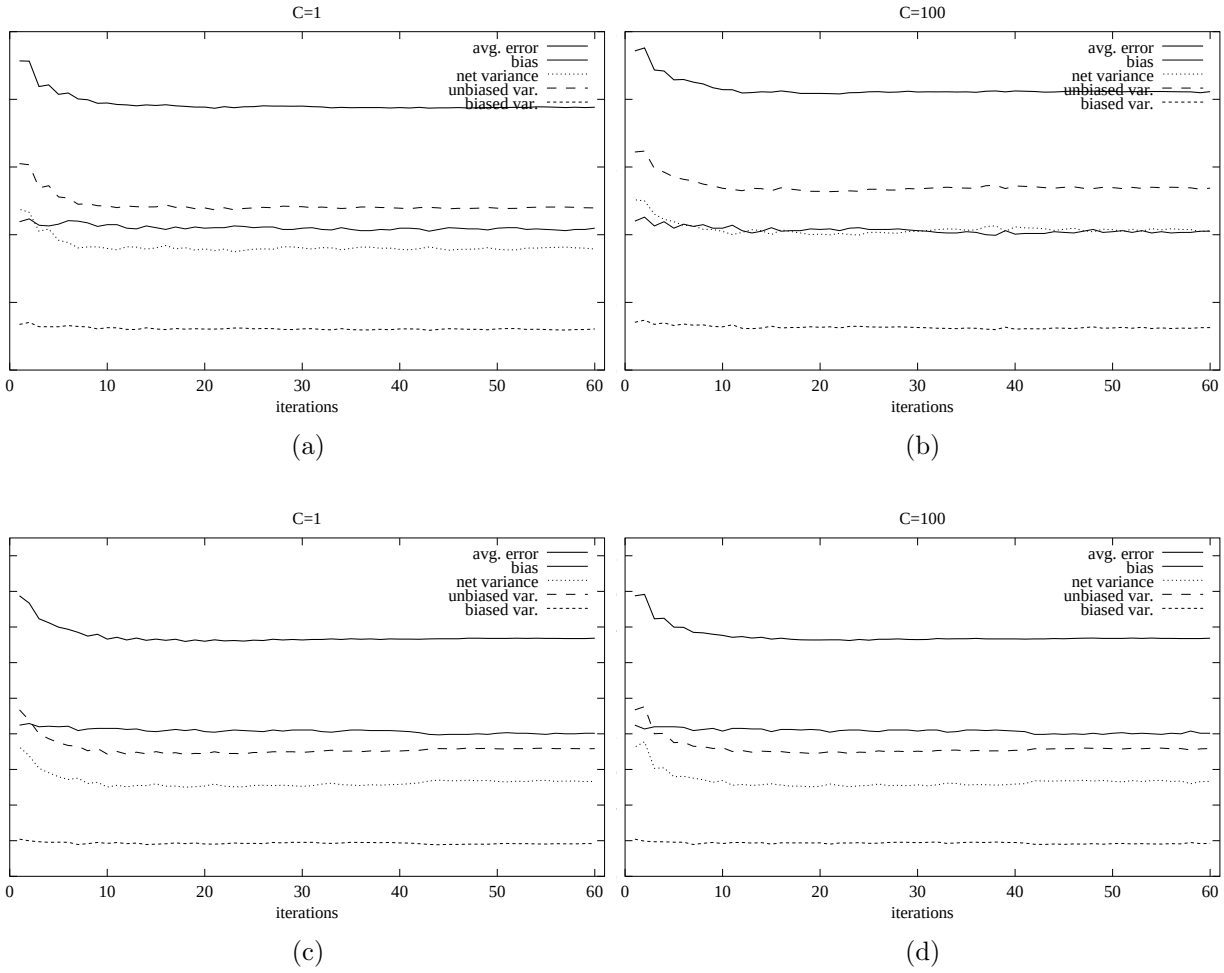


Figure 47: Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in bagged dot-product SVM with respect to the number of iterations: (a) Spam, $C = 1$, (b) Spam, $C = 100$, (c) Musk, $C = 1$, (d) Musk, $C = 100$.

4 Comparison of bias–variance decomposition in single and bagged SVMs.

Here are reported the graphics comparing bias–variance decomposition in single SVMs and bagged ensembles of SVMs. In all graphics of this section the data referred to single SVMs are labeled with crosses, while bagged SVMs are labeled with triangles. The corresponding quantities (e.g. bias, net-variance, etc.) are represented with the same type of line both in single and bagged SVMs.

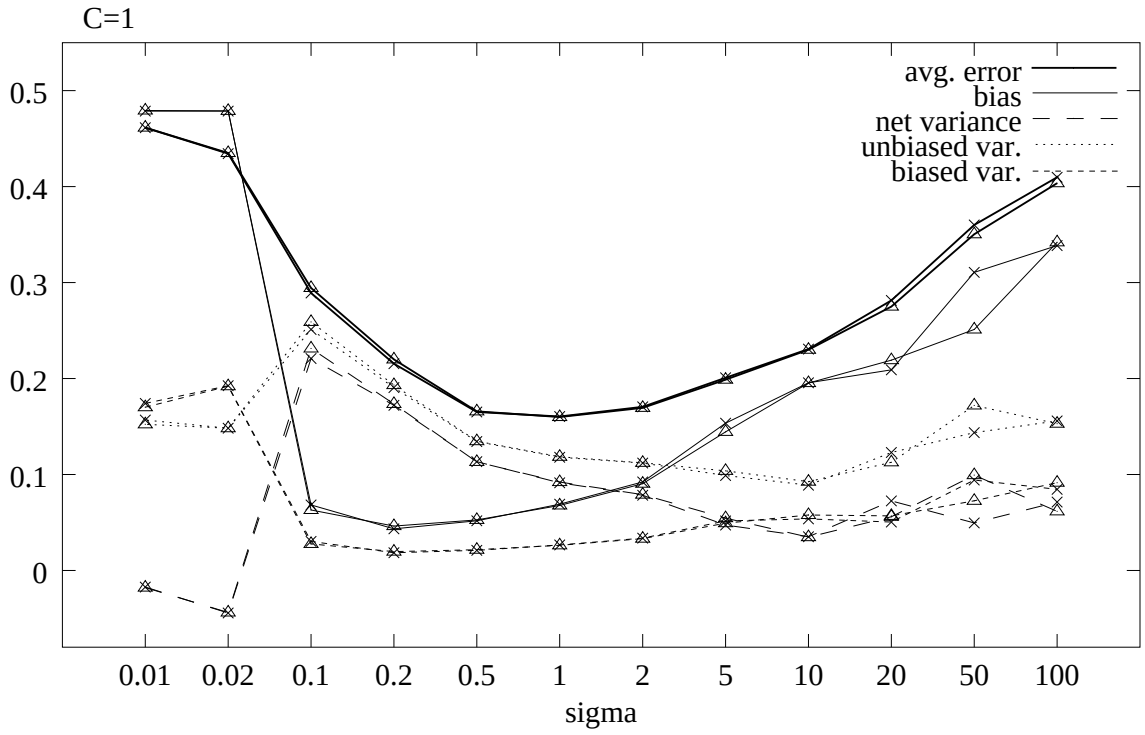
4.1 Comparison between single and bagged RBF-SVM

We analyze the relationships between bias-variance decompostion of the error in single and bagged RBF-SVMs for each different region that characterizes the bias-variance decompostion itself.

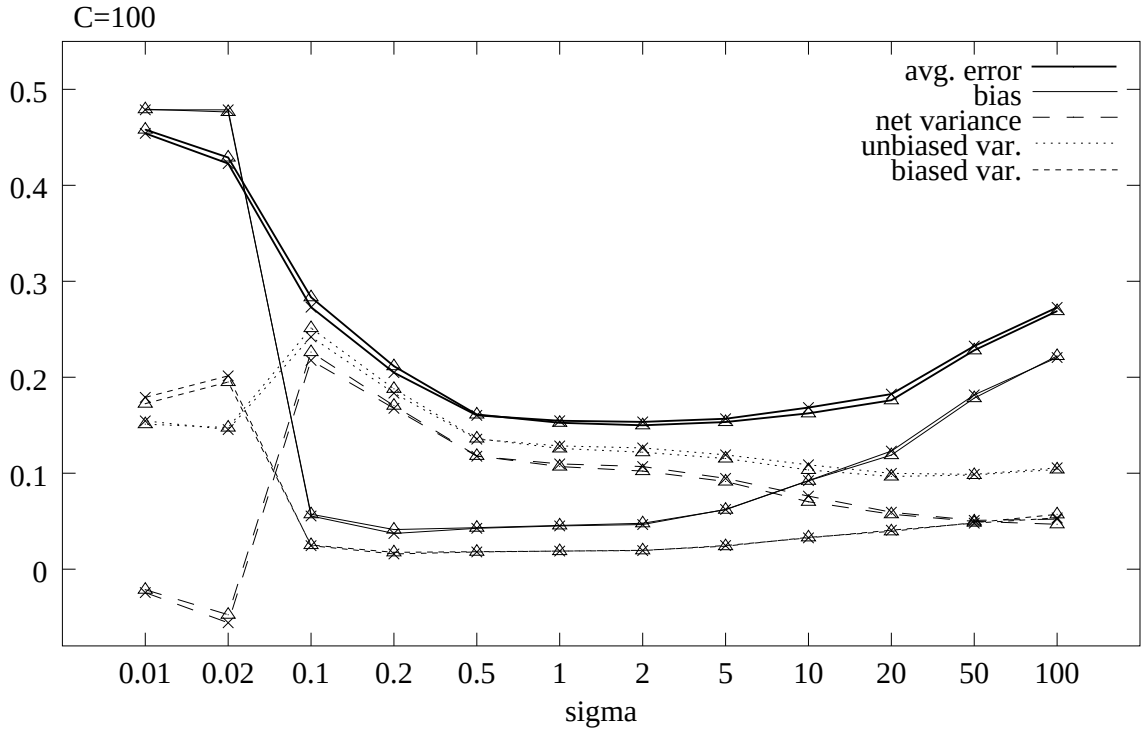
High bias region. In this region the the error of single and bagged SVMs is about equal, and it is characterized by a very high bias. The net-variance is close to 0, because biased variance is about equal to the unbiased variance. In most cases they are both close to 0. In some cases they are equal but greater than 0 with significantly larger values in single that in in bagged SVMs (Fig. 49).

Transition region. The bias goes down in the transition region at about the same rate in single and bagged SVM ensembles. The net-variance mainitains the wave-shape also in bagged SVMs, but it is lower. In some data sets (Fig. 48, 52) the net-variance remains low with no significant variations also for small values of σ. By there reasons the error drops down more quickly in bagged SVMs, and the error of the ensemble is about equal to the bias.

Stabilized region. The net-variance stabilizes, but at lower values compared with net-variance of single SVMs. Hence we have a reduction of the error for bagged SVM ensembles in this region. Note that the reduction of the error dpends heavily on the level of the unbiased variance in the stabilized region. If it is sufficiently high, we can achieve substantial reduction of the error in bagged SVM enesmbles.

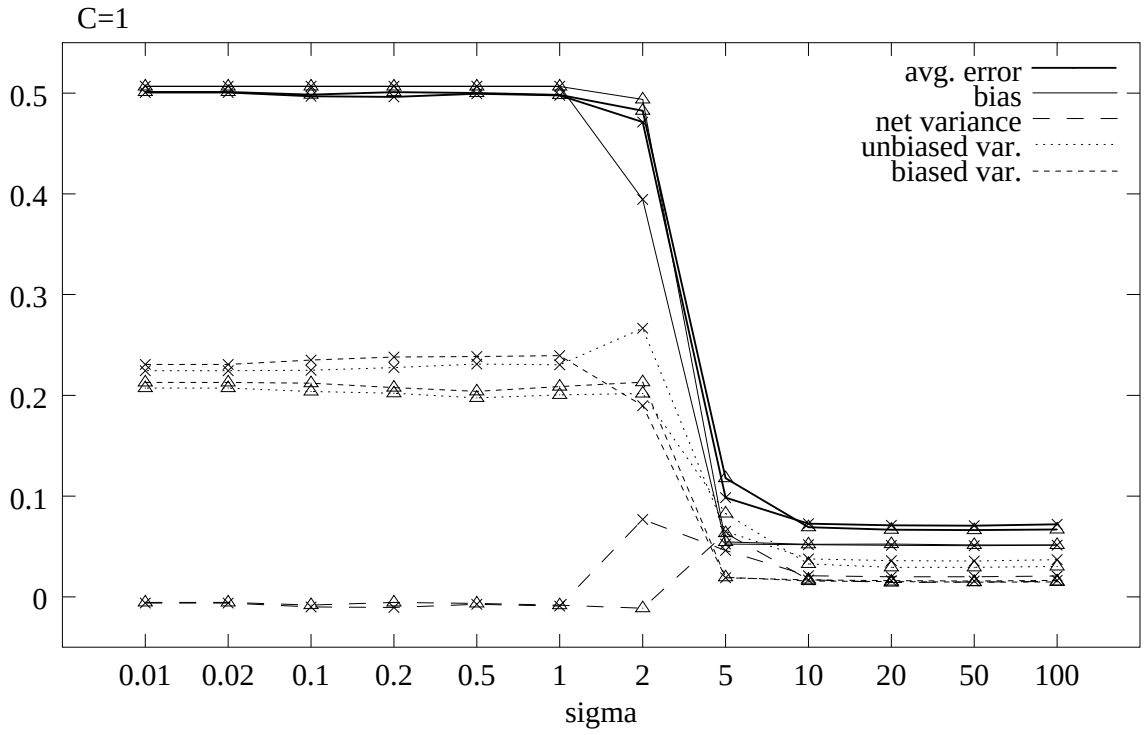


(a)

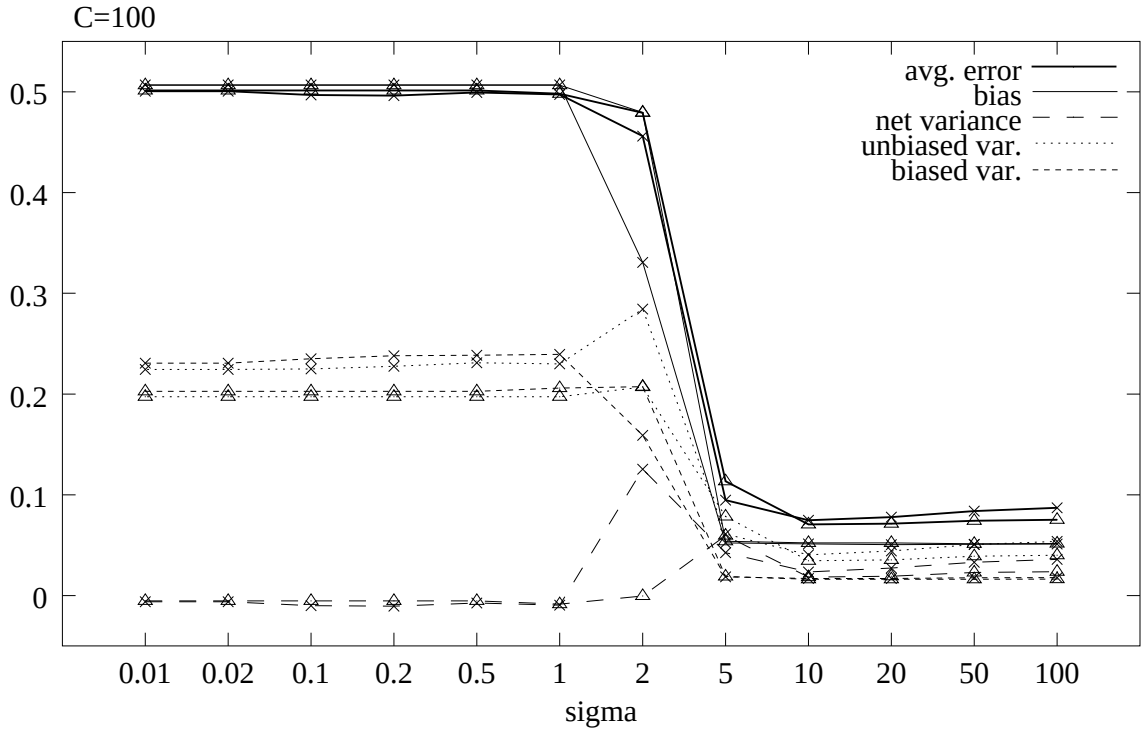


(b)

Figure 48: P2 data set. Comparison between bias-variance decomposition between single RBF-SVMs (lines labeled with crosses) and bagged SVM RBF ensembles (lines labeled with triangles), while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

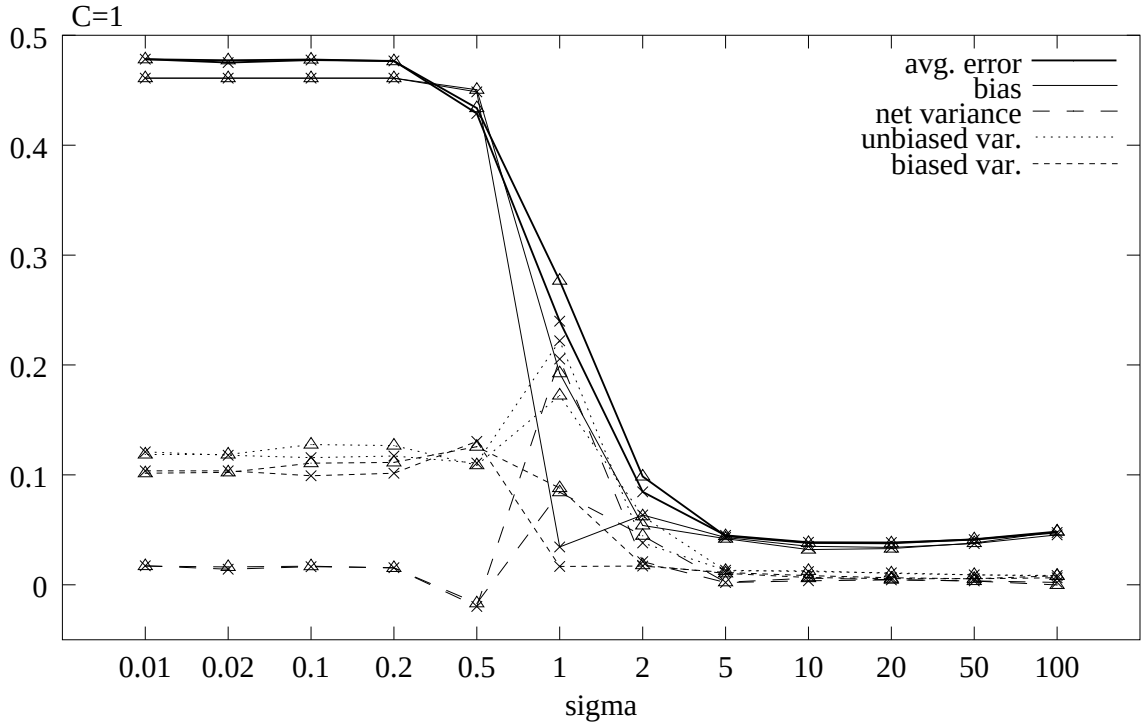


(a)

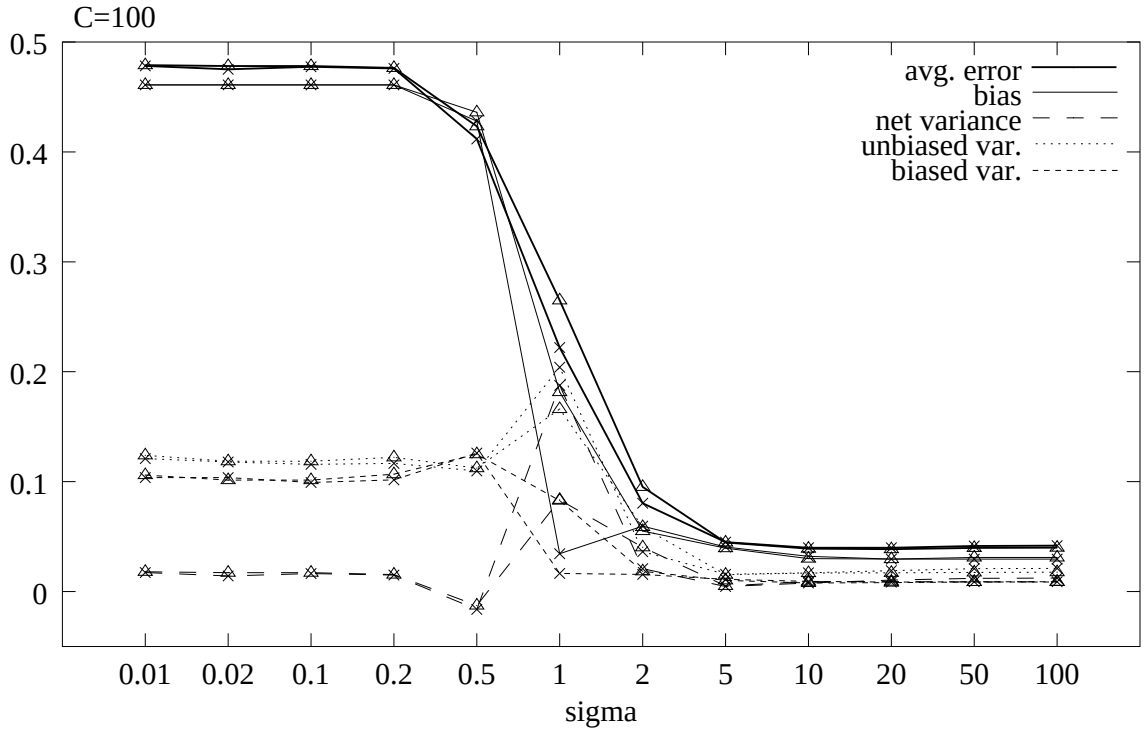


(b)

Figure 49: Waveform data set. Comparison of bias-variance decomposition between single RBF-SVMs (lines labeled with crosses) and bagged SVM RBF ensembles (lines labeled with triangles), while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

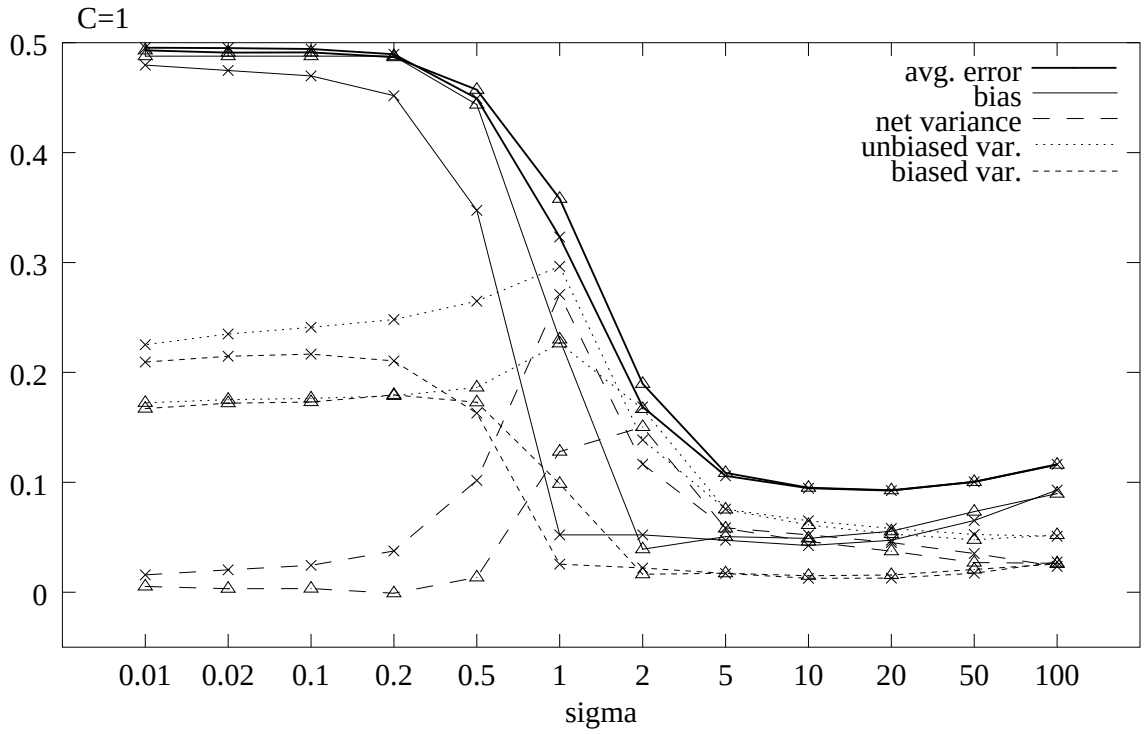


(a)

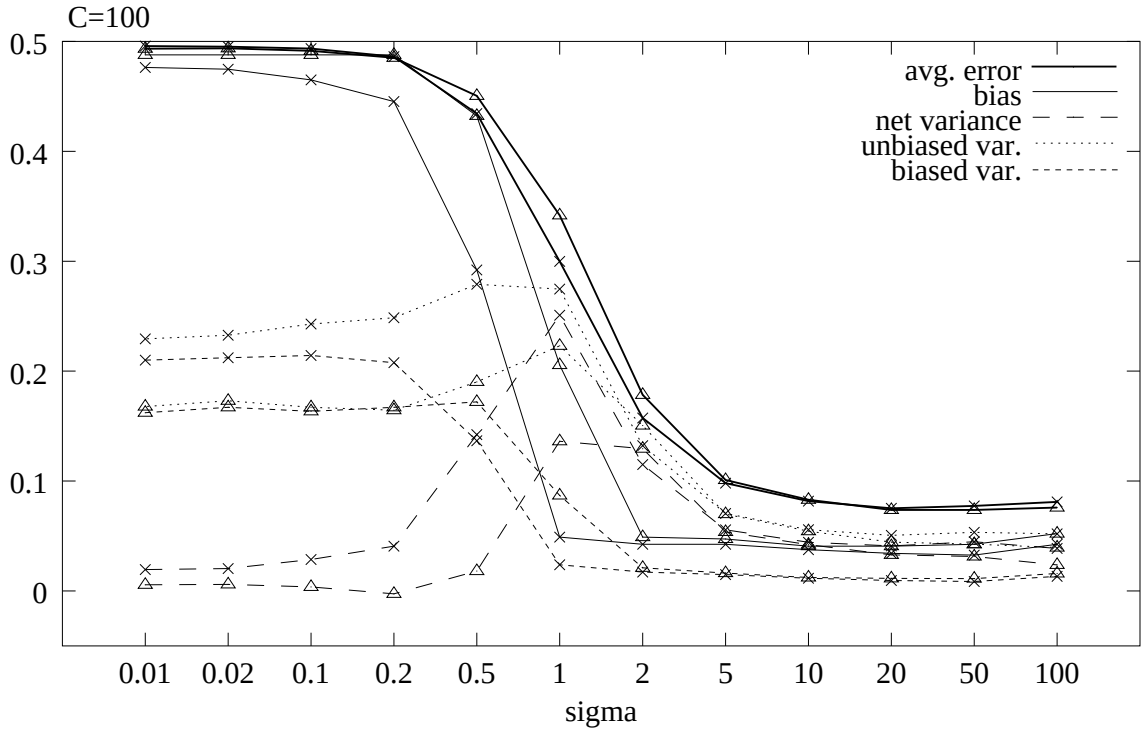


(b)

Figure 50: Grey-Landsat data set. Comparison of bias-variance decomposition between single RBF-SVMs (lines labeled with crosses) and bagged SVM RBF ensembles (lines labeled with triangles), while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

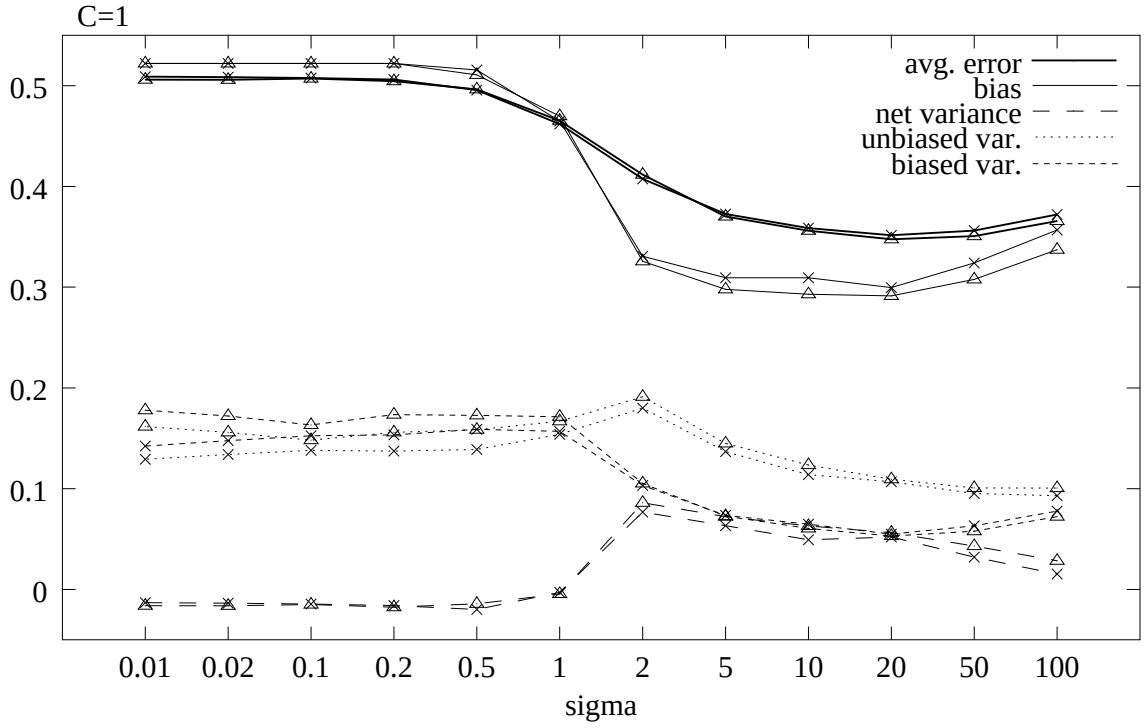


(a)

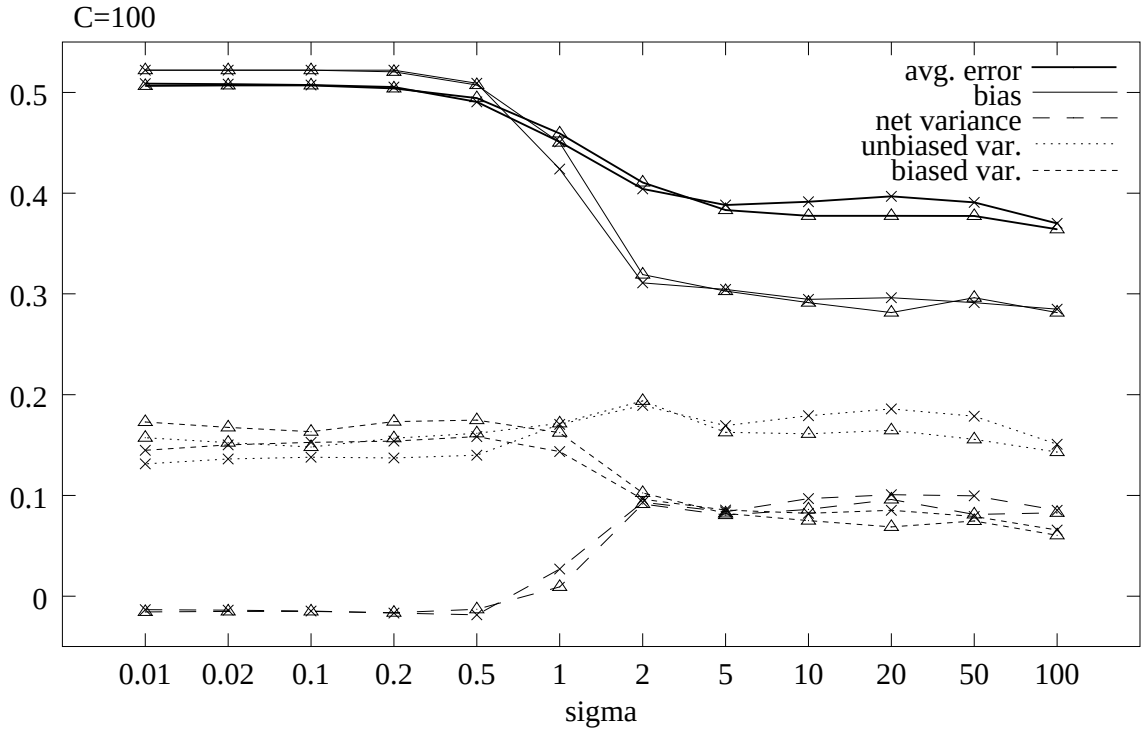


(b)

Figure 51: Letter-Two data set. Comparison of bias-variance decomposition between single RBF-SVMs (lines labeled with crosses) and bagged SVM RBF ensembles (lines labeled with triangles), while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

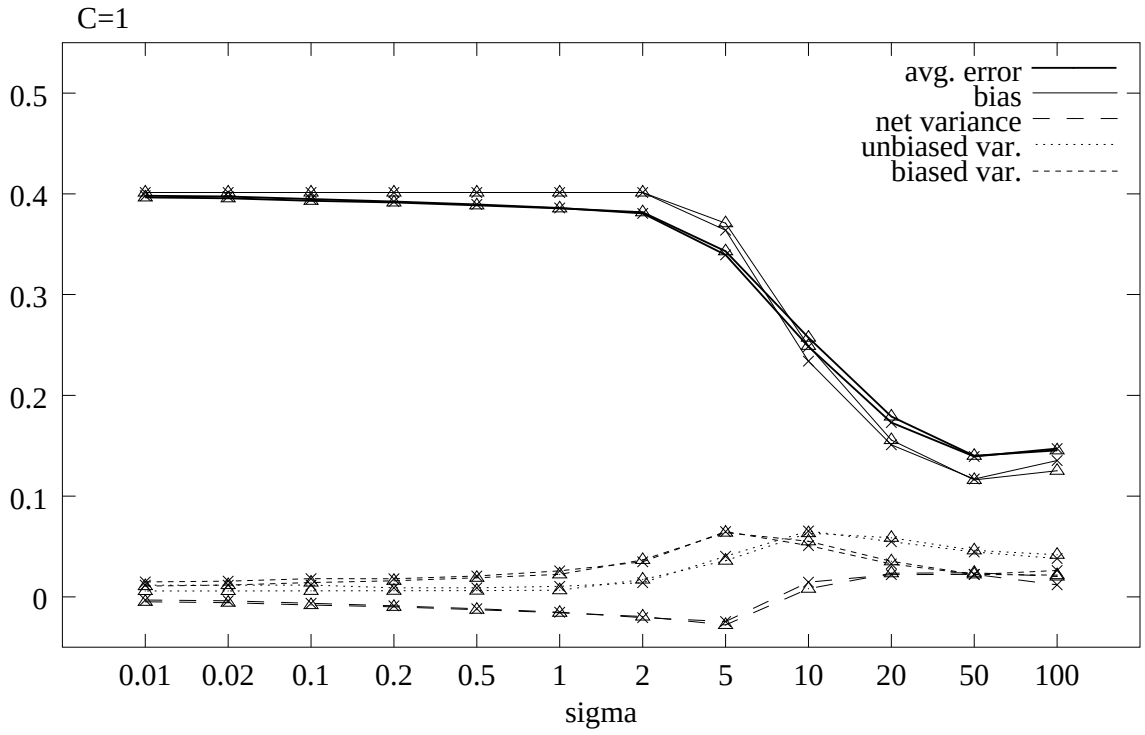


(a)

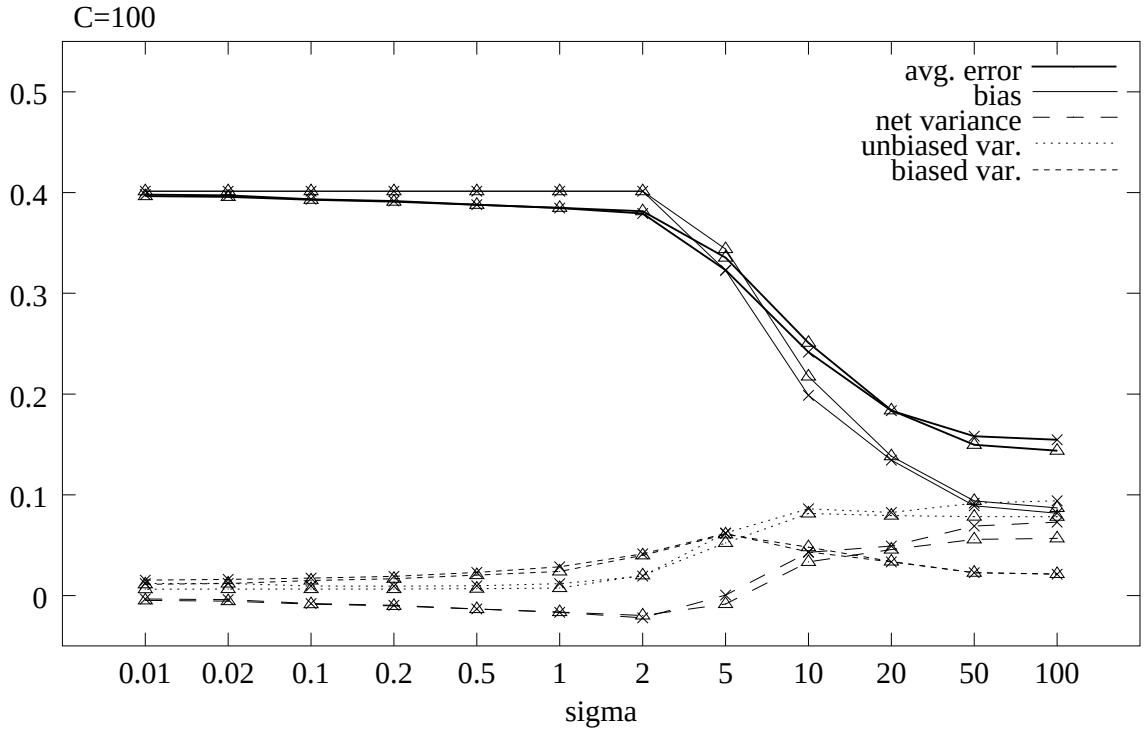


(b)

Figure 52: Letter-Two with noise data set. Comparison of bias-variance decomposition between single RBF-SVMs (lines labeled with crosses) and bagged SVM RBF ensembles (lines labeled with triangles), while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

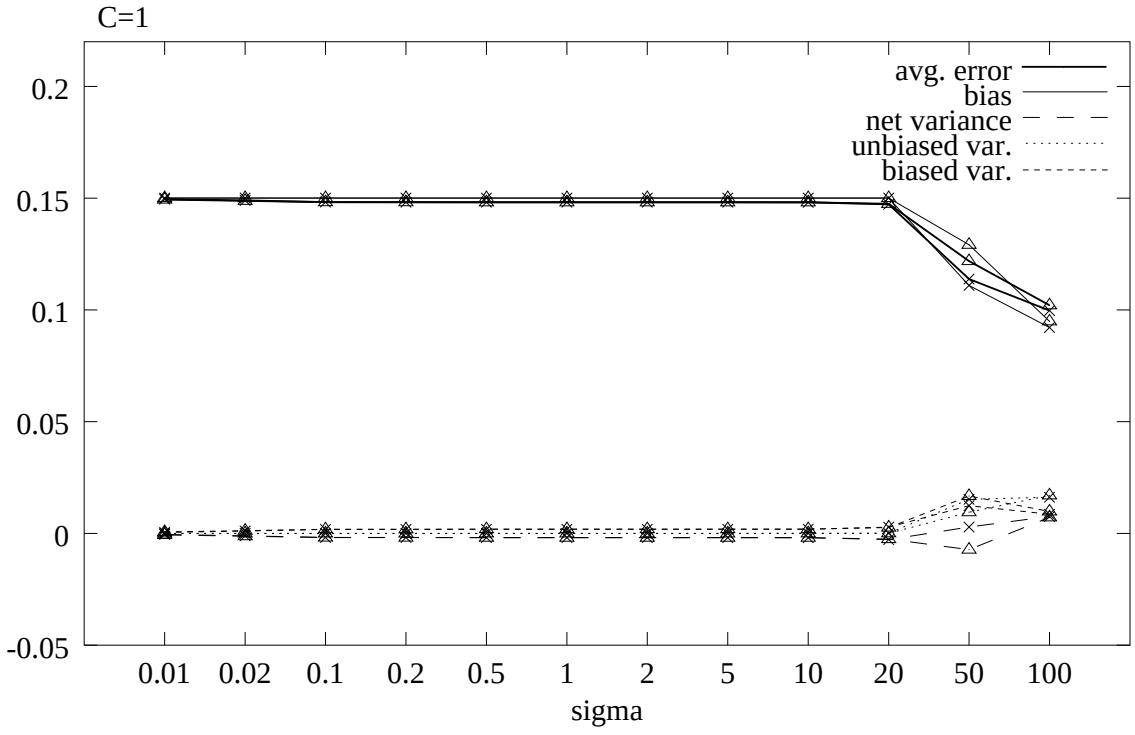


(a)

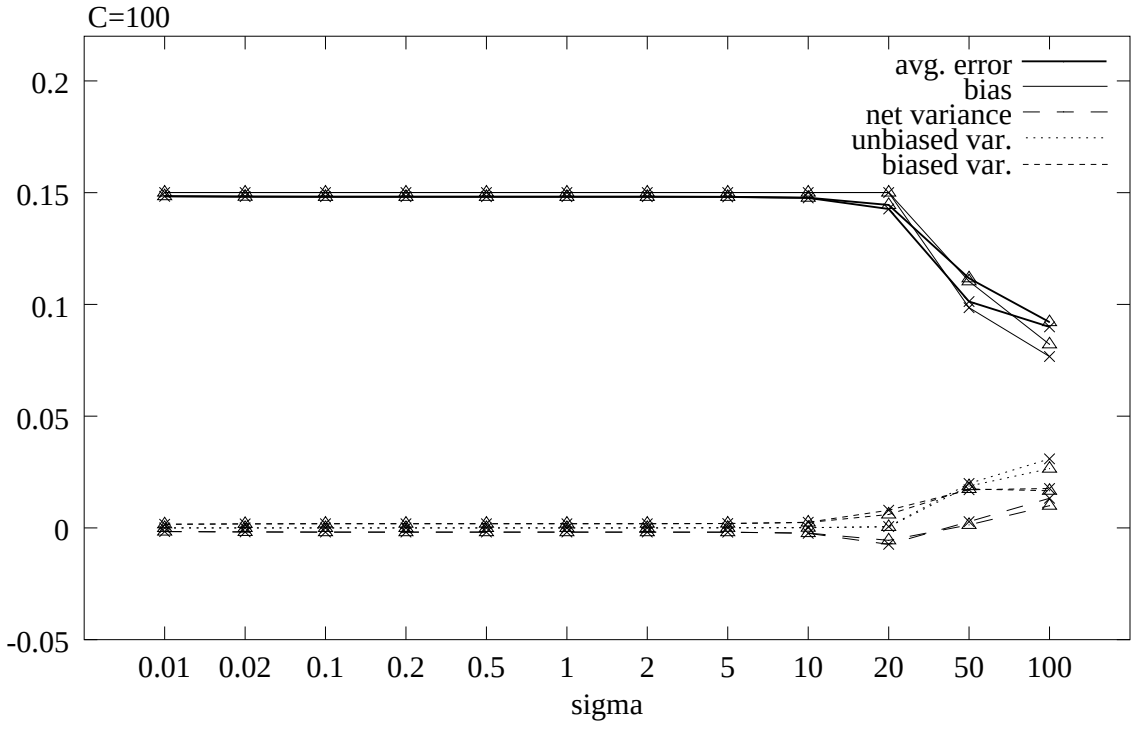


(b)

Figure 53: Spam data set. Comparison of bias-variance decomposition between single RBF-SVMs (lines labeled with crosses) and bagged SVM RBF ensembles (lines labeled with triangles), while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.



(a)



(b)

Figure 54: Musk data set. Comparison of bias-variance decomposition between single RBF-SVMs (lines labeled with crosses) and bagged SVM RBF ensembles (lines labeled with triangles), while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

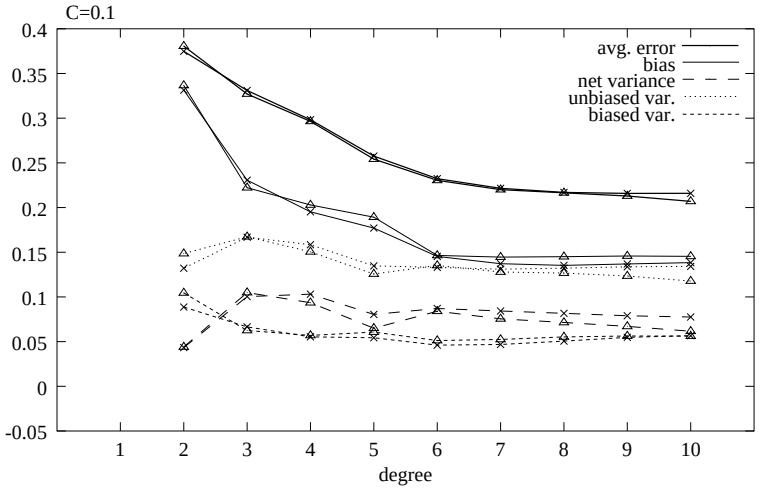
4.2 Comparison between single and bagged polynomial SVM

In bagged polynomial SVMs the error is due almost entirely to the bias. The bias component is about equal in bagged and single SVMs.

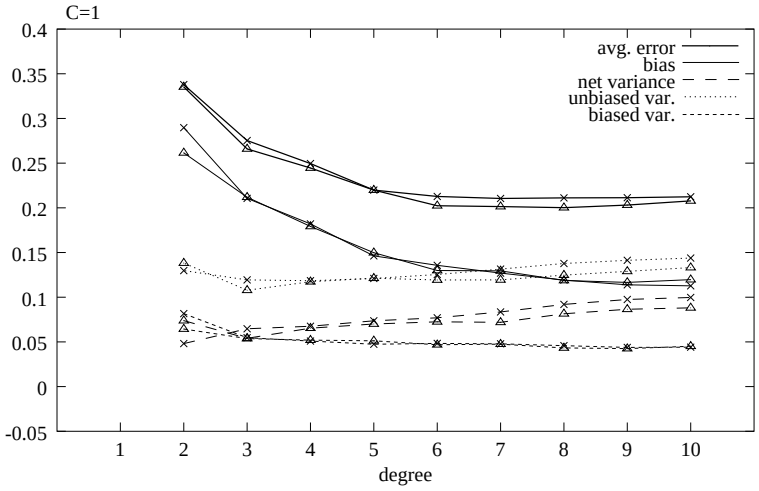
The variance component is close to 0: its contribute to the error is almost removed.

In bagged SVMs, the shape of the error with respect to the degree depends on the shape of the bias. Consequently the error curve shape is bias-dependent,, while in single SVMs it is variance or bias-variance dependent.

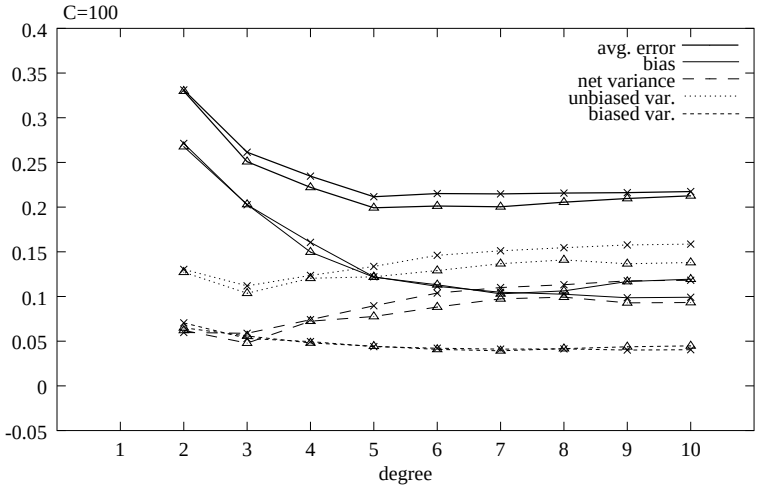
The general shape of the error with respect to the degree resembles an "U" curve, or can be flatted in dependence of the bias trend.



(a)

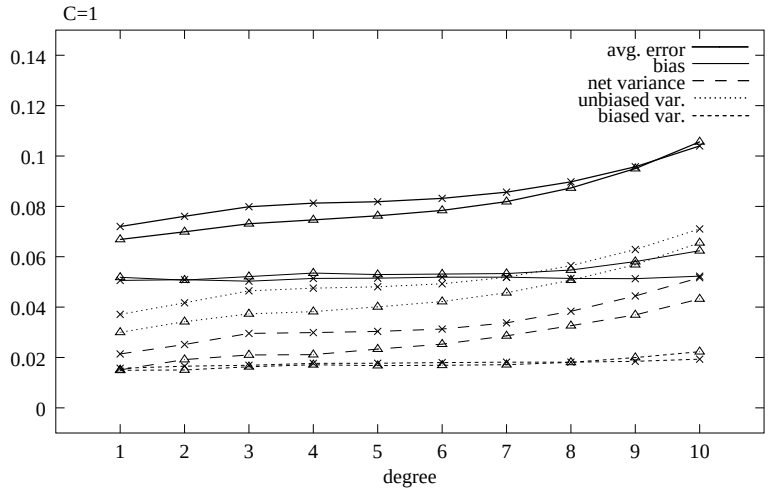


(b)

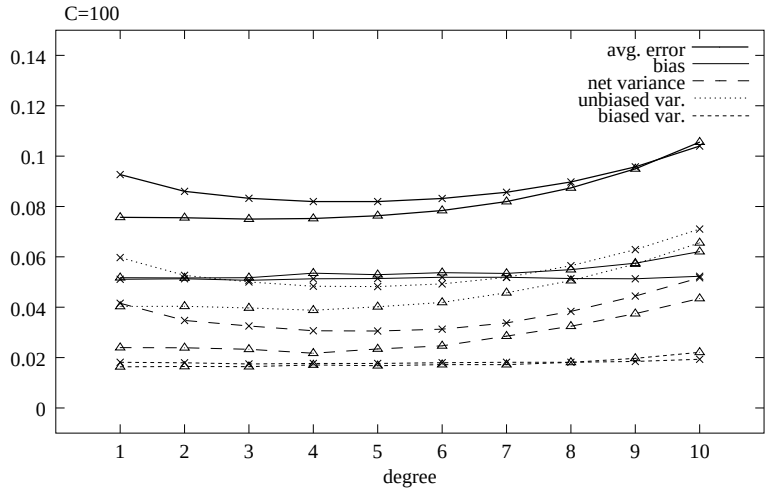


(c)

Figure 55: P2 data set. Comparison of bias-variance decomposition between single polynomial SVMs (lines labeled with crosses) and bagged polynomial SVM ensembles (lines labeled with triangles), while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 1$, (c) $C = 100$.

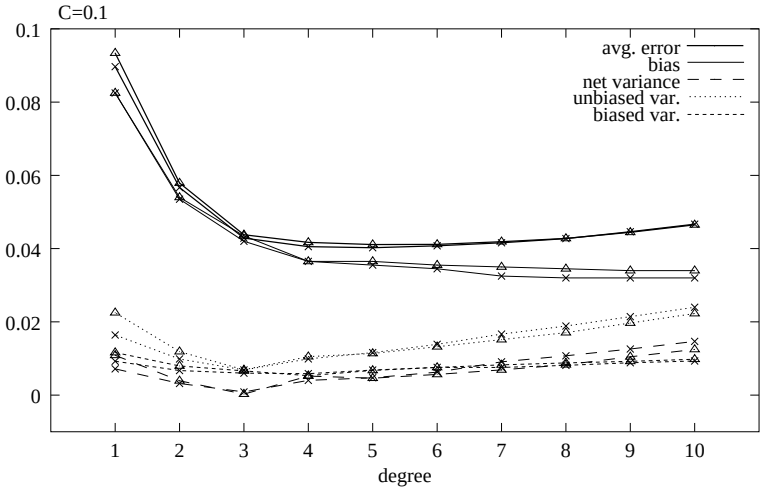


(a)

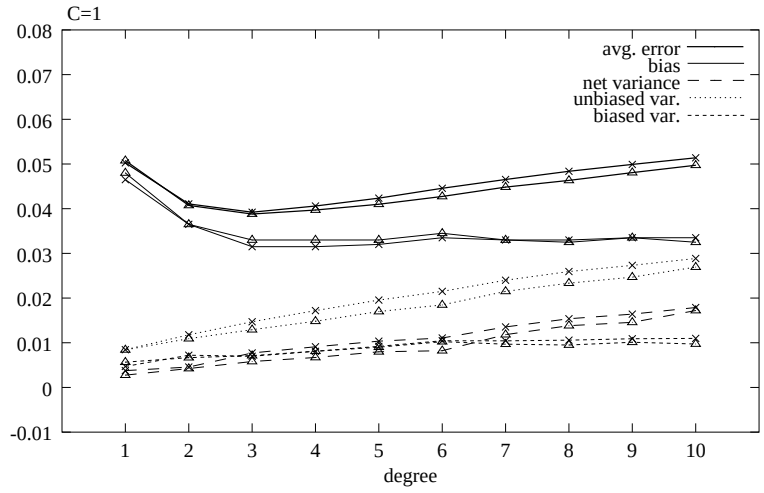


(b)

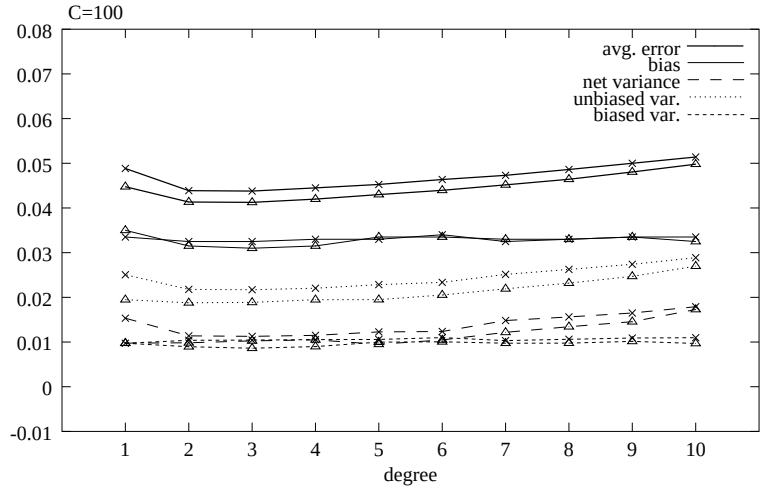
Figure 56: Waveform data set. Comparison of bias-variance decomposition between single polynomial SVMs (lines labeled with crosses) and bagged polynomial SVM ensembles (lines labeled with triangles), while varying the degree and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.



(a)

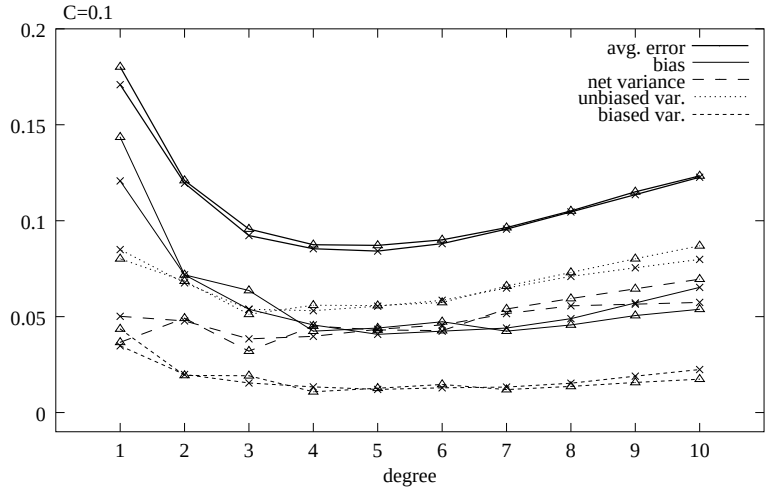


(b)

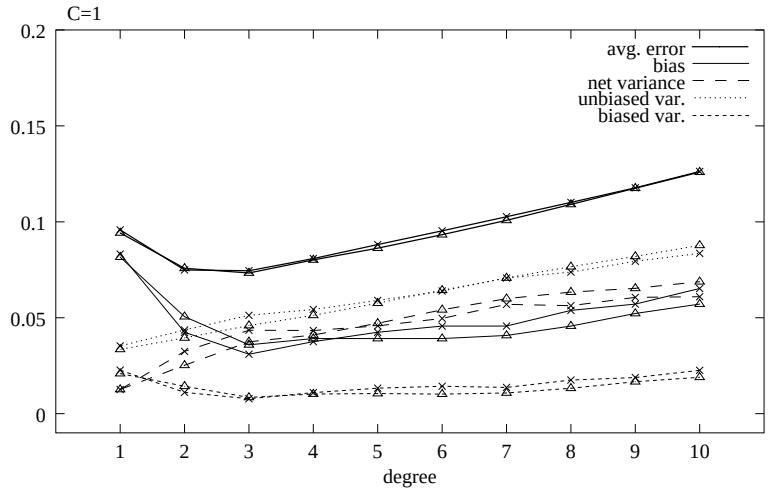


(c)

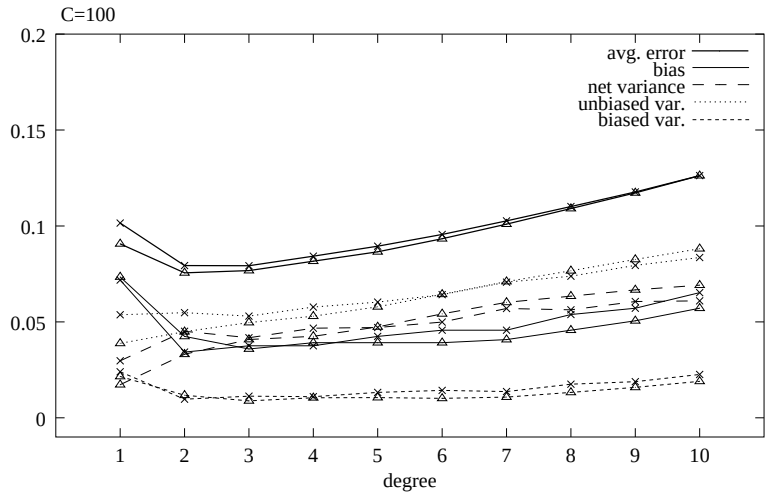
Figure 57: Grey-Landsat data set. Comparison of bias-variance decomposition between single polynomial SVMs (lines labeled with crosses) and bagged polynomial SVM ensembles (lines labeled with triangles), while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 1$, (c) $C = 100$.



(a)

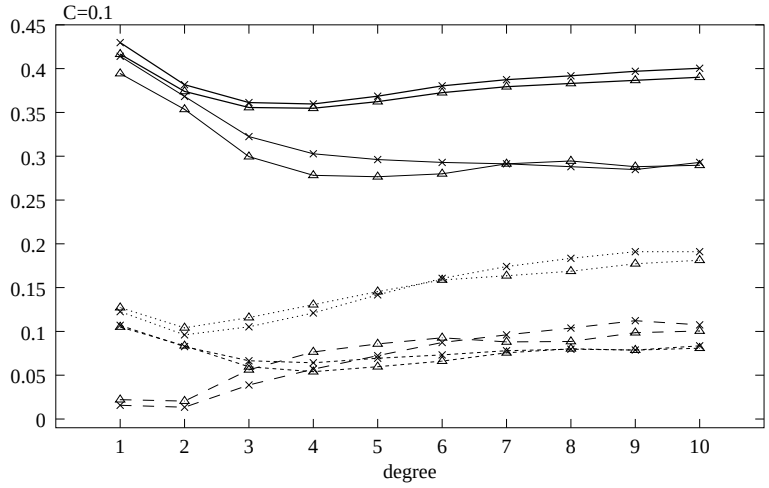


(b)

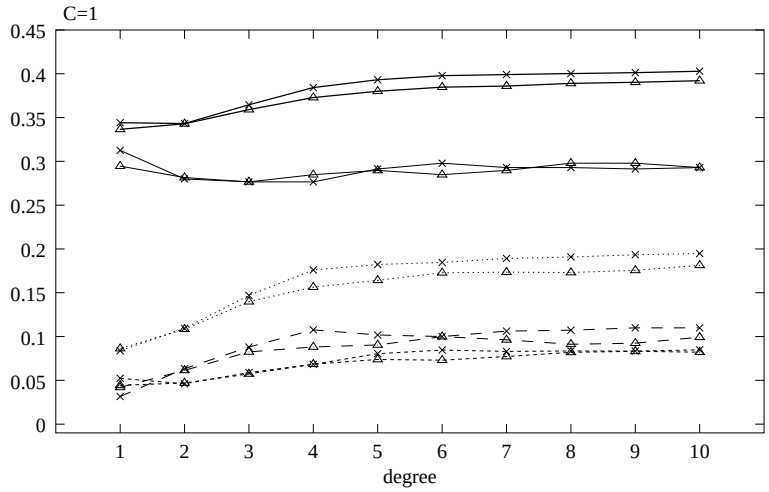


(c)

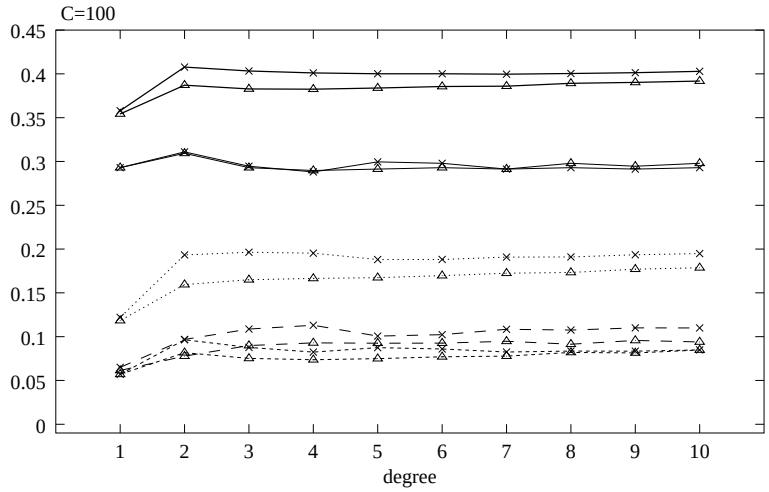
Figure 58: Letter-Two data set. Comparison of bias-variance decomposition between single polynomial SVMs (lines labeled with crosses) and bagged polynomial SVM ensembles (lines labeled with triangles), while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 1$, (c) $C = 100$.



(a)

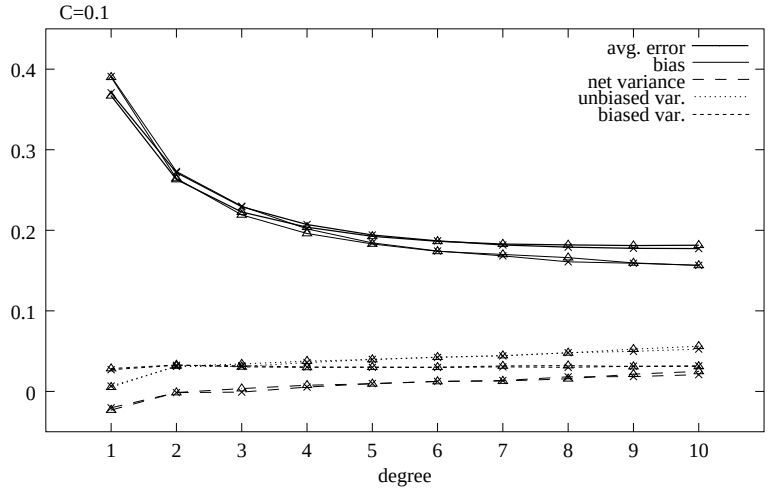


(b)

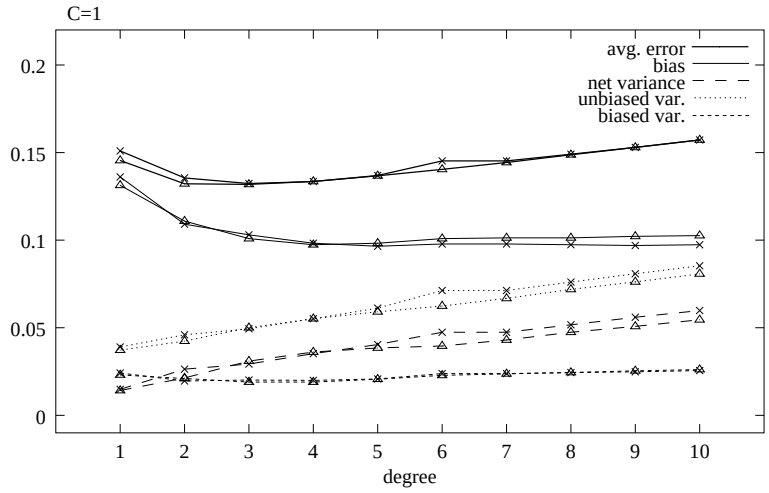


(c)

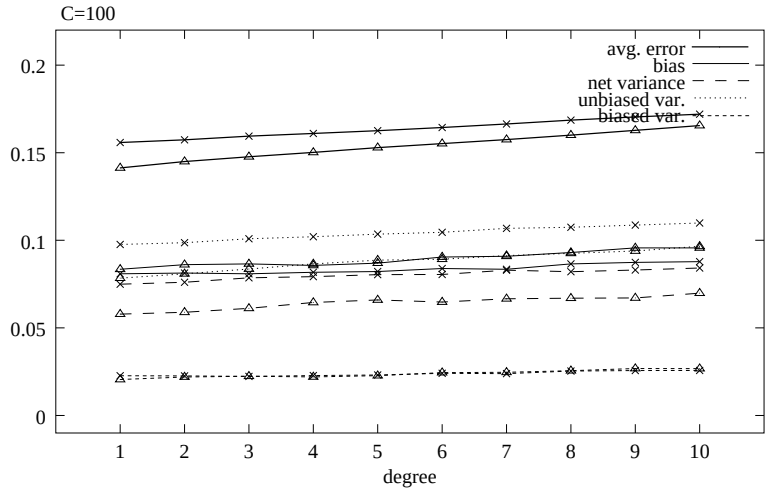
Figure 59: Letter-Two with noise data set. Comparison of bias-variance decomposition between single polynomial SVMs (lines labeled with crosses) and bagged polynomial SVM ensembles (lines labeled with triangles), while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 1$, (c) $C = 100$.



(a)

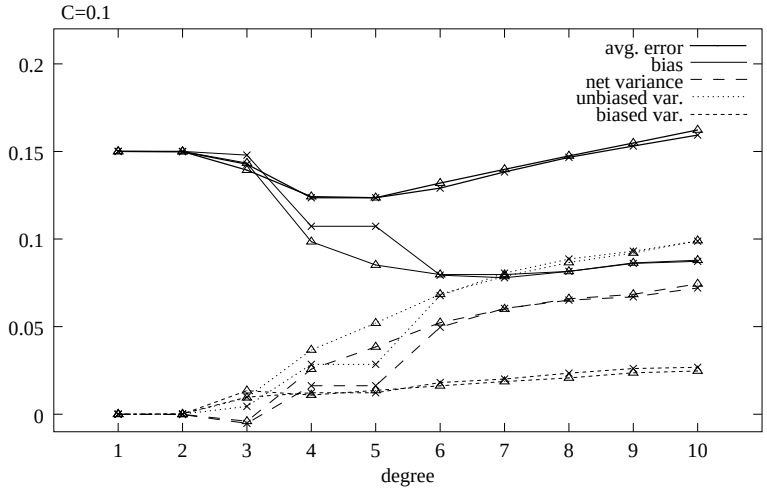


(b)

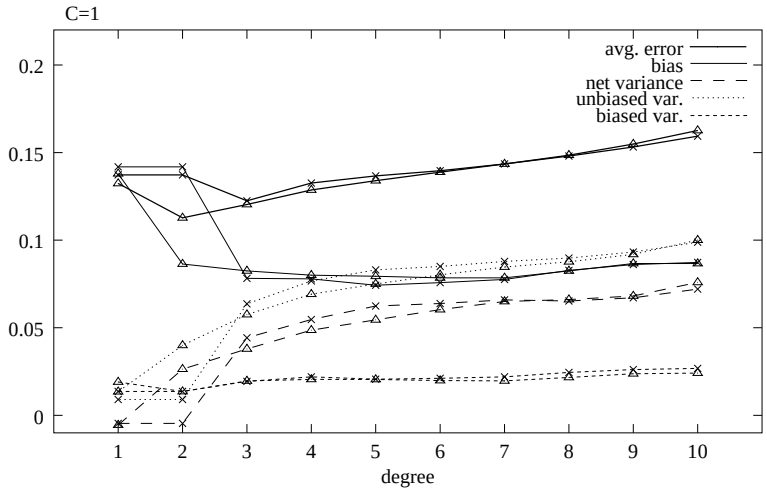


(c)

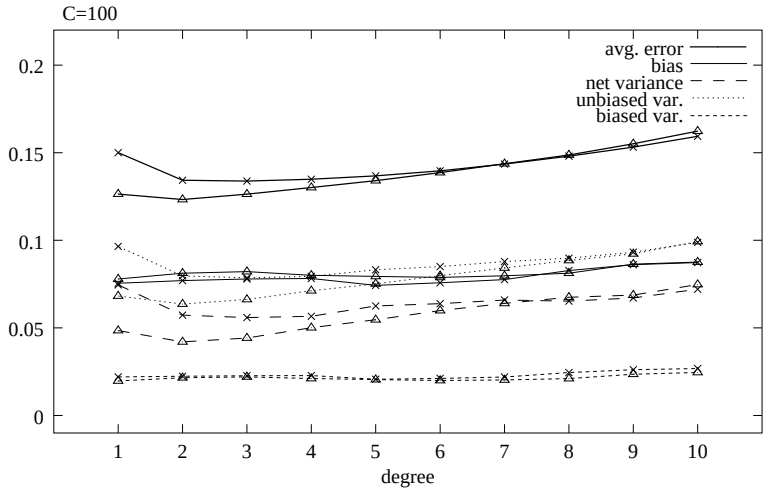
Figure 60: Spam data set. Comparison of bias-variance decomposition between single polynomial SVMs (lines labeled with crosses) and bagged polynomial SVM ensembles (lines labeled with triangles), while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 1$, (c) $C = 100$.



(a)



(b)



(c)

Figure 61: Musk data set. Comparison of bias-variance decomposition between single polynomial SVMs (lines labeled with crosses) and bagged polynomial SVM ensembles (lines labeled with triangles), while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 1$, (c) $C = 100$.

4.3 Comparison between single and bagged dot-product SVM

In all cases the error is about equal to the bias, that remains unchanged with respect to the single SVMs. As a consequence the error shape is equal to the shape of the bias and it is independent of the C values, at least for $C \geq 1$.

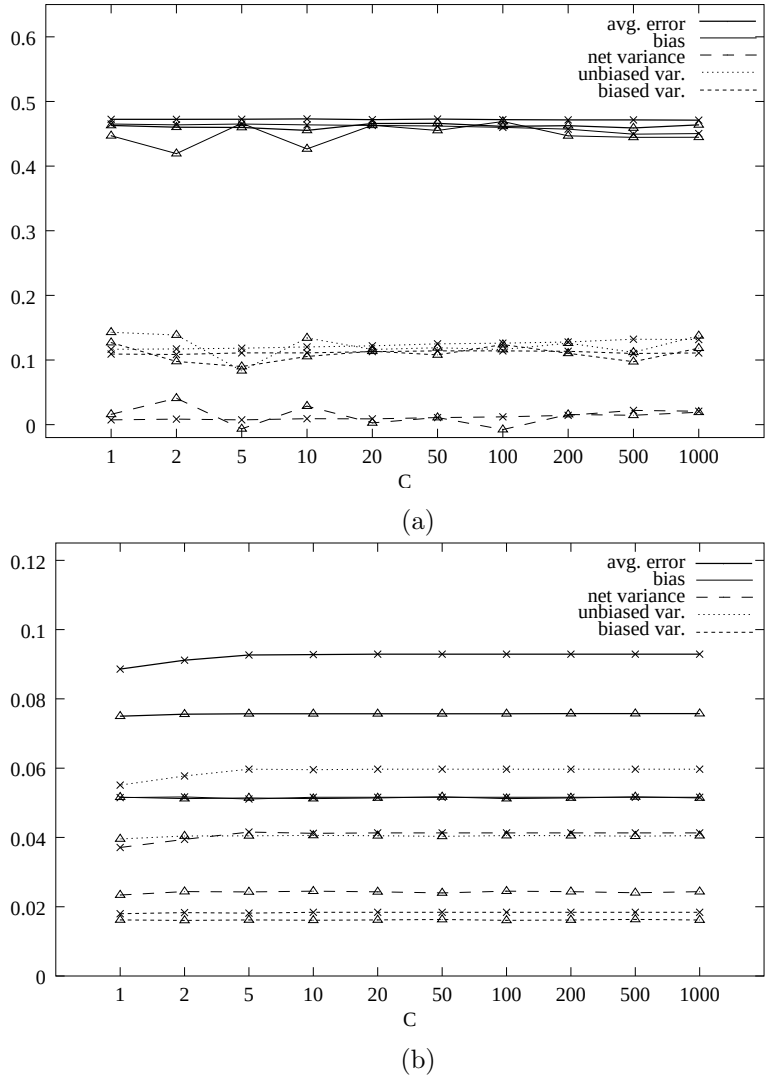
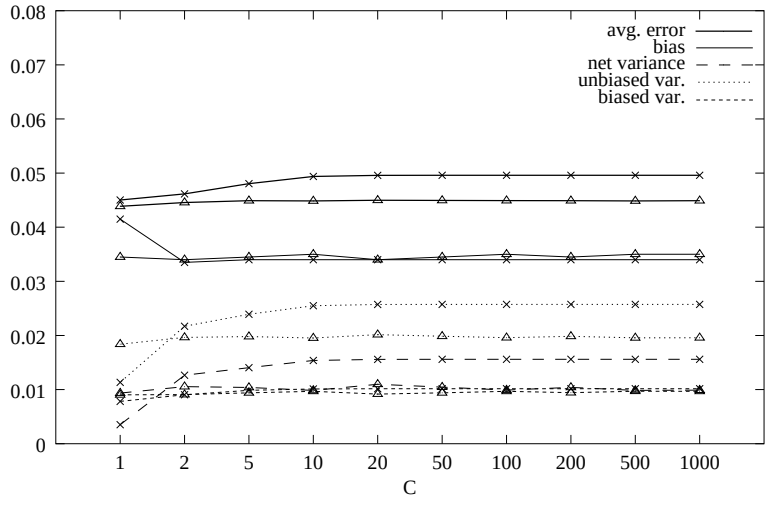
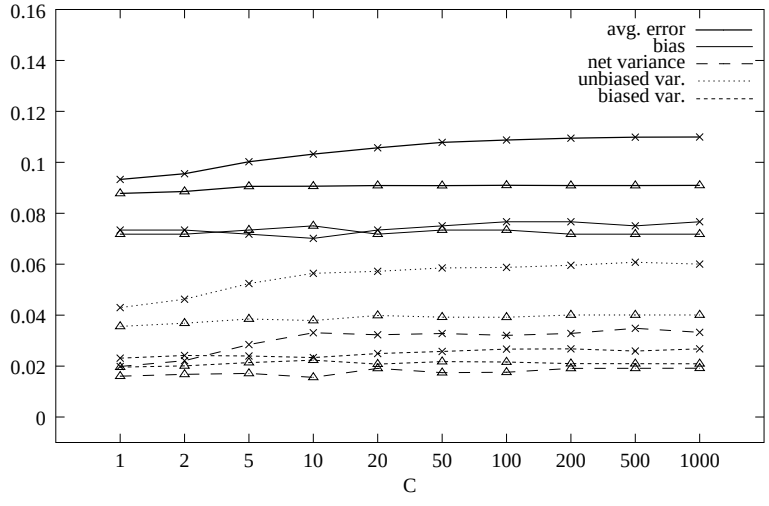


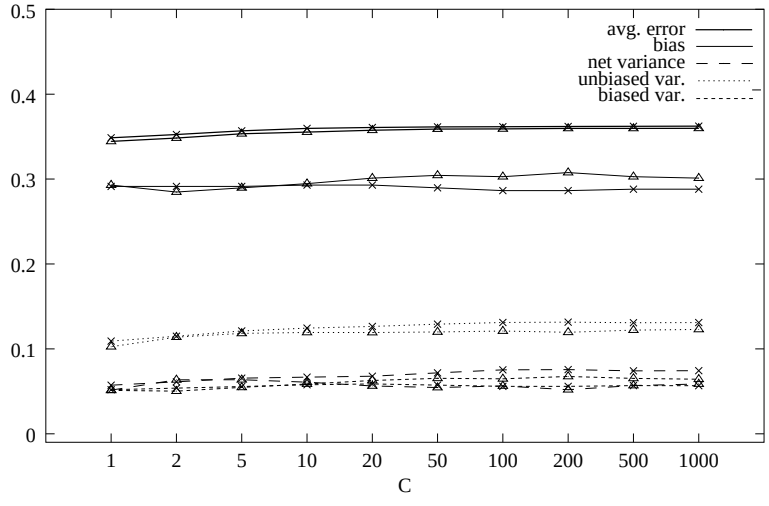
Figure 62: Comparison of bias-variance decomposition between single dot-product SVMs (lines labeled with crosses) and bagged dot-product SVM ensembles (lines labeled with triangles), while varying C: (a) P2, (b) Waveform.



(a)

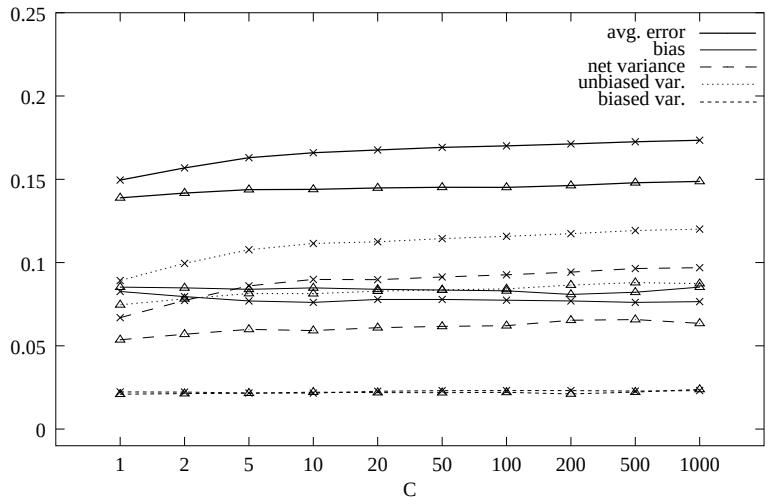


(b)

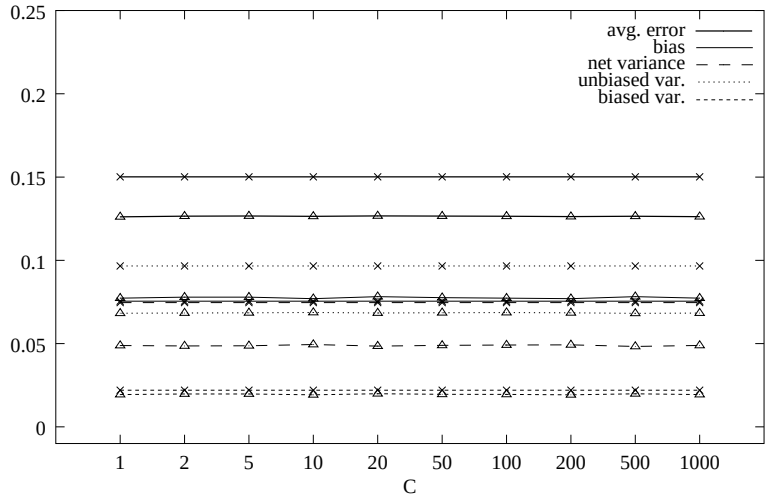


(c)

Figure 63: Comparison of bias-variance decomposition between single dot-product SVMs (lines labeled with crosses) and bagged dot-product SVM ensembles (lines labeled with triangles), while varying C: (a) Grey-Landsat, (b) Letter-Two, (c) Letter-Two with noise,



(a)



(b)

Figure 64: Comparison of bias-variance decomposition between single dot-product SVMs (lines labeled with crosses) and bagged dot-product SVM ensembles (lines labeled with triangles), while varying C: (a) Spam, (b) Musk.

5 Comparative performance between single SVM and bagged SVM ensembles: tables

In the following tables are summarized the compared results of bias–variance decomposition between single SVMs and bagged SVM ensembles, varying the parameters of the kernel. E_{SVM} stands for the estimated error of single SVMs, E_{bag} for the estimated error of bagged ensembles of SVMs, $\%\ Error\ reduction$ stands for the percent error reduction of the error between single and bagged ensembles, that is:

$$\%Error\ reduction = \frac{E_{SVM} - E_{bag}}{E_{SVM}}$$

$\%\ Bias\ reduction$, $\%\ NetVar\ reduction$ and $\%\ UnbVar\ reduction$ corresponds respectively to the percent bias, net–variance and unbiased variance reduction between single and bagged ensemble of SVMs. The negative signs means that we have a larger error in the bagged ensemble. Note that sometimes the decrement of the net–variance can be larger than 100 %: recall that net–variance can be negative (when the biased variance is larger than the inbiased variance).

5.1 Comparison between single and bagged RBF-SVMs

Table 1: Comparison of the best results between single and bagged SVMs.

	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
Data set *P2*						
RBF-SVM	0.1517	0.1500	1.14	-2.64	3.18	2.19
Poly-SVM	0.2088	0.1985	4.95	4.85	5.08	5.91
D-prod SVM	0.4715	0.4590	2.65	1.11	34.09	15.28
Data set *Waveform*						
RBF-SVM	0.0707	0.0662	6.30	-1.41	26.03	17.82
Poly-SVM	0.0761	0.0699	8.11	0.36	23.78	17.94
D-prod SVM	0.0886	0.0750	15.37	-0.22	37.00	28.20
Data set *Grey-Landsat*						
RBF-SVM	0.0384	0.0378	1.74	2.94	-7.46	3.94
Poly-SVM	0.0392	0.0388	1.05	-4.76	24.80	12.06
D-prod SVM	0.0450	0.0439	2.58	16.87	-165.72	-62.21
Data set *Letter-Two*						
RBF-SVM	0.0745	0.0736	1.20	-25.00	21.63	12.29
Poly-SVM	0.0745	0.0733	1.55	-15.79	13.92	10.41
D-prod SVM	0.0955	0.0878	8.09	2.22	27.55	23.06
Data set *Letter-Two with added noise*						
RBF-SVM	0.3362	0.3345	0.49	1.75	-5.78	0.40
Poly-SVM	0.3432	0.3429	0.09	-0.58	3.06	0.91
D-prod SVM	0.3486	0.3444	1.21	-0.56	10.23	6.09
Data set *Spam*						
RBF-SVM	0.1292	0.1290	0.14	-0.48	1.57	2.22
Poly-SVM	0.1323	0.1318	0.35	2.11	-5.83	-1.19
D-prod SVM	0.1495	0.1389	7.15	-3.16	19.87	16.38
Data set *Musk*						
RBF-SVM	0.1373	0.1324	3.57	2.78	20.77	-49.49
Poly-SVM	0.1225	0.1128	7.92	-10.49	38.17	37.26
D-prod SVM	0.1501	0.1261	15.97	-2.41	34.56	29.38

Table 2: Comparison of the bias variance decomposition between single RBF SVM and bagged RBF SVM, with the P2 data set.

Parameters σ and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
σ	**C = 1**					
0.01	0.4614	0.4615	-0.03	-0.15	3.31	2.83
0.02	0.4344	0.4351	-0.16	-0.04	-1.16	-0.00
0.1	0.2892	0.2944	-1.80	7.76	-4.75	-3.13
0.2	0.2154	0.2199	-2.06	-6.68	-0.90	-1.41
0.5	0.1651	0.1658	-0.41	-1.93	0.28	-0.33
1	0.1604	0.1599	0.30	1.88	-0.89	-0.42
2	0.1705	0.1694	0.63	1.95	-0.92	-0.18
5	0.2008	0.1989	0.97	5.93	-15.11	-4.92
10	0.2306	0.2300	0.29	0.10	1.34	-4.11
20	0.2814	0.2750	2.29	-4.83	22.83	8.26
50	0.3601	0.3504	2.69	19.16	-100.75	-19.63
100	0.4098	0.4035	1.52	-1.03	13.67	1.83
	C = 10					
0.01	0.4542	0.4581	-0.86	-0.15	-13.03	2.22
0.02	0.4228	0.4285	-1.36	0.54	-14.96	-1.42
0.1	0.2732	0.2835	-3.74	-3.24	-3.87	-3.73
0.2	0.2049	0.2113	-3.13	-10.11	-1.57	-2.30
0.5	0.1599	0.1611	-0.77	-3.05	0.07	-0.39
1	0.1532	0.1526	0.42	-1.31	1.16	0.85
2	0.1517	0.1500	1.14	-2.64	3.18	2.19
5	0.1624	0.1580	2.72	1.17	4.44	3.32
10	0.1850	0.1781	3.72	3.05	5.08	4.04
20	0.2172	0.2126	2.12	1.99	2.66	-0.10
50	0.2814	0.2755	2.11	-0.31	11.74	3.81
100	0.3295	0.3149	4.44	-5.89	30.52	13.91
	C = 100					
0.01	0.4542	0.4581	-0.84	-0.15	-12.80	2.25
0.02	0.4228	0.4290	-1.47	0.46	-15.02	-1.40
0.1	0.2730	0.2834	-3.81	-3.25	-3.95	-3.81
0.2	0.2048	0.2116	-3.31	-10.75	-1.65	-2.46
0.5	0.1602	0.1612	-0.64	-1.41	-0.36	-0.51
1	0.1546	0.1526	1.33	-1.33	2.43	2.03
2	0.1535	0.1500	2.27	-2.80	4.47	3.55
5	0.1568	0.1534	2.16	0.32	3.37	3.29
10	0.1684	0.1624	3.58	0.22	7.67	4.94
20	0.1824	0.1760	3.51	3.49	3.54	3.26
50	0.2325	0.2281	1.88	1.82	2.12	0.56
100	0.2728	0.2689	1.41	-0.68	10.26	1.63

Table 3: Comparison of the bias variance decomposition between single RBF SVM and bagged RBF SVM, with the Waveform data set.

Parameters σ and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
σ	**C = 1**					
0.01	0.5006	0.5011	-0.09	0.04	-10.73	7.65
0.02	0.5006	0.5011	-0.09	0.04	-10.73	7.65
0.1	0.4968	0.4984	-0.33	0.04	-18.79	9.37
0.2	0.4962	0.5010	-0.96	0.04	-47.11	11.17
0.5	0.4994	0.5000	-0.14	0.04	-12.26	14.54
1	0.4976	0.4983	-0.12	0.04	-8.96	12.97
2	0.4713	0.4824	-2.37	-25.15	114.68	24.31
5	0.0984	0.1178	-19.63	-4.01	-37.41	-27.27
10	0.0726	0.0691	4.78	-0.80	18.59	13.04
20	0.0710	0.0666	6.21	-2.57	28.80	18.28
50	0.0707	0.0662	6.30	-1.41	26.03	17.82
100	0.0720	0.0669	7.15	-0.42	25.90	18.29
	C = 10					
0.01	0.5006	0.5013	-0.14	0.04	-14.98	12.05
0.02	0.5006	0.5013	-0.14	0.04	-14.98	12.05
0.1	0.4968	0.5013	-0.91	0.04	-47.58	12.25
0.2	0.4962	0.5013	-1.03	0.04	-50.34	13.24
0.5	0.4993	0.5013	-0.39	0.04	-29.29	14.56
1	0.4973	0.4981	-0.17	0.04	-11.20	14.09
2	0.4560	0.4786	-4.95	-44.43	99.10	26.78
5	0.0947	0.1134	-19.75	-4.22	-38.81	-28.03
10	0.0747	0.0706	5.46	-1.78	21.33	14.38
20	0.0778	0.0713	8.32	-3.19	29.62	20.03
50	0.0802	0.0729	9.16	-1.59	28.10	20.00
100	0.0774	0.0708	8.60	-1.20	27.73	19.36
	C = 100					
0.01	0.5006	0.5013	-0.14	0.04	-14.98	12.05
0.02	0.5006	0.5013	-0.14	0.04	-14.98	12.05
0.1	0.4968	0.5013	-0.91	0.04	-47.58	12.25
0.2	0.4962	0.5013	-1.03	0.04	-50.34	13.24
0.5	0.4993	0.5013	-0.39	0.04	-29.29	14.56
1	0.4973	0.4980	-0.15	0.04	-10.26	14.18
2	0.4560	0.4791	-5.05	-45.03	100.32	27.23
5	0.0947	0.1134	-19.67	-3.27	-39.80	-28.30
10	0.0747	0.0706	5.53	-1.78	21.56	14.53
20	0.0778	0.0714	8.29	-3.39	29.87	20.16
50	0.0839	0.0742	11.58	-0.82	30.71	22.89
100	0.0870	0.0752	13.53	-1.01	34.21	25.62

Table 4: Comparison of the bias variance decomposition between single RBF SVM and bagged RBF SVM, with the Grey landsat data set.

Parameters σ and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
σ	**C = 1**					
0.01	0.4782	0.4781	0.03	0.00	0.74	1.97
0.02	0.4752	0.4773	-0.43	0.00	-14.46	-0.55
0.1	0.4774	0.4780	-0.11	0.00	-3.13	-10.29
0.2	0.4766	0.4764	0.05	0.00	1.47	-8.14
0.5	0.4288	0.4337	-1.13	-0.45	-14.53	1.72
1	0.2399	0.2768	-15.37	-457.97	58.98	22.45
2	0.0845	0.0986	-16.61	14.96	-111.83	-63.75
5	0.0450	0.0443	1.49	2.33	-16.19	3.84
10	0.0387	0.0381	1.42	8.57	-66.96	-5.55
20	0.0384	0.0378	1.74	2.94	-7.46	3.94
50	0.0411	0.0415	-1.08	-1.33	1.62	0.11
100	0.0476	0.0484	-1.66	-6.59	103.76	3.94
	C = 10					
0.01	0.4782	0.4782	0.02	0.00	0.48	1.97
0.02	0.4752	0.4789	-0.79	0.00	-26.29	-5.12
0.1	0.4774	0.4781	-0.15	0.00	-4.28	-2.51
0.2	0.4760	0.4763	-0.06	0.00	-1.97	-4.70
0.5	0.4117	0.4243	-3.05	-1.87	-28.12	-0.08
1	0.2220	0.2649	-19.32	-424.64	55.25	18.48
2	0.0803	0.0949	-18.16	7.56	-91.75	-60.35
5	0.0450	0.0446	0.88	0.00	8.70	5.91
10	0.0398	0.0391	1.74	7.69	-24.85	-0.24
20	0.0393	0.0381	3.07	0.00	12.95	8.89
50	0.0389	0.0381	2.06	3.28	-2.36	4.05
100	0.0385	0.0380	1.22	-8.06	39.63	15.00
	C = 100					
0.01	0.4782	0.4789	-0.15	0.00	-4.04	-2.48
0.02	0.4752	0.4782	-0.62	0.00	-20.80	-0.55
0.1	0.4774	0.4781	-0.15	0.00	-4.40	-2.51
0.2	0.4760	0.4764	-0.09	0.00	-2.97	-4.75
0.5	0.4117	0.4233	-2.81	-1.87	-21.91	-2.28
1	0.2220	0.2649	-19.29	-426.09	55.54	18.61
2	0.0803	0.0953	-18.66	7.56	-93.70	-61.41
5	0.0450	0.0446	0.92	2.47	-12.89	2.47
10	0.0398	0.0391	1.84	6.25	-16.23	1.43
20	0.0399	0.0385	3.37	0.00	12.97	9.78
50	0.0415	0.0397	4.27	-5.08	27.30	16.36
100	0.0417	0.0400	4.25	-5.08	26.77	16.27

Table 5: Comparison of the bias variance decomposition between single RBF SVM and bagged RBF SVM, with the Letter-Two data set.

Parameters σ and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
σ	**C = 1**					
0.01	0.4954	0.4930	0.48	-1.70	66.63	23.48
0.02	0.4950	0.4910	0.83	-2.75	84.28	25.45
0.1	0.4943	0.4911	0.64	-3.82	86.20	26.83
0.2	0.4894	0.4869	0.52	-7.94	102.39	27.98
0.5	0.4493	0.4572	-1.77	-27.70	86.73	29.66
1	0.3231	0.3579	-10.77	-340.62	52.78	23.59
2	0.1687	0.1896	-12.41	25.00	-29.18	-20.32
5	0.1059	0.1086	-2.51	-6.90	1.03	0.45
10	0.0946	0.0951	-0.55	-15.38	11.50	5.71
20	0.0925	0.0928	-0.33	-17.24	17.37	8.54
50	0.1005	0.1004	0.14	-12.50	23.53	9.09
100	0.1164	0.1161	0.31	3.51	-12.38	-2.44
	C = 10					
0.01	0.4958	0.4933	0.49	-2.40	71.28	26.85
0.02	0.4951	0.4937	0.29	-2.75	70.78	25.60
0.1	0.4935	0.4915	0.40	-4.91	86.76	31.20
0.2	0.4861	0.4852	0.19	-9.52	106.20	30.28
0.5	0.4343	0.4508	-3.80	-48.60	88.15	33.04
1	0.2998	0.3416	-13.95	-320.00	45.76	18.36
2	0.1574	0.1783	-13.29	3.85	-19.61	-16.85
5	0.0981	0.1009	-2.93	-15.38	6.57	2.04
10	0.0818	0.0832	-1.71	-8.70	4.22	2.21
20	0.0745	0.0736	1.20	-25.00	21.63	12.29
50	0.0747	0.0736	1.49	-18.18	19.68	12.34
100	0.0784	0.0788	-0.47	-22.58	39.65	16.88
	C = 100					
0.01	0.4958	0.4933	0.49	-2.40	71.37	26.85
0.02	0.4951	0.4937	0.29	-2.75	70.86	25.62
0.1	0.4935	0.4915	0.41	-4.91	87.04	31.19
0.2	0.4861	0.4852	0.20	-9.52	106.36	33.82
0.5	0.4343	0.4504	-3.71	-48.04	87.28	31.84
1	0.2998	0.3417	-13.97	-320.00	45.73	18.79
2	0.1574	0.1784	-13.33	-15.38	-12.58	-13.87
5	0.0981	0.1010	-2.98	-11.54	3.55	0.76
10	0.0818	0.0829	-1.43	-8.70	4.74	2.47
20	0.0753	0.0737	2.10	-19.05	19.78	12.20
50	0.0774	0.0737	4.72	-30.00	30.03	20.09
100	0.0811	0.0759	6.44	-23.08	38.82	23.86

Table 6: Comparison of the bias variance decomposition between single RBF SVM and bagged RBF SVM, with the Letter-Two with noise data set.

Parameters σ and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
σ	**C = 1**					
0.01	0.5089	0.5059	0.59	0.00	22.86	-25.16
0.02	0.5084	0.5059	0.50	0.00	18.38	-16.36
0.1	0.5076	0.5070	0.11	0.00	3.89	-7.31
0.2	0.5061	0.5044	0.33	0.00	10.26	-13.47
0.5	0.4956	0.4963	-0.14	0.95	-28.19	-14.09
1	0.4618	0.4652	-0.73	-1.06	50.82	-8.42
2	0.4074	0.4117	-1.06	1.49	-12.03	-6.22
5	0.3726	0.3701	0.67	3.70	-14.14	-5.99
10	0.3586	0.3560	0.72	5.29	-28.00	-8.38
20	0.3515	0.3475	1.13	2.73	-8.09	-2.23
50	0.3561	0.3507	1.53	5.05	-34.05	-5.81
100	0.3720	0.3655	1.74	5.50	-86.26	-8.19
	C = 10					
0.01	0.5087	0.5065	0.44	0.00	16.65	-19.68
0.02	0.5081	0.5070	0.22	0.00	8.03	-12.18
0.1	0.5073	0.5069	0.10	0.00	3.27	-7.32
0.2	0.5053	0.5040	0.27	0.31	-1.66	-14.54
0.5	0.4905	0.4948	-0.88	0.64	-41.15	-14.15
1	0.4509	0.4596	-1.94	-5.79	58.56	-1.38
2	0.4041	0.4102	-1.51	-2.11	0.47	-2.62
5	0.3873	0.3829	1.11	1.62	-0.71	2.45
10	0.3842	0.3754	2.31	2.25	2.49	5.41
20	0.3729	0.3647	2.18	0.57	7.52	6.52
50	0.3456	0.3430	0.76	-1.20	8.18	3.65
100	0.3362	0.3345	0.49	1.75	-5.78	0.40
	C = 100					
0.01	0.5087	0.5065	0.44	0.00	16.77	-19.75
0.02	0.5081	0.5070	0.23	0.00	8.38	-11.90
0.1	0.5073	0.5070	0.07	0.00	2.50	-7.37
0.2	0.5053	0.5040	0.26	0.31	-1.85	-14.35
0.5	0.4905	0.4943	-0.76	0.32	-29.14	-15.53
1	0.4509	0.4593	-1.87	-6.18	65.83	-0.61
2	0.4042	0.4106	-1.59	-2.63	1.89	-2.38
5	0.3882	0.3832	1.27	0.54	3.95	3.95
10	0.3914	0.3775	3.57	1.11	11.06	10.05
20	0.3969	0.3775	4.89	4.97	4.63	11.40
50	0.3909	0.3773	3.47	-1.69	18.56	12.87
100	0.3699	0.3642	1.56	1.15	2.94	5.16

Table 7: Comparison of the bias variance decomposition between single RBF SVM and bagged RBF SVM, with the Spam data set.

Parameters σ and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
σ	**C = 1**					
0.01	0.3981	0.3964	0.42	0.00	51.75	49.55
0.02	0.3973	0.3955	0.46	0.00	45.30	49.14
0.1	0.3948	0.3931	0.45	0.00	27.15	48.62
0.2	0.3923	0.3914	0.24	0.00	10.28	32.65
0.5	0.3895	0.3885	0.25	0.00	8.12	30.77
1	0.3860	0.3855	0.12	0.00	2.87	37.57
2	0.3806	0.3817	-0.30	0.00	-5.56	-21.96
5	0.3394	0.3430	-1.07	-1.91	13.50	10.87
10	0.2484	0.2572	-3.53	-6.51	44.39	3.29
20	0.1730	0.1790	-3.50	-3.17	-5.73	-7.18
50	0.1392	0.1400	-0.54	0.74	-7.28	-3.72
100	0.1471	0.1453	1.20	7.40	-69.25	-9.54
	C = 10					
0.01	0.3979	0.3964	0.37	0.00	42.59	47.73
0.02	0.3971	0.3955	0.38	0.00	35.73	47.55
0.1	0.3931	0.3925	0.15	0.00	7.30	33.14
0.2	0.3914	0.3908	0.16	0.00	6.10	33.35
0.5	0.3879	0.3877	0.04	0.00	1.07	30.94
1	0.3845	0.3848	-0.10	0.00	-2.30	38.48
2	0.3790	0.3817	-0.71	0.00	-11.99	-5.22
5	0.3211	0.3349	-4.30	-7.43	1548.57	14.90
10	0.2375	0.2491	-4.89	-10.84	23.67	4.95
20	0.1766	0.1788	-1.20	-4.89	10.18	3.84
50	0.1391	0.1386	0.36	-0.46	2.06	2.87
100	0.1292	0.1290	0.14	-0.48	1.57	2.22
	C = 100					
0.01	0.3979	0.3964	0.37	0.00	43.28	47.26
0.02	0.3972	0.3957	0.38	0.00	36.81	46.72
0.1	0.3933	0.3928	0.14	0.00	6.98	31.56
0.2	0.3916	0.3910	0.16	0.00	6.37	31.83
0.5	0.3880	0.3878	0.04	0.00	1.11	29.57
1	0.3846	0.3849	-0.08	0.00	-1.82	37.18
2	0.3792	0.3815	-0.61	0.00	-10.47	-4.16
5	0.3230	0.3354	-3.84	-6.60	2111.70	15.75
10	0.2416	0.2508	-3.83	-9.41	22.02	5.42
20	0.1834	0.1839	-0.25	-2.91	7.05	3.84
50	0.1582	0.1496	5.38	-5.37	19.26	14.49
100	0.1546	0.1438	7.04	-6.38	22.08	16.90

Table 8: Comparison of the bias variance decomposition between single RBF SVM and bagged RBF SVM, with the Musk data set.

Parameters σ and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
σ	**C = 1**					
0.01	0.1496	0.1494	0.13	0.00	42.49	0.00
0.02	0.1490	0.1488	0.10	0.00	13.40	0.00
0.1	0.1482	0.1482	0.03	0.00	2.18	0.00
0.2	0.1482	0.1482	0.03	0.00	2.49	0.00
0.5	0.1482	0.1482	0.03	0.00	2.64	0.00
1	0.1482	0.1482	0.03	0.00	2.64	0.00
2	0.1482	0.1482	0.03	0.00	2.39	0.00
5	0.1482	0.1482	0.03	0.00	2.13	0.00
10	0.1482	0.1481	0.01	0.00	0.72	-450.00
20	0.1473	0.1475	-0.15	0.00	-7.98	-84.62
50	0.1138	0.1219	-7.09	-16.39	353.12	37.93
100	0.0997	0.1021	-2.38	-2.96	4.74	-5.20
	C = 10					
0.01	0.1484	0.1484	-0.02	0.00	-1.38	0.00
0.02	0.1483	0.1482	0.02	0.00	1.69	0.00
0.1	0.1482	0.1482	0.03	0.00	2.64	0.00
0.2	0.1482	0.1482	0.03	0.00	2.64	0.00
0.5	0.1482	0.1482	0.03	0.00	2.39	0.00
1	0.1482	0.1482	0.03	0.00	2.22	0.00
2	0.1482	0.1482	0.03	0.00	2.22	0.00
5	0.1481	0.1481	-0.01	0.00	-1.10	-33.33
10	0.1476	0.1477	-0.07	0.00	-4.34	-1.30
20	0.1427	0.1445	-1.25	0.00	-24.20	21.74
50	0.1013	0.1117	-10.23	-11.38	30.65	4.27
100	0.0899	0.0920	-2.36	-6.72	22.91	13.67
	C = 100					
0.01	0.1484	0.1484	-0.02	0.00	-1.38	0.00
0.02	0.1483	0.1482	0.02	0.00	1.69	0.00
0.1	0.1482	0.1482	0.03	0.00	2.64	0.00
0.2	0.1482	0.1482	0.03	0.00	2.64	0.00
0.5	0.1482	0.1482	0.03	0.00	2.39	0.00
1	0.1482	0.1482	0.03	0.00	2.22	0.00
2	0.1482	0.1482	0.03	0.00	2.22	0.00
5	0.1481	0.1481	-0.02	0.00	-1.25	-40.00
10	0.1476	0.1477	-0.07	0.00	-4.34	-3.90
20	0.1427	0.1445	-1.25	0.00	-24.20	20.72
50	0.1013	0.1117	-10.23	-12.00	52.28	5.81
100	0.0899	0.0921	-2.37	-7.11	25.09	14.18

5.2 Comparison between single and bagged polynomial SVMs

Table 9: Comparison of the bias variance decomposition between single polynomial SVM and bagged polynomial SVM, with the P2 data set.

Parameters **degree** and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
degree	**C = 1**					
2	0.3379	0.3353	0.75	9.73	-53.27	-6.49
3	0.2753	0.2659	3.41	-0.62	16.52	9.77
4	0.2495	0.2446	1.98	1.54	3.17	0.96
5	0.2199	0.2198	0.02	-2.39	4.83	-0.25
6	0.2127	0.2024	4.87	4.27	5.91	4.80
7	0.2105	0.2014	4.33	-1.97	13.91	9.13
8	0.2111	0.2002	5.21	0.42	11.42	9.44
9	0.2114	0.2031	3.92	-2.37	11.28	8.76
10	0.2124	0.2078	2.18	-6.21	11.66	7.38
	C = 10					
2	0.3314	0.3300	0.42	-0.07	2.66	1.43
3	0.2621	0.2513	4.13	-5.84	36.36	11.94
4	0.2364	0.2218	6.19	4.67	9.86	6.08
5	0.2109	0.2077	1.51	1.91	0.86	0.42
6	0.2088	0.1985	4.95	4.85	5.08	5.91
7	0.2104	0.2031	3.46	-4.05	11.88	7.88
8	0.2120	0.2043	3.63	2.02	5.34	4.56
9	0.2131	0.2083	2.22	-4.10	8.64	6.56
10	0.2145	0.2061	3.92	-10.14	17.28	11.19
	C = 100					
2	0.3310	0.3298	0.38	1.33	-3.93	2.44
3	0.2614	0.2509	4.03	-0.20	18.60	7.32
4	0.2346	0.2221	5.36	6.79	2.26	2.55
5	0.2116	0.1992	5.87	0.33	13.43	8.66
6	0.2152	0.2013	6.47	-1.62	15.14	11.65
7	0.2148	0.2004	6.71	1.72	11.47	9.57
8	0.2157	0.2056	4.71	-3.71	12.33	8.86
9	0.2162	0.2096	3.01	-18.46	21.03	13.38
10	0.2174	0.2126	2.19	-20.36	21.11	13.02

Table 10: Comparison of the bias variance decomposition between single polynomial SVM and bagged polynomial SVM, with the Waveform data set.

Parameters degree and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
degree	**C = 1**					
1	0.0720	0.0669	7.11	-2.40	29.58	19.25
2	0.0761	0.0699	8.11	0.36	23.78	17.94
3	0.0799	0.0731	8.46	-3.60	28.98	19.72
4	0.0813	0.0746	8.16	-4.10	29.25	19.58
5	0.0819	0.0762	6.86	-2.74	23.14	16.56
6	0.0832	0.0784	5.76	-2.33	19.18	14.33
7	0.0856	0.0819	4.38	-2.71	15.29	11.88
8	0.0897	0.0873	2.70	-6.43	14.94	10.24
9	0.0957	0.0950	0.76	-13.24	16.93	9.58
10	0.1040	0.1056	-1.57	-19.26	16.34	7.72
	C = 10					
1	0.0842	0.0737	12.48	-0.03	31.90	24.08
2	0.0859	0.0755	12.19	-0.42	30.78	23.16
3	0.0832	0.0751	9.81	-2.00	28.22	20.50
4	0.0820	0.0753	8.13	-3.14	26.97	18.92
5	0.0820	0.0764	6.83	-3.13	23.60	16.72
6	0.0832	0.0784	5.75	-2.71	19.80	14.46
7	0.0856	0.0820	4.23	-3.29	15.79	11.89
8	0.0897	0.0872	2.78	-6.04	14.60	10.21
9	0.0957	0.0950	0.70	-12.65	16.13	9.27
10	0.1040	0.1056	-1.58	-19.65	16.71	7.99
	C = 100					
1	0.0927	0.0757	18.33	-1.20	42.33	32.49
2	0.0860	0.0755	12.22	-0.61	31.18	23.31
3	0.0832	0.0750	9.90	-2.00	28.43	20.63
4	0.0820	0.0752	8.21	-4.30	29.15	19.72
5	0.0820	0.0763	6.86	-2.94	23.35	16.65
6	0.0832	0.0784	5.77	-3.48	21.12	14.95
7	0.0856	0.0819	4.32	-2.91	15.44	11.86
8	0.0897	0.0874	2.64	-6.81	15.32	10.45
9	0.0957	0.0949	0.80	-12.07	15.68	9.09
10	0.1040	0.1055	-1.50	-18.69	15.91	7.67

Table 11: Comparison of the bias variance decomposition between single polynomial SVM and bagged polynomial SVM, with the GreyLandsat data set.

Parameters degree and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
degree	**C = 1**					
1	0.0503	0.0507	-0.98	-3.23	26.81	0.59
2	0.0411	0.0407	0.83	0.00	7.43	7.41
3	0.0392	0.0388	1.05	-4.76	24.80	12.06
4	0.0406	0.0397	2.22	-4.76	26.39	13.90
5	0.0424	0.0410	3.23	-3.12	22.89	13.13
6	0.0446	0.0427	4.12	-2.99	25.63	14.16
7	0.0465	0.0448	3.69	0.00	12.68	10.35
8	0.0484	0.0463	4.22	1.52	10.04	10.08
9	0.0499	0.0481	3.67	0.00	11.17	9.61
10	0.0514	0.0497	3.25	2.99	3.73	6.69
	C = 10					
1	0.0442	0.0432	2.33	1.30	9.27	7.31
2	0.0422	0.0406	3.73	-1.59	19.41	13.09
3	0.0433	0.0410	5.24	4.69	6.81	11.35
4	0.0445	0.0420	5.64	3.03	13.15	12.93
5	0.0453	0.0430	5.08	-3.03	26.89	15.81
6	0.0464	0.0440	5.11	0.00	19.17	13.33
7	0.0473	0.0451	4.59	-3.08	21.42	13.78
8	0.0486	0.0464	4.49	1.52	10.77	10.52
9	0.0500	0.0480	3.94	0.00	11.94	9.97
10	0.0514	0.0497	3.37	4.48	1.30	5.96
	C = 100					
1	0.0488	0.0448	8.35	-4.48	36.39	22.24
2	0.0439	0.0413	5.83	3.08	13.70	13.71
3	0.0438	0.0413	5.75	4.62	9.02	13.00
4	0.0445	0.0420	5.65	4.55	8.80	11.63
5	0.0453	0.0430	5.01	-1.52	22.58	14.71
6	0.0464	0.0440	5.19	1.47	15.41	12.15
7	0.0473	0.0452	4.52	-1.54	17.81	12.87
8	0.0486	0.0464	4.52	0.00	14.07	11.71
9	0.0500	0.0481	3.88	0.00	11.76	9.84
10	0.0514	0.0498	3.13	2.99	3.39	6.57

Table 12: Comparison of the bias variance decomposition between single polynomial SVM and bagged polynomial SVM, with the Letter-Two data set.

Parameters degree and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
degree	**C = 1**					
1	0.0958	0.0942	1.67	1.96	-0.26	4.97
2	0.0749	0.0758	-1.20	-19.23	22.35	9.34
3	0.0745	0.0733	1.55	-15.79	13.92	10.41
4	0.0809	0.0801	0.88	-4.35	5.40	5.63
5	0.0881	0.0863	2.14	7.69	-3.01	2.39
6	0.0953	0.0933	2.11	14.29	-9.09	-0.60
7	0.1027	0.1008	1.86	10.71	-5.24	-0.09
8	0.1101	0.1091	0.96	15.15	-12.61	-3.80
9	0.1177	0.1175	0.17	8.57	-7.74	-3.16
10	0.1262	0.1259	0.28	12.50	-12.80	-5.01
	C = 10					
1	0.0911	0.0862	5.36	4.35	10.09	14.23
2	0.0783	0.0753	3.84	-22.73	26.33	17.30
3	0.0792	0.0768	3.13	0.00	5.94	7.88
4	0.0842	0.0816	3.11	-4.35	9.09	8.28
5	0.0895	0.0866	3.20	3.85	2.62	5.26
6	0.0956	0.0935	2.17	14.29	-8.93	-0.58
7	0.1027	0.1009	1.73	10.71	-5.47	-0.30
8	0.1101	0.1089	1.08	15.15	-12.38	-3.78
9	0.1177	0.1176	0.13	11.43	-10.51	-4.27
10	0.1262	0.1259	0.26	12.50	-12.83	-5.11
	C = 100					
1	0.1015	0.0907	10.65	-2.27	41.86	27.74
2	0.0793	0.0756	4.76	-23.81	26.47	18.07
3	0.0792	0.0767	3.19	4.35	2.15	6.34
4	0.0842	0.0816	3.09	-4.35	9.06	8.28
5	0.0895	0.0865	3.31	7.69	-0.64	3.99
6	0.0956	0.0934	2.30	14.29	-8.67	-0.31
7	0.1027	0.1010	1.65	10.71	-5.61	-0.44
8	0.1101	0.1091	0.95	15.15	-12.64	-3.85
9	0.1177	0.1172	0.41	11.43	-9.97	-3.86
10	0.1262	0.1262	0.01	12.50	-13.36	-5.48

Table 13: Comparison of the bias variance decomposition between single polynomial SVM and bagged polynomial SVM, with the Letter-Two with noise data set.

Parameters degree and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
degree	**C = 1**					
1	0.3441	0.3366	2.18	5.76	-33.28	-2.88
2	0.3432	0.3429	0.09	-0.58	3.06	0.91
3	0.3646	0.3590	1.53	0.00	6.34	4.88
4	0.3843	0.3729	2.96	-2.96	18.17	11.10
5	0.3932	0.3800	3.33	0.56	11.27	9.93
6	0.3978	0.3847	3.31	4.40	0.07	6.35
7	0.3991	0.3860	3.29	1.12	9.30	8.30
8	0.4002	0.3890	2.79	-1.68	15.00	9.26
9	0.4013	0.3903	2.73	-2.25	15.93	9.20
10	0.4029	0.3921	2.67	0.00	9.80	6.93
	C = 10					
1	0.3454	0.3410	1.26	-0.56	11.81	6.91
2	0.3860	0.3755	2.71	1.10	8.14	9.03
3	0.3995	0.3821	4.36	0.00	15.63	14.18
4	0.4005	0.3824	4.53	1.12	13.81	13.42
5	0.4001	0.3839	4.04	2.19	9.54	11.35
6	0.4001	0.3857	3.61	2.20	7.74	9.37
7	0.3997	0.3862	3.38	0.00	12.48	9.55
8	0.4005	0.3891	2.83	-1.68	15.09	9.39
9	0.4014	0.3904	2.74	-1.12	12.96	8.42
10	0.4029	0.3921	2.67	-1.68	14.25	8.13
	C = 100					
1	0.3579	0.3543	1.01	0.00	5.58	2.97
2	0.4078	0.3871	5.07	0.53	19.64	17.49
3	0.4033	0.3828	5.09	0.56	17.38	15.97
4	0.4011	0.3826	4.62	-0.57	17.83	14.78
5	0.4001	0.3839	4.06	2.73	8.03	10.98
6	0.4001	0.3856	3.64	1.65	9.44	9.85
7	0.3997	0.3860	3.43	0.00	12.64	9.61
8	0.4005	0.3892	2.81	-1.68	15.03	9.27
9	0.4014	0.3902	2.78	-1.12	13.10	8.51
10	0.4029	0.3918	2.75	-1.68	14.56	8.28

Table 14: Comparison of the bias variance decomposition between single polynomial SVM and bagged polynomial SVM, with the Spam data set.

Parameters degree and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
degree	**C = 1**					
1	0.1509	0.1455	3.61	3.51	4.53	4.58
2	0.1355	0.1322	2.45	-1.59	19.19	7.94
3	0.1323	0.1318	0.35	2.11	-5.83	-1.19
4	0.1335	0.1335	-0.04	0.88	-2.62	0.00
5	0.1369	0.1367	0.14	-1.80	4.78	3.42
6	0.1453	0.1404	3.33	-3.11	16.61	12.41
7	0.1453	0.1443	0.63	-3.56	9.27	6.33
8	0.1491	0.1487	0.25	-4.02	8.27	5.56
9	0.1529	0.1529	0.00	-5.38	9.33	5.74
10	0.1572	0.1572	0.00	-5.36	8.73	5.44
	C = 10					
1	0.1325	0.1276	3.74	-0.94	14.43	10.62
2	0.1370	0.1344	1.88	-4.64	12.34	9.00
3	0.1461	0.1406	3.73	-3.63	13.66	10.80
4	0.1521	0.1451	4.60	-4.71	15.79	12.63
5	0.1565	0.1483	5.21	-10.05	22.09	15.58
6	0.1593	0.1516	4.84	-10.99	22.08	15.16
7	0.1615	0.1549	4.08	-11.17	21.31	14.42
8	0.1636	0.1575	3.73	-8.87	18.49	13.12
9	0.1660	0.1603	3.44	-6.70	15.71	11.41
10	0.1678	0.1630	2.89	-8.57	16.56	11.28
	C = 100					
1	0.1558	0.1413	9.29	-3.23	22.80	19.69
2	0.1573	0.1450	7.86	-5.88	22.55	18.00
3	0.1595	0.1477	7.38	-6.99	22.15	17.15
4	0.1610	0.1502	6.71	-4.79	18.57	15.19
5	0.1626	0.1529	5.95	-5.82	17.98	14.34
6	0.1644	0.1552	5.60	-7.77	19.54	14.67
7	0.1664	0.1575	5.35	-8.85	19.65	14.54
8	0.1686	0.1600	5.09	-7.54	18.40	13.87
9	0.1704	0.1628	4.51	-9.45	19.19	13.61
10	0.1720	0.1655	3.79	-8.91	17.03	12.12

Table 15: Comparison of the bias variance decomposition between single polynomial SVM and bagged polynomial SVM, with the Musk data set.

Parameters degree and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
degree	**C = 1**					
1	0.1373	0.1324	3.57	2.78	20.77	-49.49
2	0.1373	0.1128	17.84	39.10	-672.05	-342.32
3	0.1225	0.1204	1.73	-5.43	14.37	9.59
4	0.1326	0.1286	3.01	-2.72	11.16	9.82
5	0.1367	0.1339	2.01	-6.94	12.65	9.69
6	0.1396	0.1389	0.56	-3.60	5.49	5.64
7	0.1435	0.1435	0.02	-1.17	1.41	3.69
8	0.1480	0.1485	-0.33	0.37	-1.21	2.32
9	0.1532	0.1549	-1.09	-0.70	-1.60	1.42
10	0.1594	0.1626	-2.03	0.69	-5.33	-1.18
	C = 10					
1	0.1210	0.1173	2.99	-5.45	18.26	11.20
2	0.1340	0.1228	8.36	-5.93	27.49	20.70
3	0.1338	0.1263	5.59	-6.61	22.60	16.25
4	0.1348	0.1300	3.56	-1.16	10.09	10.02
5	0.1396	0.1340	4.03	-4.40	14.04	11.49
6	0.1396	0.1386	0.70	-3.60	5.80	5.88
7	0.1435	0.1437	-0.13	-2.34	2.47	3.97
8	0.1480	0.1487	-0.47	0.00	-1.08	2.24
9	0.1532	0.1549	-1.15	-0.35	-2.18	1.12
10	0.1594	0.1624	-1.93	-0.69	-3.44	-0.40
	C = 100					
1	0.1500	0.1264	15.75	-3.21	34.94	29.36
2	0.1343	0.1233	8.21	-5.51	26.66	20.20
3	0.1338	0.1264	5.56	-5.45	20.89	15.70
4	0.1348	0.1301	3.48	-2.33	11.50	10.34
5	0.1368	0.1341	1.99	-6.94	12.60	9.66
6	0.1396	0.1387	0.69	-4.00	6.26	6.09
7	0.1435	0.1437	-0.13	-2.73	2.93	4.14
8	0.1480	0.1487	-0.45	1.83	-3.34	1.41
9	0.1532	0.1551	-1.26	-0.35	-2.42	1.01
10	0.1594	0.1623	-1.87	-0.35	-3.71	-0.45

5.3 Comparison between single and bagged dot-product SVMs

Table 16: Comparison of the bias variance decomposition between single dot-prod SVM and bagged dot-prod SVM, with the P2 data set.

ParameterC	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
1	0.4725	0.4630	2.01	3.89	-116.69	-22.64
2	0.4725	0.4603	2.58	9.59	-376.70	-18.33
5	0.4725	0.4598	2.68	-0.26	187.69	29.56
10	0.4730	0.4553	3.74	8.02	-211.59	-11.55
20	0.4719	0.4659	1.27	-0.04	69.60	4.63
50	0.4729	0.4661	1.43	1.49	-1.50	4.81
100	0.4719	0.4615	2.21	-2.00	163.65	7.80
200	0.4715	0.4626	1.89	2.25	-9.54	1.34
500	0.4715	0.4590	2.65	1.11	34.09	15.28
1000	0.4711	0.4639	1.53	1.27	7.35	-4.39

Table 17: Comparison of the bias variance decomposition between single dot-prod SVM and bagged dot-prod SVM, with the Waveform data set.

ParameterC	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
1	0.0886	0.0750	15.37	-0.22	37.00	28.20
2	0.0912	0.0756	17.09	0.94	38.25	30.01
5	0.0927	0.0757	18.31	-0.62	41.58	32.22
10	0.0928	0.0757	18.40	0.74	40.52	31.88
20	0.0929	0.0757	18.52	0.36	41.20	32.20
50	0.0929	0.0757	18.52	-0.22	41.92	32.47
100	0.0929	0.0757	18.53	0.74	40.74	32.09
200	0.0929	0.0758	18.46	0.36	41.07	32.16
500	0.0929	0.0757	18.47	-0.22	41.82	32.37
1000	0.0929	0.0758	18.46	0.36	41.07	32.13

Table 18: Comparison of the bias variance decomposition between single dot-prod SVM and bagged dot-prod SVM, with the GreyLandsat data set.

ParameterC	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
1	0.0450	0.0439	2.58	16.87	-165.72	-62.21
2	0.0462	0.0446	3.48	-1.49	16.63	9.54
5	0.0481	0.0449	6.56	-1.47	25.99	17.22
10	0.0494	0.0449	9.16	-2.94	35.91	23.38
20	0.0496	0.0450	9.27	0.00	29.51	21.71
50	0.0496	0.0450	9.35	-1.47	32.94	22.88
100	0.0496	0.0449	9.43	-2.94	36.40	23.89
200	0.0496	0.0449	9.45	-1.47	33.26	22.92
500	0.0496	0.0449	9.55	-2.94	36.79	24.04
1000	0.0496	0.0449	9.46	-2.94	36.50	23.95

Table 19: Comparison of the bias variance decomposition between single dot-prod SVM and bagged dot-prod SVM, with the Letter-Two data set.

ParameterC	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
1	0.0955	0.0878	8.09	2.22	27.55	23.06
2	0.1002	0.0885	11.67	0.00	41.11	29.71
5	0.1032	0.0905	12.30	-4.65	48.24	31.79
10	0.1057	0.0906	14.26	-2.22	51.74	33.76
20	0.1078	0.0909	15.73	4.35	41.78	31.88
50	0.1087	0.0908	16.47	4.26	45.68	33.27
100	0.1095	0.0910	16.89	4.26	46.41	34.26
200	0.1099	0.0909	17.30	4.35	45.18	34.07
500	0.1099	0.0909	17.31	6.38	42.51	33.28
1000	0.1099	0.0909	17.27	6.38	42.36	33.25

Table 20: Comparison of the bias variance decomposition between single dot-prod SVM and bagged dot-prod SVM, with the Letter-Two with noise data set.

ParameterC	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
1	0.3486	0.3444	1.21	-0.56	10.23	6.09
2	0.3525	0.3483	1.18	2.25	-3.89	0.96
5	0.3568	0.3533	0.98	0.56	2.84	2.36
10	0.3597	0.3554	1.21	-0.56	8.99	4.17
20	0.3608	0.3575	0.89	-2.79	16.82	5.76
50	0.3615	0.3589	0.70	-5.08	24.05	7.14
100	0.3617	0.3591	0.71	-5.71	25.14	7.69
200	0.3620	0.3596	0.65	-7.43	31.27	9.09
500	0.3622	0.3597	0.69	-5.11	23.25	6.64
1000	0.3624	0.3598	0.72	-4.55	21.10	6.21

Table 21: Comparison of the bias variance decomposition between single dot-prod SVM and bagged dot-prod SVM, with the Spam data set.

ParameterC	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
1	0.1495	0.1389	7.15	-3.16	19.87	16.38
2	0.1568	0.1417	9.63	-6.56	26.30	21.33
5	0.1630	0.1438	11.79	-9.04	30.42	24.43
10	0.1659	0.1439	13.26	-11.43	34.16	27.04
20	0.1676	0.1448	13.59	-7.82	32.16	26.37
50	0.1692	0.1452	14.18	-7.26	32.46	26.95
100	0.1701	0.1451	14.65	-7.30	32.98	27.29
200	0.1712	0.1462	14.60	-5.08	30.66	26.30
500	0.1725	0.1479	14.26	-8.00	31.83	26.23
1000	0.1734	0.1487	14.24	-11.36	34.47	27.23

Table 22: Comparison of the bias variance decomposition between single dot-prod SVM and bagged dot-prod SVM, with the Musk data set.

ParameterC	$\mathbf{E_{SVM}}$	$\mathbf{E_{bag}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
1	0.1501	0.1261	15.97	-2.41	34.56	29.38
2	0.1501	0.1265	15.72	-3.21	34.87	29.27
5	0.1501	0.1266	15.66	-3.21	34.75	29.15
10	0.1501	0.1264	15.79	-2.01	33.78	28.95
20	0.1501	0.1267	15.61	-3.61	35.06	29.23
50	0.1501	0.1265	15.70	-2.81	34.42	29.11
100	0.1501	0.1264	15.77	-2.41	34.15	28.97
200	0.1501	0.1262	15.90	-2.01	34.02	29.13
500	0.1501	0.1265	15.74	-3.61	35.31	29.47
1000	0.1501	0.1262	15.92	-2.41	34.46	29.32