Bias–Variance decomposition of the error in Random aggregated 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 **3**

2 Bias–variance decomposition of the error in RA SVM ensembles **3**

2.1 Decomposition in RA RBF-SVMs 4
2.2 Decomposition in RA Polynomial SVMs 19
2.3 Decomposition in RA Dot–product SVMs 34

3 Decomposition of bias–variance with respect to the number of base learners **39**

3.1 Decomposition with respect to the number of base learners in RBF-SVM RA ensembles 39
3.2 Decomposition with respect to the number of base learners in Polynomial SVM RA ensembles 53
3.3 Decomposition with respect to the number of base learners in Dot-product SVM RA ensembles 68

4 Comparison of bias–variance decomposition in single and RA SVMs. **71**

4.1 Comparison between single and RA RBF-SVM 71
4.2 Comparison between single and RA polynomial SVM 79
4.3 Comparison between single and RA dot-product SVM 87

5 Comparative performance between single SVM and RA SVM ensembles: tables **92**

5.1 Comparison between single and RA RBF-SVMs 92
5.2 Comparison between single and RA polynomial SVMs 100

5.3 Comparison between single and RA dot-product SVMs 108

1 Experimental setup

In order to estimate the decompostion of the error in bias, unbiased and biased variance with random aggregated (RA) ensembles of SVMs, we drew from each data set 100 sets of samples using a uniform probability distribution. Each set is composed by 60 samples with size 100 (that is, each sample is composed by 100 patterns).

The for each SVM model (specified by a choice of a kernel and its associated parameters) we trained 100 RA 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 RA 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 RA SVM ensembles and a total of about 5 millions of single SVMs.

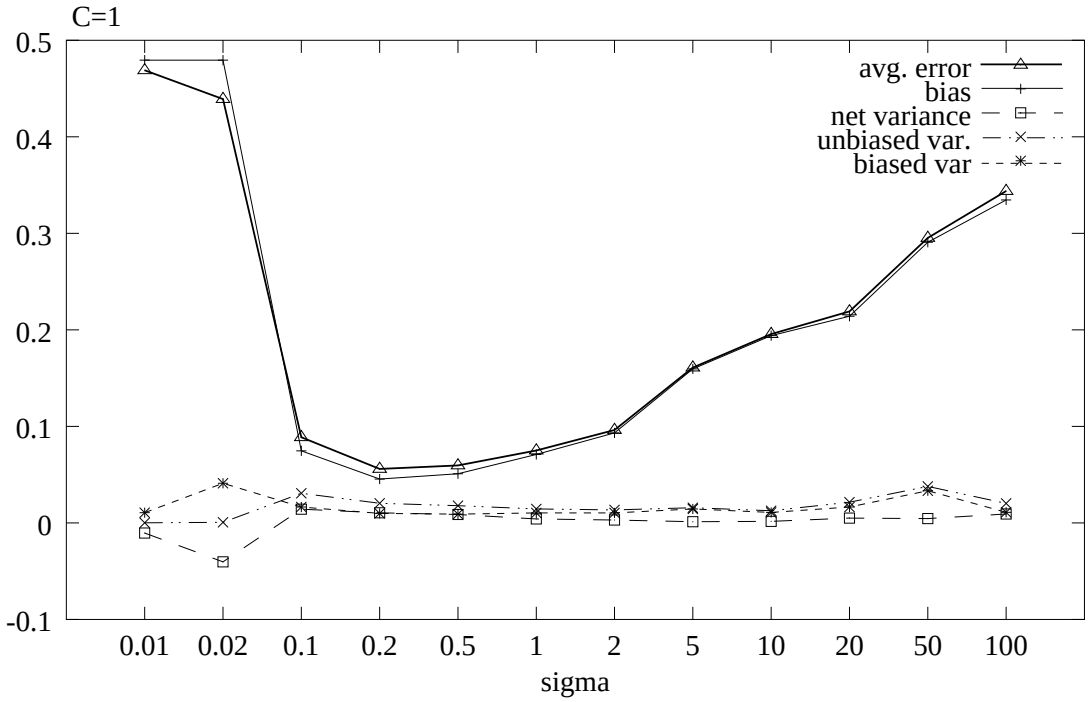
The most important fact we can observe from the results is that the variance of the RA ensembles is always reduced close to 0, independently of the type of kernel used. The reduction is due promarily to the unbiased variance reduction, while the bias remains unchanged with respect to the single SVMs.

2.1 Decomposition in RA RBF-SVMs

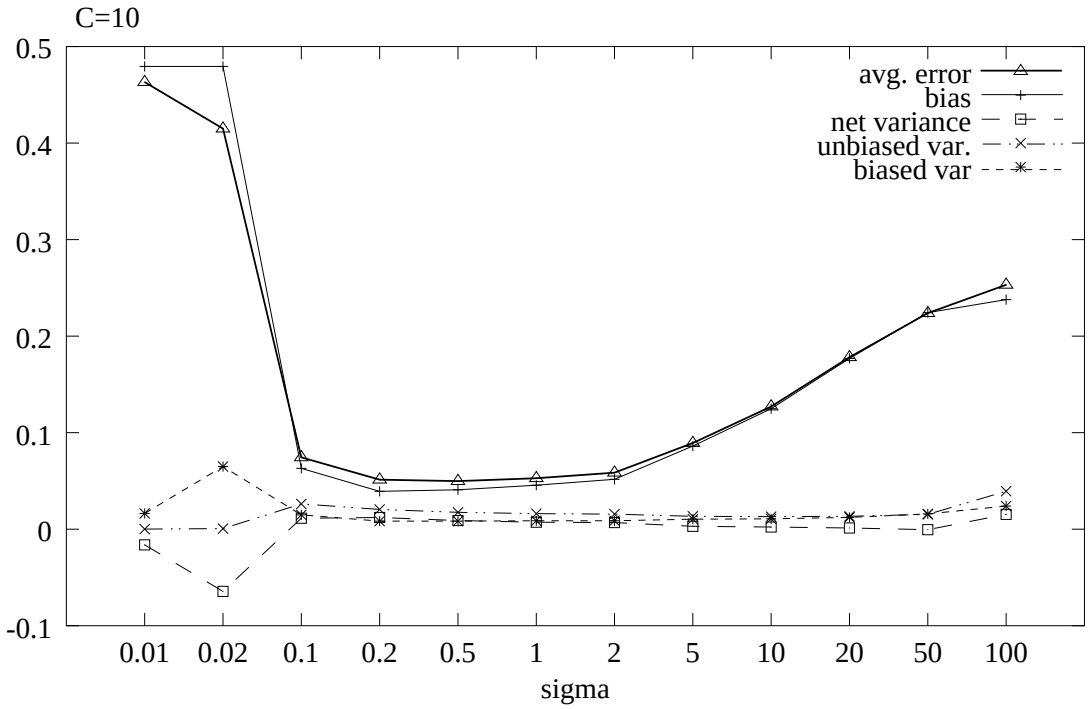
The decomposition of the error is represented with respect to different values of σ and for fixed values of C.

Schematically we can observe the following facts:

- In all the data sets the net-variance is about 0 for all the values of σ. As a consequence the wave-shape of the net-variance is very reduced in some data sets (Waveform, Grey-Landsat, Letter-Two), and in the other data sets is in practice absent.
- The net-variance is 0 as both biased and unbiased variance are very low close to 0. Only in the wafeform data set unbiased and biased variance are both quite large in the "high bias" region (and partially also in Letter-Two).
- The net-variance is always 0 in the "stabilized region" in all the considered data sets.
- The error is determinated almost totally by the bias.

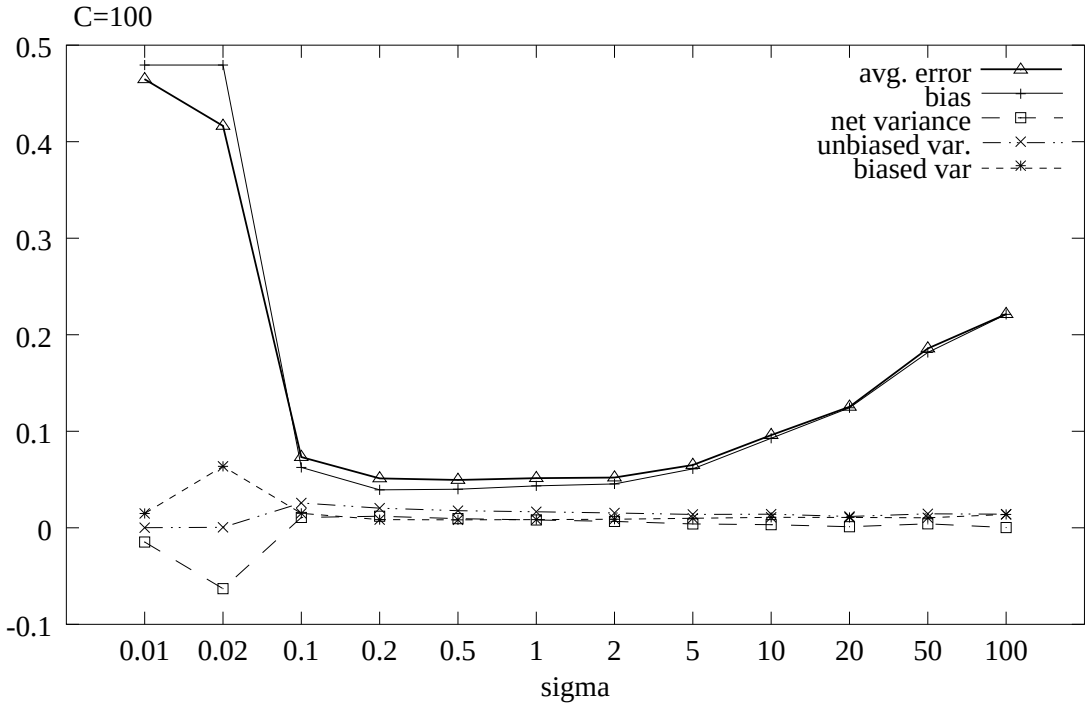


(a)

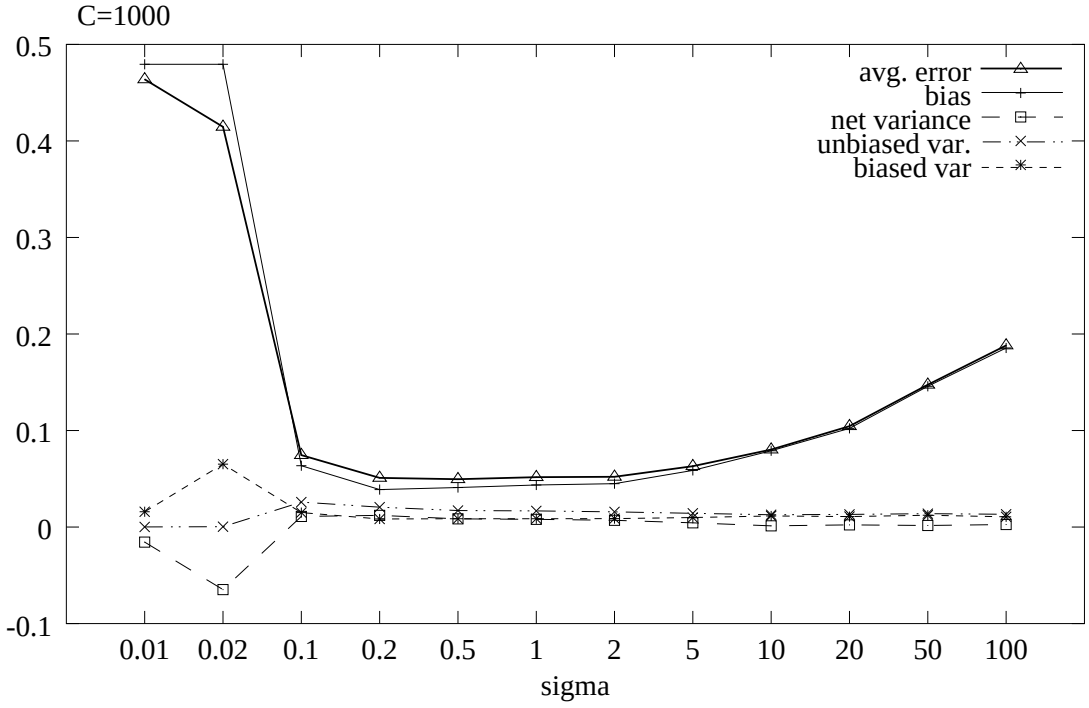


(b)

Figure 1: P2 data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF, while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 10$.

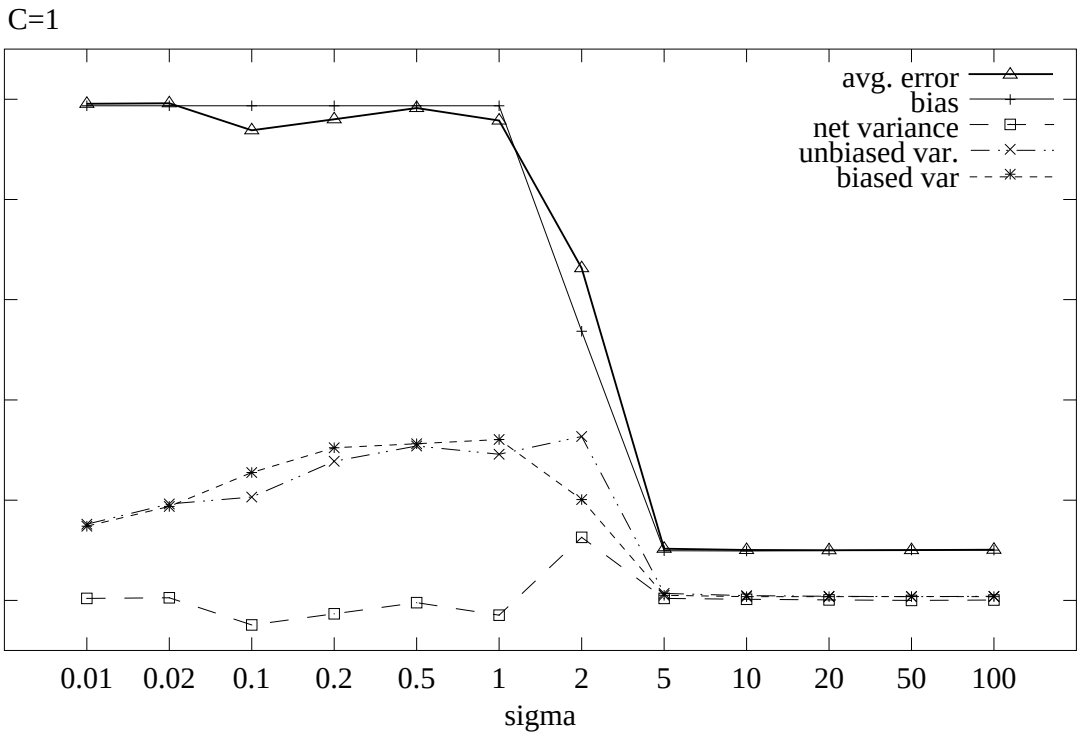


(a)

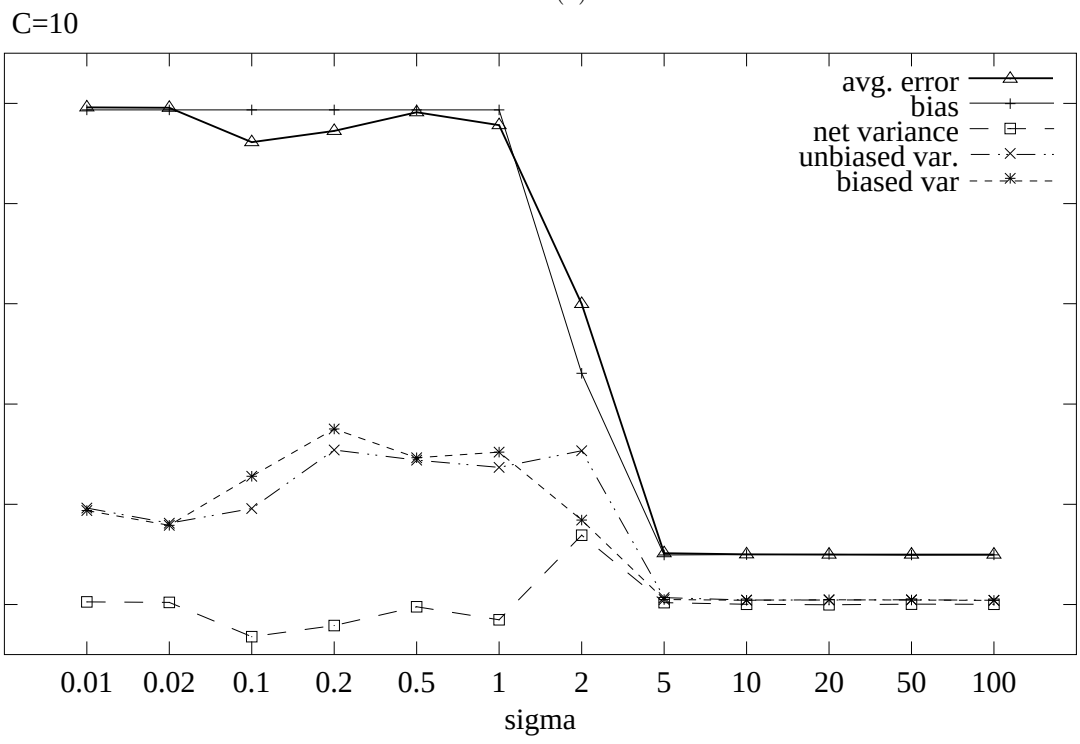


(b)

Figure 2: P2 data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF, while varying σ and for some fixed values of C: (a) $C = 100$, (b) $C = 1000$.

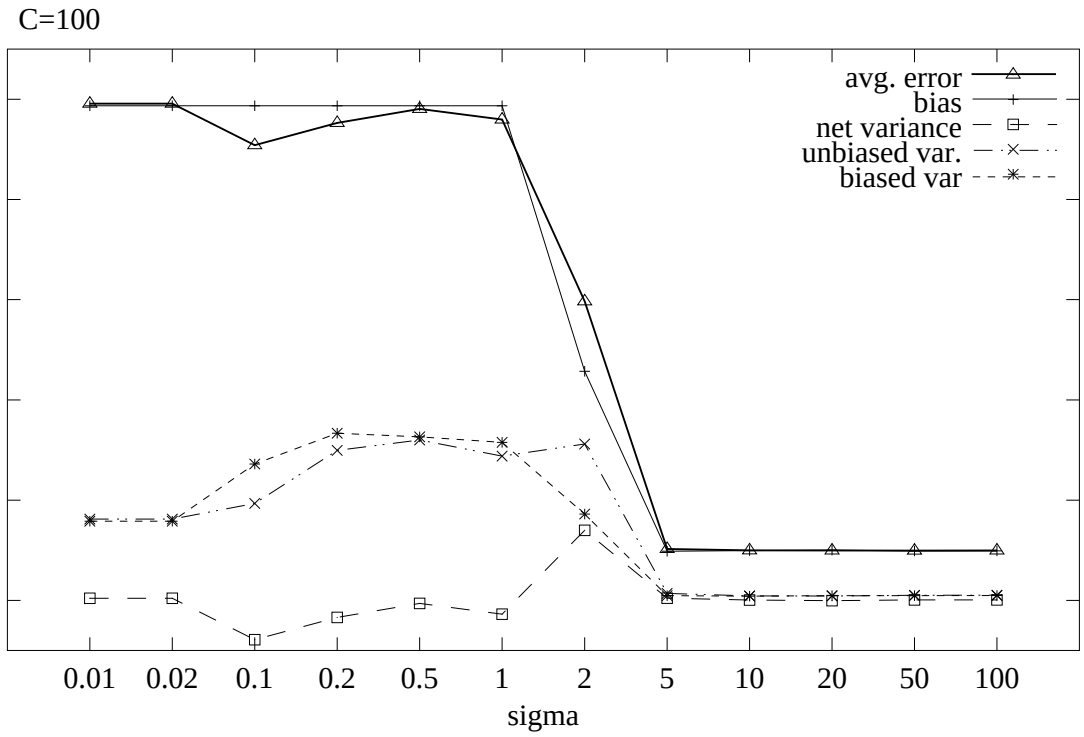


(a)

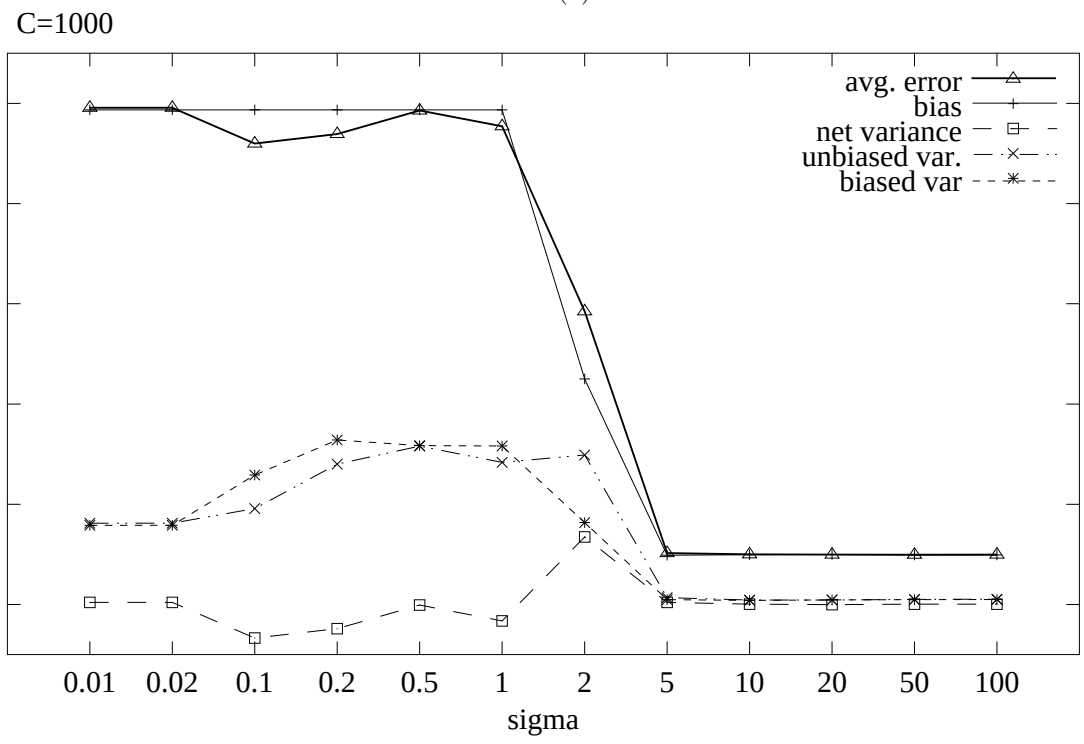


(b)

Figure 3: Waveform data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF, while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 10$.

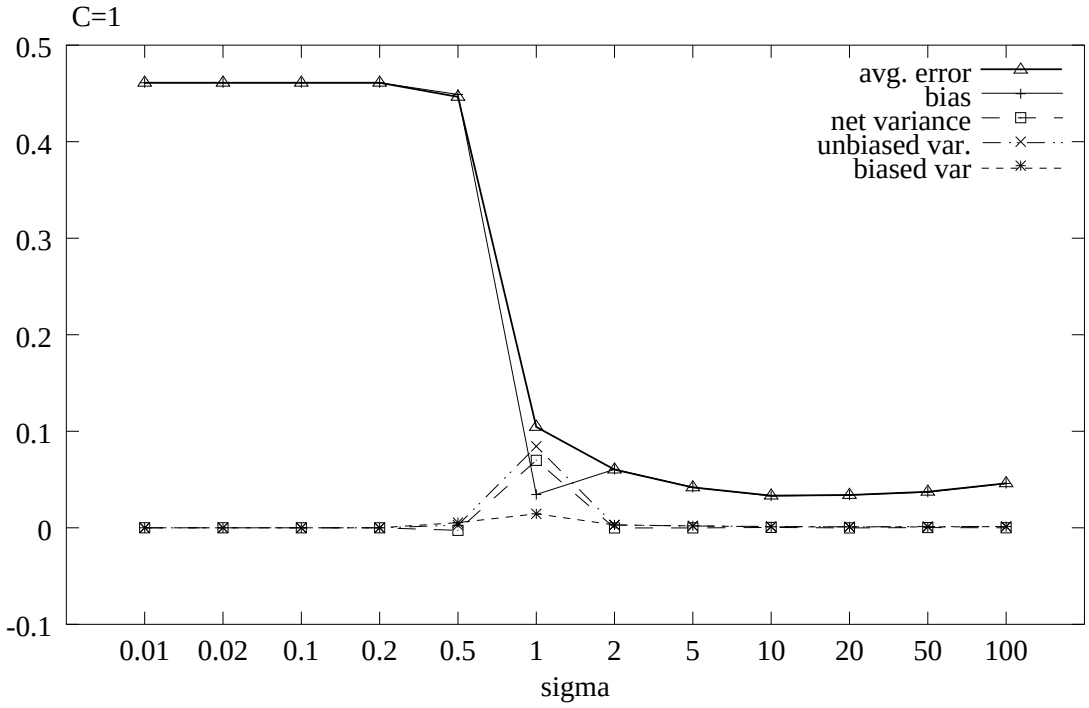


(a)

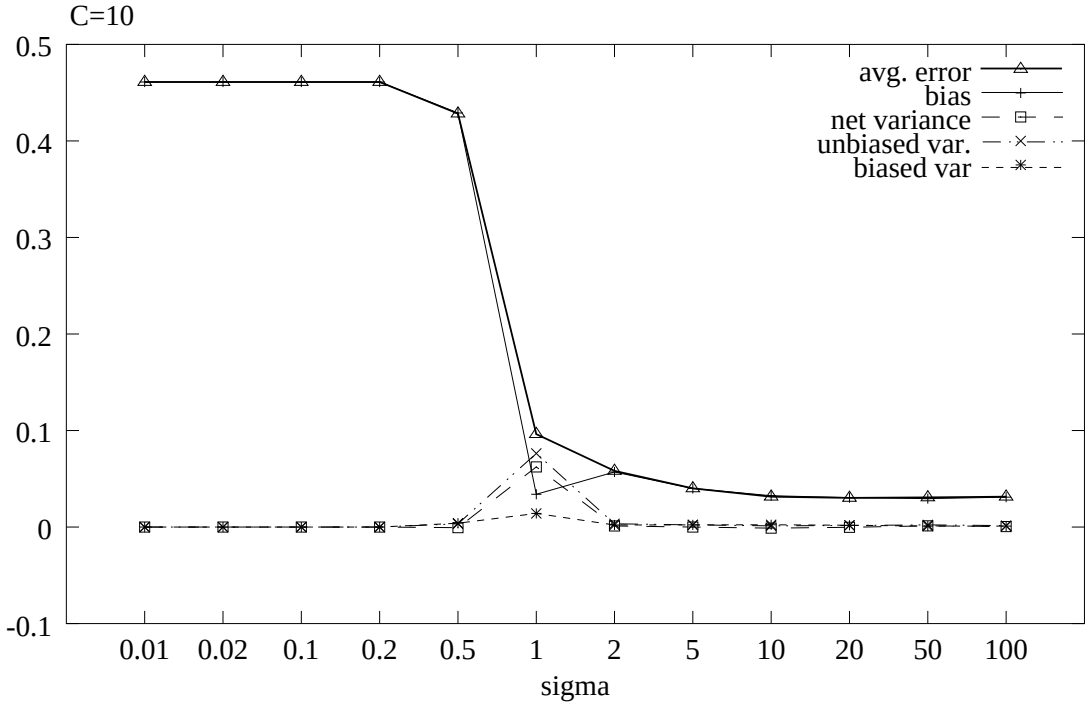


(b)

Figure 4: Waveform data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF, while varying σ and for some fixed values of C: (a) $C = 100$, (b) $C = 1000$.

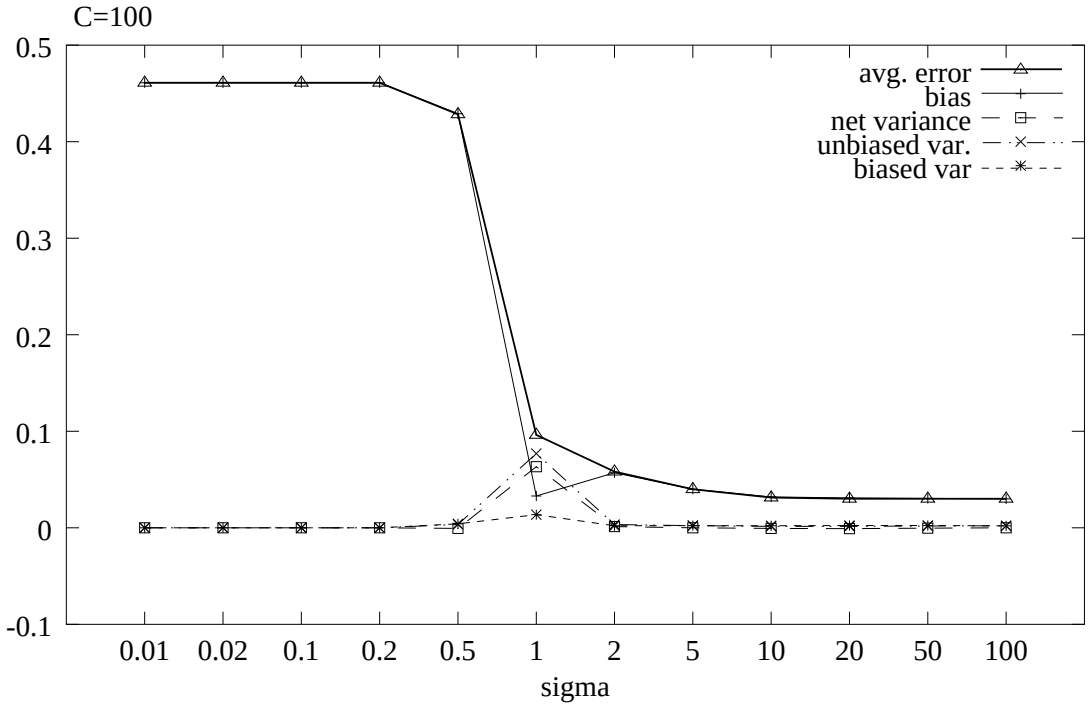


(a)

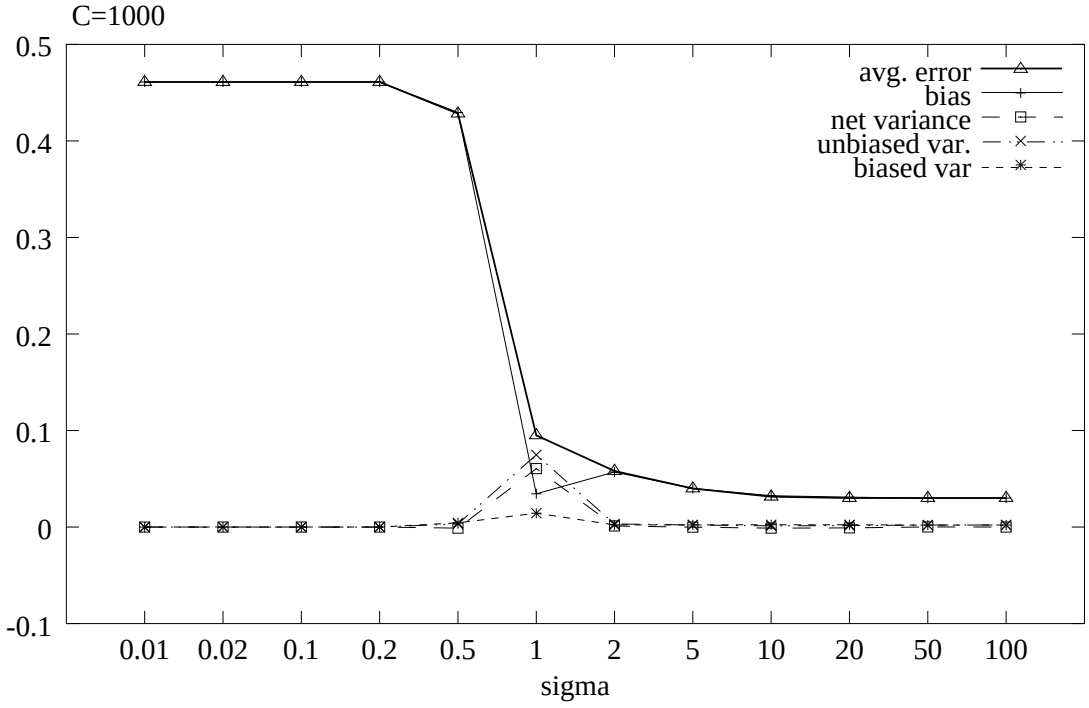


(b)

Figure 5: Grey-Landsat data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF, while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 10$.

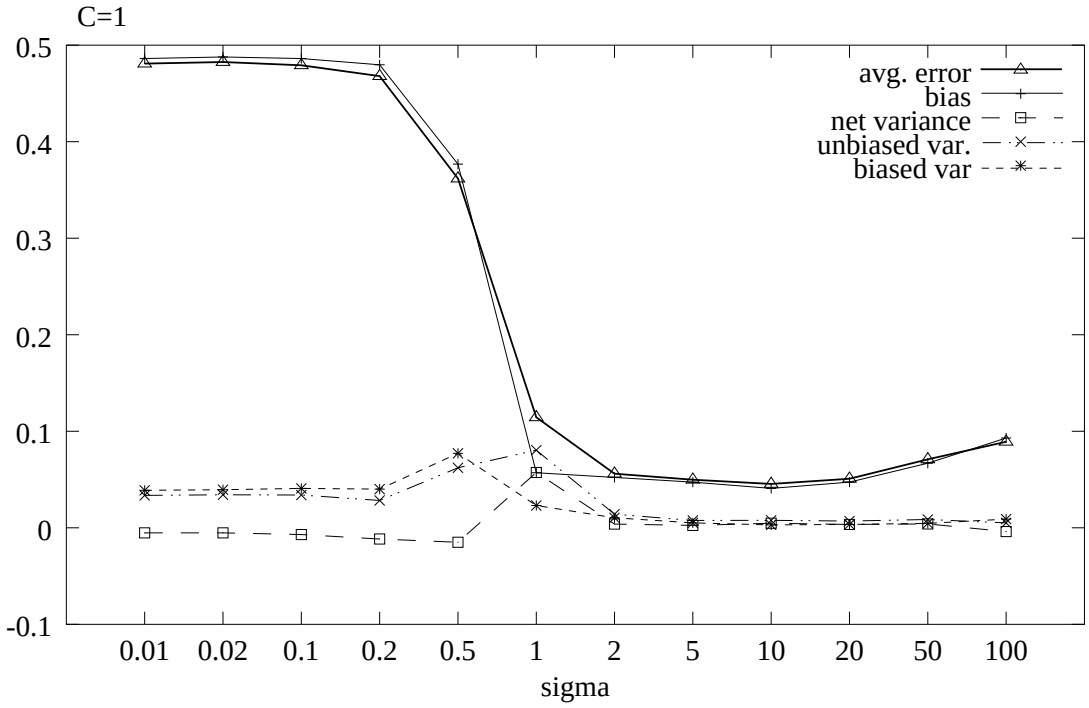


(a)

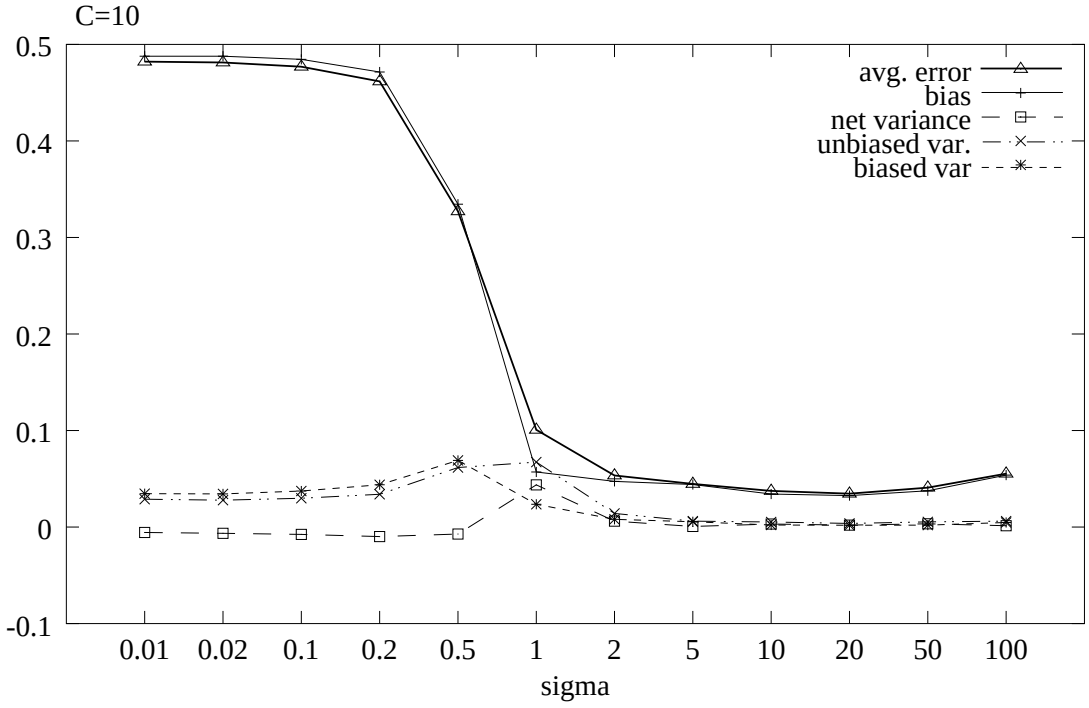


(b)

Figure 6: Grey-Landsat data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF, while varying σ and for some fixed values of C: (a) $C = 100$, (b) $C = 1000$.

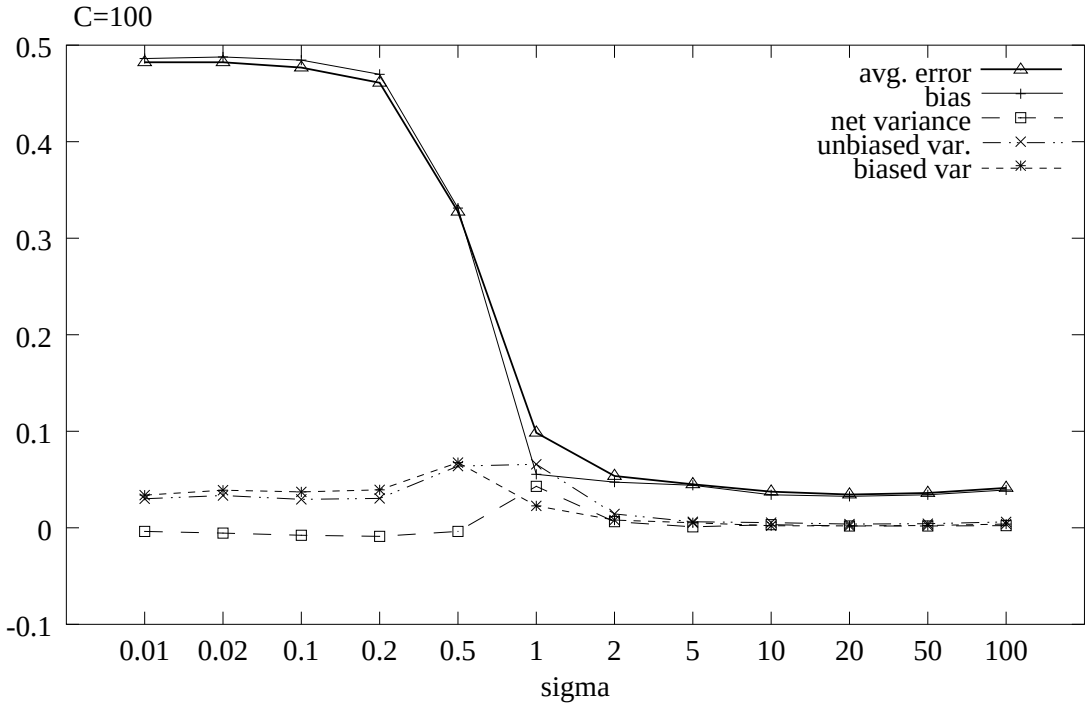


(a)

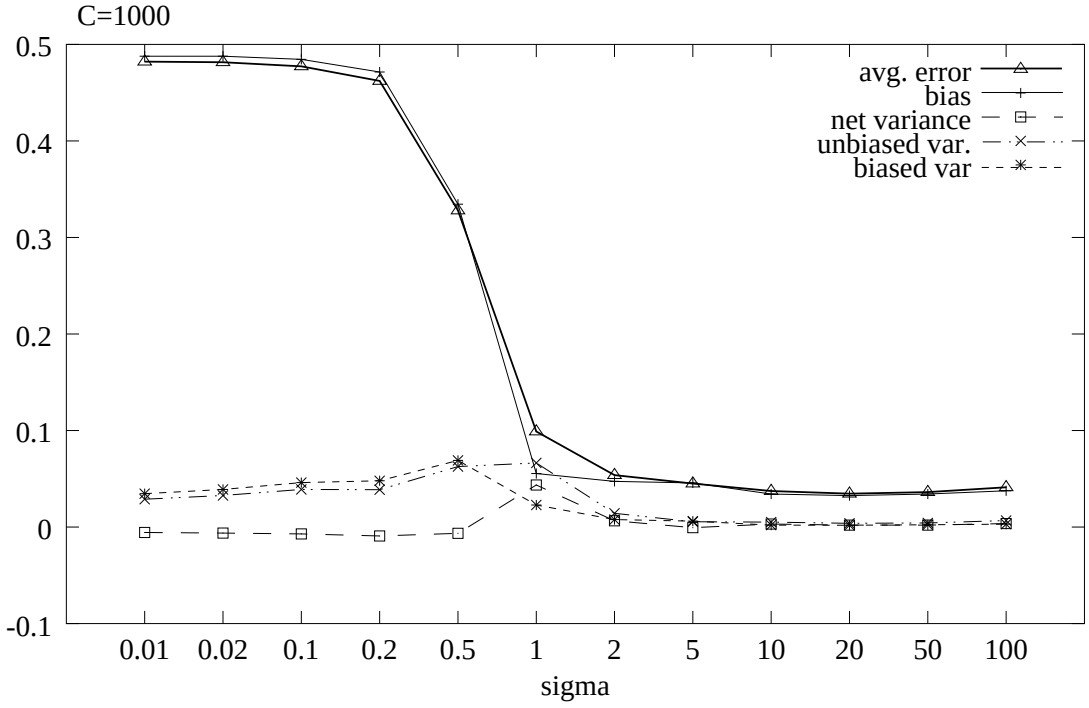


(b)

Figure 7: Letter-Two data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF, while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 10$.

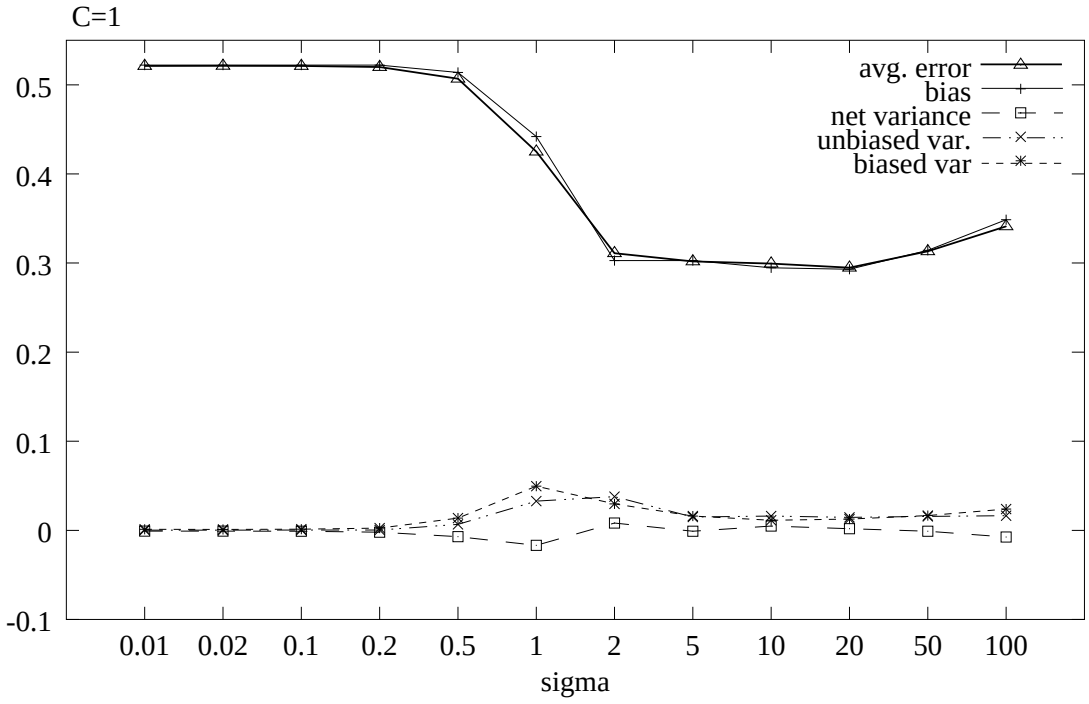


(a)

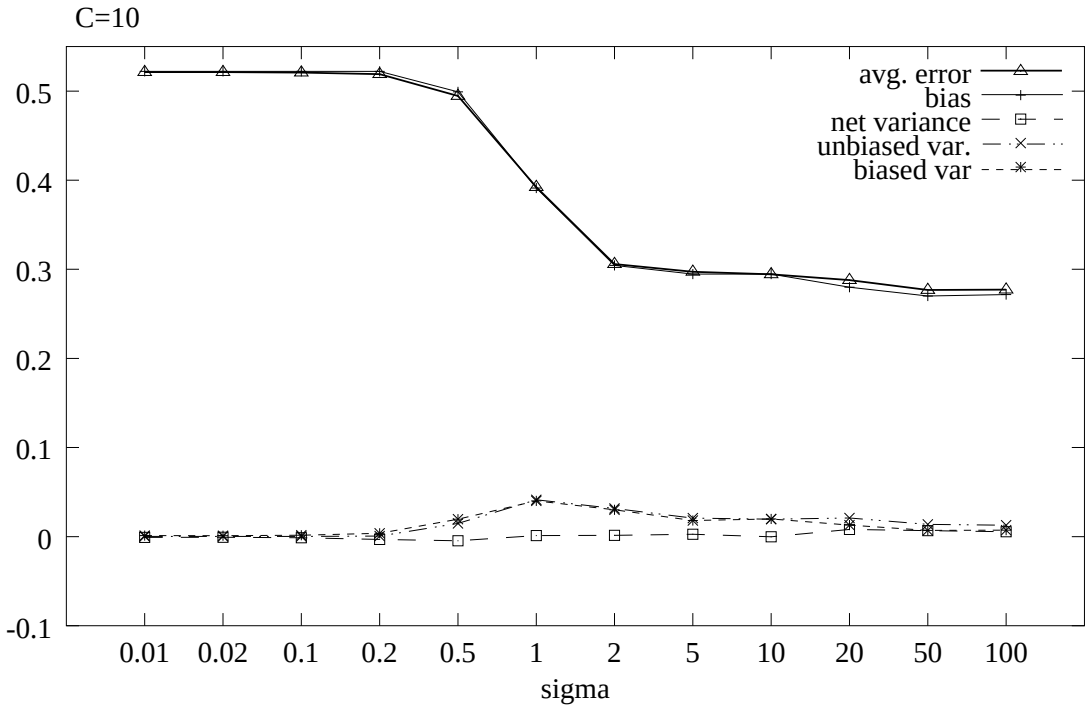


(b)

Figure 8: Letter-Two data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF, while varying σ and for some fixed values of C: (a) $C = 100$, (b) $C = 1000$.

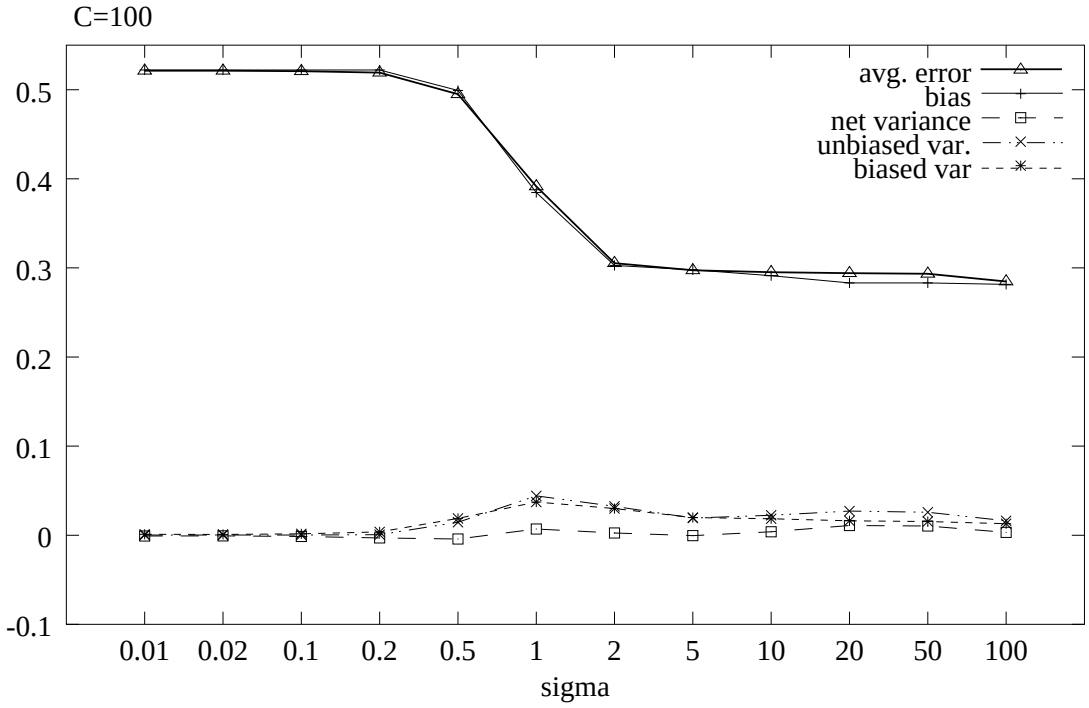


(a)

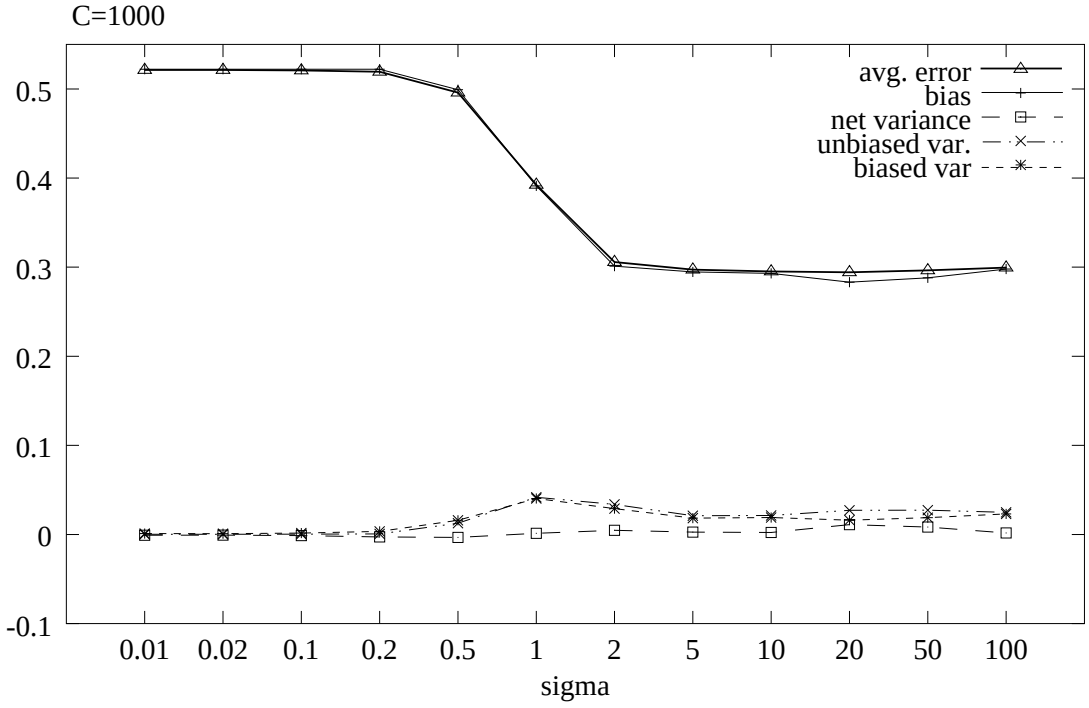


(b)

Figure 9: Letter-Two with noise data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF, while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 10$.

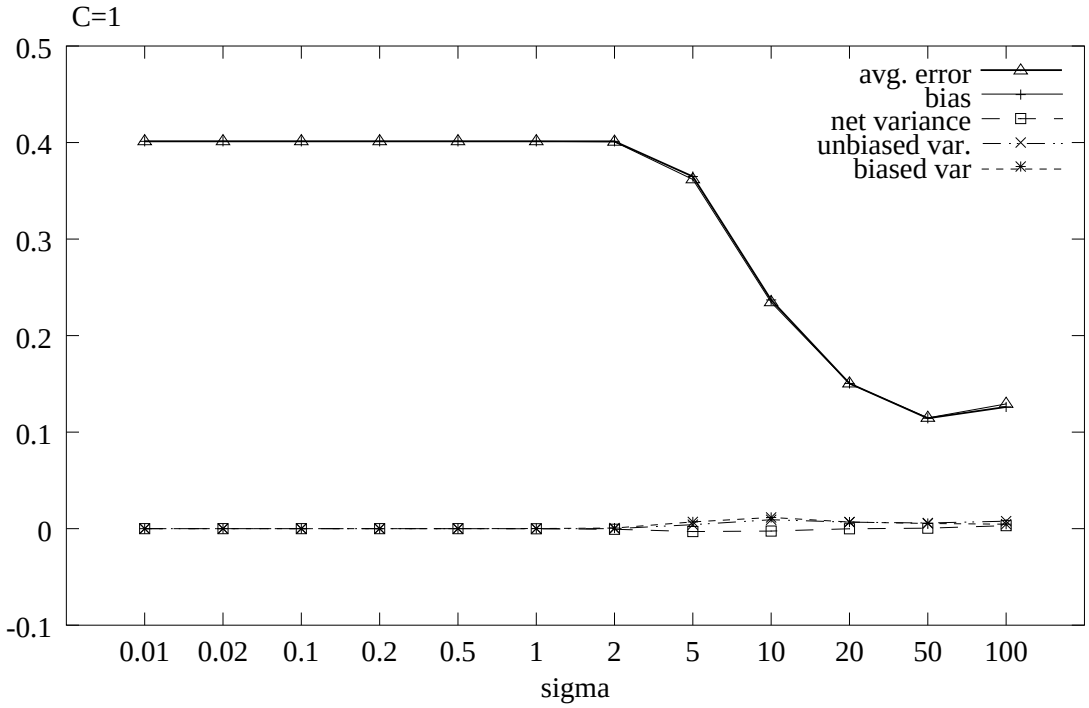


(a)

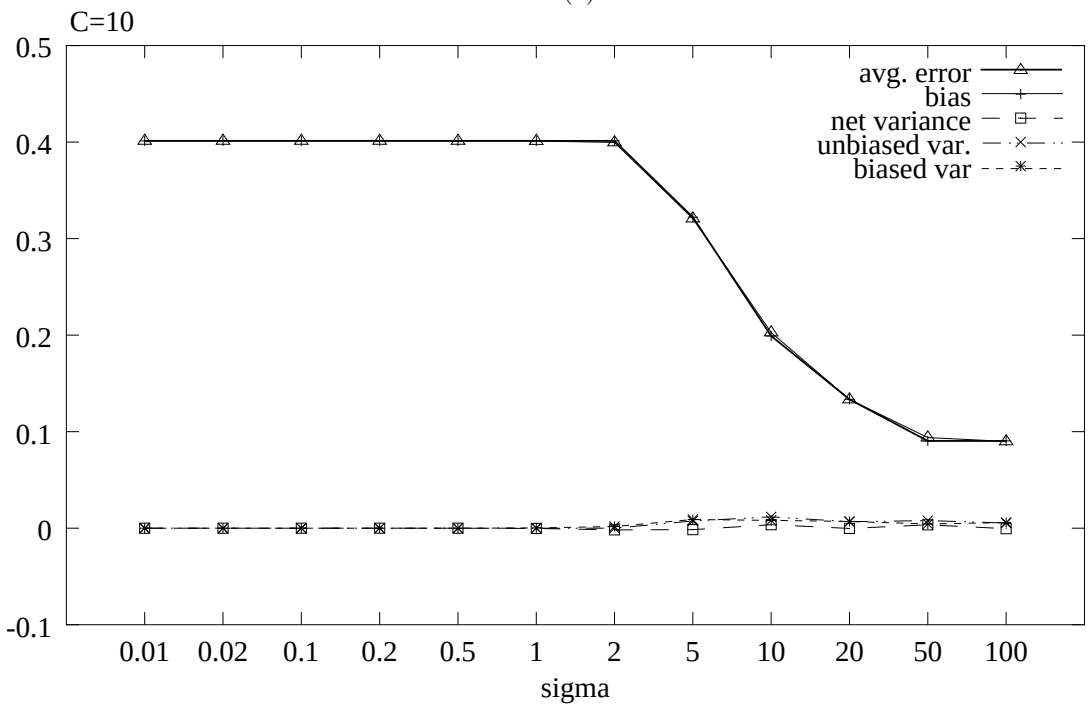


(b)

Figure 10: Letter-Two with noise data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF, while varying σ and for some fixed values of C: (a) $C = 100$, (b) $C = 1000$.

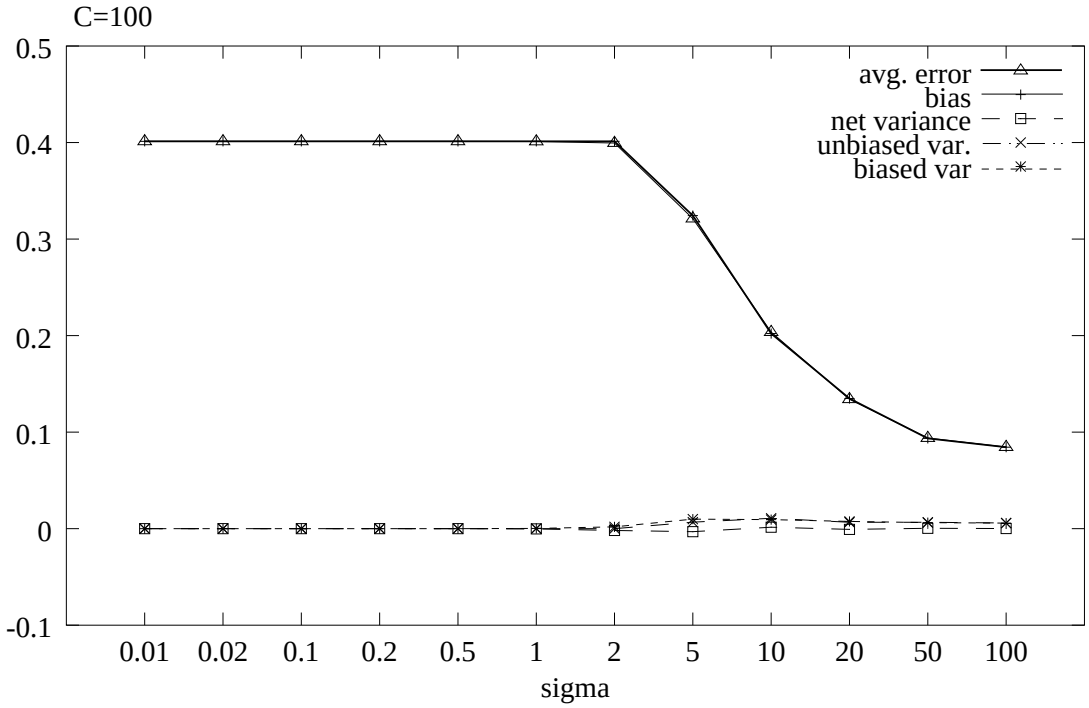


(a)

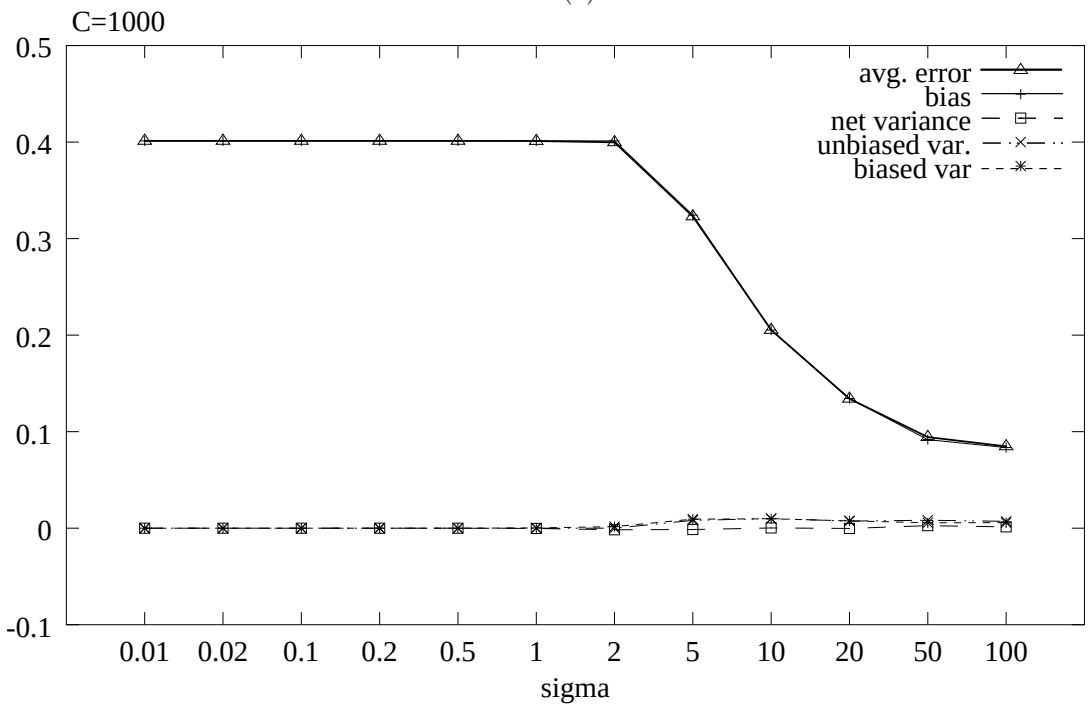


(b)

Figure 11: Spam data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF, while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 10$.

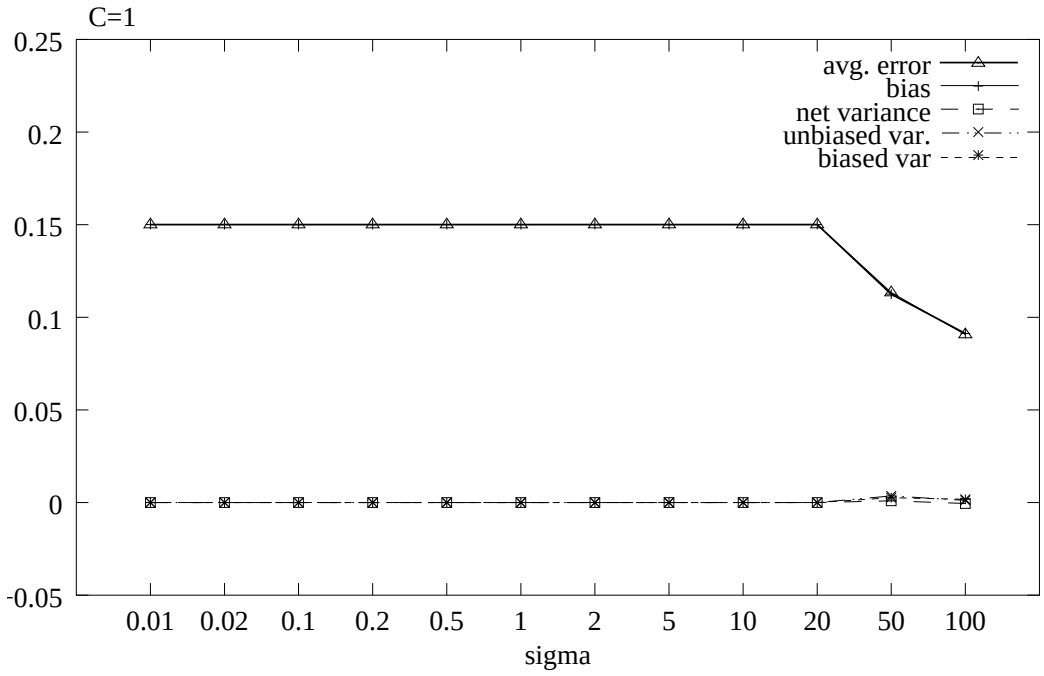


(a)

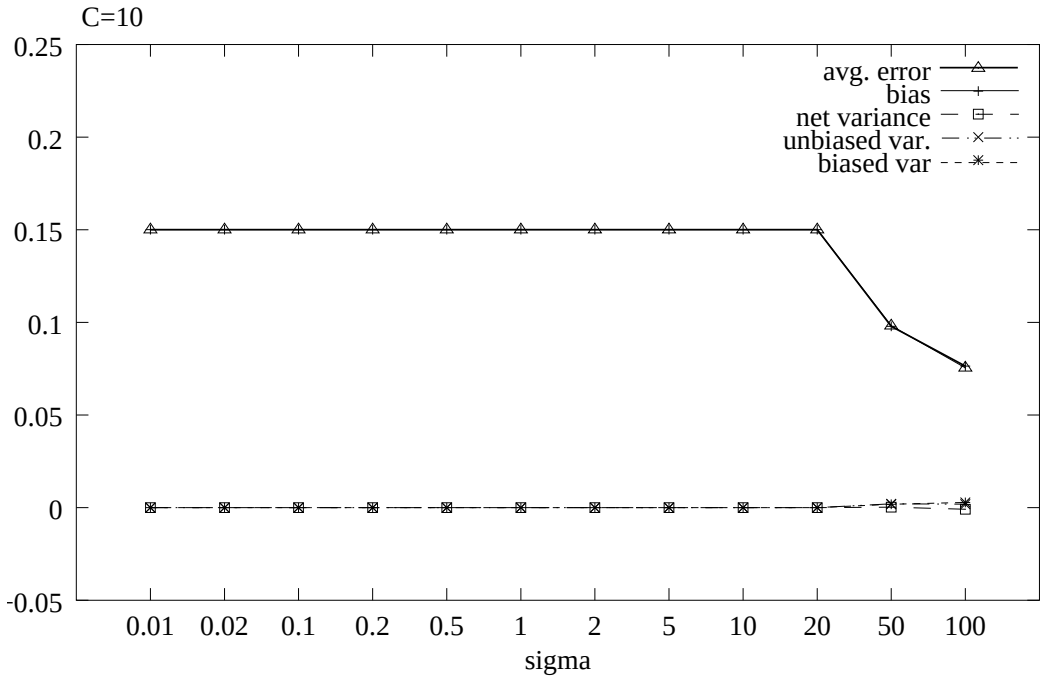


(b)

Figure 12: Spam data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF, while varying σ and for some fixed values of C: (a) $C = 100$, (b) $C = 1000$.

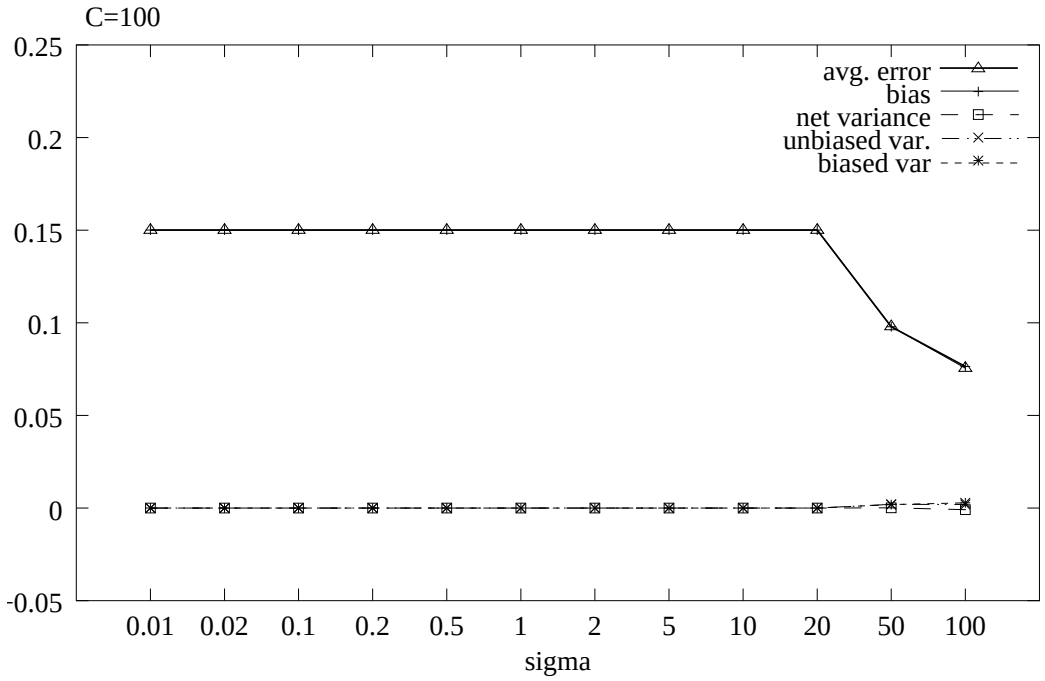


(a)

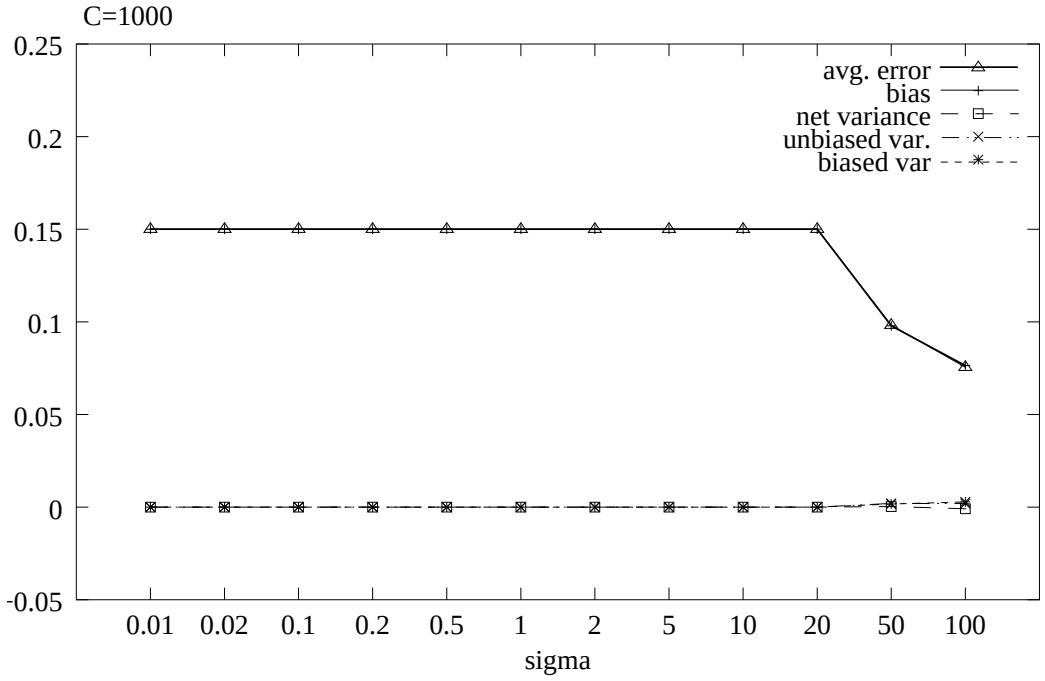


(b)

Figure 13: Musk data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF, while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 10$.



(a)



(b)

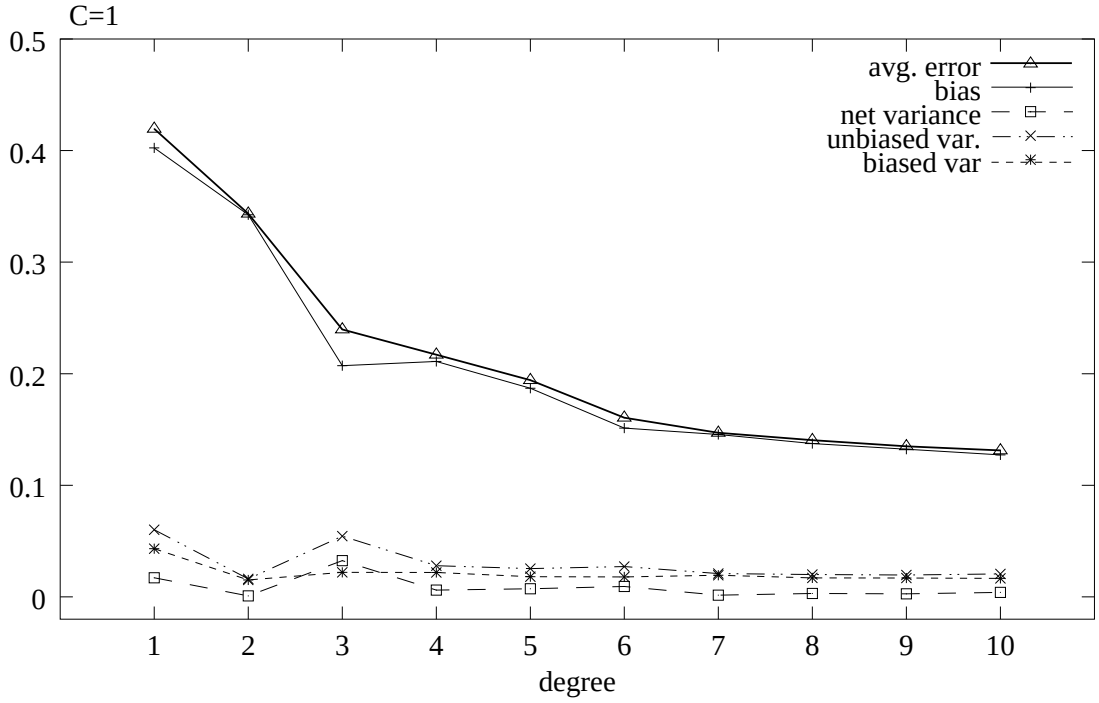
Figure 14: Musk data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF, while varying σ and for some fixed values of C: (a) $C = 100$, (b) $C = 1000$.

2.2 Decomposition in RA Polynomial SVMs

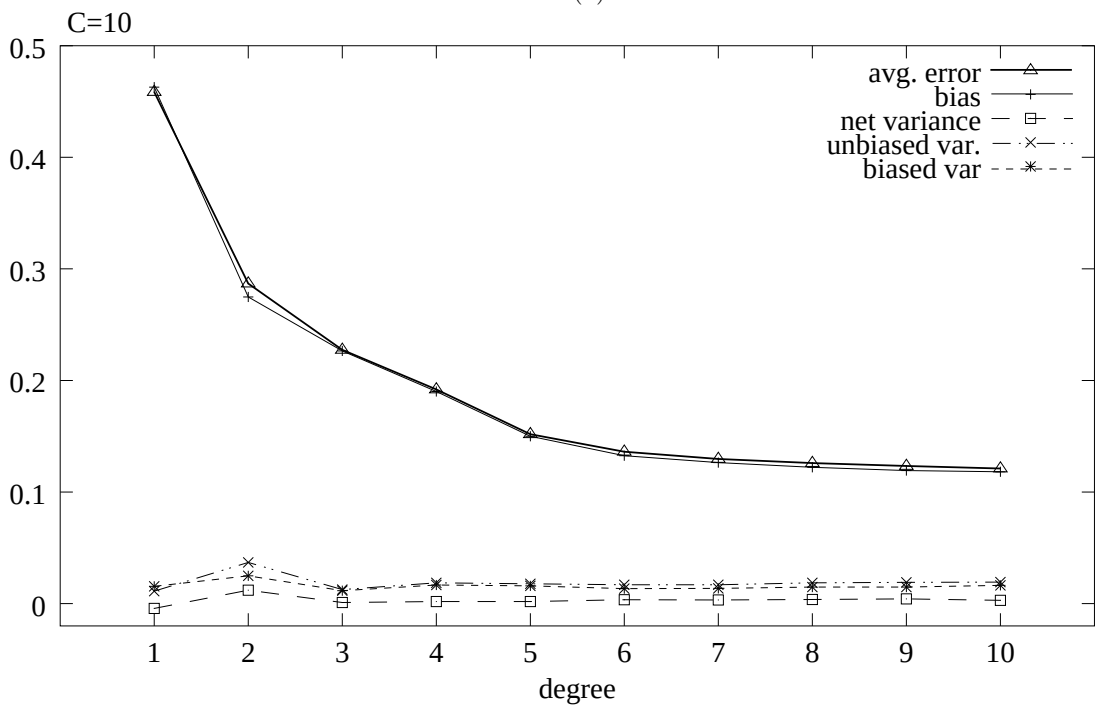
The decomposition of the error is represented with respect to different values of the polynomial degree and for fixed values of C.

Schematically we can observe the following facts:

- In all the data sets the net-variance is about 0 for all the values of polynomial degree, as both biased and unbiased variance are very low close to 0. Only in some data sets (e.g. P2), with low values of C we can observe a certain level on unbiased variance, especially with low degree polynomials.
- In almost all the considered data sets the error shows an "U" shape with respect to the degree. This shape tends to a parallel line if C is relatively large. With the P2 data set the error decreases with the degree.
- The error is determinated almost totally by the bias.

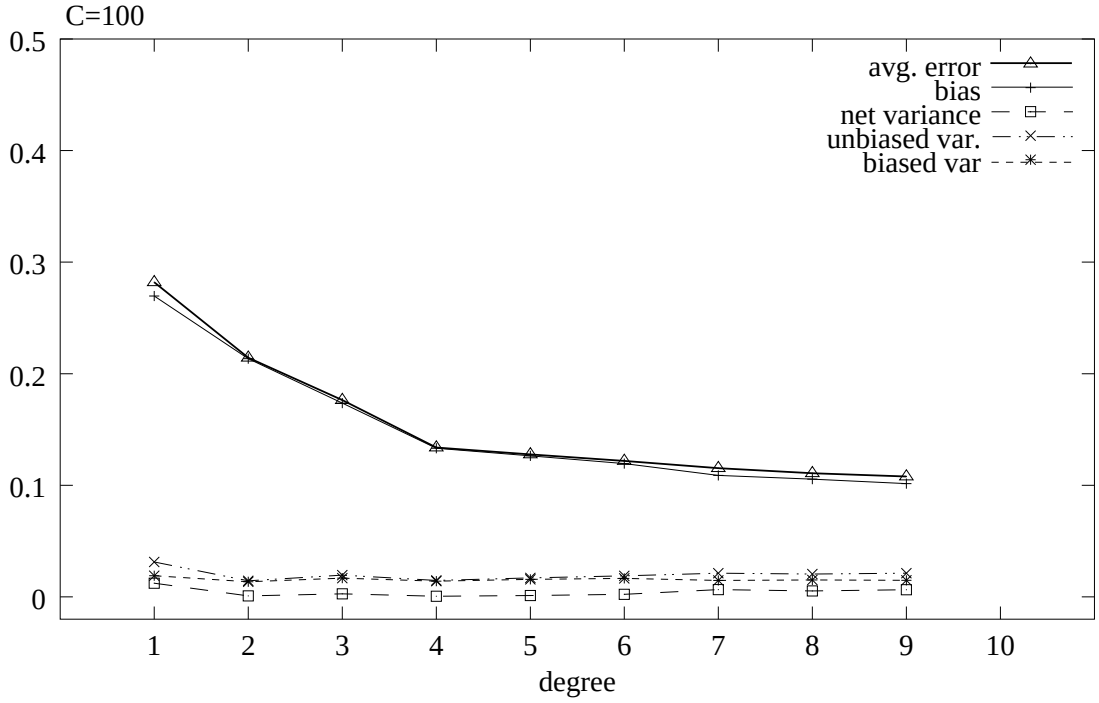


(a)

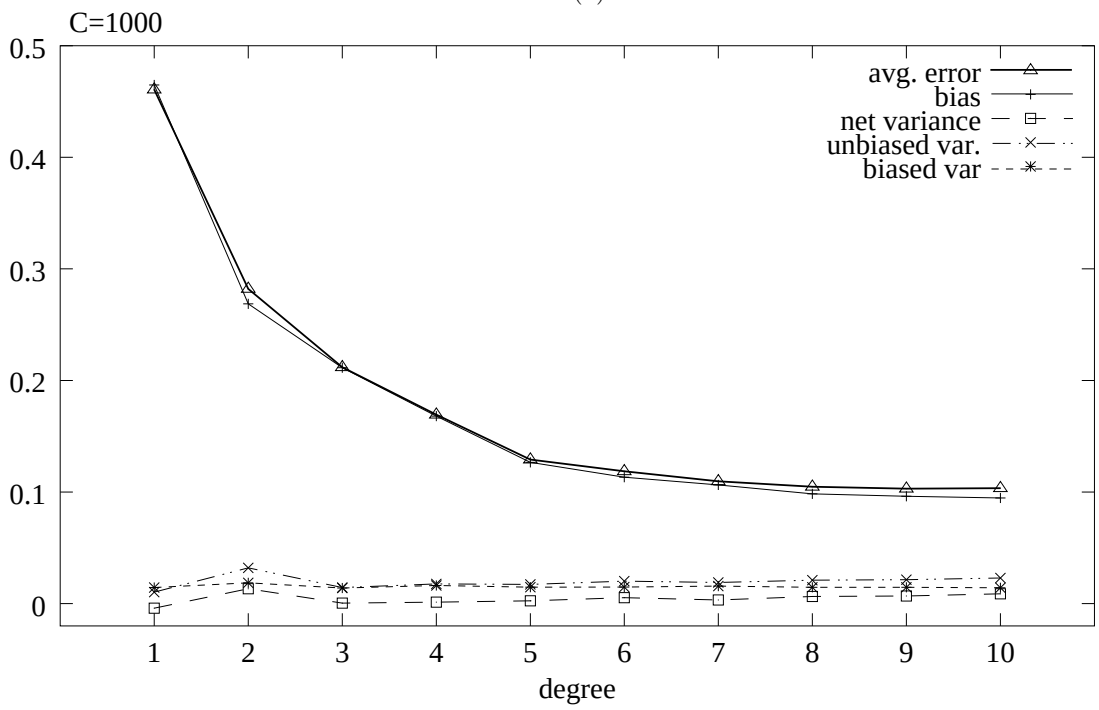


(b)

Figure 15: P2 data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 1$.

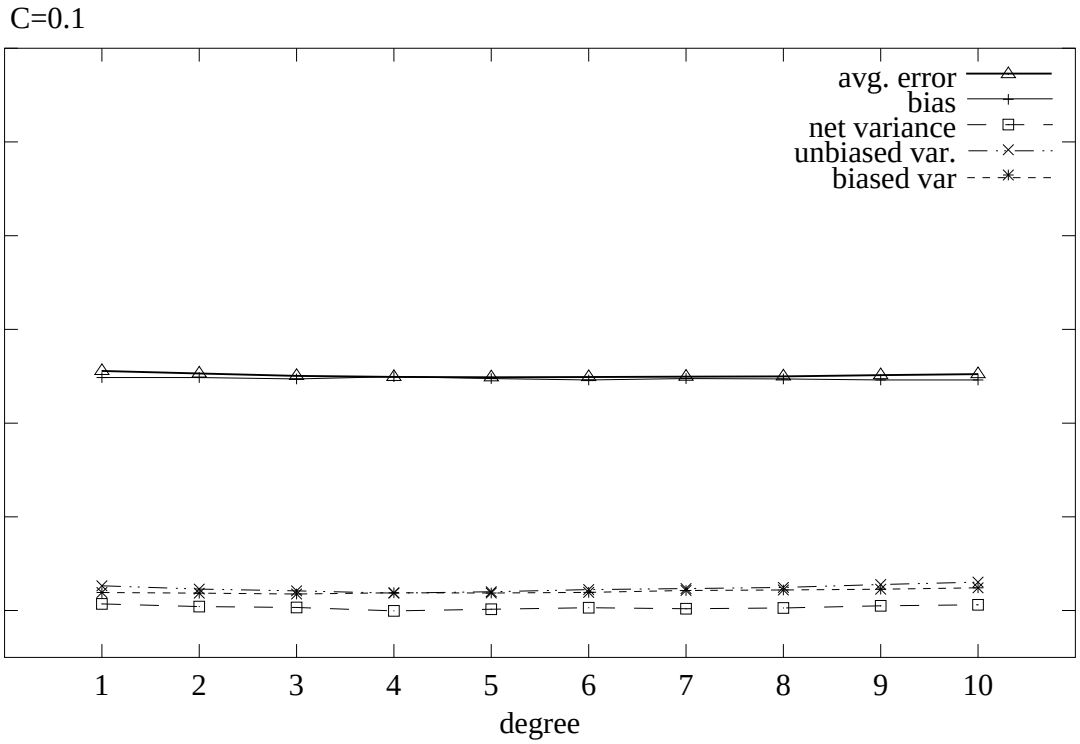


(a)

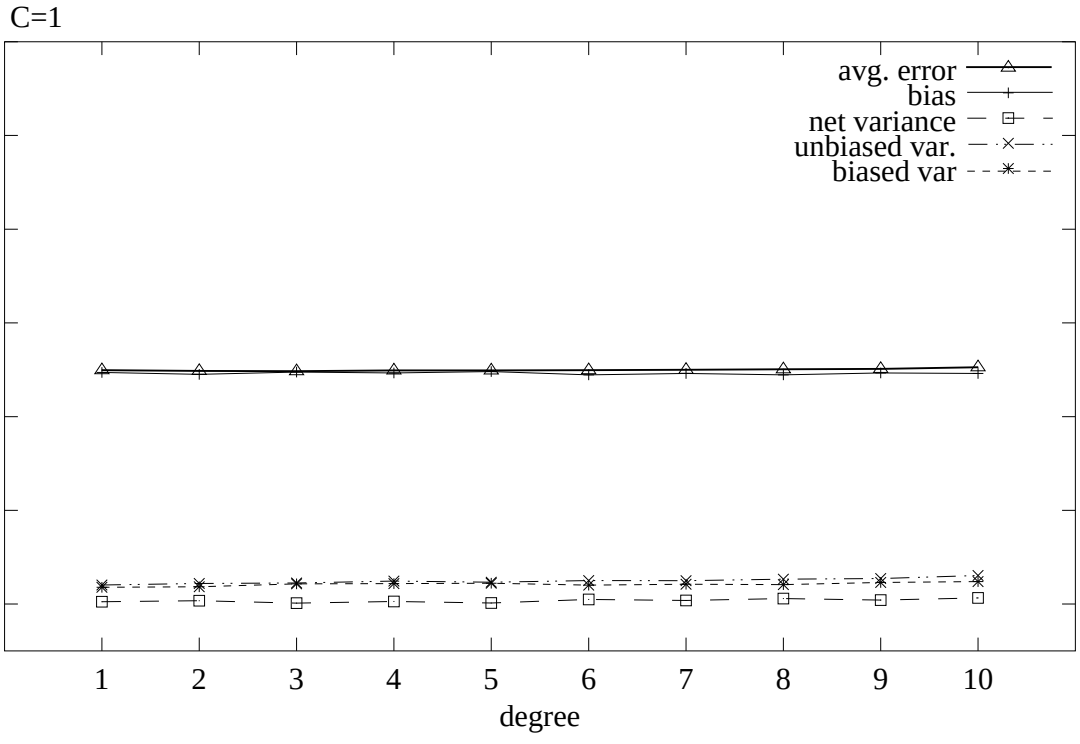


(b)

Figure 16: P2 data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 10$, (b) $C = 100$.

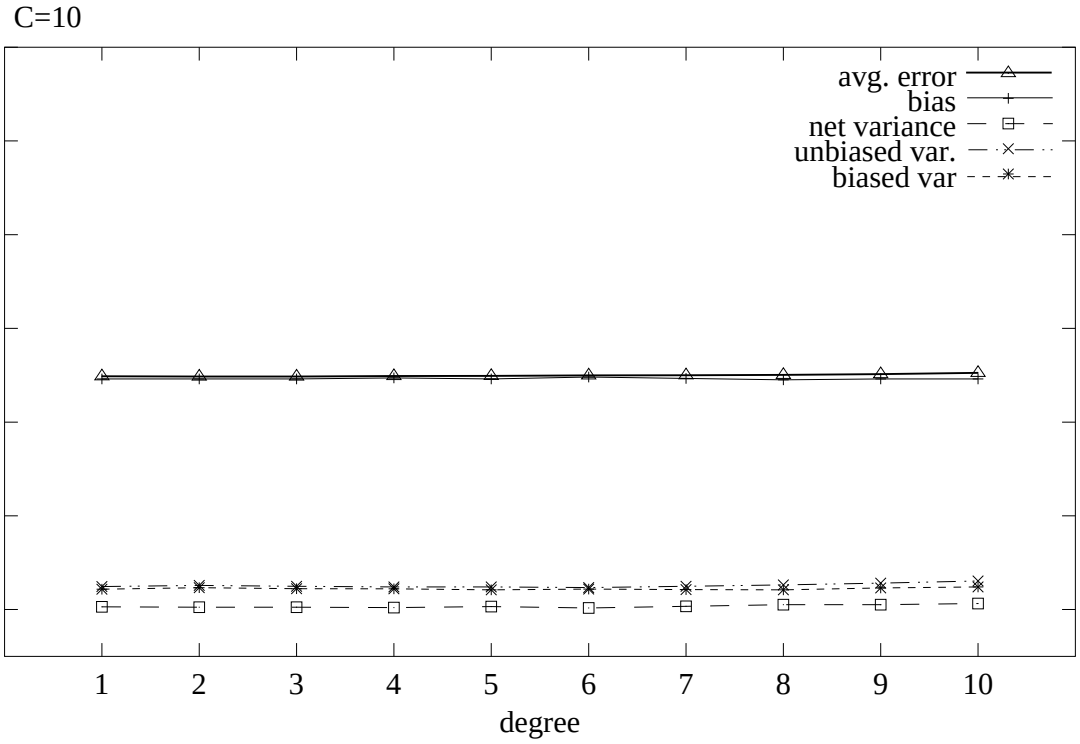


(a)

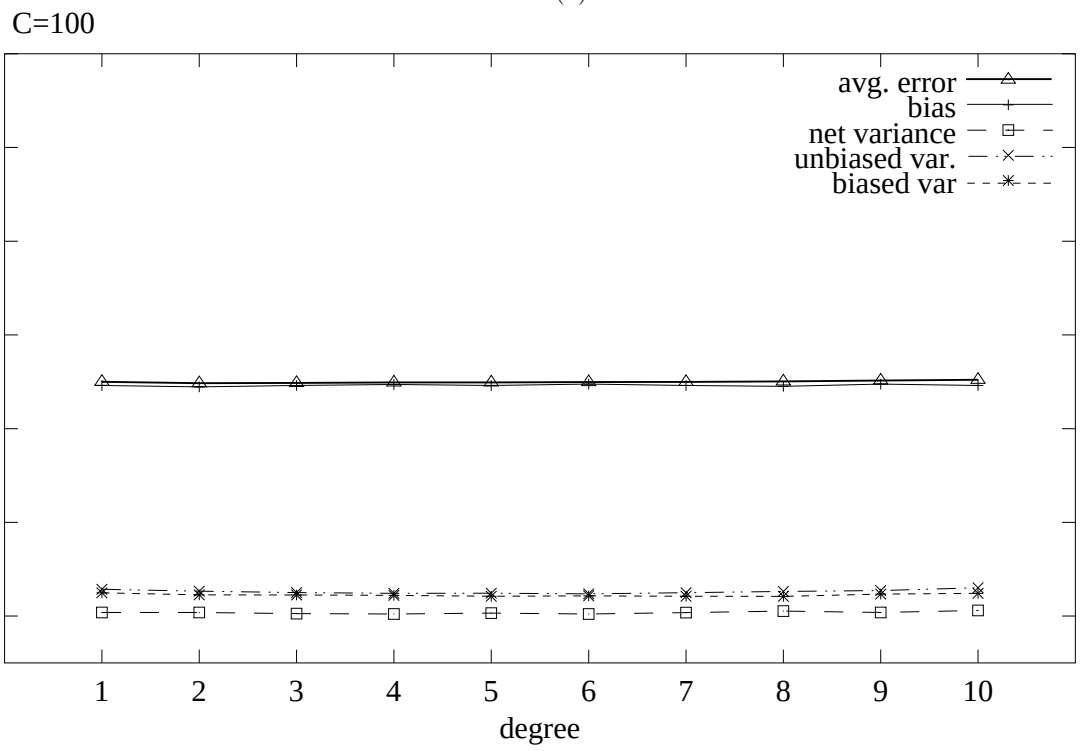


(b)

Figure 17: Waveform data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 1$.

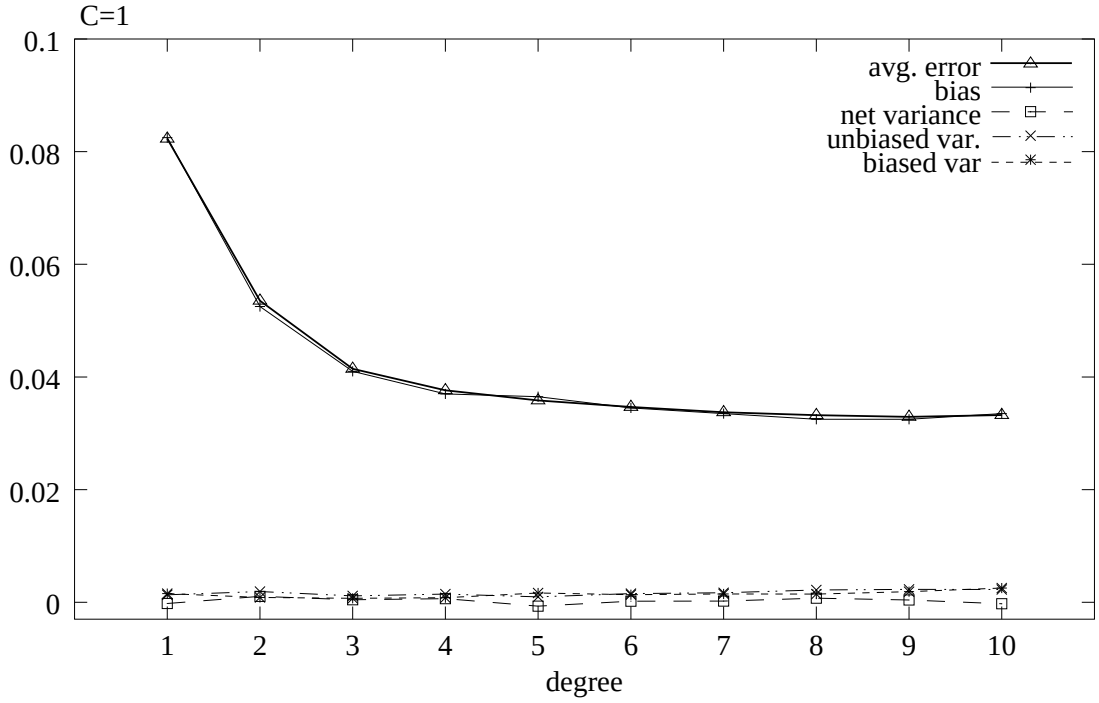


(a)

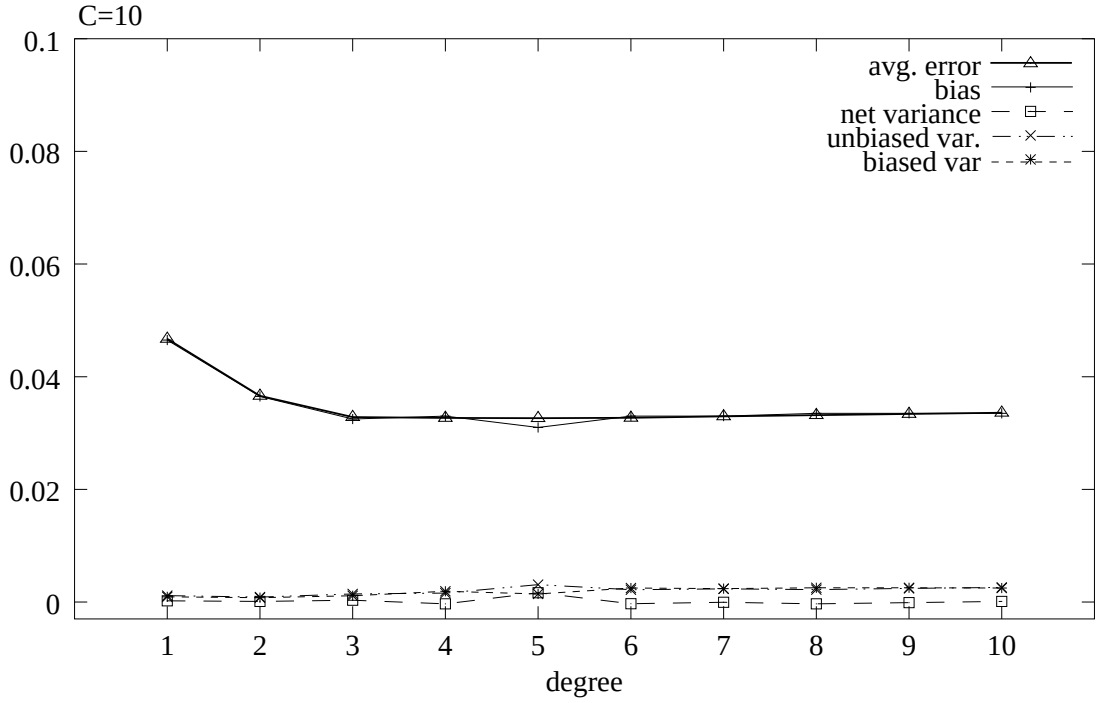


(b)

Figure 18: Waveform data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 10$, (b) $C = 100$.

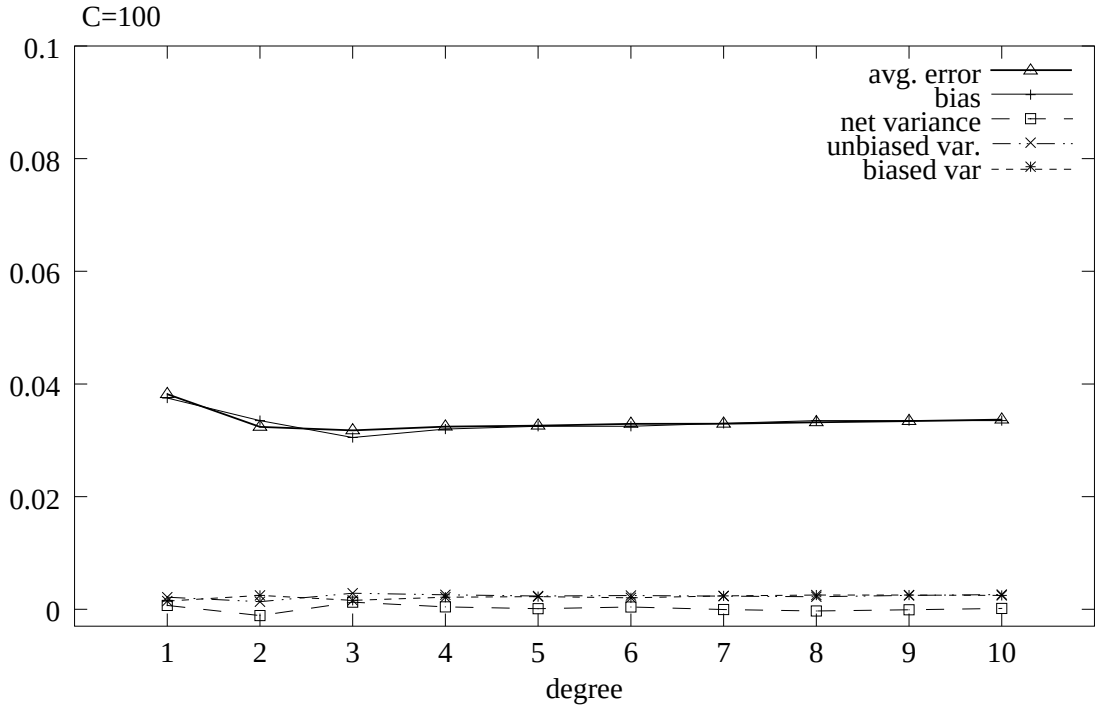


(a)

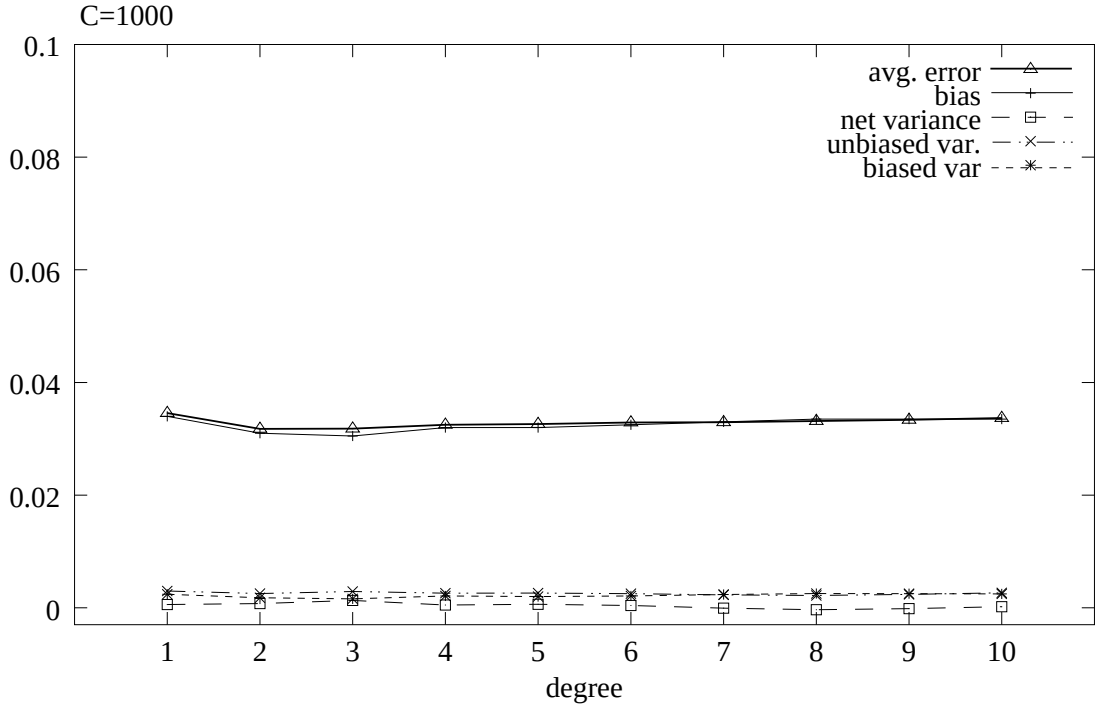


(b)

Figure 19: Grey-Landsat data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 1$.

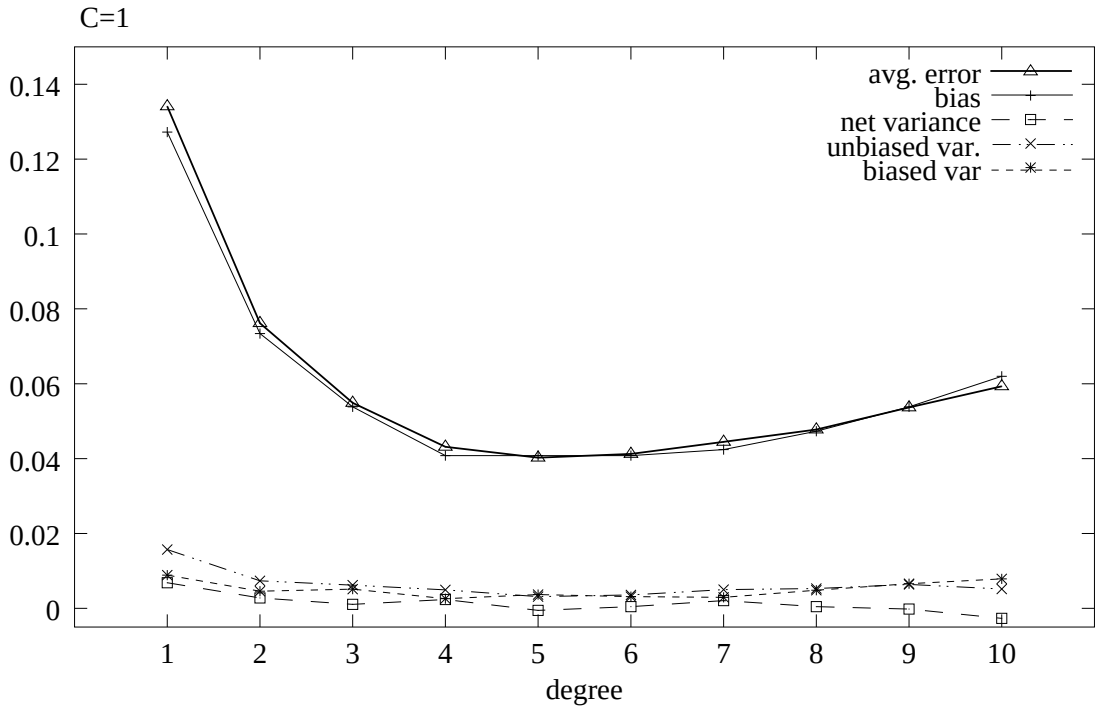


(a)

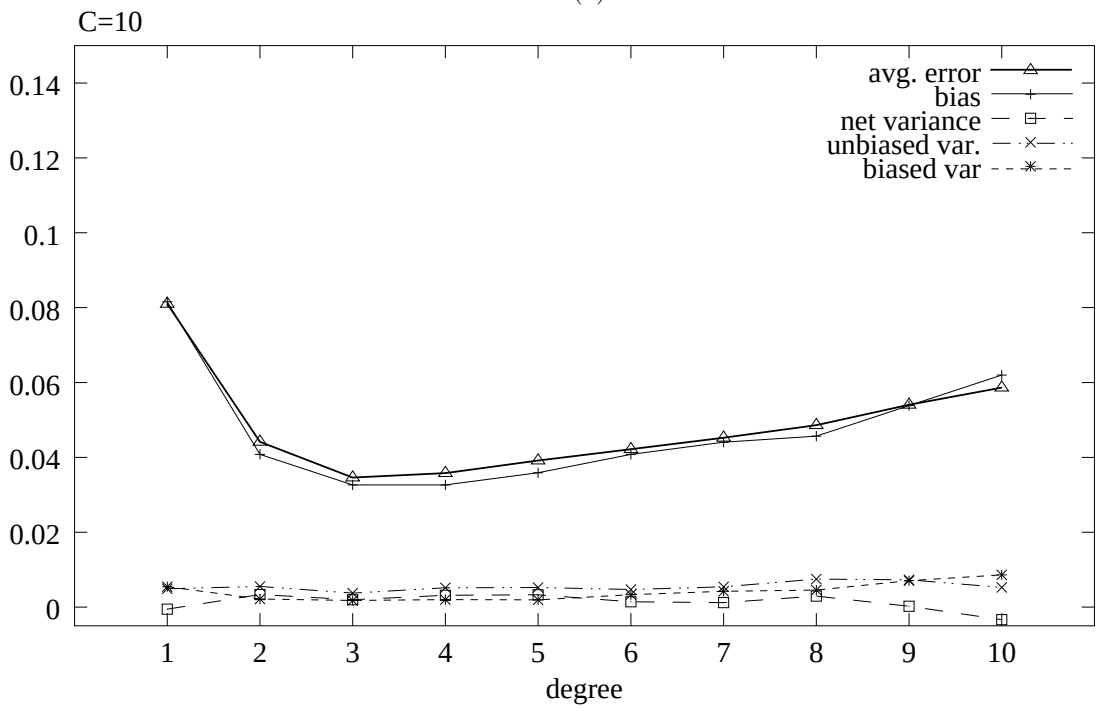


(b)

Figure 20: Grey-Landsat data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 10$, (b) $C = 100$.

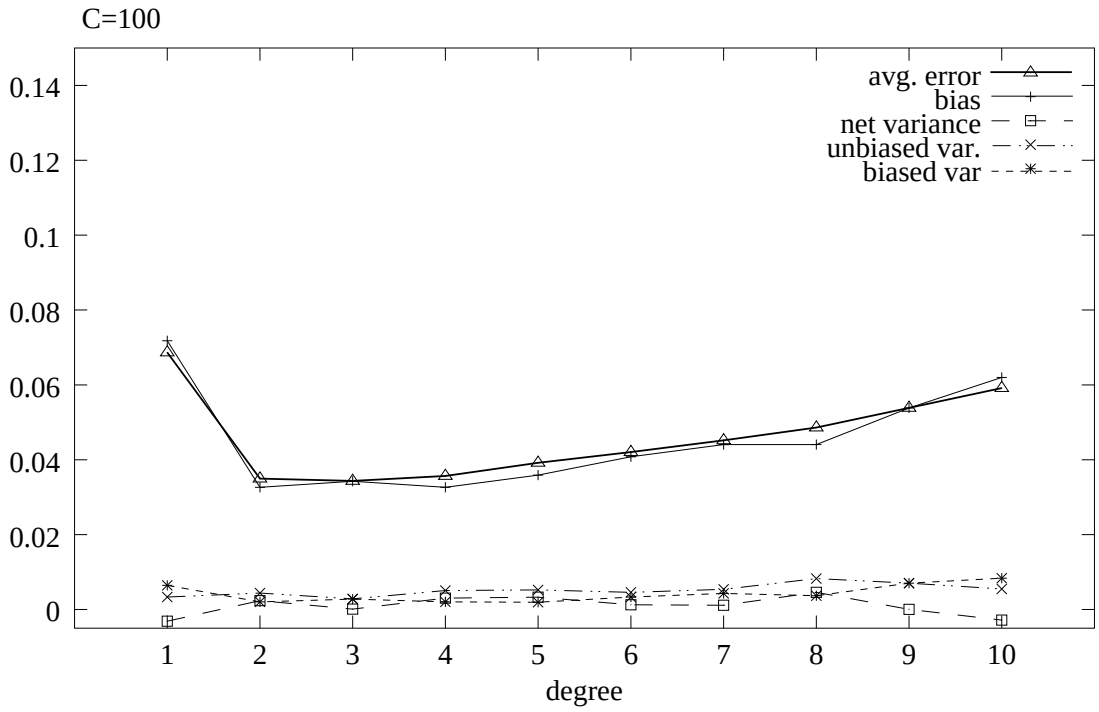


(a)

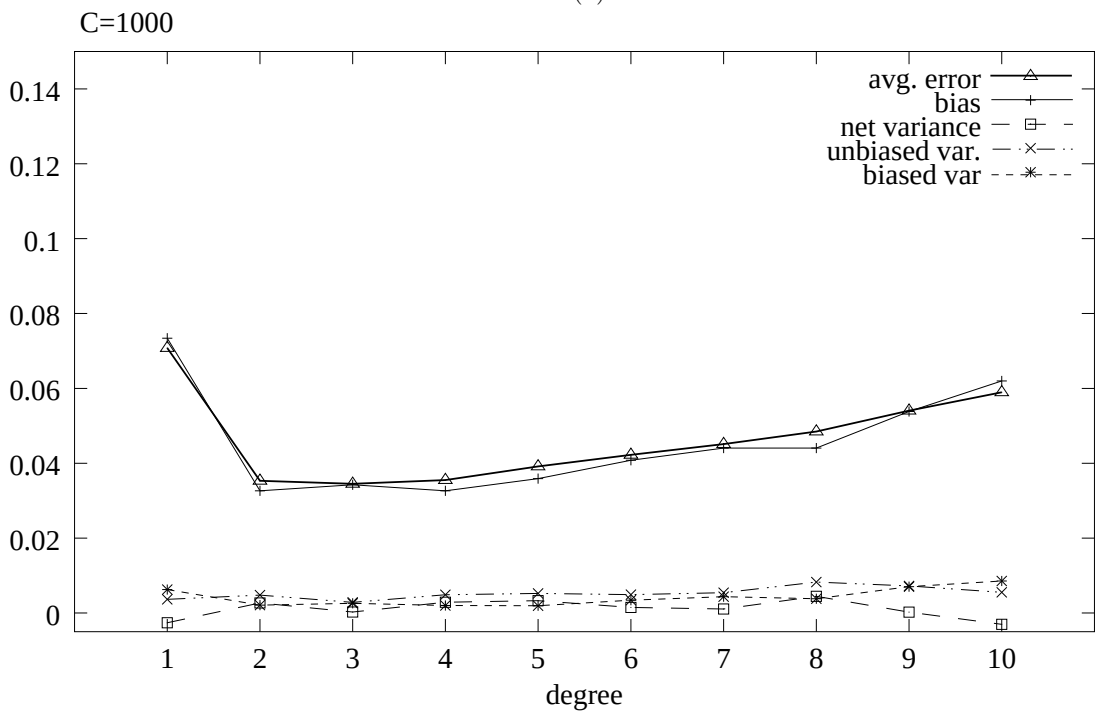


(b)

Figure 21: Letter-Two data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 1$.

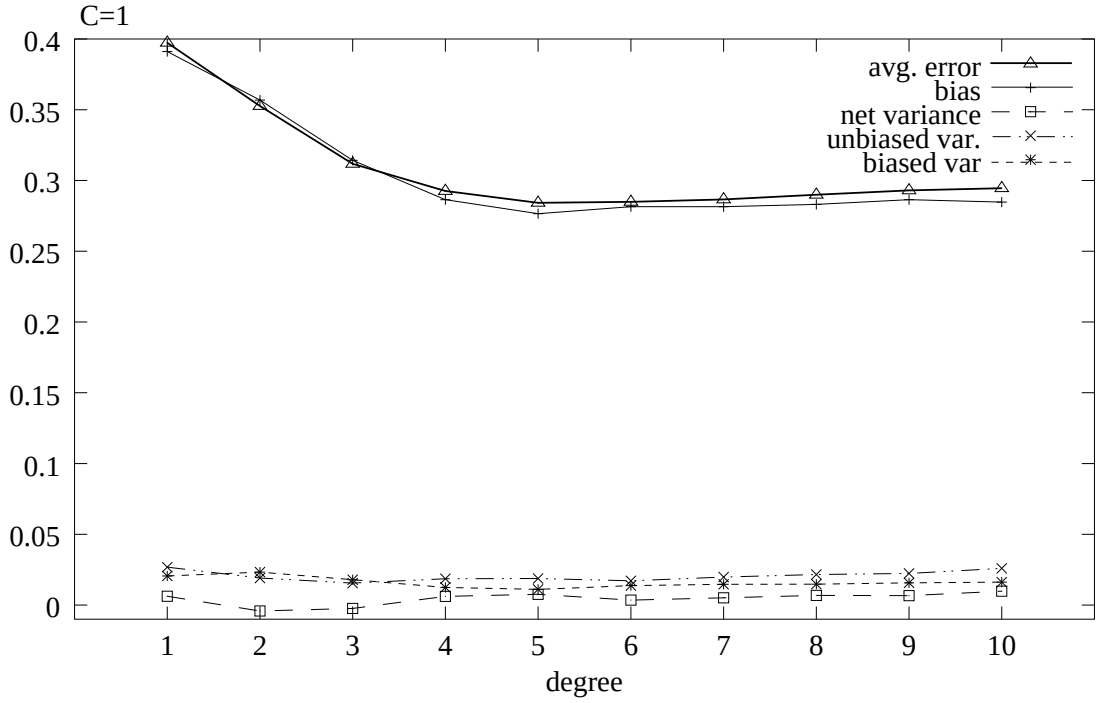


(a)

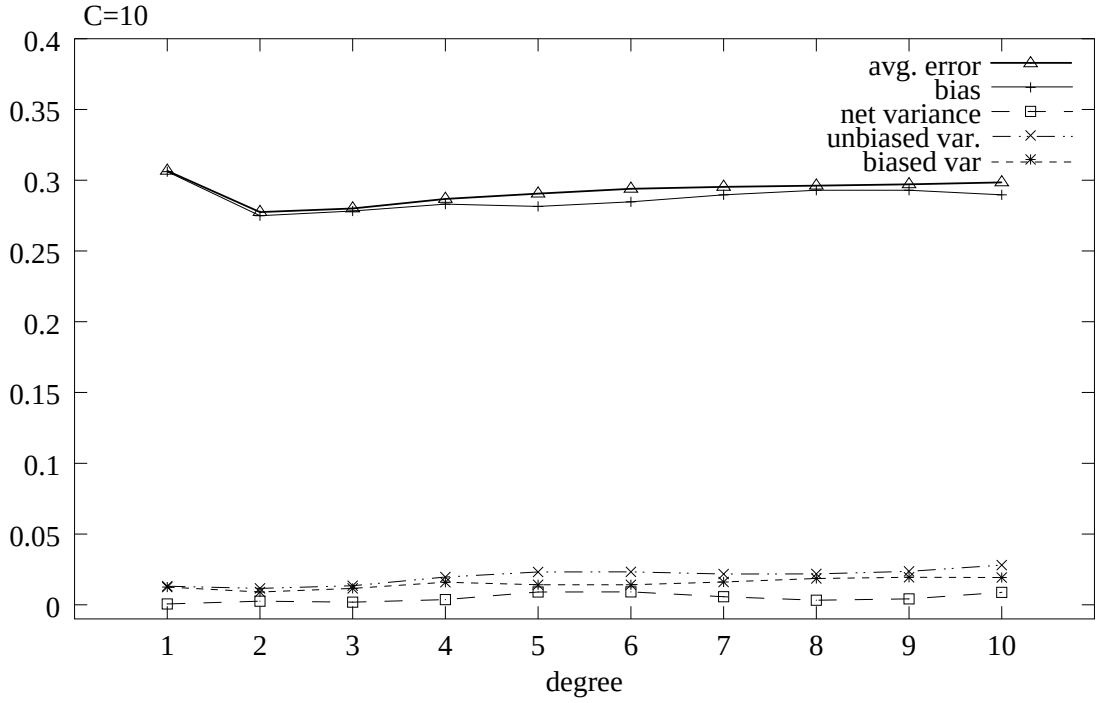


(b)

Figure 22: Letter-Two data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 10$, (b) $C = 100$.

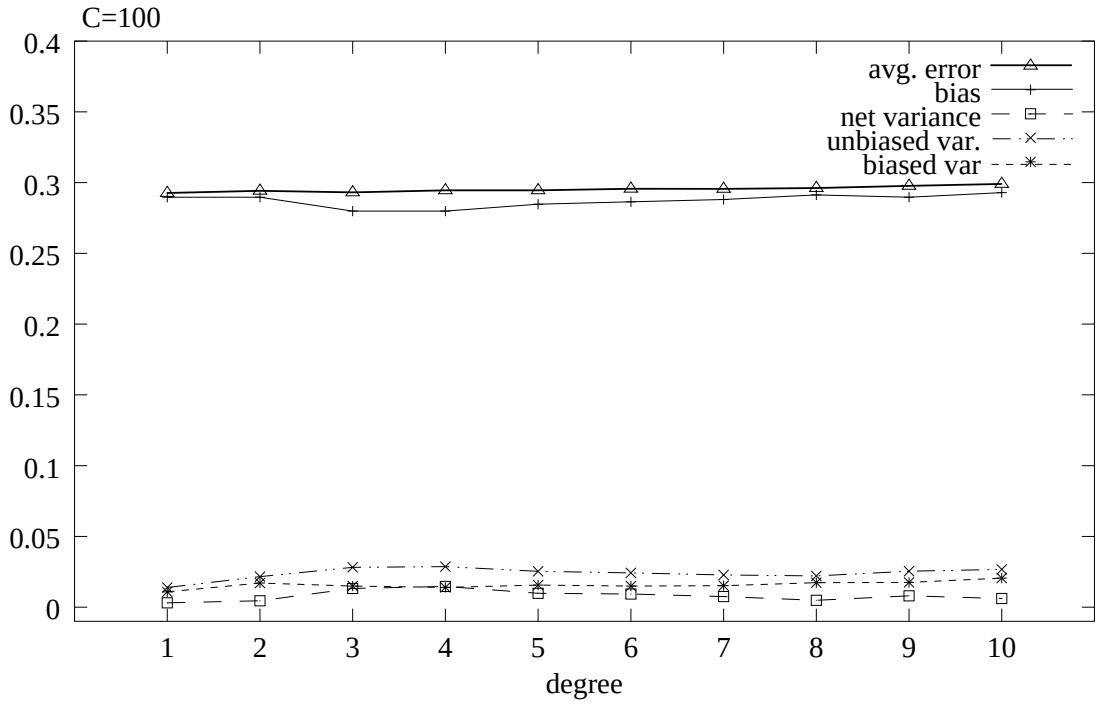


(a)

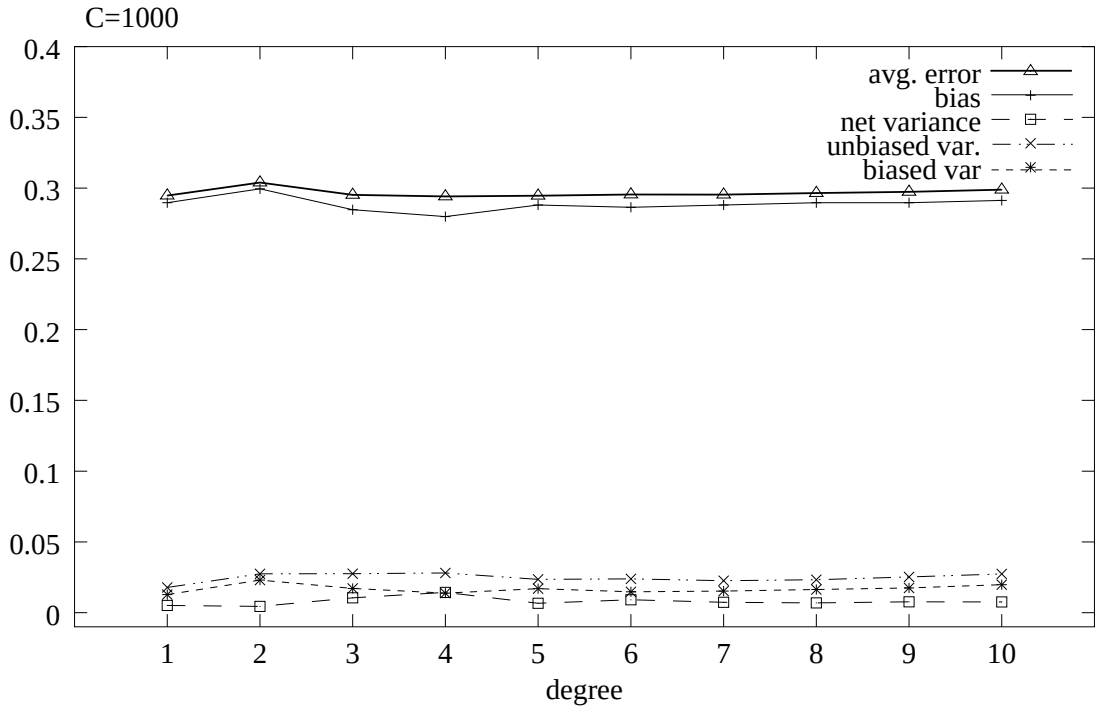


(b)

Figure 23: Letter-Two with noise data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 1$.

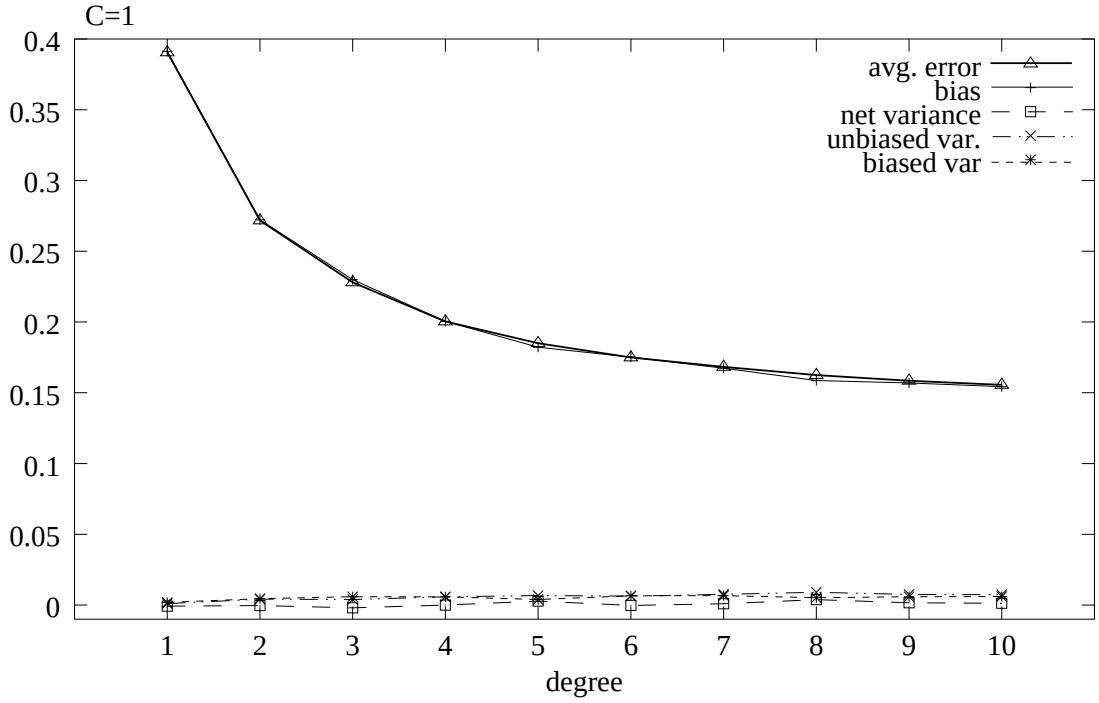


(a)

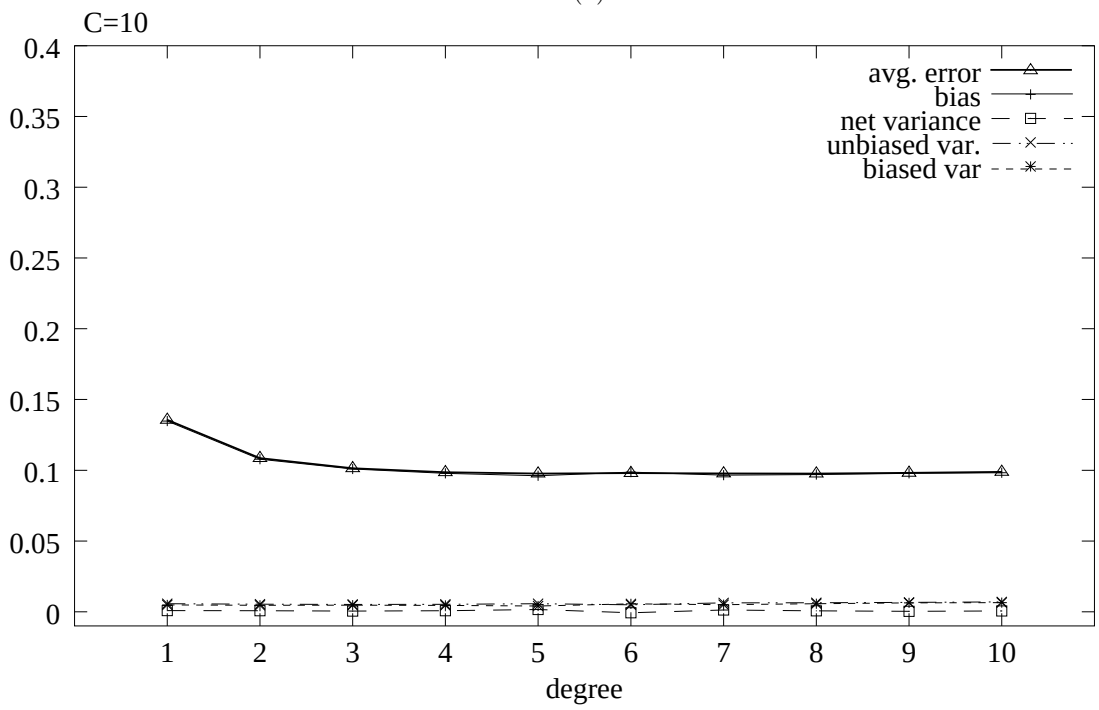


(b)

Figure 24: Letter-Two with noise data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 10$, (b) $C = 100$.

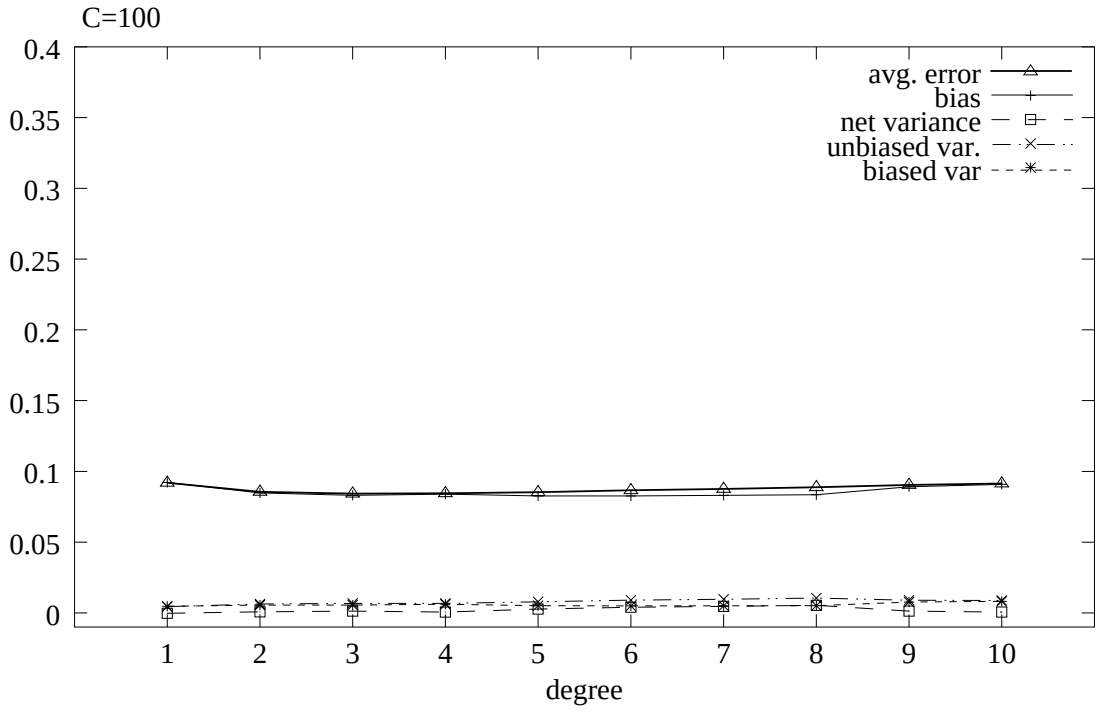


(a)

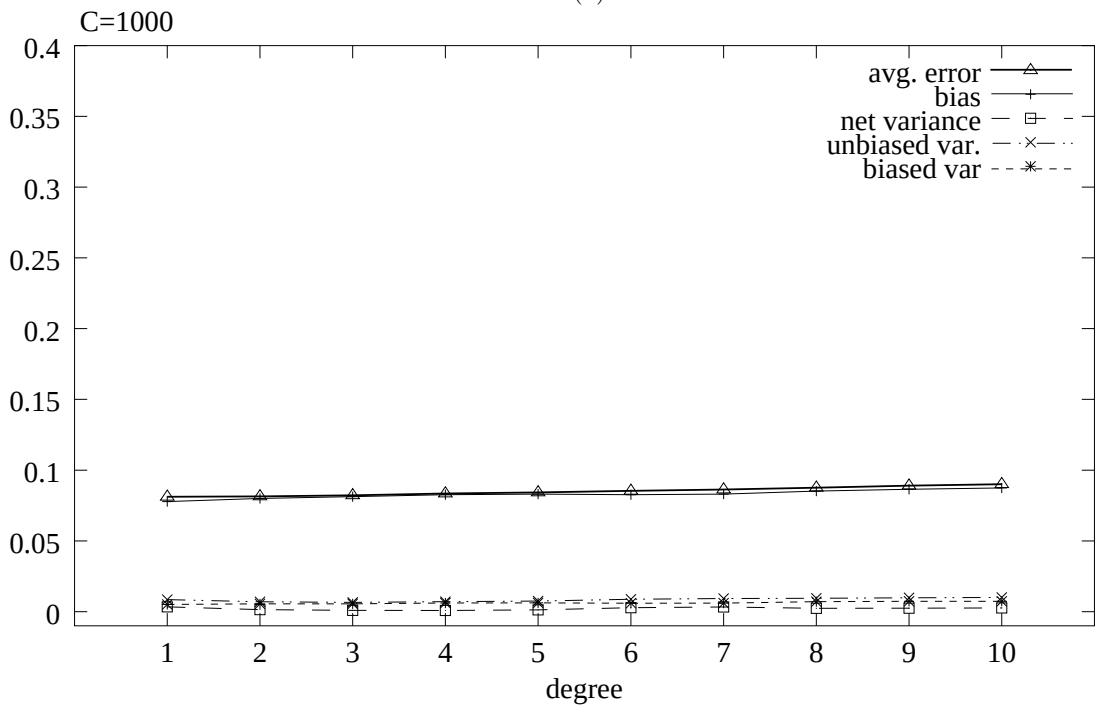


(b)

Figure 25: Spam data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 1$.

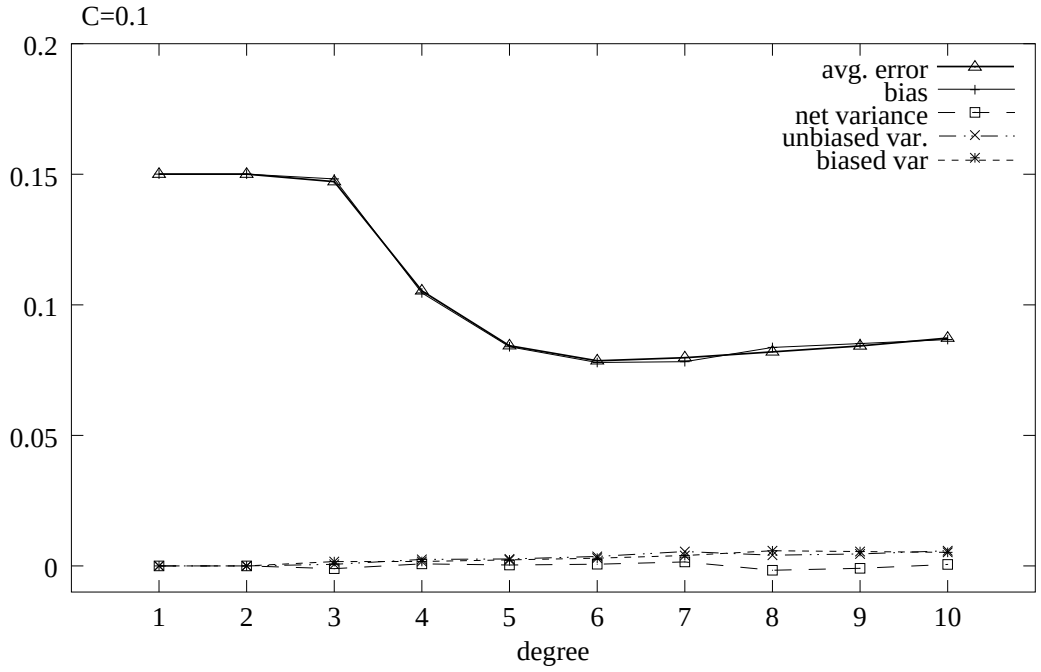


(a)

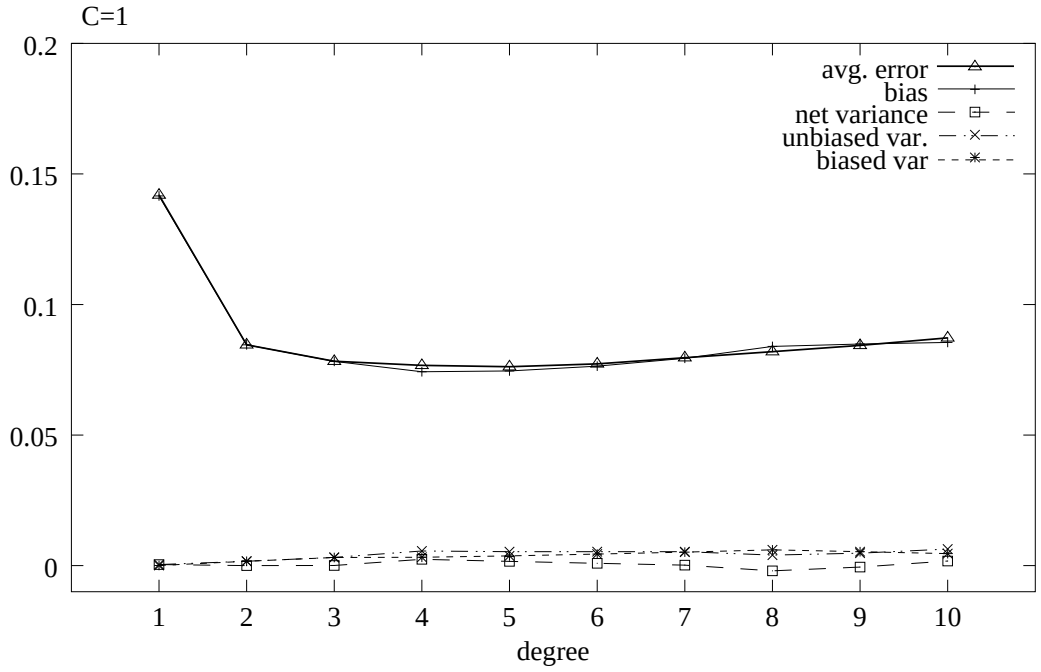


(b)

Figure 26: Spam data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 10$, (b) $C = 100$.

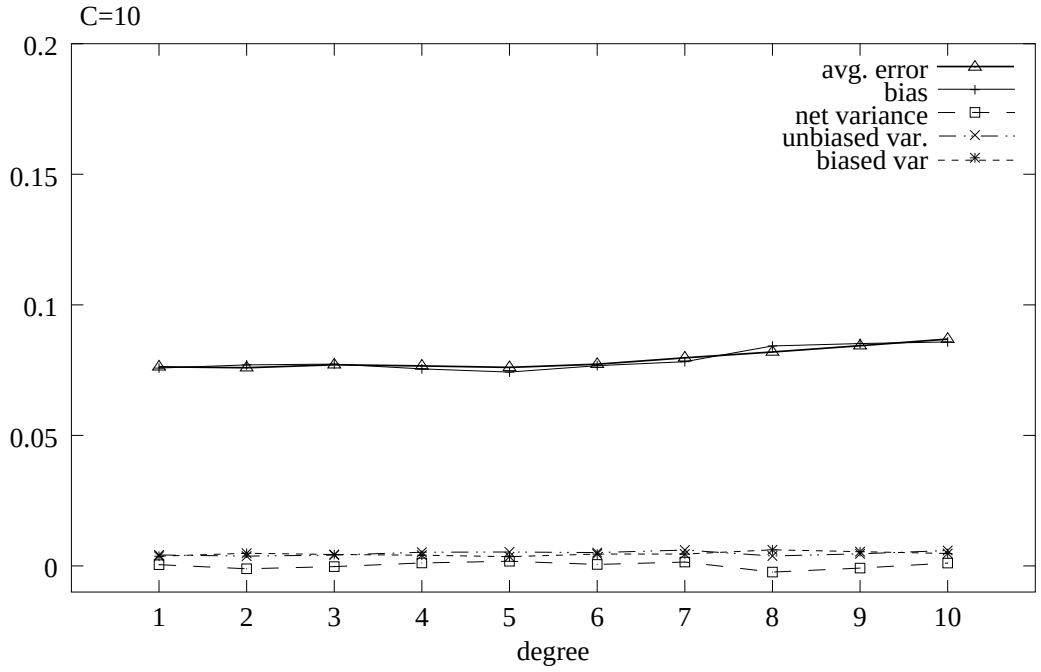


(a)

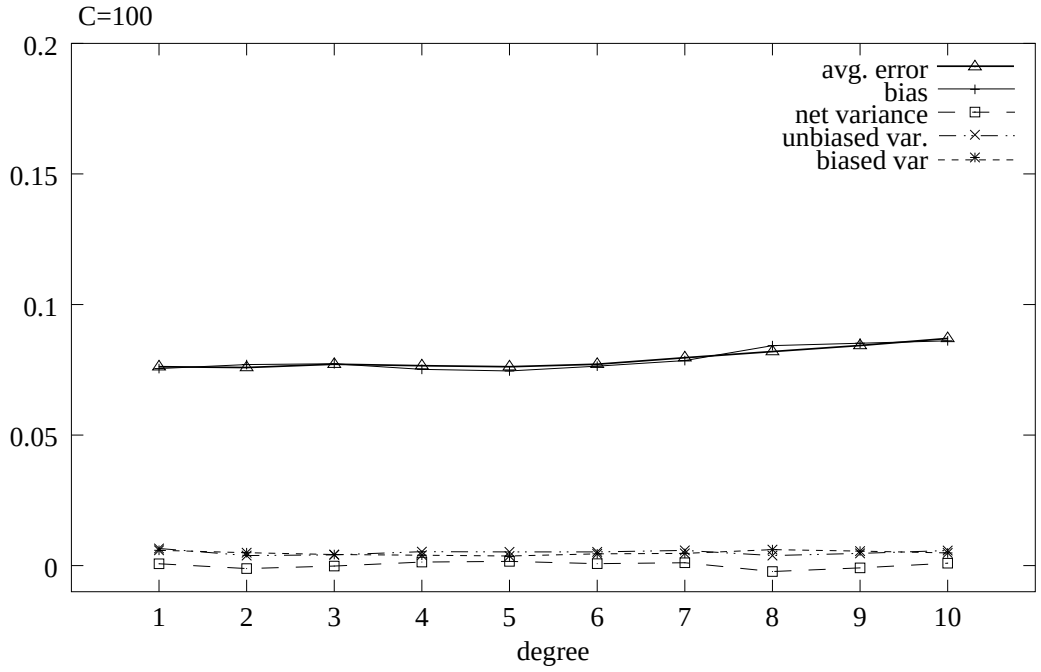


(b)

Figure 27: Musk data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 0.1$, (b) $C = 1$.



(a)



(b)

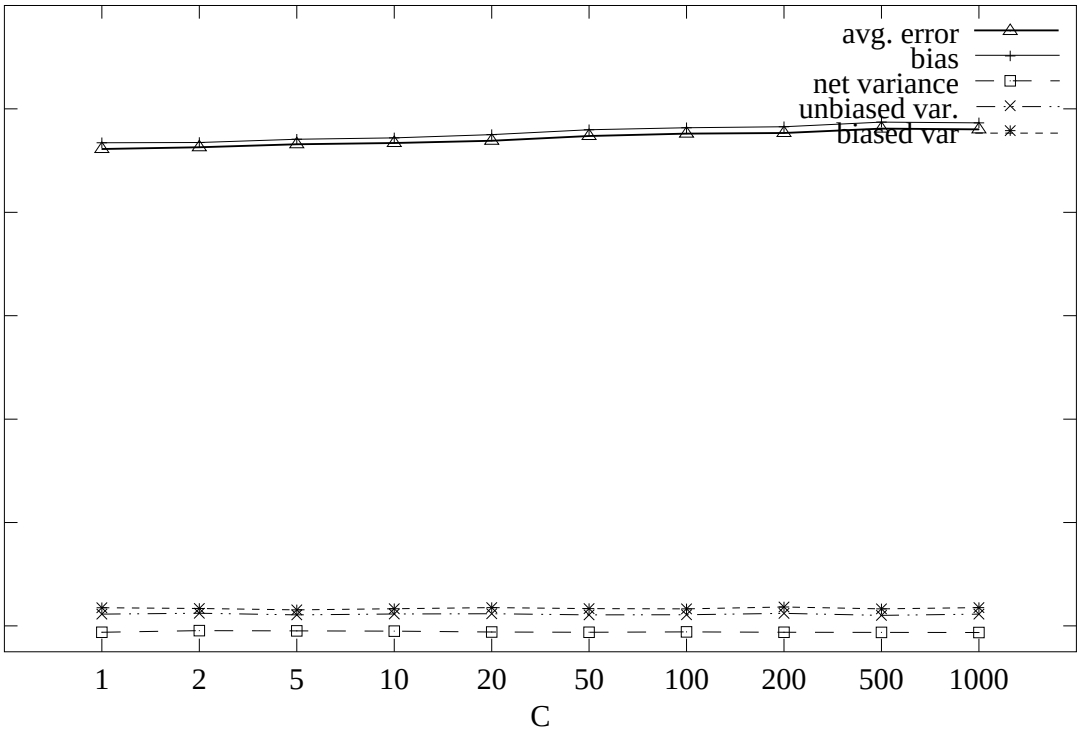
Figure 28: Musk data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA polynomial SVM, while varying the degree and for some fixed values of C: (a) $C = 10$, (b) $C = 100$.

2.3 Decomposition in RA Dot–product SVMs

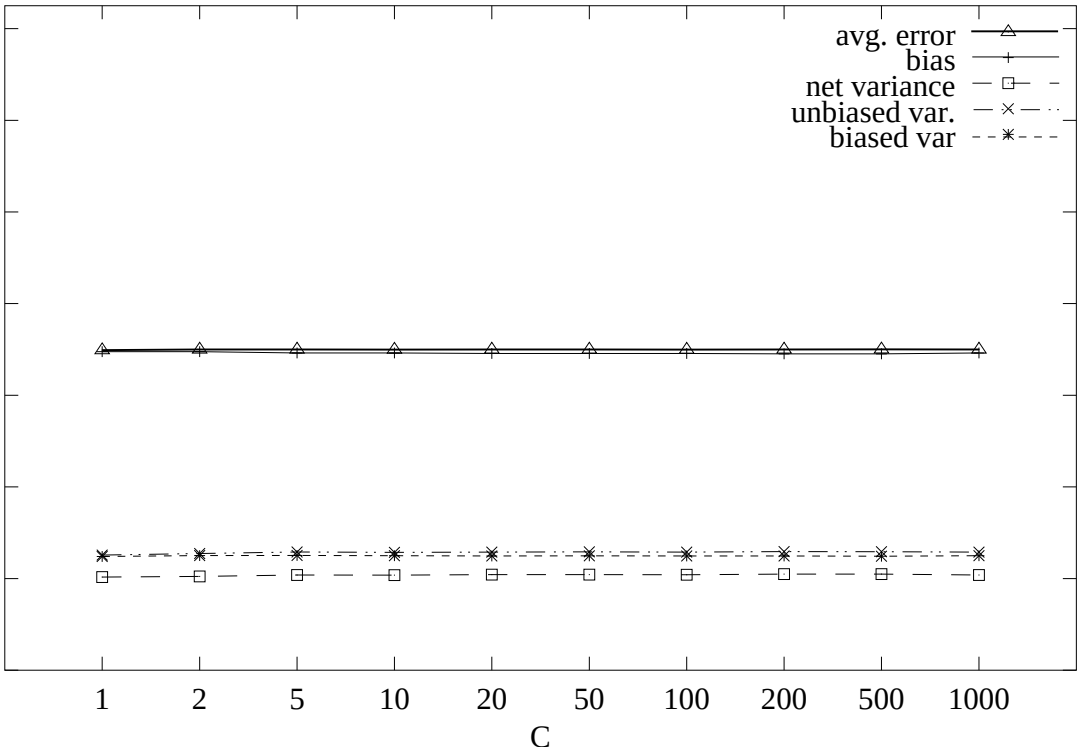
The decomposition of the error is represented with respect to different values of C.

Schematically we can observe the following facts:

- Net-variance is about 0 for all the values of C, as both biased and unbiased variance are very low close to 0.
- The error, bias and variance seem to be independent of the values of C. Anyway, note that in the experiments we used only values of $C \geq 1$.
- The error is determinated almost totally by the bias.

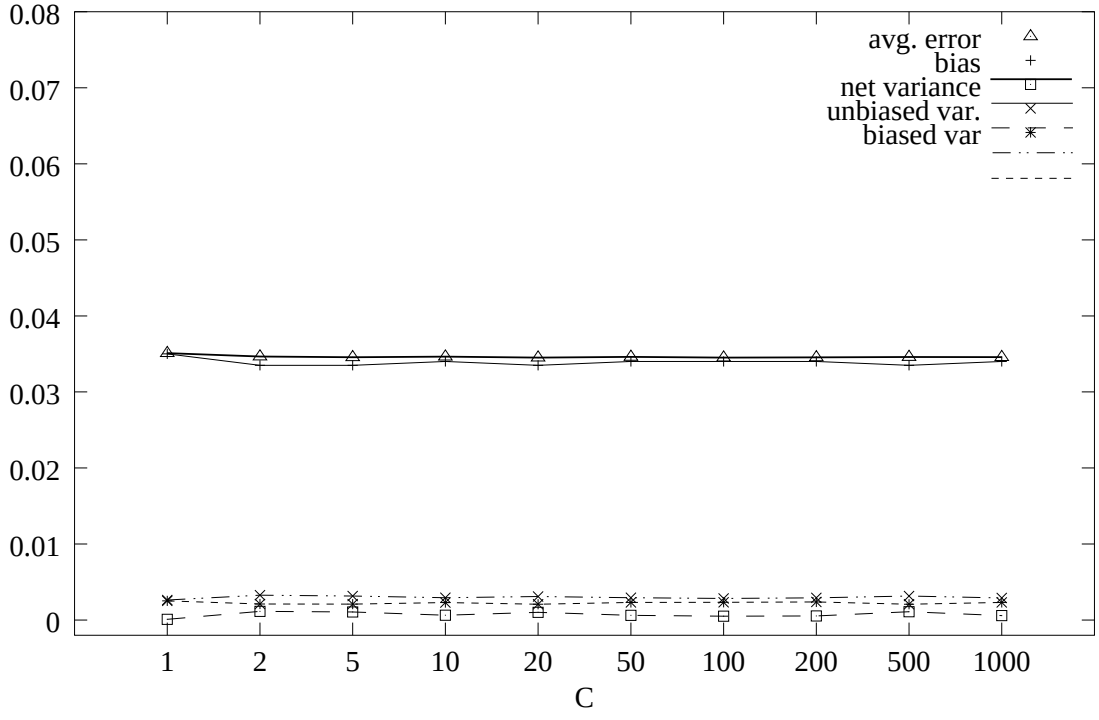


(a)

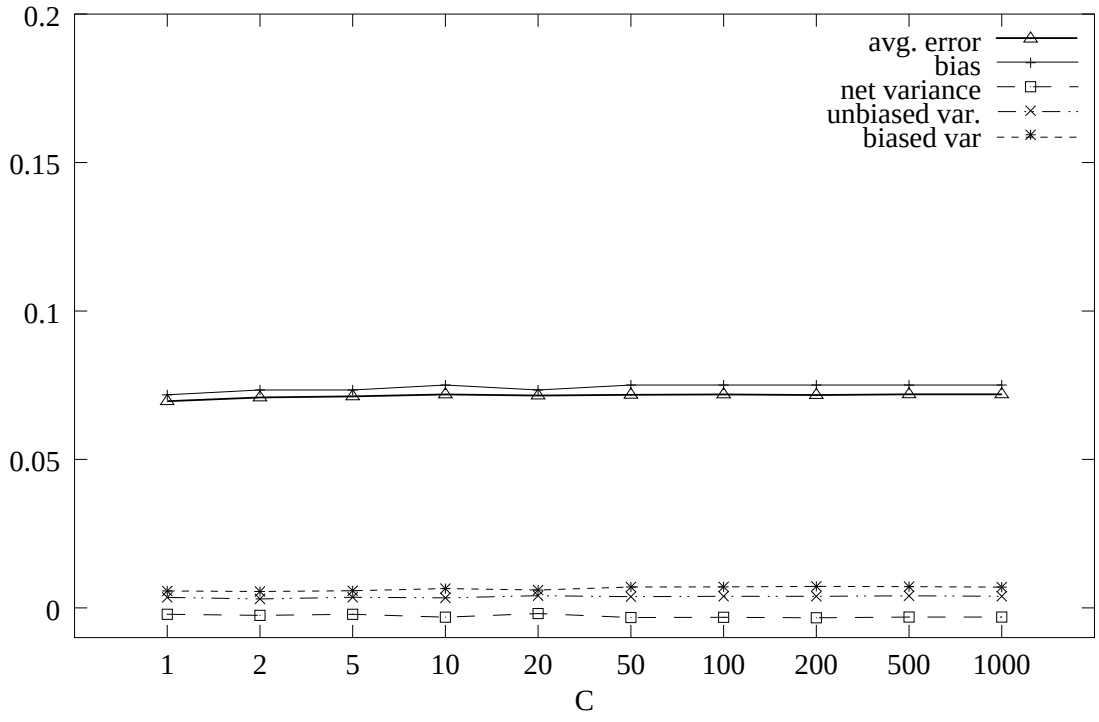


(b)

Figure 29: Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA dot-product SVM, while varying C: (a) P2, (b) Waveform.

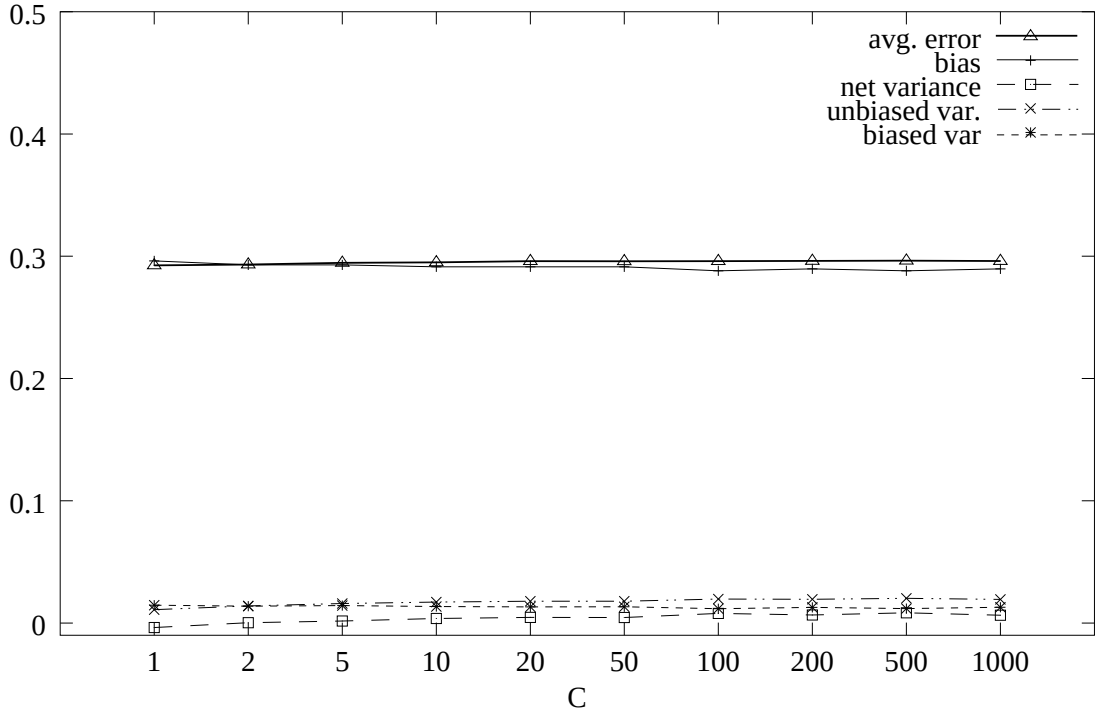


(a)

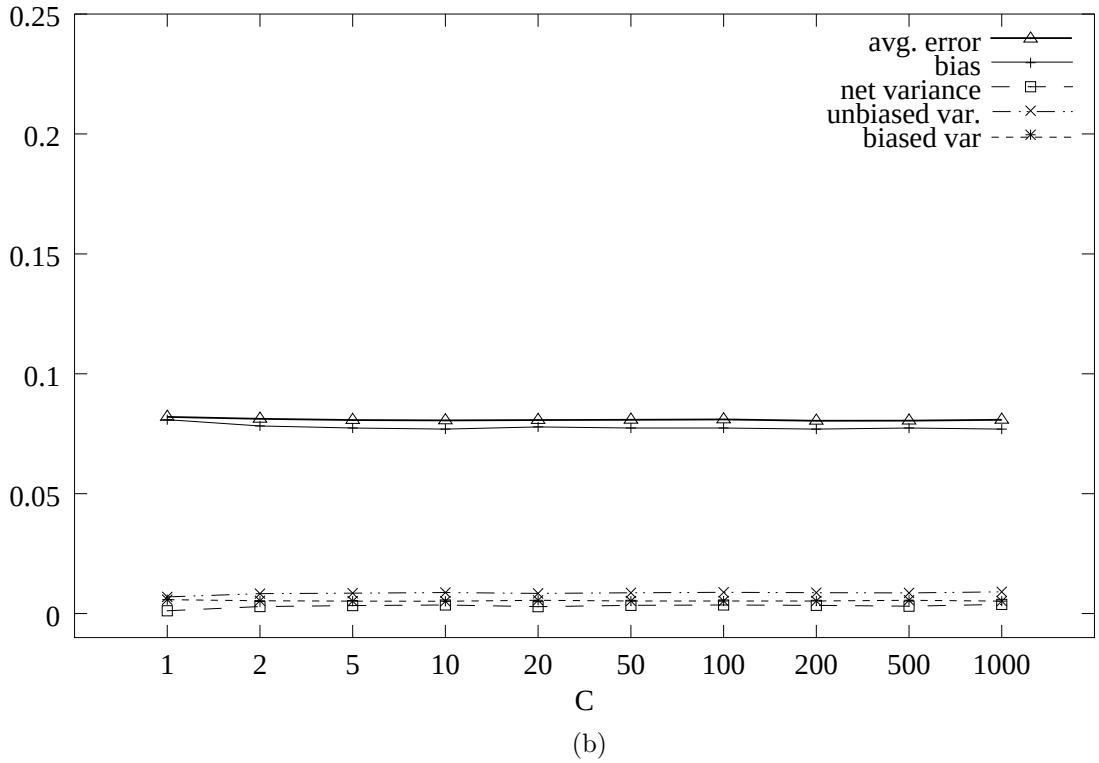


(b)

Figure 30: Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA dot-product SVM, while varying C: (a) Grey-Landsat, (b) Letter-Two.



(a)



(b)

Figure 31: Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA dot-product SVM, while varying C: (a) Letter-Two with noise, (b) Spam

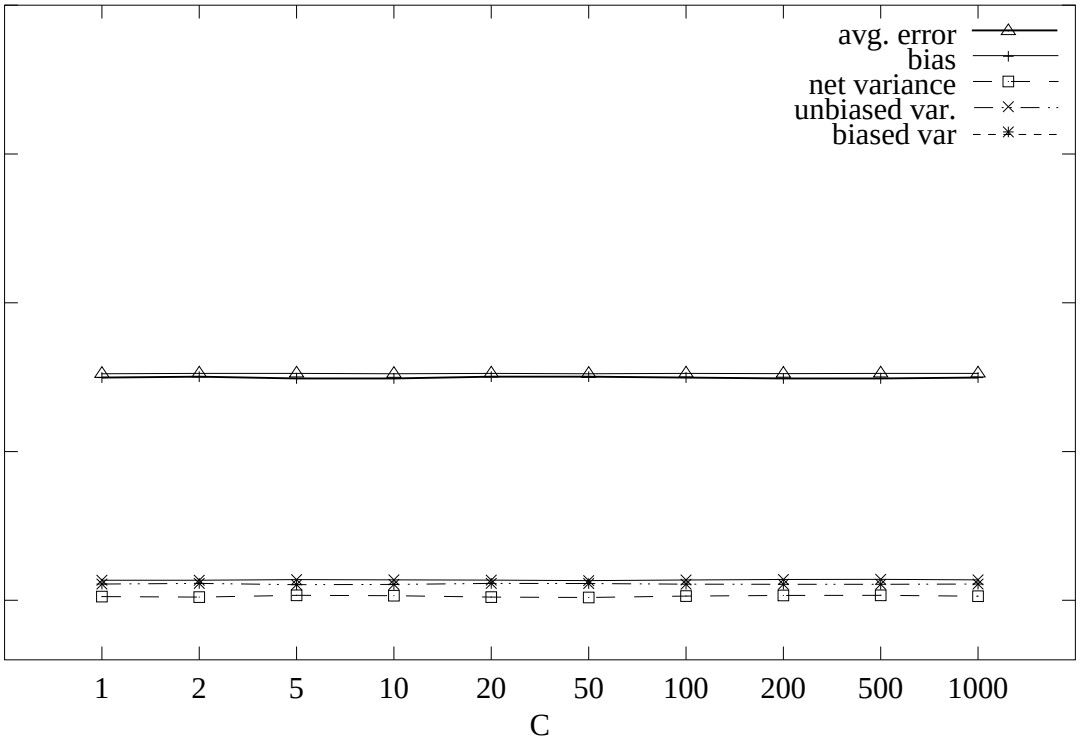


Figure 32: Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA dot-product SVM, while varying C with Musk data set

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 RA SVMs with respect to the number of base learners, that is the number of base learners used.

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

Most of the figures refer to RA SVMs of the "stabilized region". In these cases we can observe the following facts:

- Most of the decrement of the error occurs within the first iterations (from 10 to 30, depending on the data set).
- The bias remains unchanged during all the iterations
- The decrement of the error is almost entirely due to the decrement of the unbiased variance.

If we consider RA SVMs of the "transition region" we can consider the following observations:

- Most of the decrement of the error occurs within the first iterations (from 10 to 30, depending on the data set).
- The bias remains unchanged in average (higher than the bias of SVMs of the "stabilized region"), but oscillates largely, especially during the first 20 iterations
- The unbiased variance also oscillates, but tends to decrement with the iterations, lowering the error.
- The biased variance oscillates in the same way (that is with the same phase) with respect to the bias, but with a lower amplitude.
- The unbiased variance and in particular the net-variance oscillate in a specular way (opposite phase) with respect to the bias.

Note that there are no bias oscillations in the Spam data and in Letter-Two with noise data set in the transition region. Maybe it could be an effect of the noise (?).

In the musk data set (Fig. 45) the decrement of the error is limited to the first 5 iterations, and we have only relatively low variations of the error.

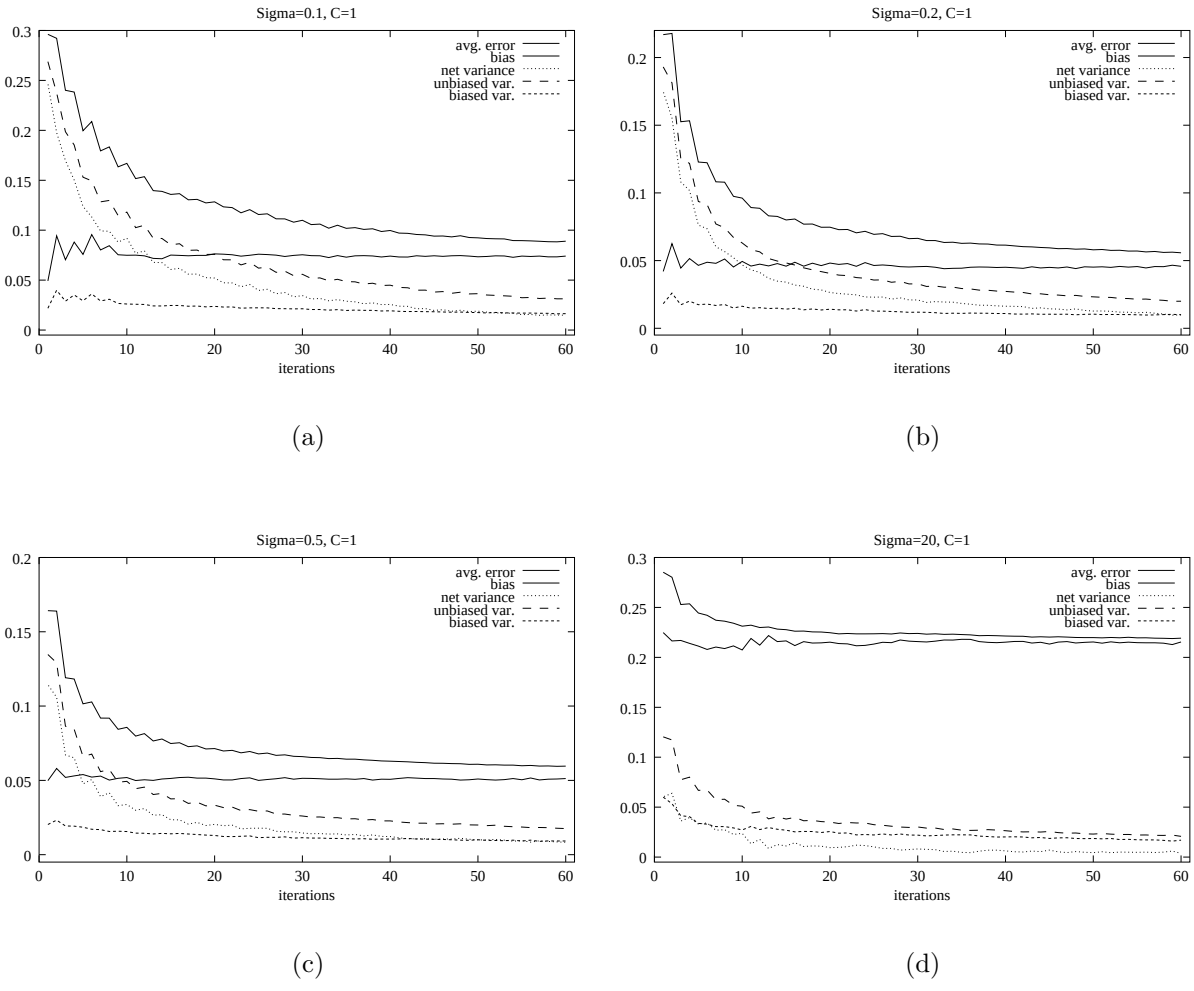


Figure 33: P2 data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF with respect to the number of iterations for some values of σ and $C = 1$: (a) $\sigma = 0.1$, (b) $\sigma = 0.2$, (c) $\sigma = 0.5$, (d) $\sigma = 20$.

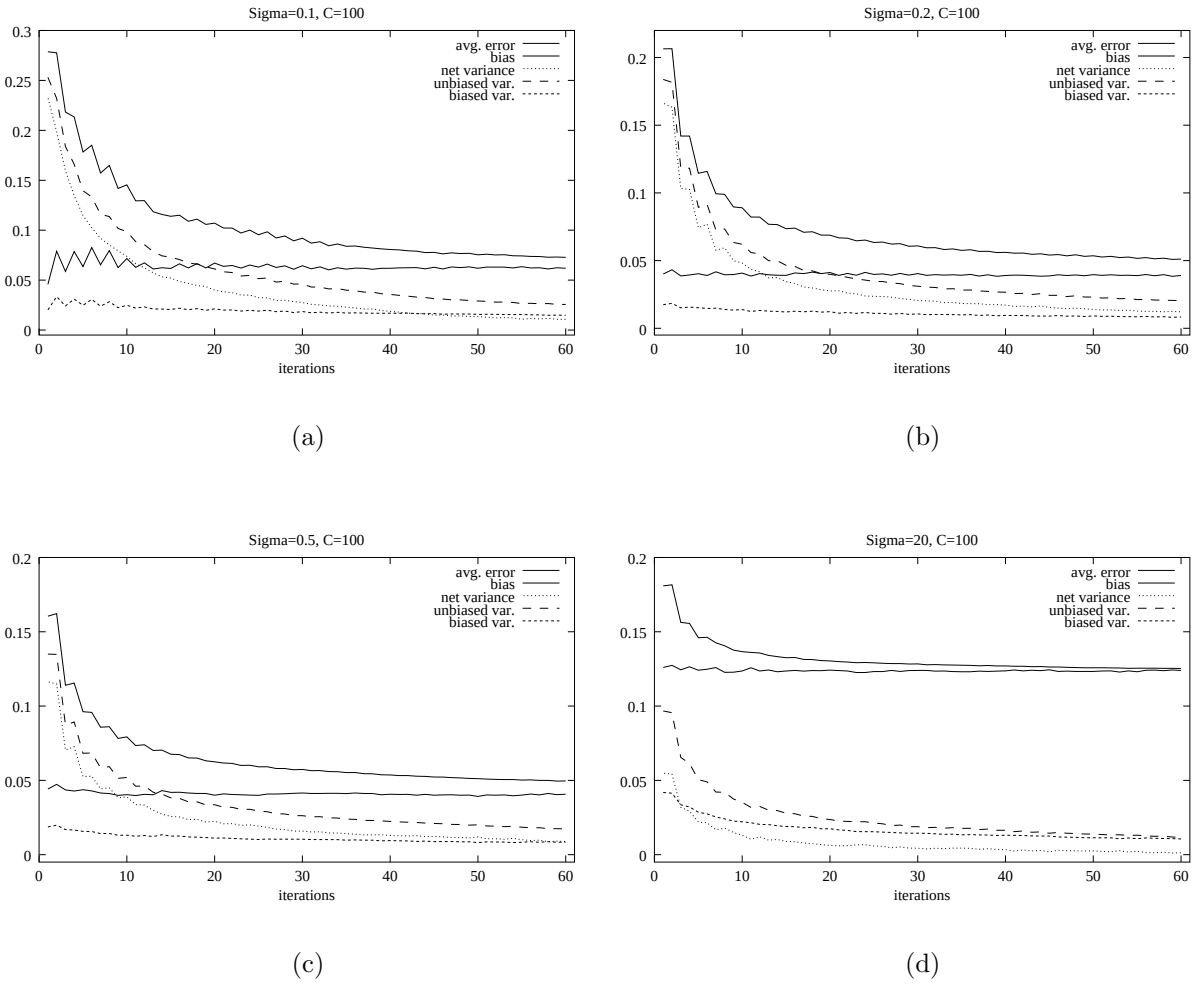


Figure 34: P2 data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF with respect to the number of iterations for some values of σ and $C = 100$: (a) $\sigma = 0.1$, (b) $\sigma = 0.2$, (c) $\sigma = 0.5$, (d) $\sigma = 20$.

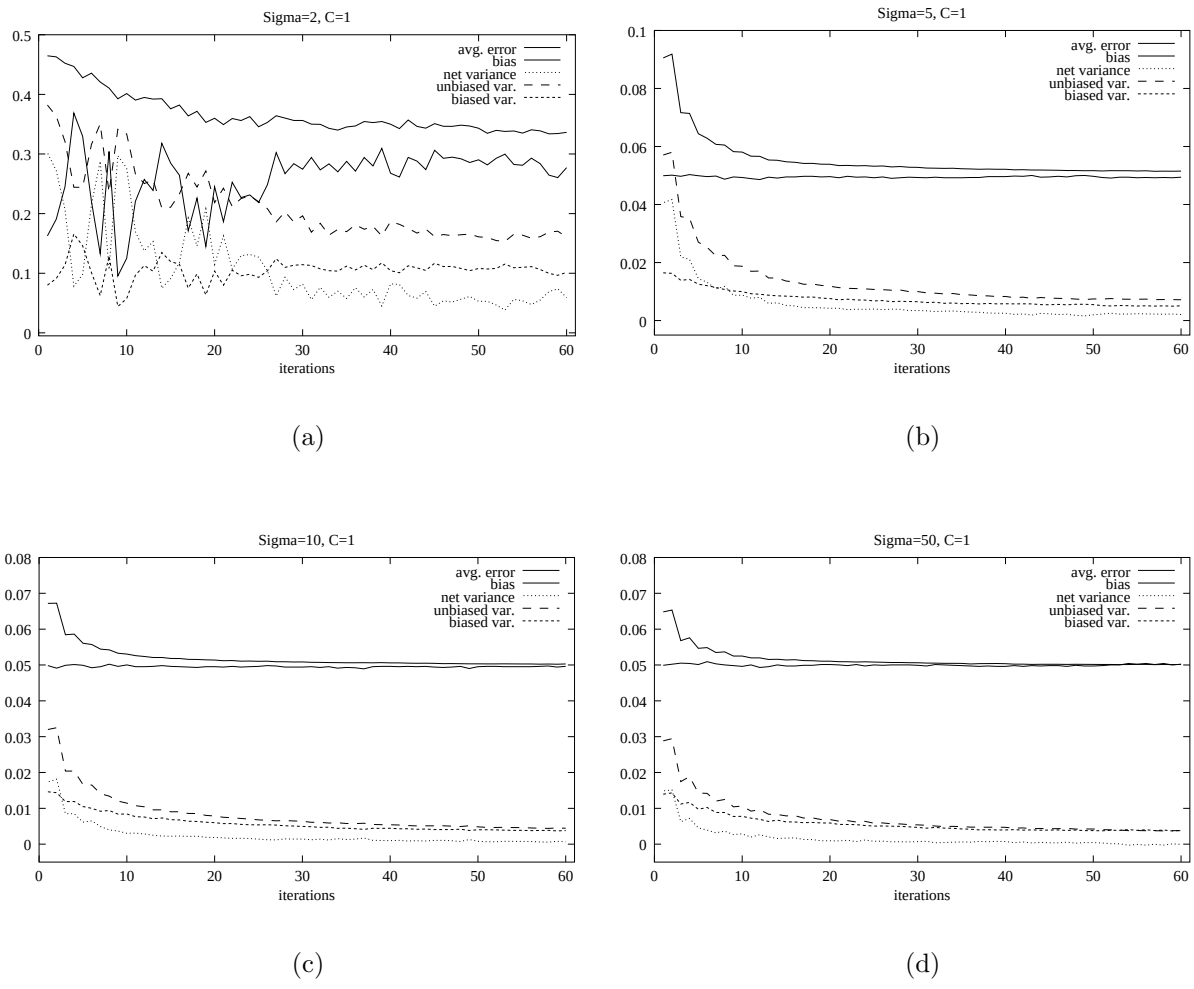


Figure 35: Waveform data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF with respect to the number of iterations for some values of σ and $C = 1$: (a) $\sigma = 2$, (b) $\sigma = 5$, (c) $\sigma = 10$, (d) $\sigma = 50$.

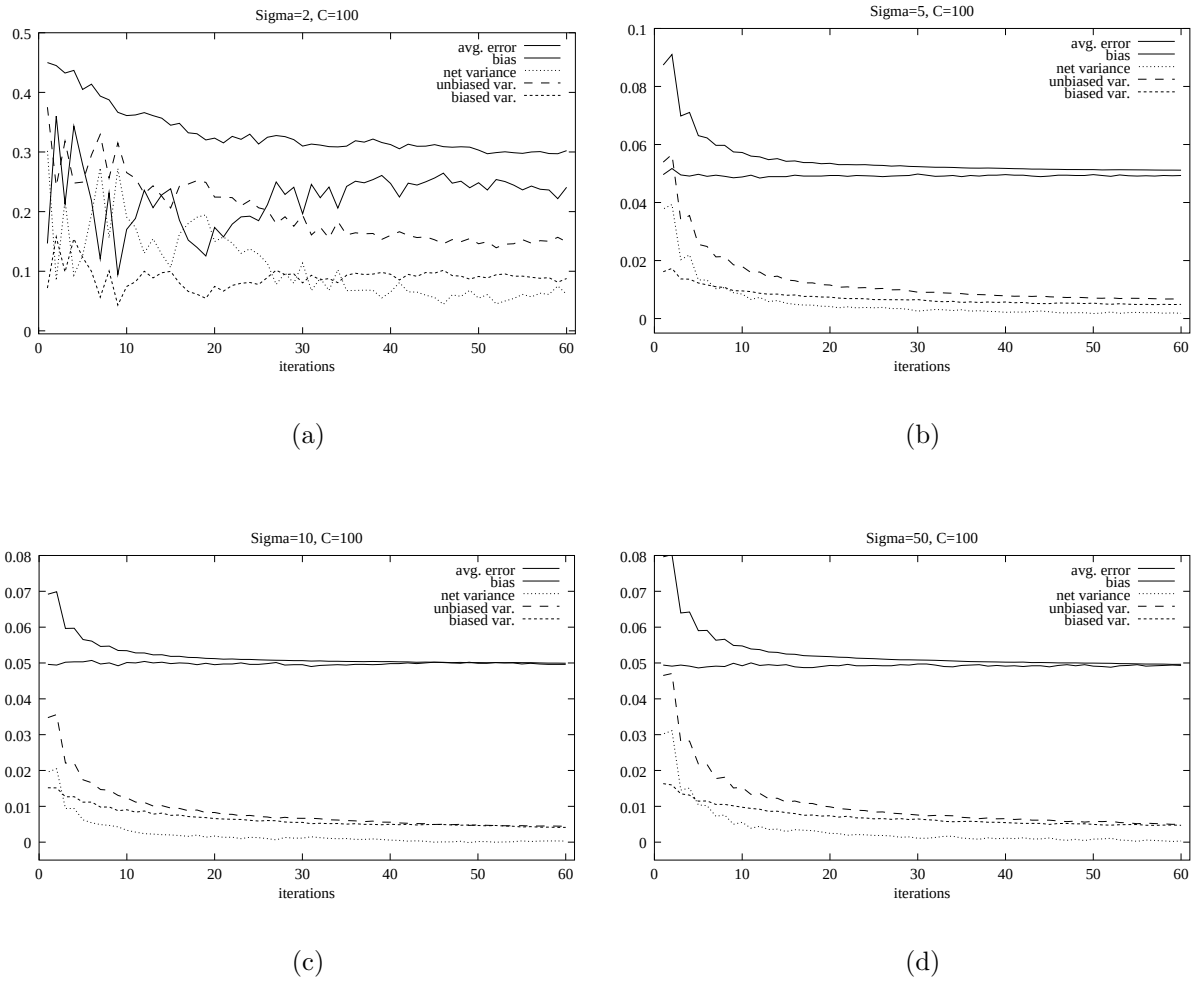


Figure 36: Waveform data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF with respect to the number of iterations for some values of σ and $C = 100$: (a) $\sigma = 2$, (b) $\sigma = 5$, (c) $\sigma = 10$, (d) $\sigma = 50$.

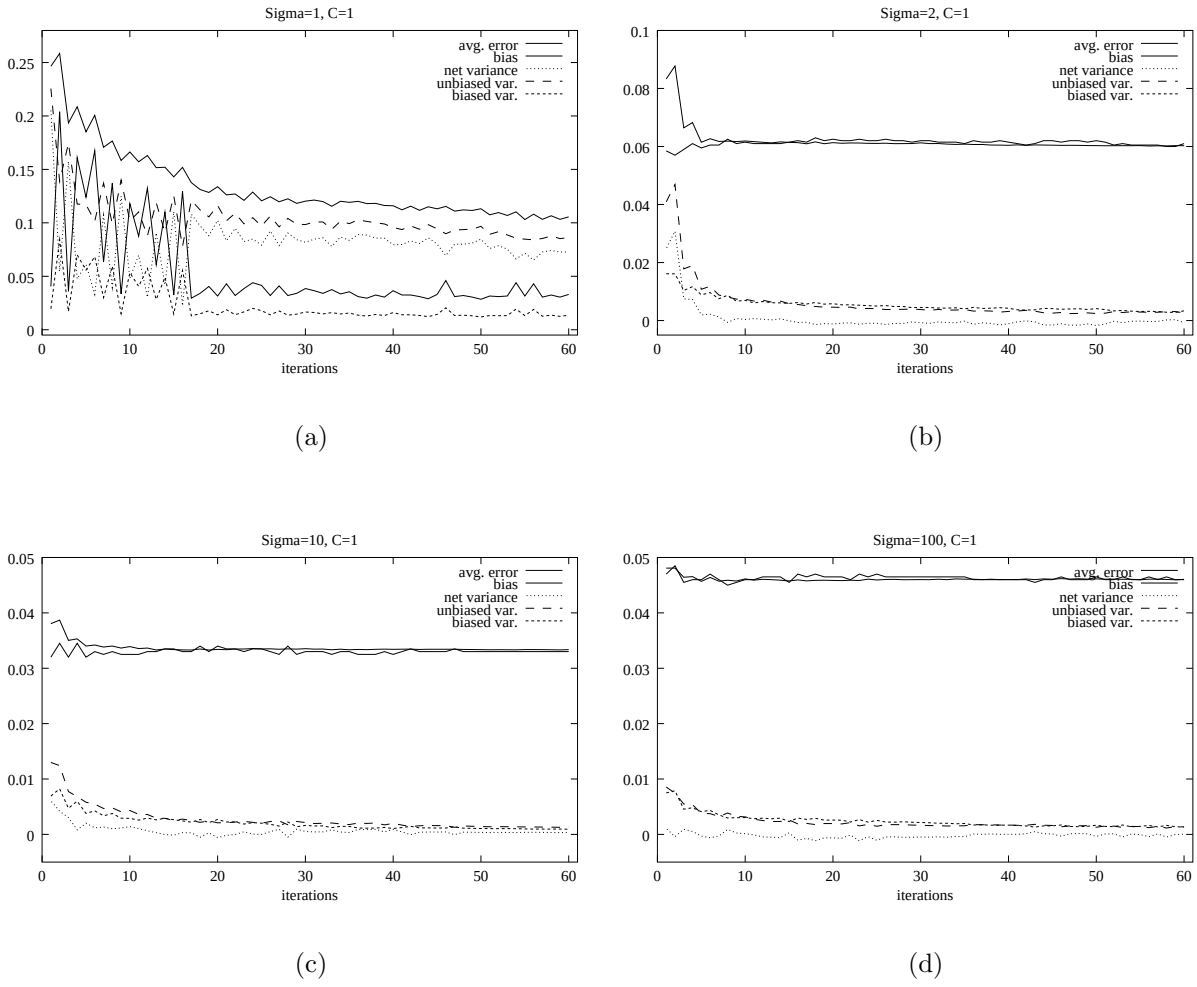


Figure 37: Grey-Landsat data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF with respect to the number of iterations for some values of σ and $C = 1$: (a) $\sigma = 1$, (b) $\sigma = 2$, (c) $\sigma = 10$, (d) $\sigma = 100$.

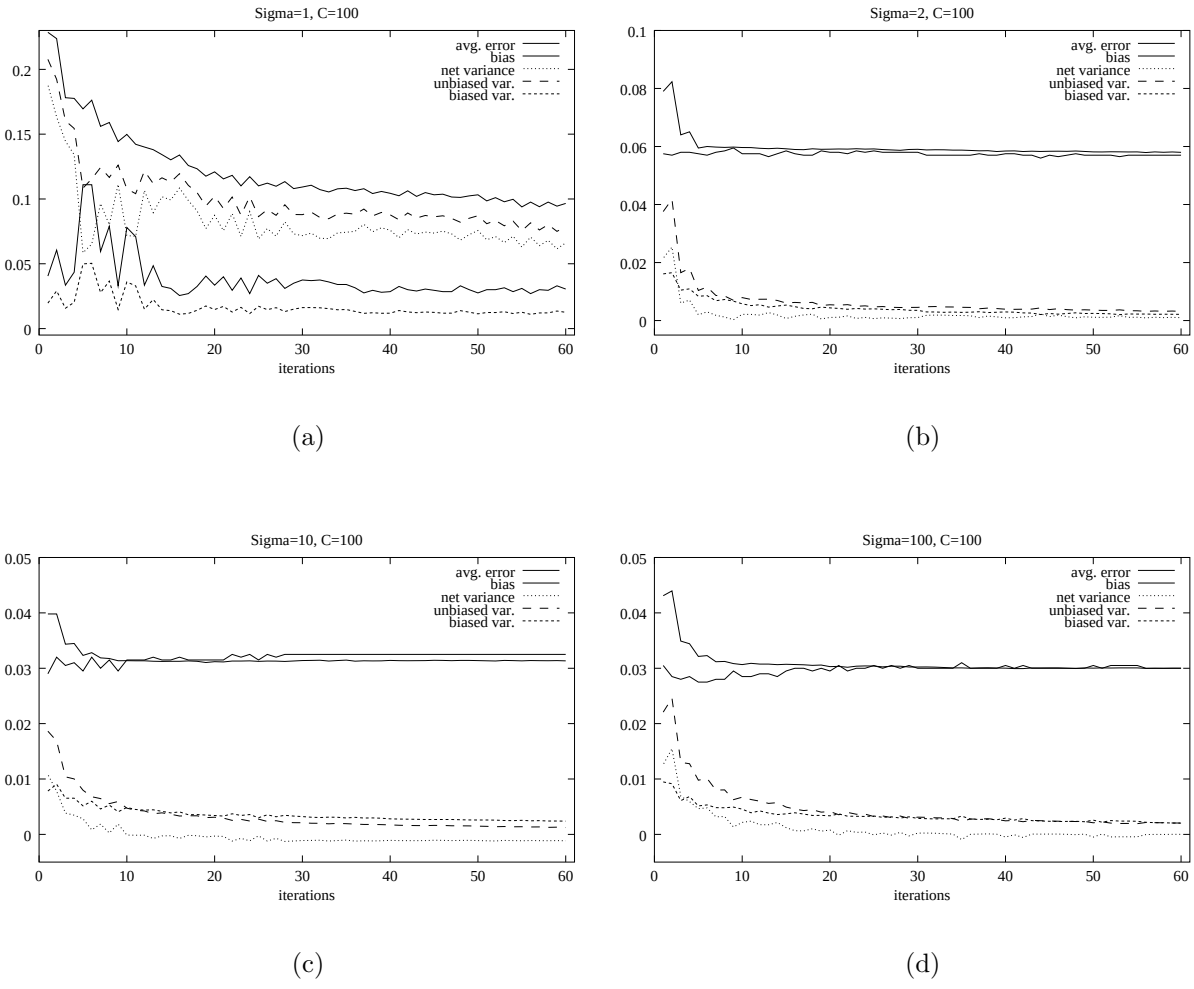


Figure 38: Grey-Landsat data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF with respect to the number of iterations for some values of σ and $C = 100$: (a) $\sigma = 1$, (b) $\sigma = 2$, (c) $\sigma = 10$, (d) $\sigma = 100$.

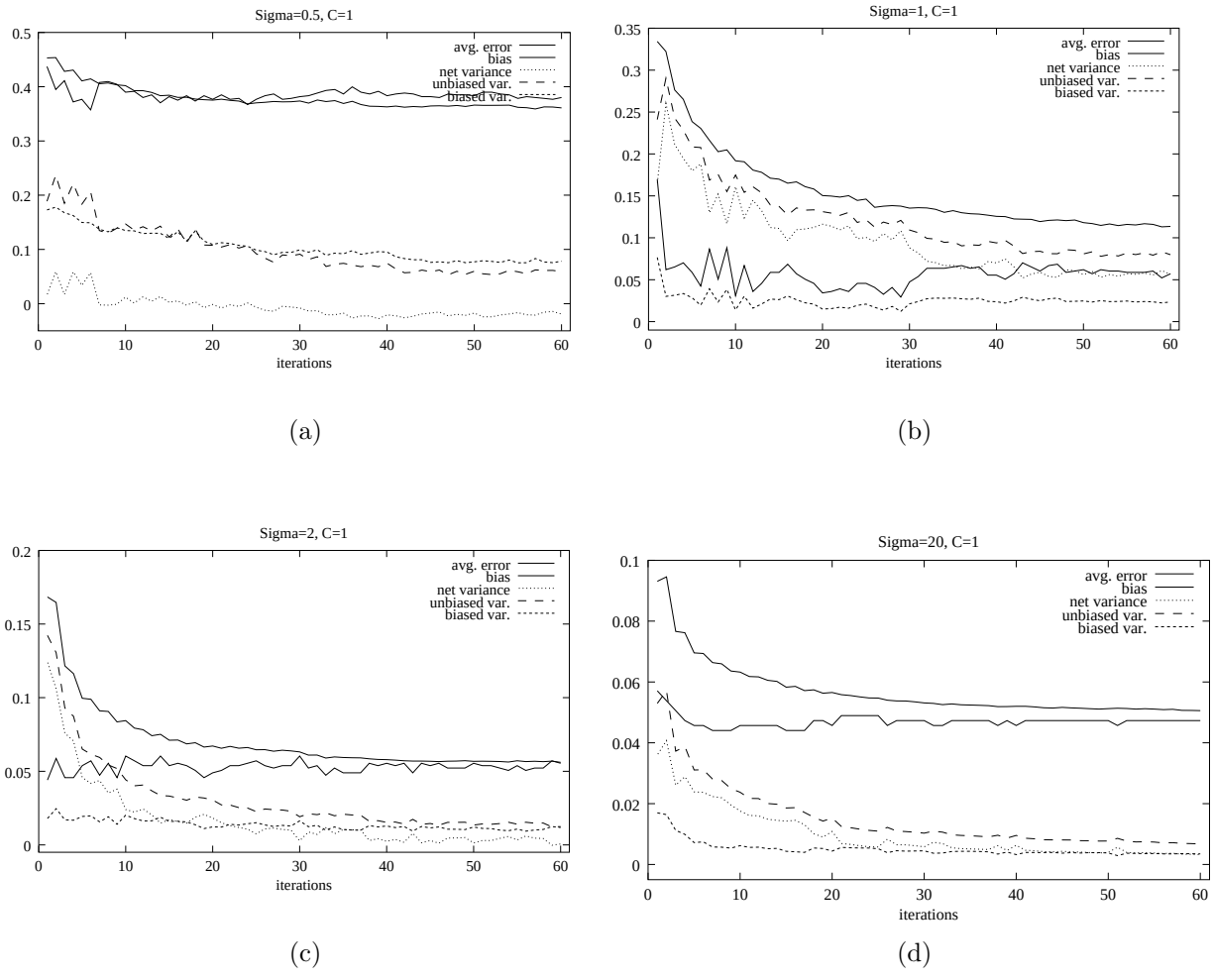


Figure 39: Letter-two data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF with respect to the number of iterations for some values of σ and $C = 1$: (a) $\sigma = 0.5$, (b) $\sigma = 1$, (c) $\sigma = 2$, (d) $\sigma = 20$.

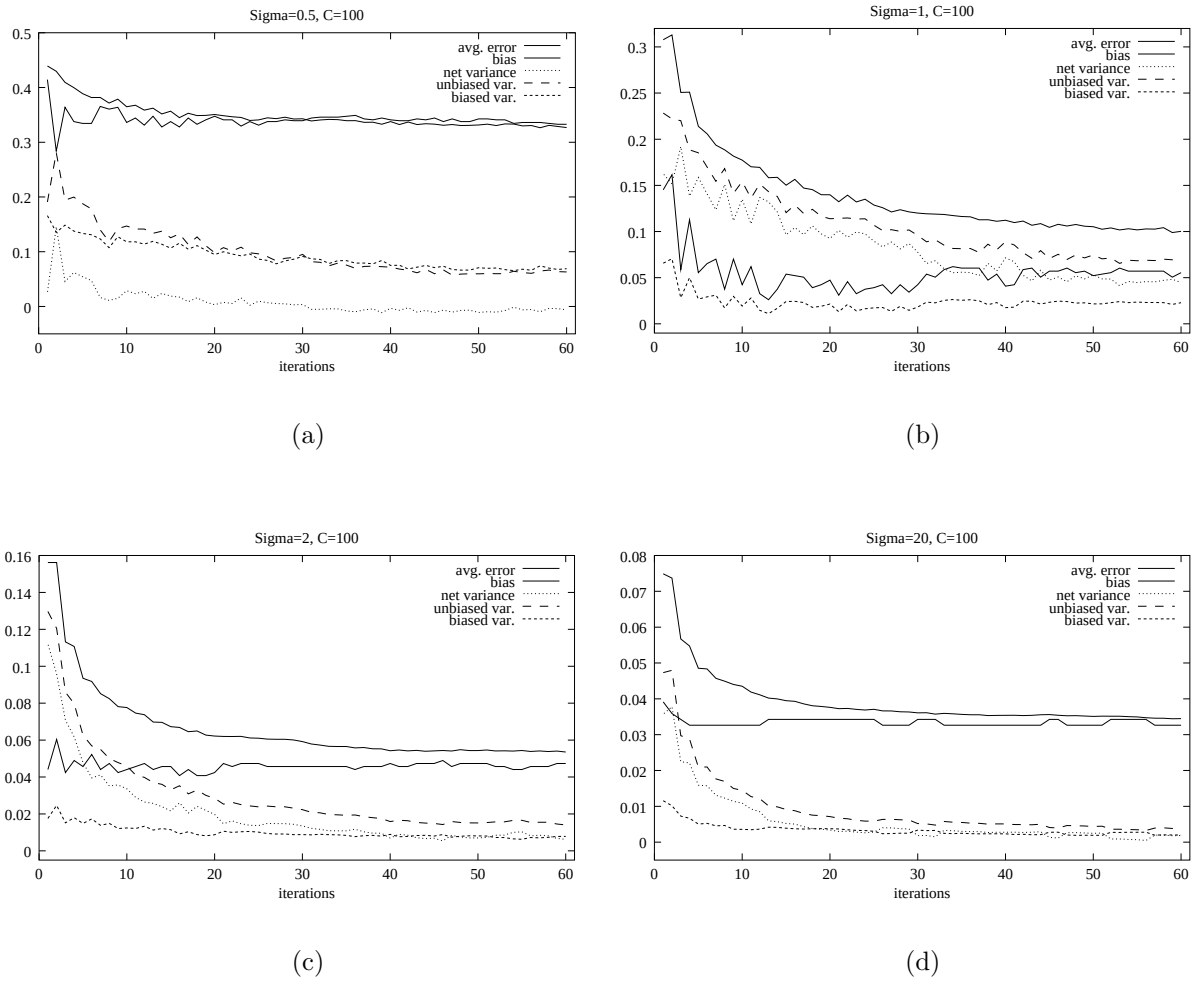


Figure 40: Letter-two data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF with respect to the number of iterations for some values of σ and $C = 100$: (a) $\sigma = 0.5$, (b) $\sigma = 1$, (c) $\sigma = 2$, (d) $\sigma = 20$.

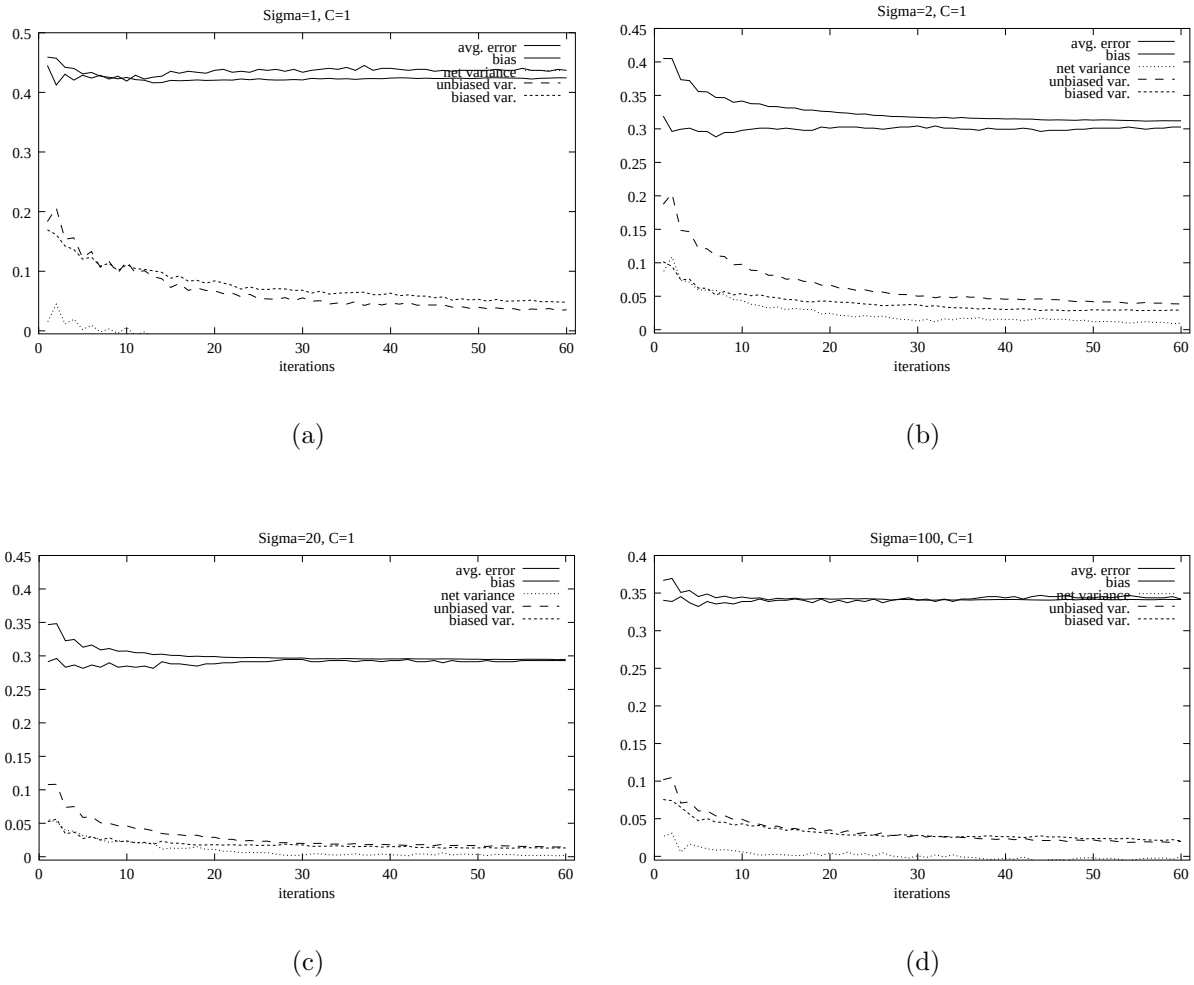


Figure 41: Letter-two with noise data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF with respect to the number of iterations for some values of σ and $C = 1$: (a) $\sigma = 1$, (b) $\sigma = 2$, (c) $\sigma = 20$, (d) $\sigma = 100$.

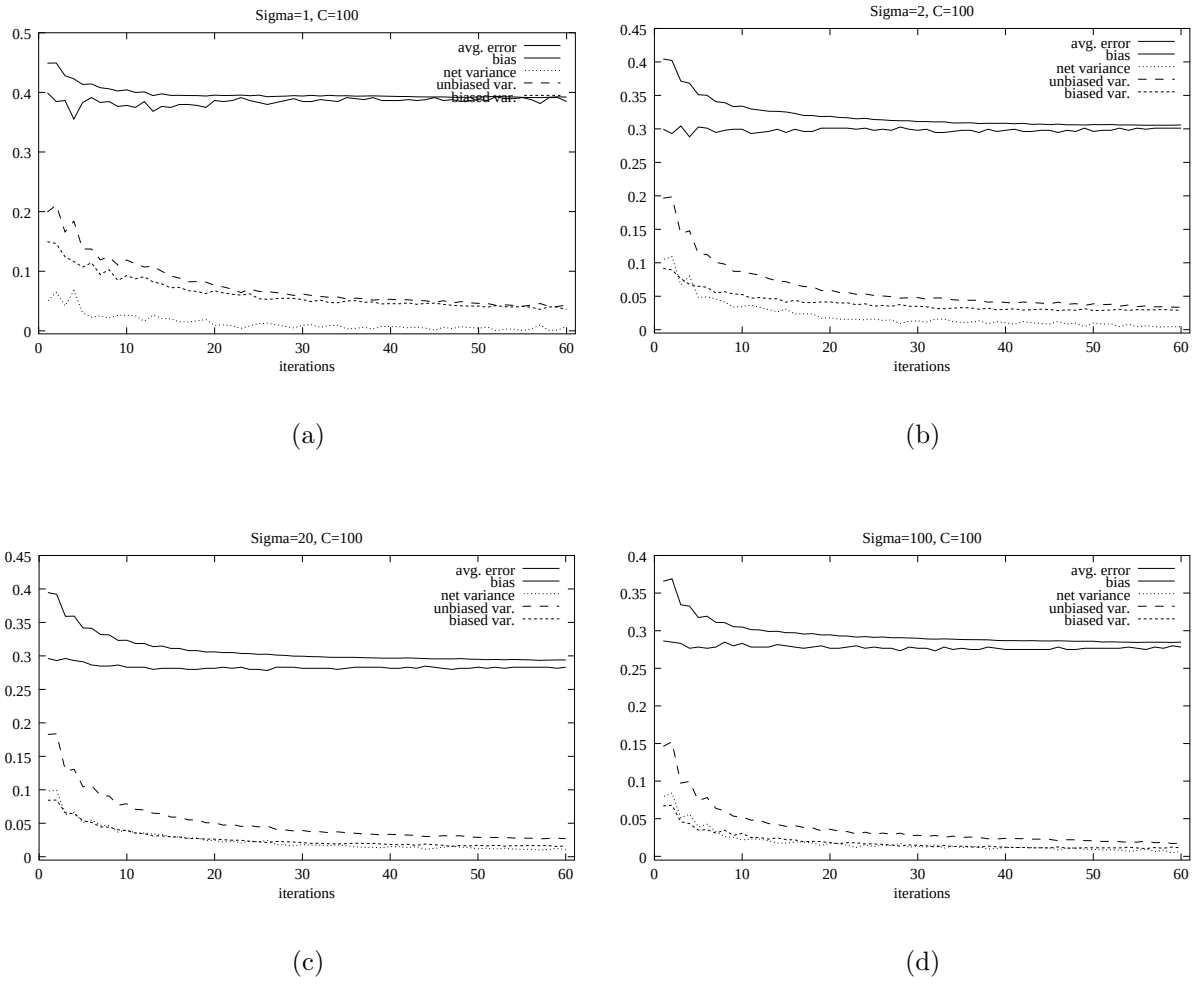


Figure 42: Letter-two with noise data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF with respect to the number of iterations for some values of σ and $C = 100$: (a) $\sigma = 1$, (b) $\sigma = 2$, (c) $\sigma = 20$, (d) $\sigma = 100$.

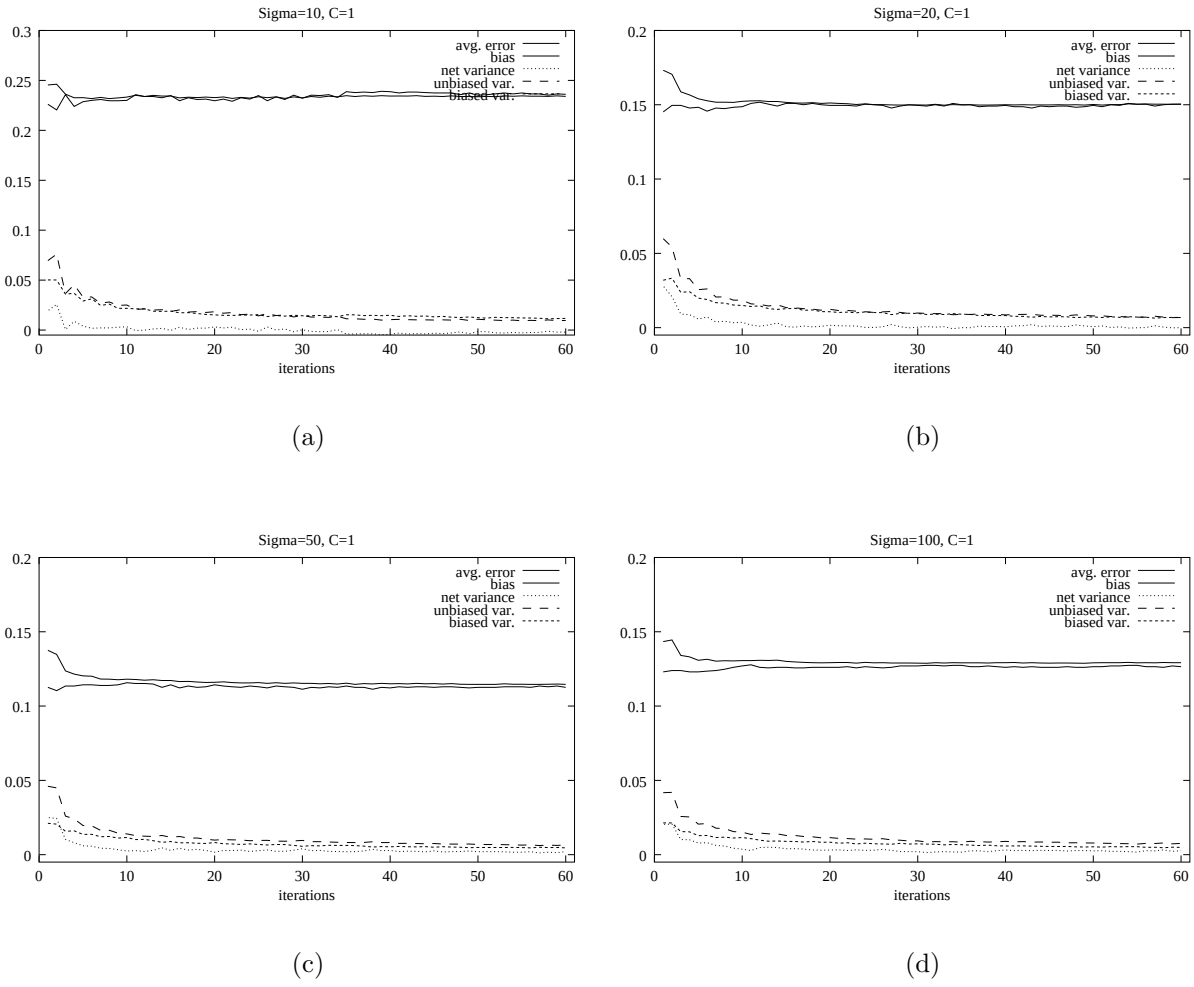


Figure 43: Spam data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF with respect to the number of iterations for some values of σ and $C = 1$: (a) $\sigma = 10$, (b) $\sigma = 20$, (c) $\sigma = 50$, (d) $\sigma = 100$.

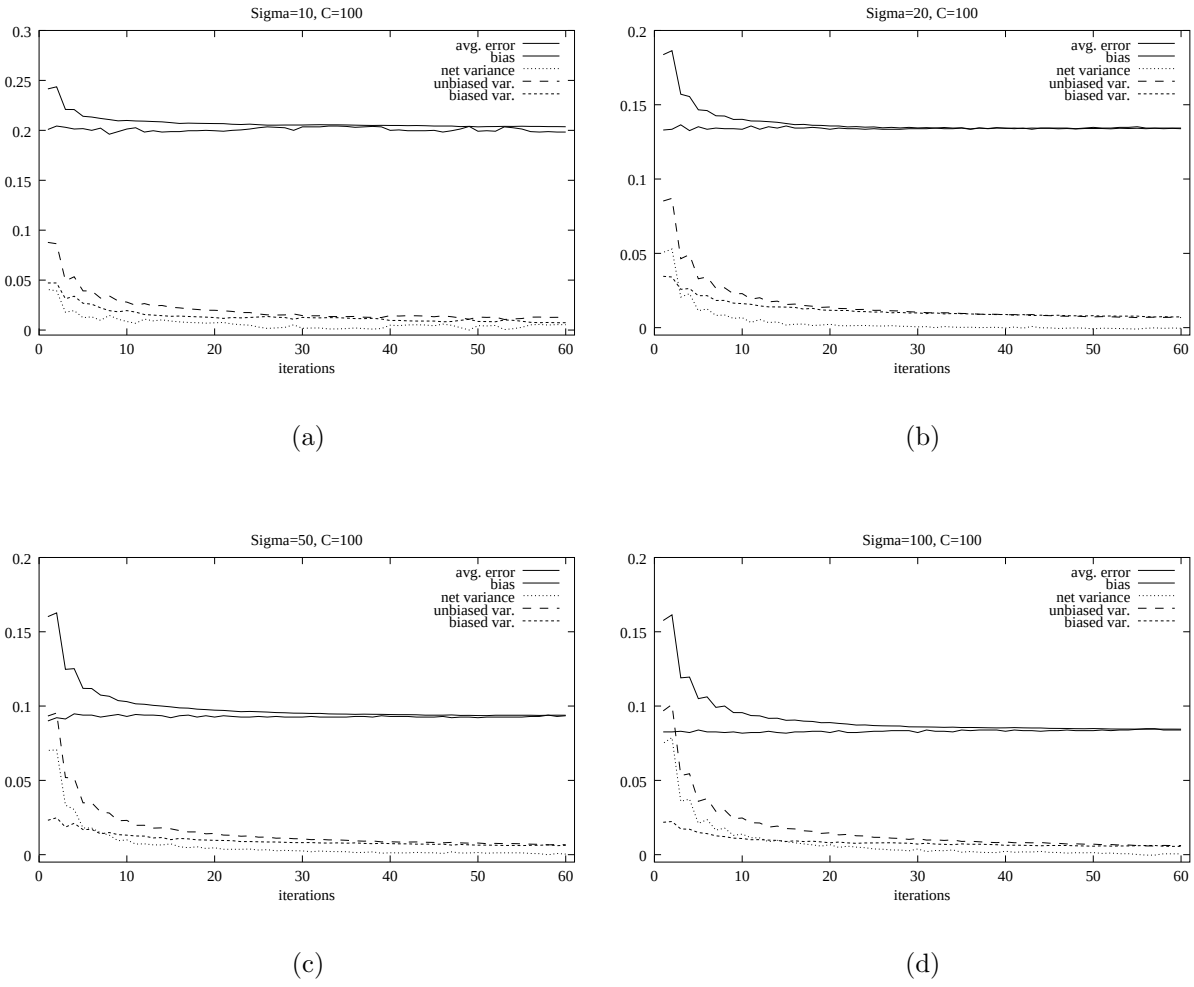


Figure 44: Spam data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF with respect to the number of iterations for some values of σ and $C = 100$: (a) $\sigma = 10$, (b) $\sigma = 20$, (c) $\sigma = 50$, (d) $\sigma = 100$.

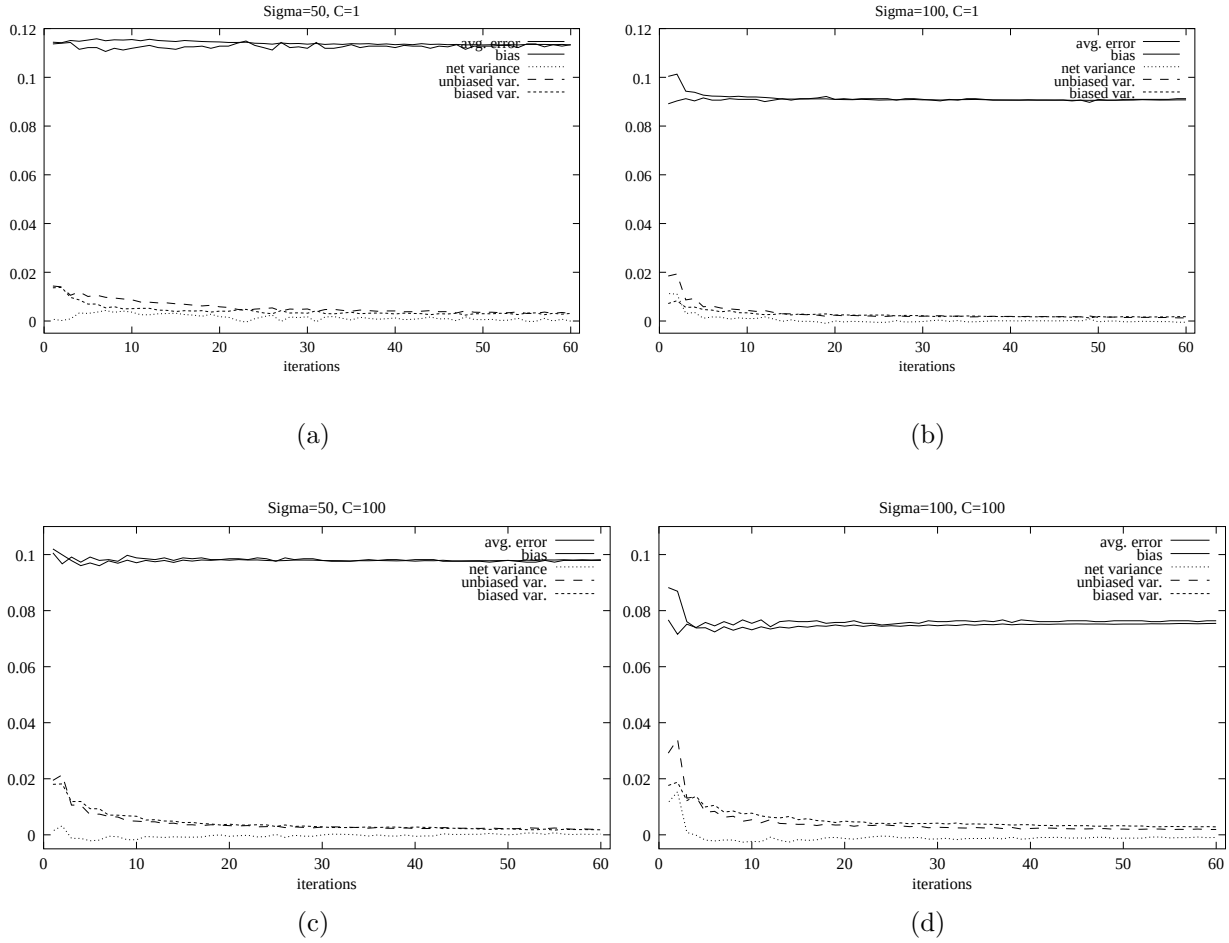


Figure 45: Musk data set. Bias-variance decomposition of error in bias, net variance, unbiased and biased variance in RA SVM RBF with respect to the number of iterations for some values of σ and C: (a) $\sigma = 50$ and $C = 1$, (b) $\sigma = 100$ and $C = 1$, (c) $\sigma = 50$ and $C = 100$, (d) $\sigma = 100$ and $C = 100$.

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

Schematically we can observe the following facts:

- The bias remains constant
- Most of the error decrement is achieved within the first 10-20 iterations
- Error decrement is due to the decrement of the unbiased variance
- The error is determinated almost totally by the bias.
- In some cases for low degree polynomials (e.g. Fig. 46 (a)) we have bias oscillations when the eror is relatively high.
- The biased variance is low and slowly decreases at each iteration
- The unbiased variance continues to decrease at each iteration, but most of its decrement occurs within the first 20 iterations.

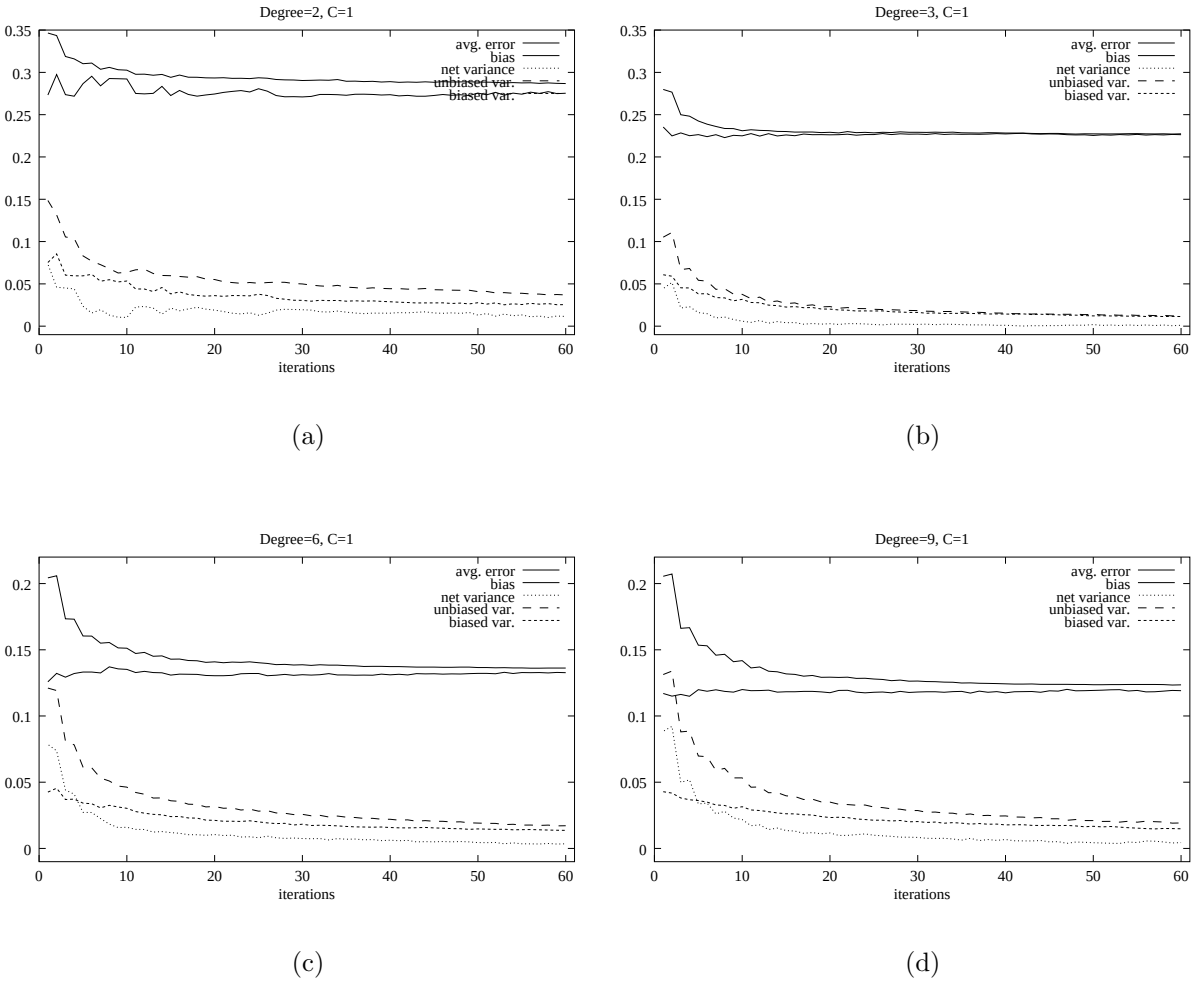


Figure 46: P2 data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 1$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

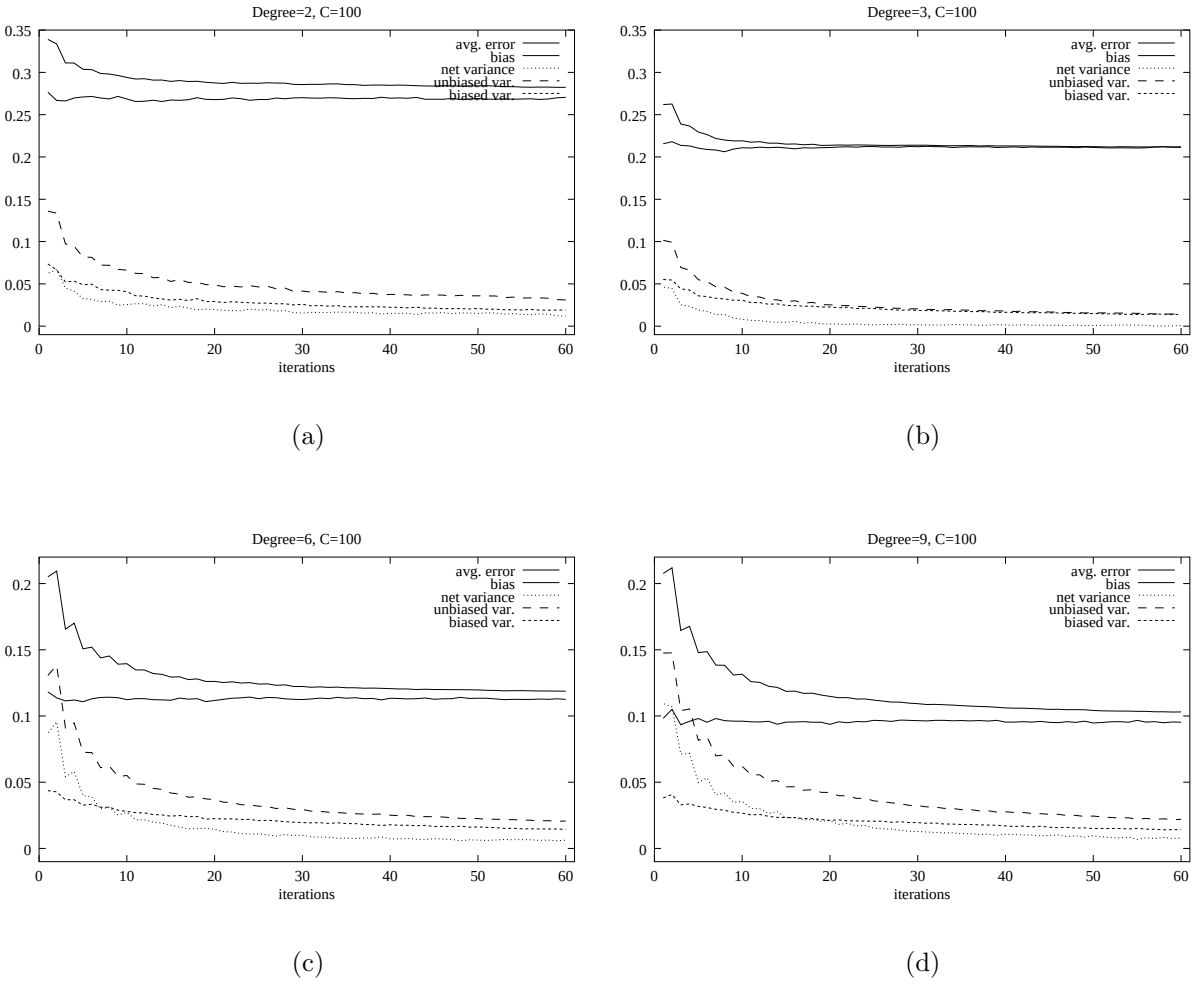


Figure 47: P2 data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 100$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

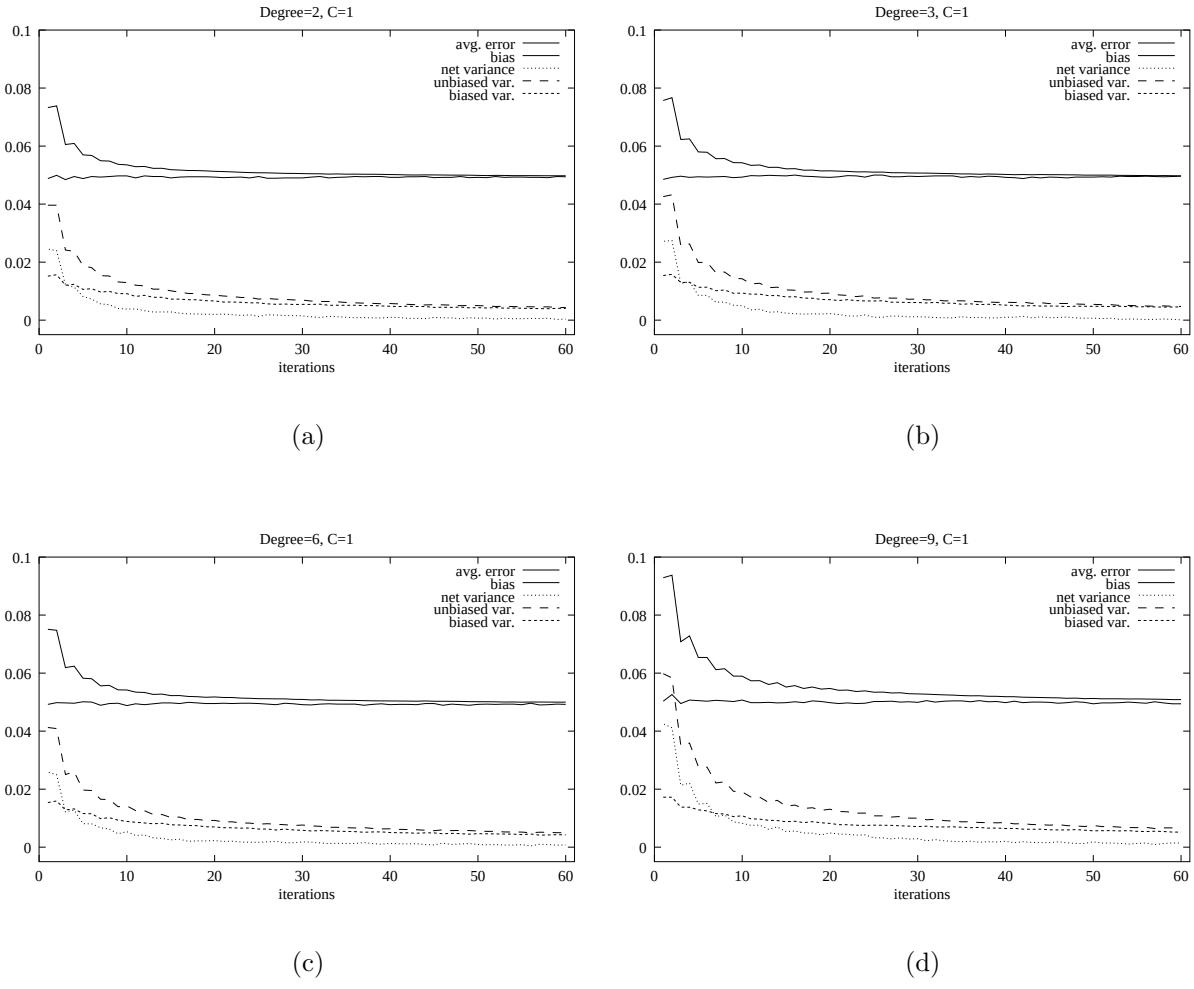


Figure 48: Waveform data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 1$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

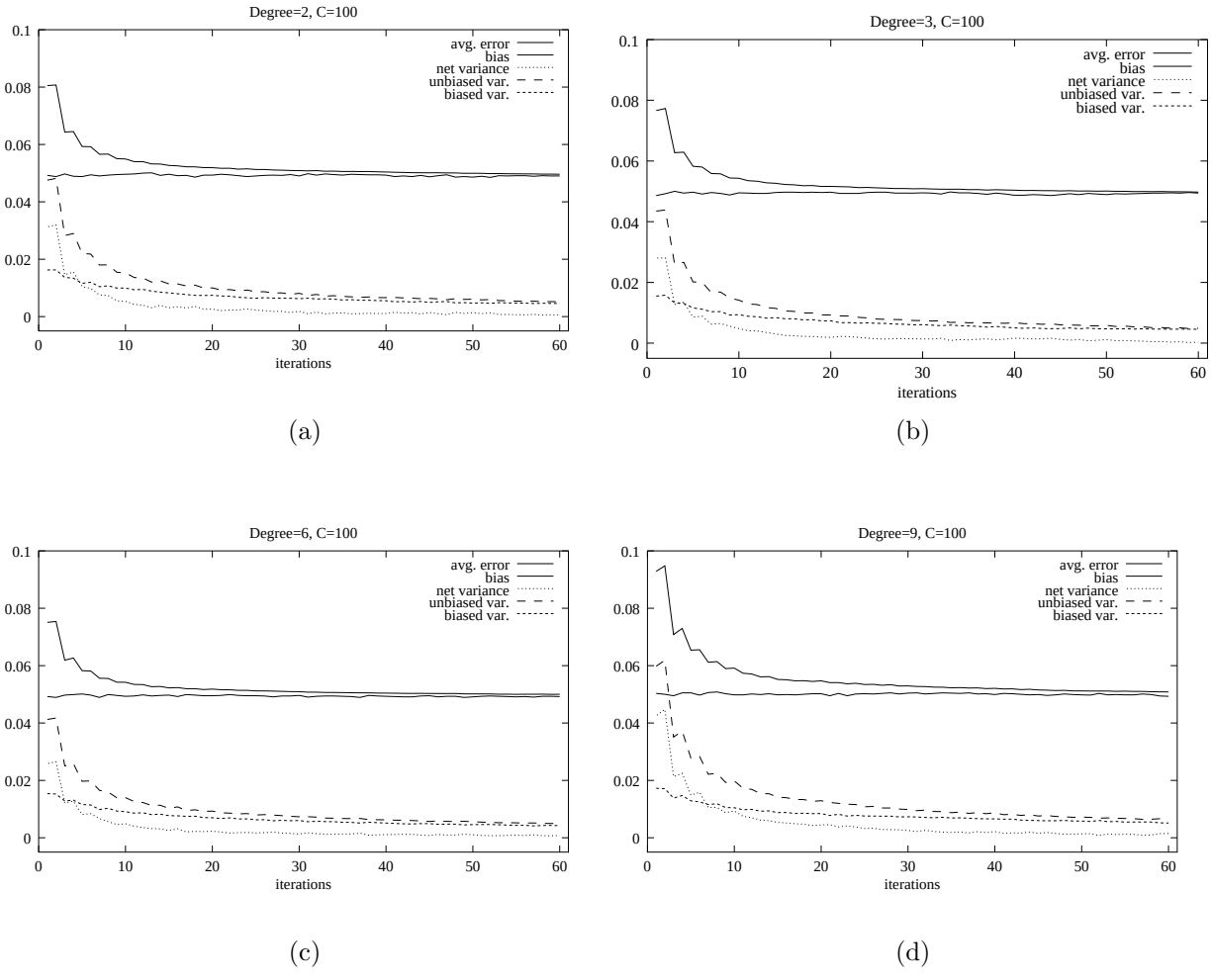


Figure 49: Waveform data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 100$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

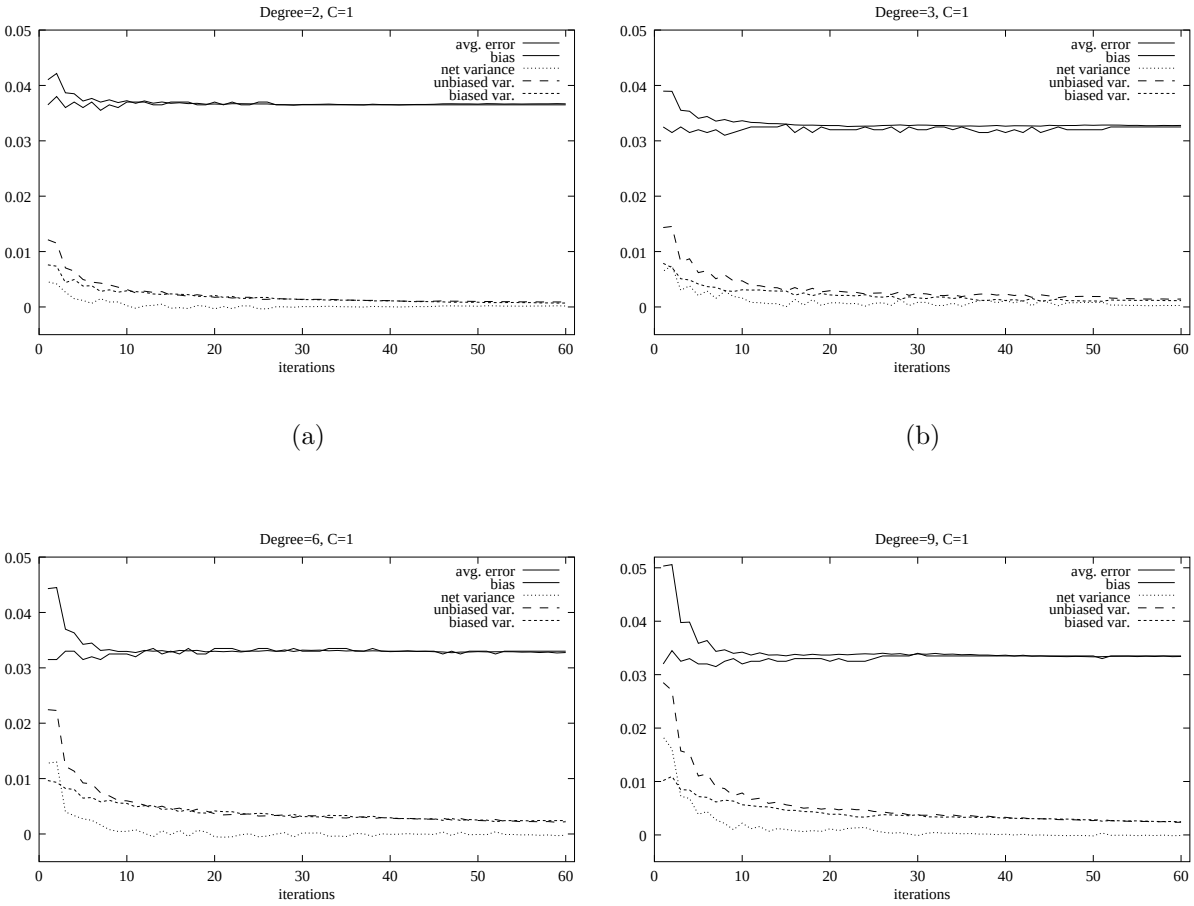


Figure 50: Grey-Landsat data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 1$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

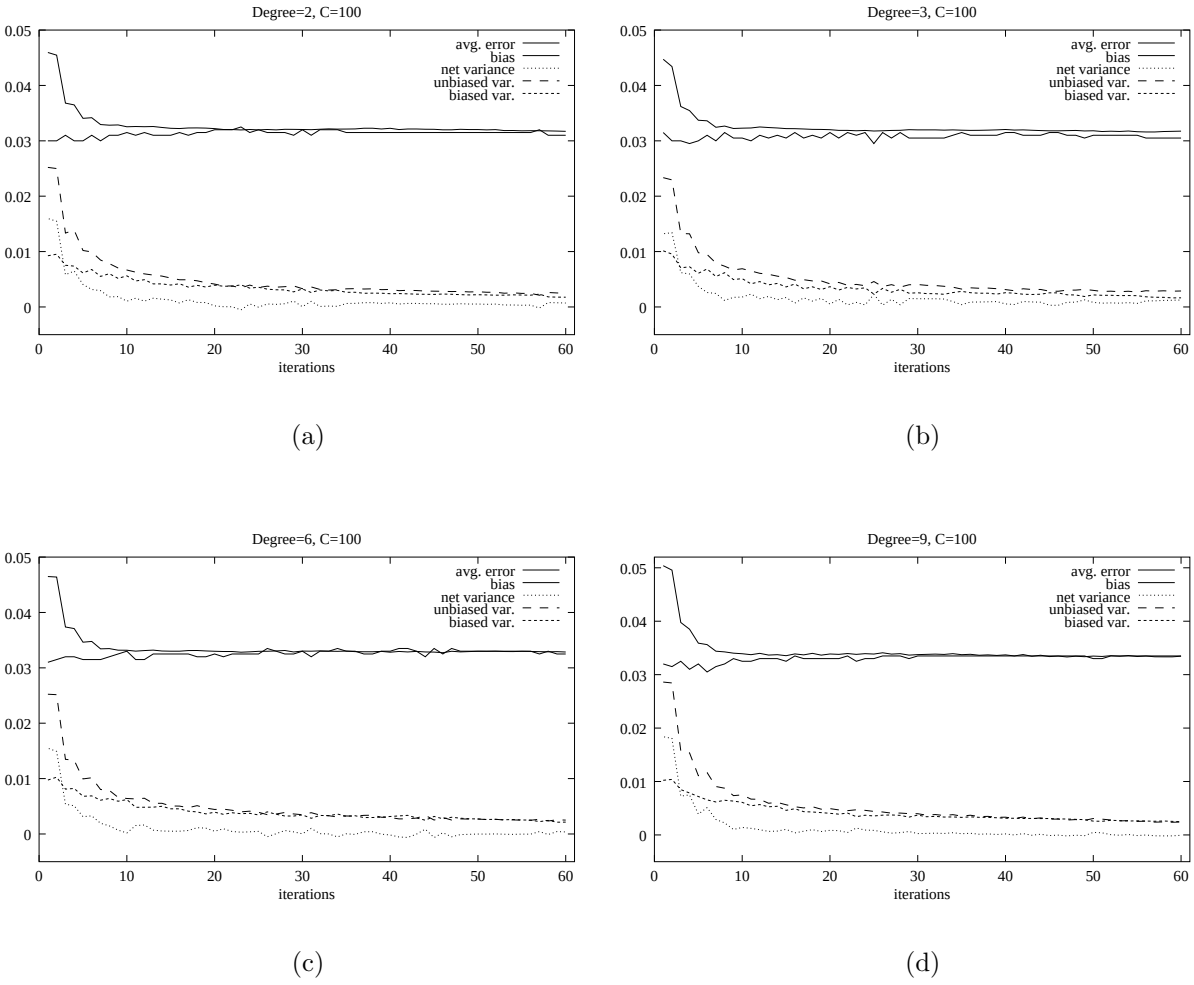


Figure 51: Grey-Landsat data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 100$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

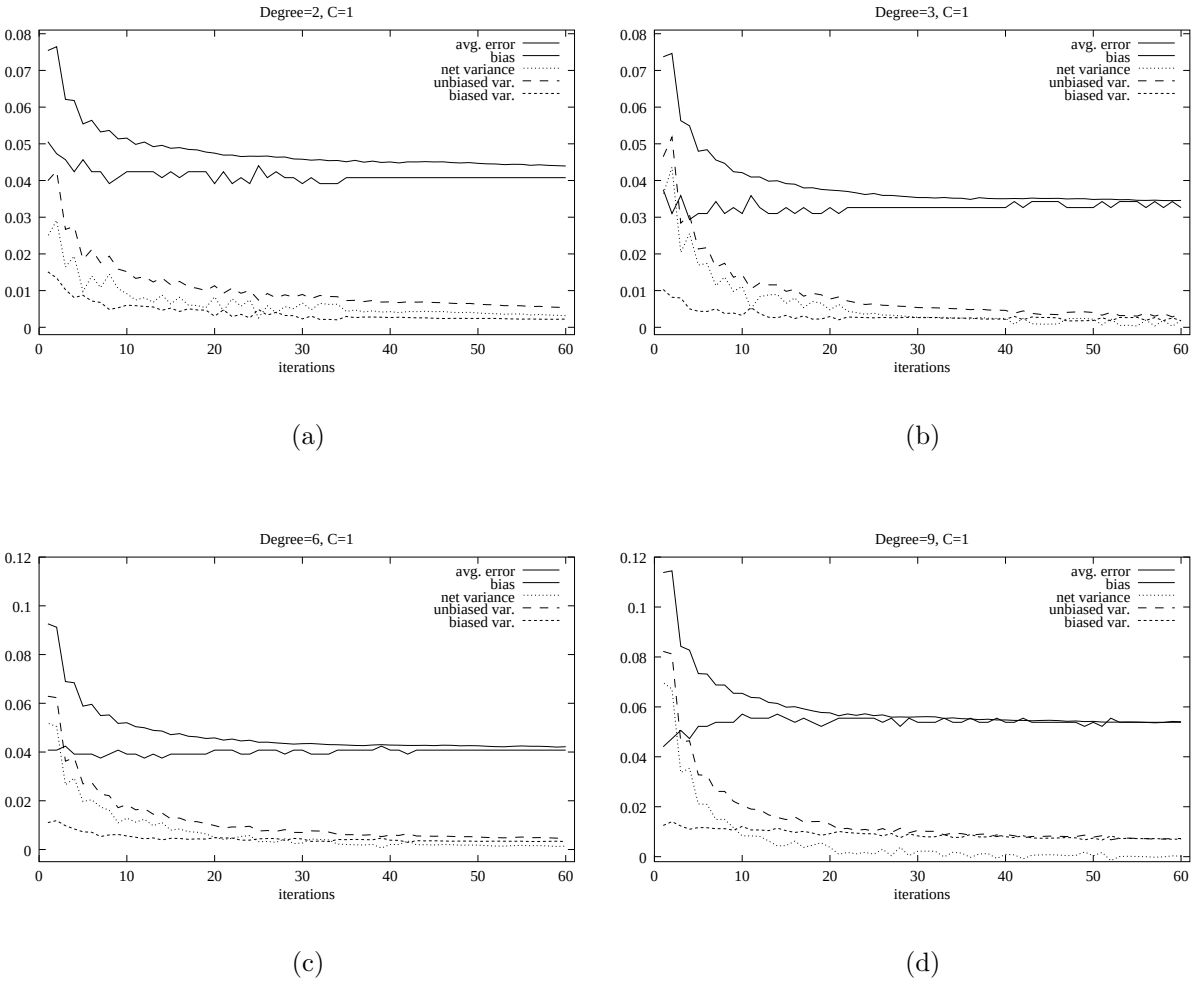


Figure 52: Letter-Two data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 1$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

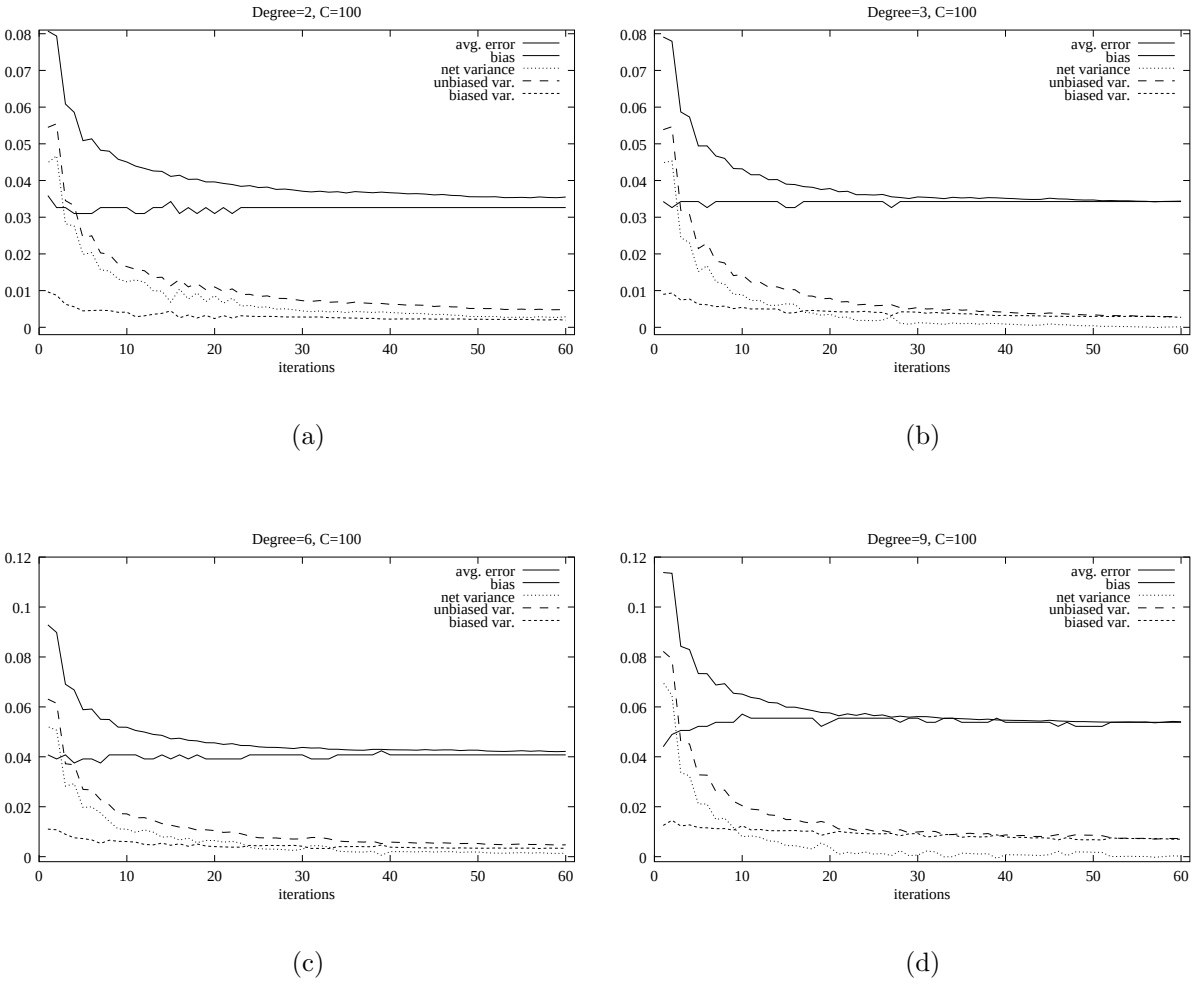


Figure 53: Letter-Two data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 100$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

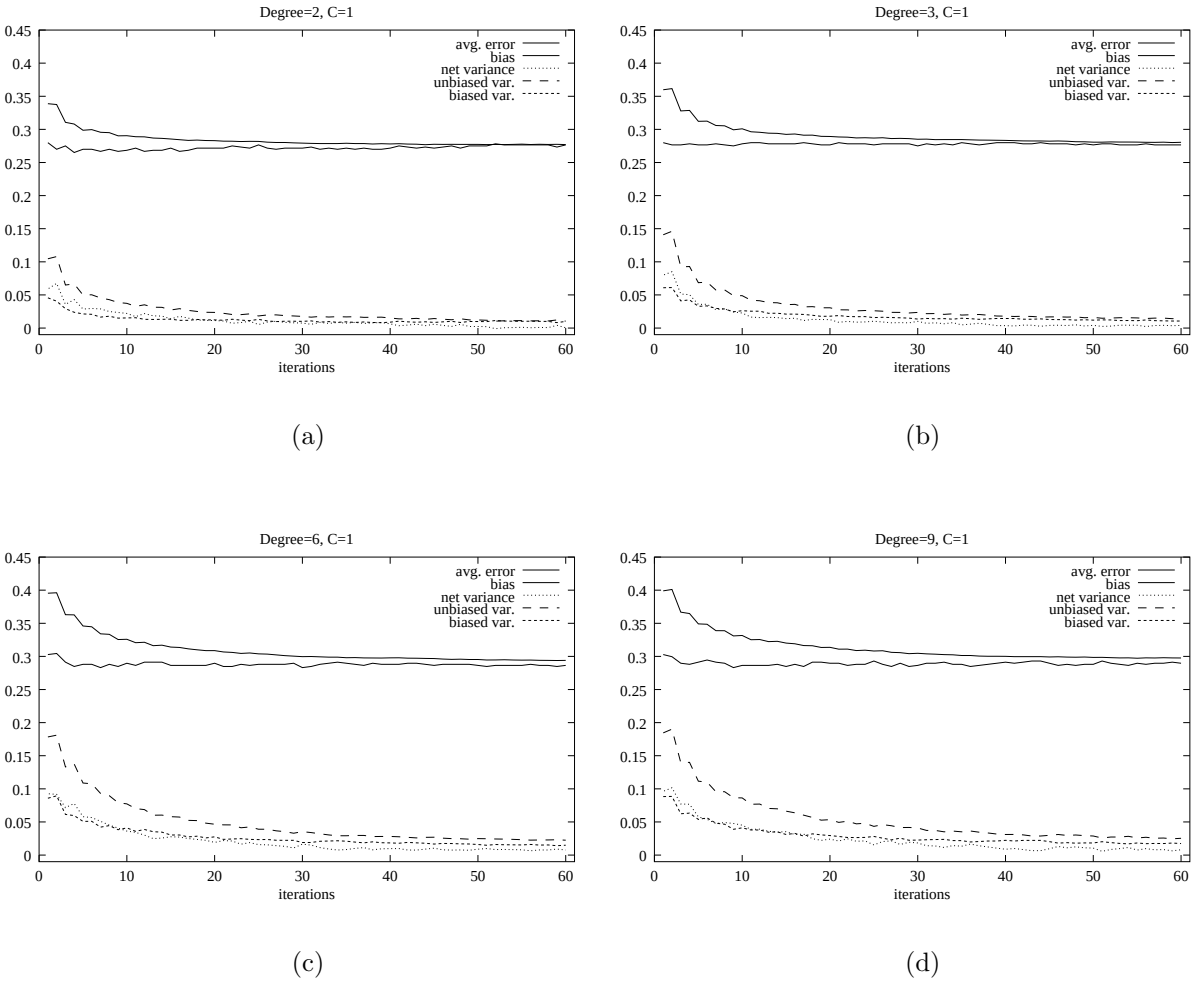


Figure 54: Letter-Two with noise data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 1$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

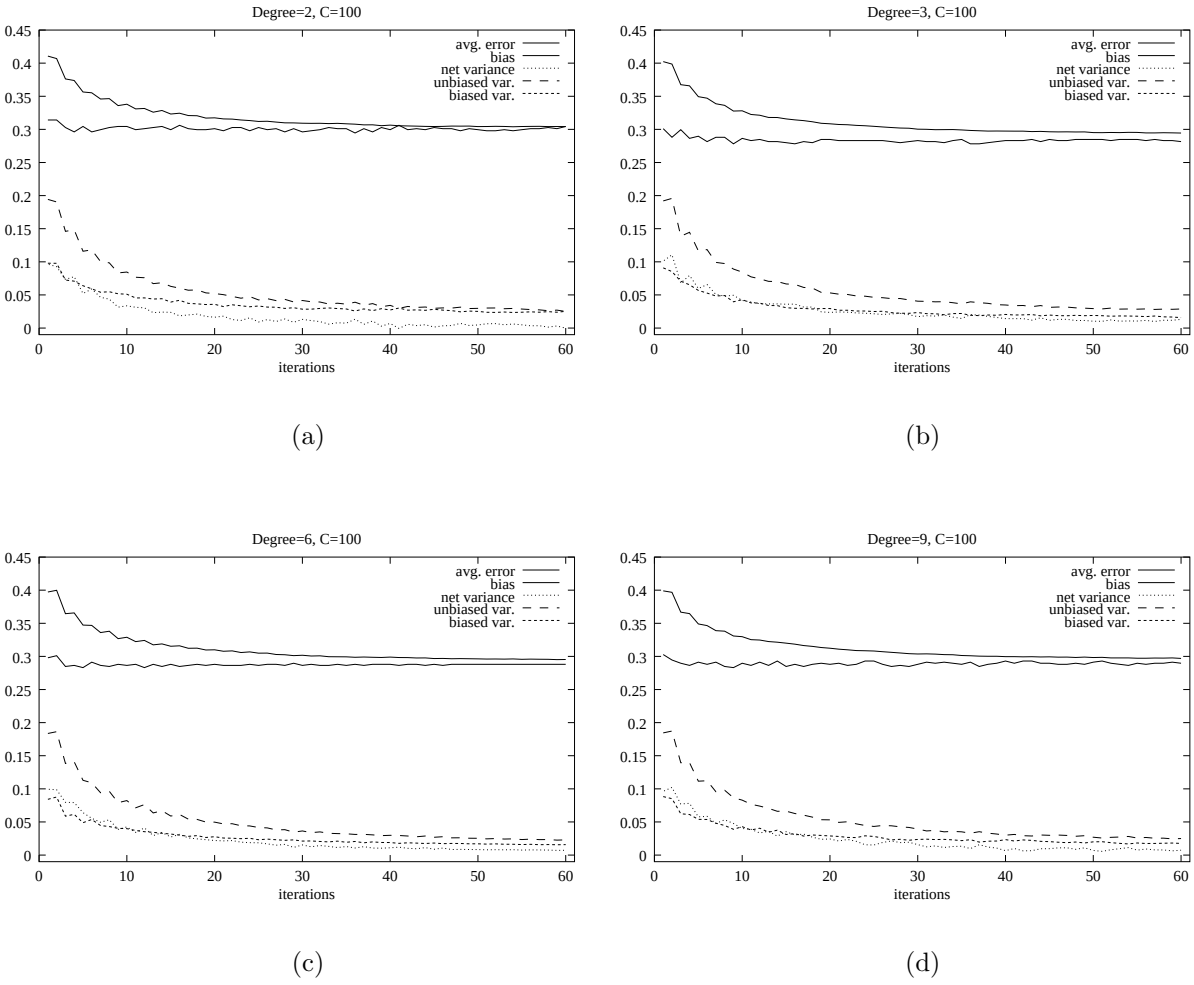


Figure 55: Letter-Two with noise data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 100$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

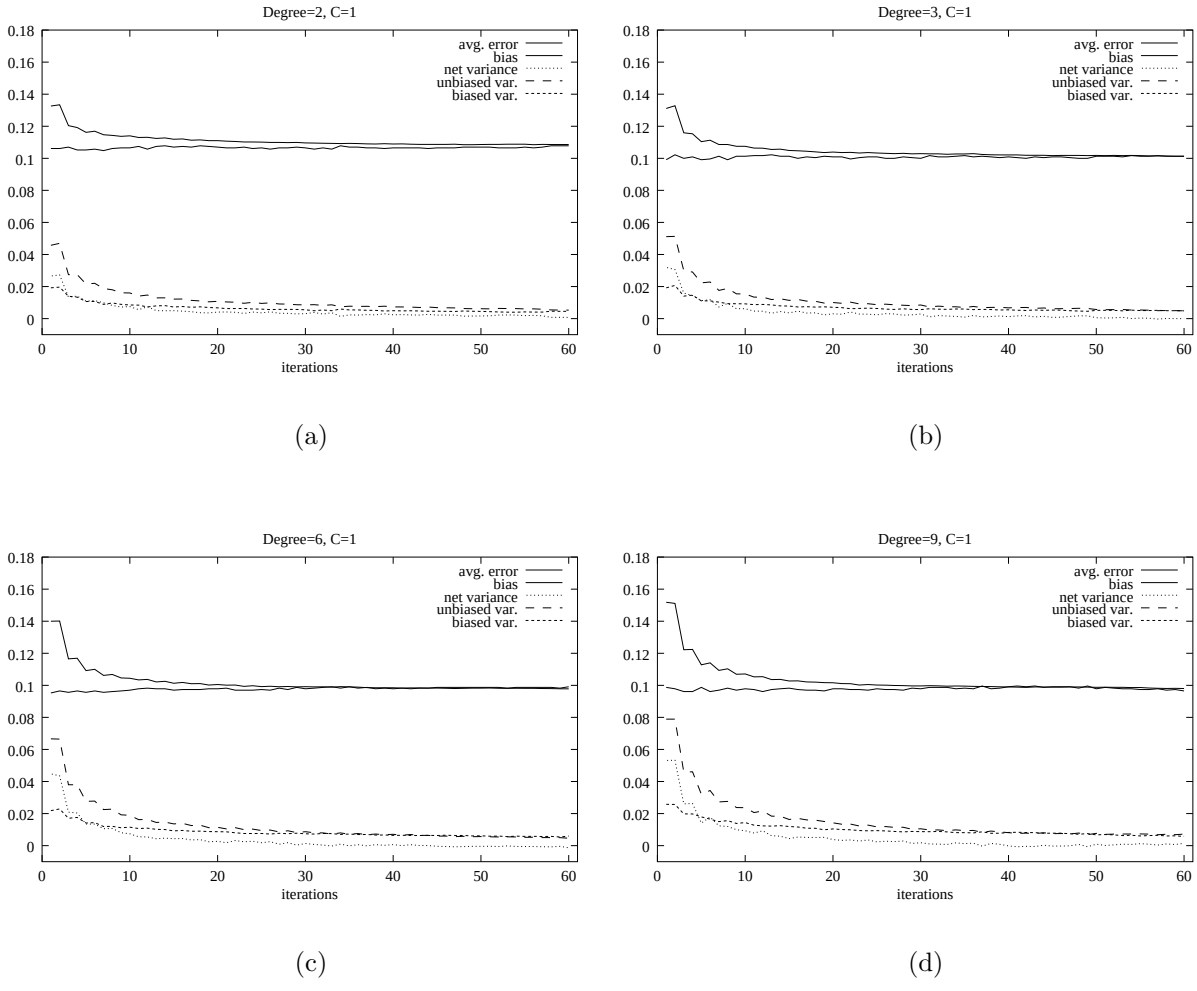


Figure 56: Spam data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 1$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

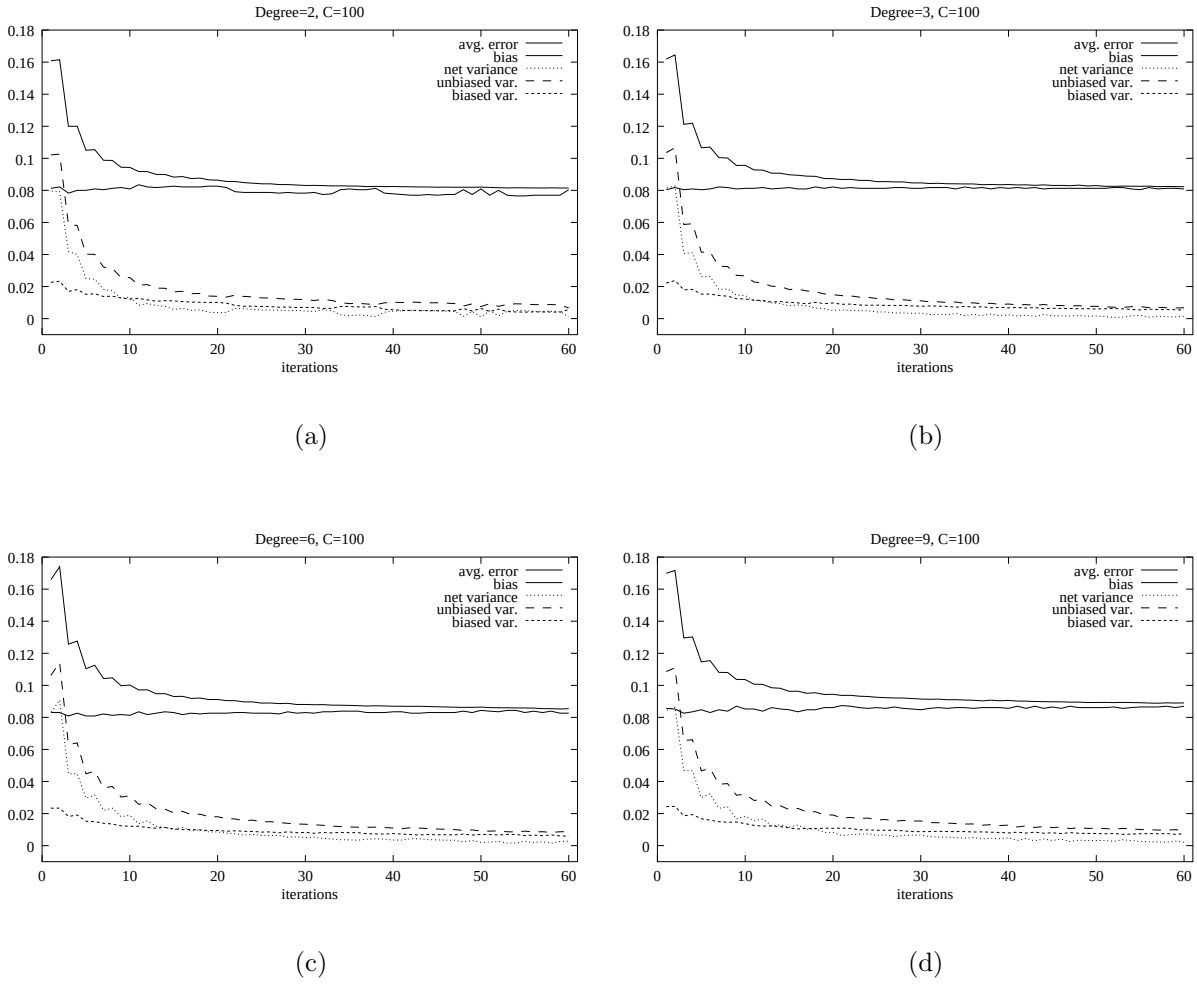


Figure 57: Spam data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 100$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

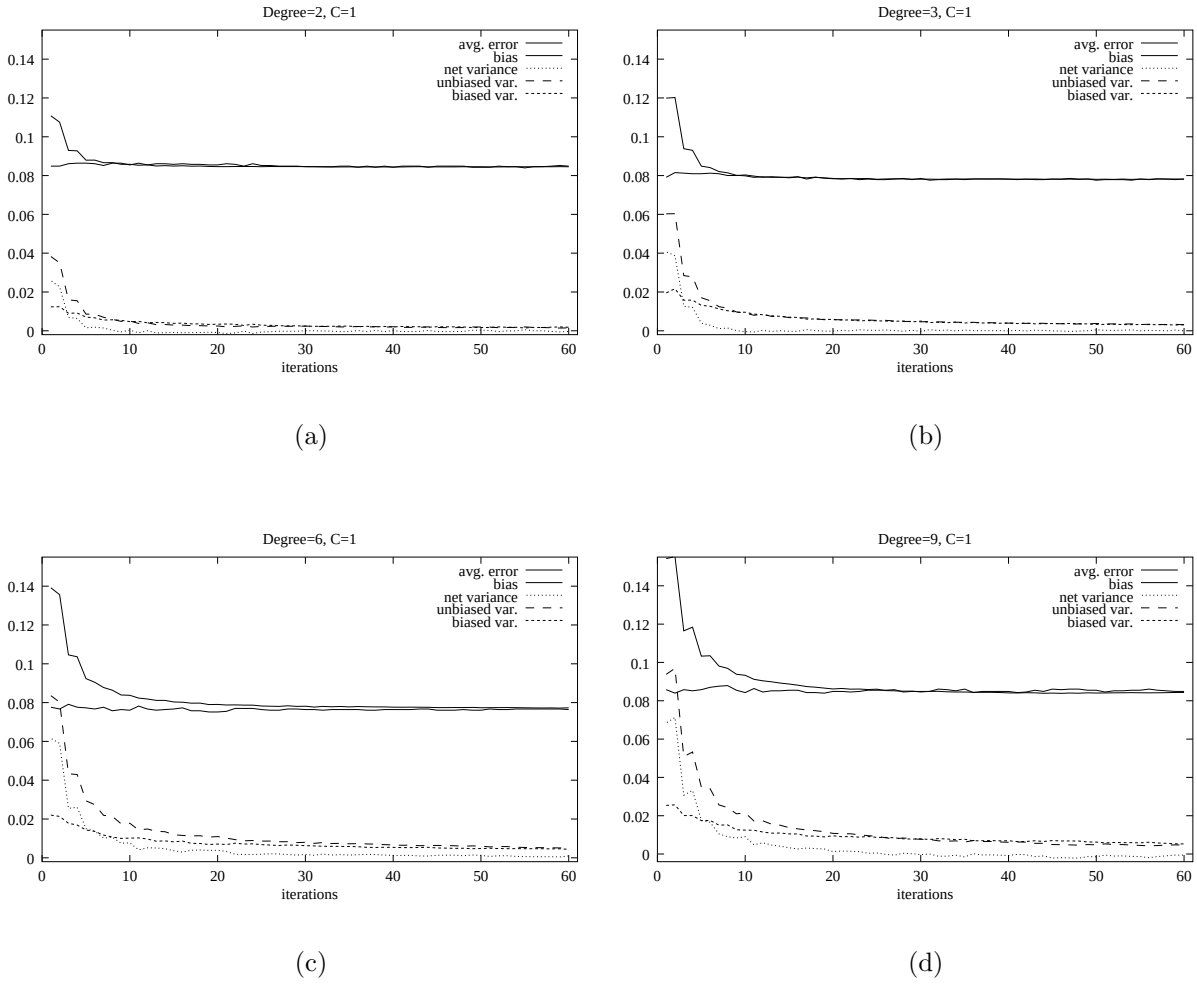


Figure 58: Musk data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 1$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

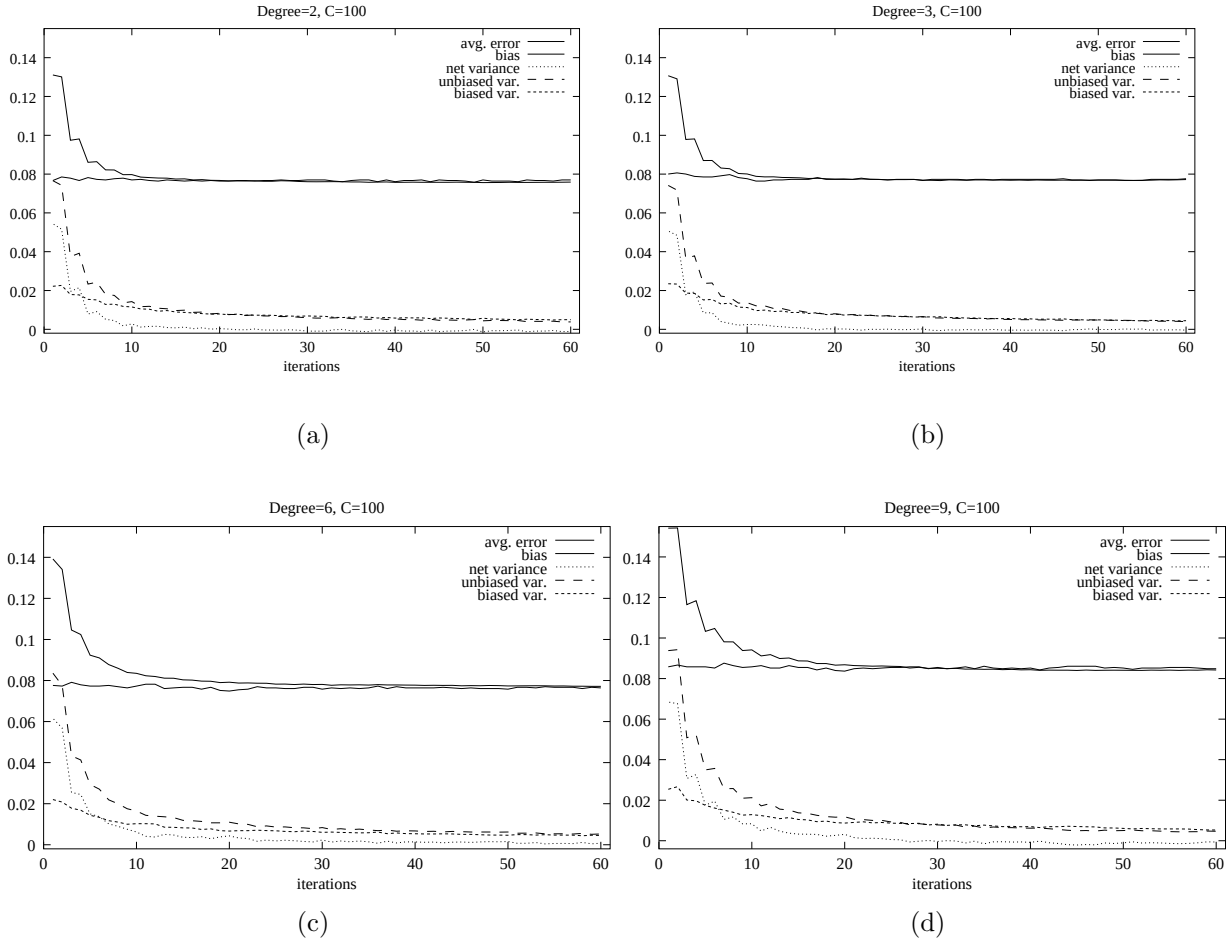


Figure 59: Musk data set. Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA Polynomial SVM with respect to the number of iterations for different values of the degree and $C = 100$: (a) degree = 2, (b) degree = 3, (c) degree = 6, (d) degree = 9.

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

The most important facts with dot-product SVM RA ensembles are the folloeing:

- The bias remains constant
- Most of the error decrement is achieved within the first 10-20 iterations
- Error decrement is due to the decrement of the unbiased variance
- The error is determinated almost totally by the bias.
- The biased variance slowly decreases at each iteration

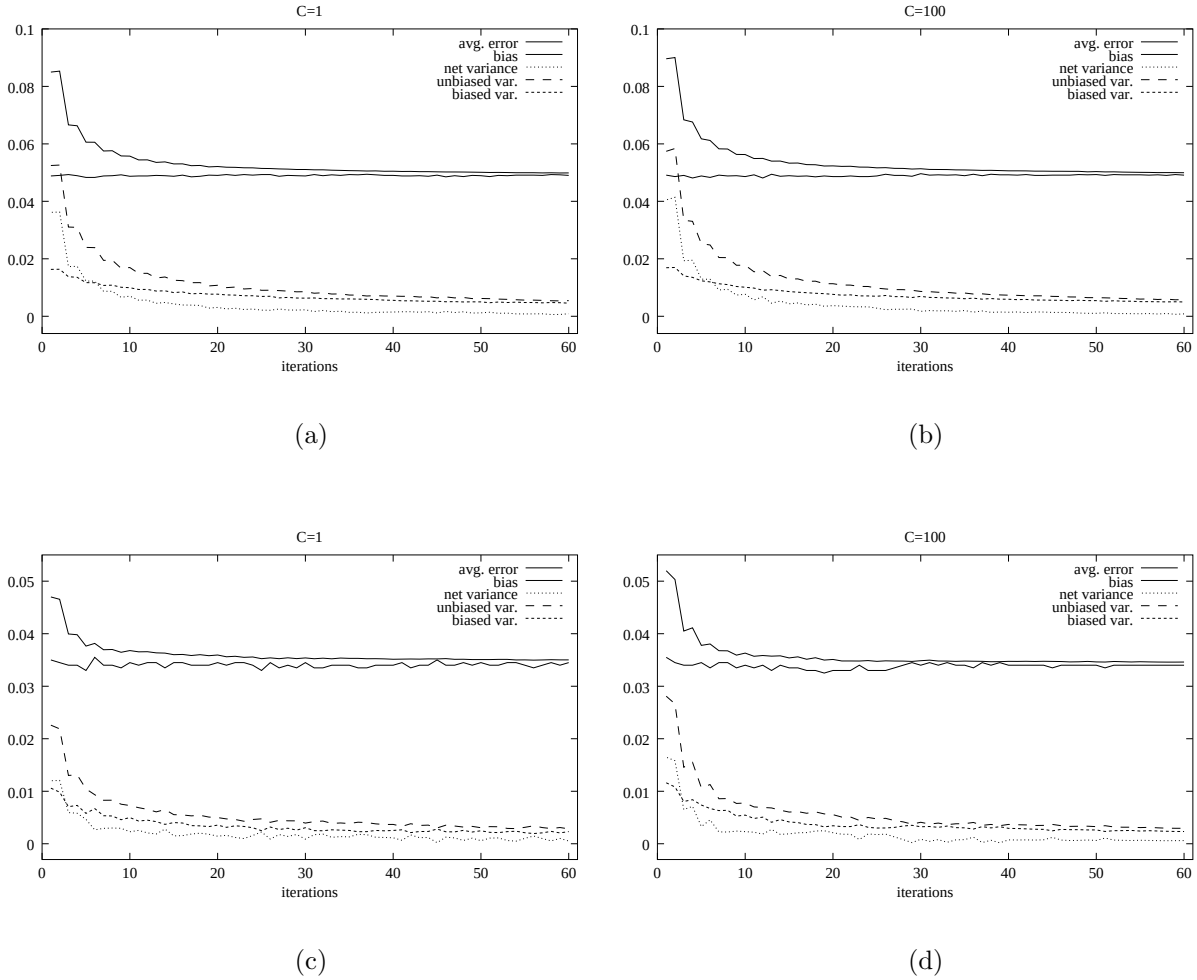


Figure 60: Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA dot-product SVM with respect to the number of iterations: (a) Waveform, $C = 1$, (b) Waveform, $C = 100$, (c) Grey-Landsat, $C = 1$, (d) Grey-Landsat, $C = 100$.

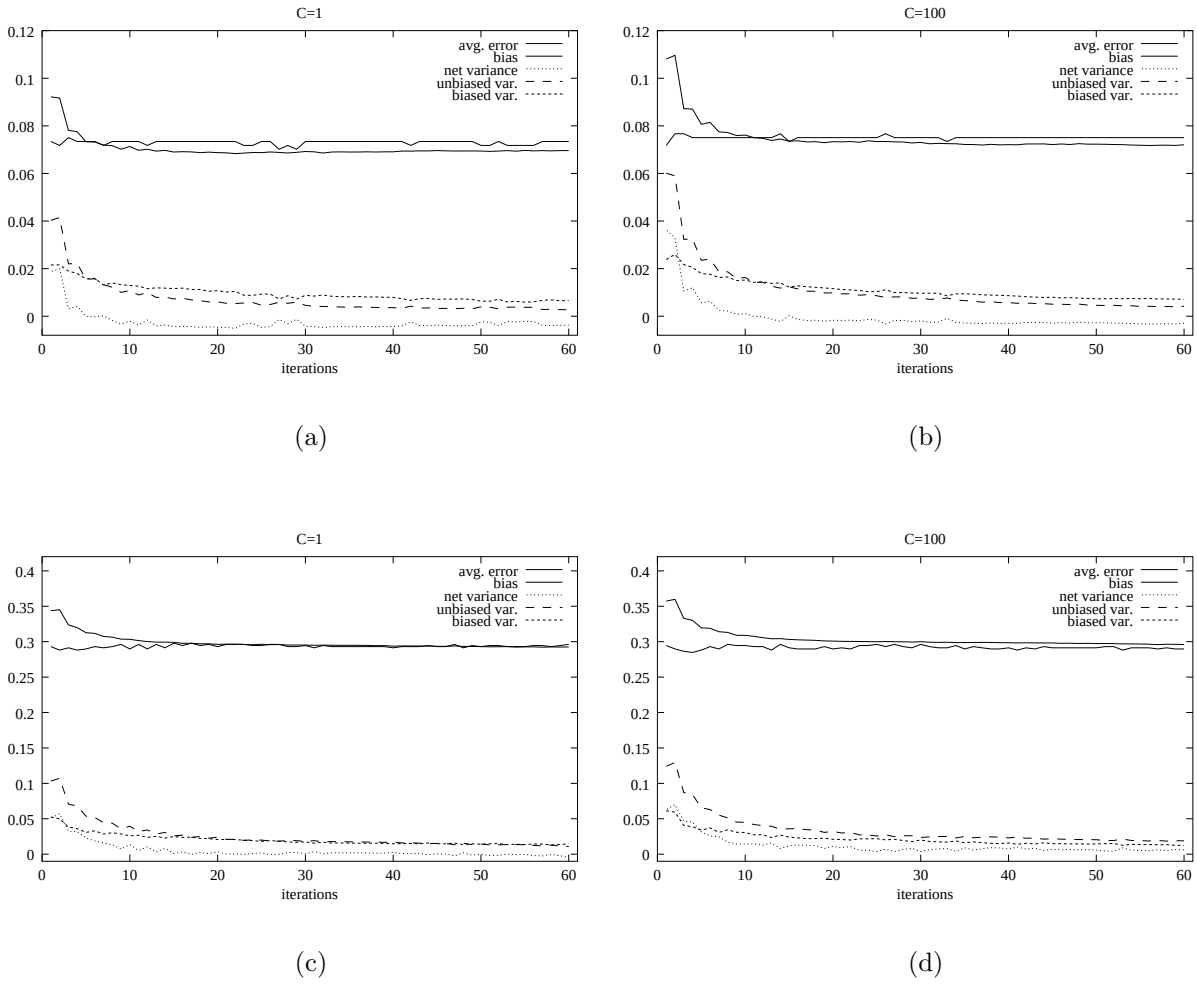


Figure 61: Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA dot-product SVM with respect to the number of iterations: (a) Letter-Two, $C = 1$, (b) Letter-Two, $C = 100$, (c) Letter-Two woth noise, $C = 1$, (d) Letter-Two woth noise, $C = 100$.

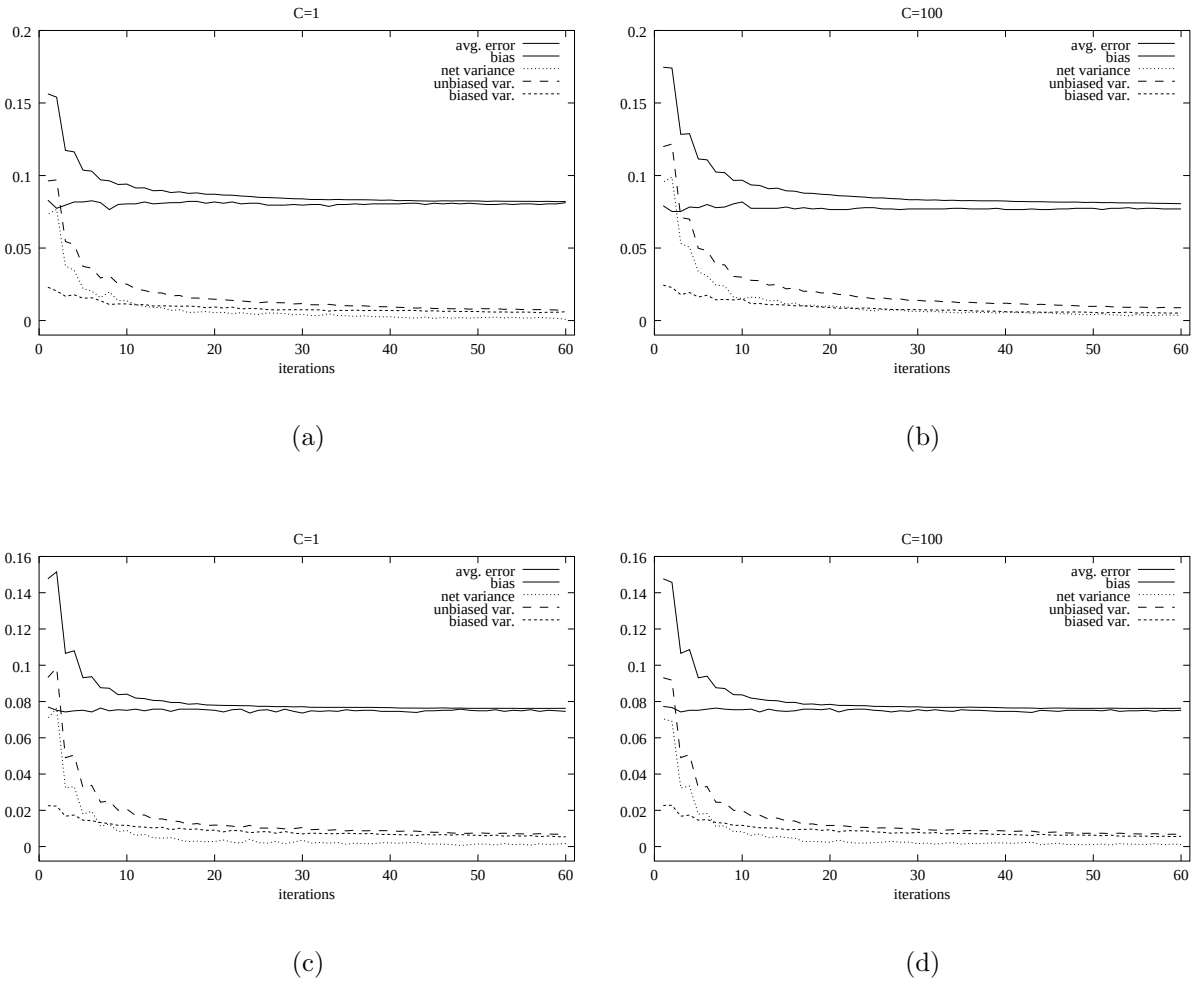


Figure 62: Bias-variance decomposition of the error in bias, net variance, unbiased and biased variance in RA dot-product SVM with respect to the number of iterations: (a) Spam, $C = 1$, (b) Spam, $C = 100$, (c) Musk, $C = 1$, (d) Musk, $C = 100$.

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

Here are reported the graphics comparing bias–variance decomposition in single SVMs and RA ensembles of SVMs. In all graphics of this section the data referred to single SVMs are labeled with crosses, while RA 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 RA SVMs.

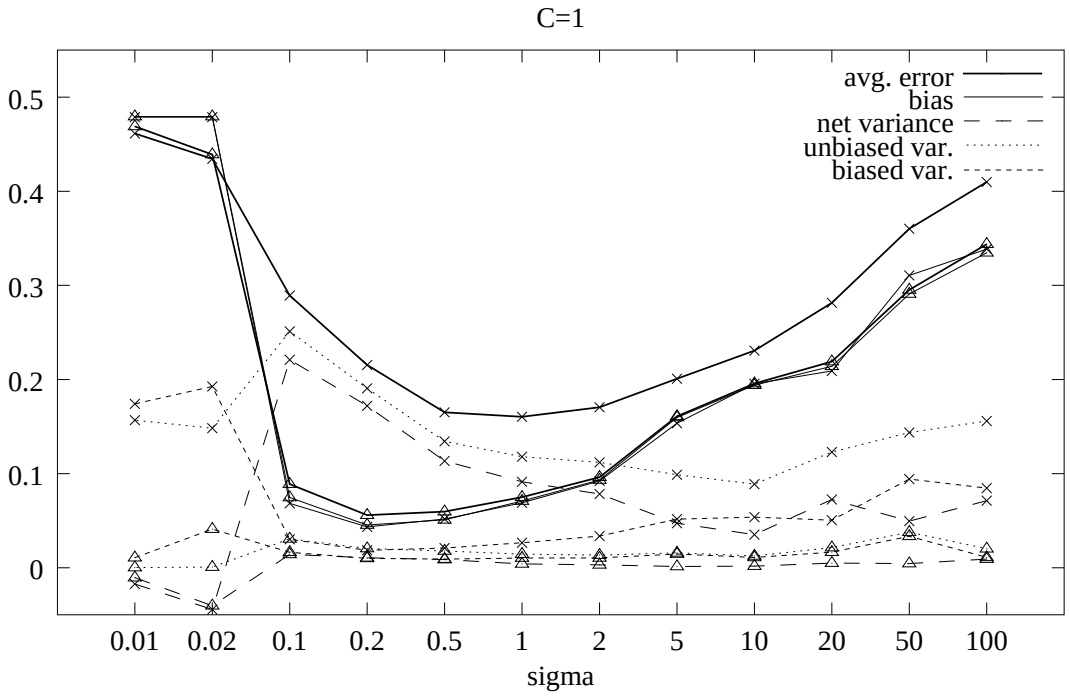
4.1 Comparison between single and RA RBF-SVM

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

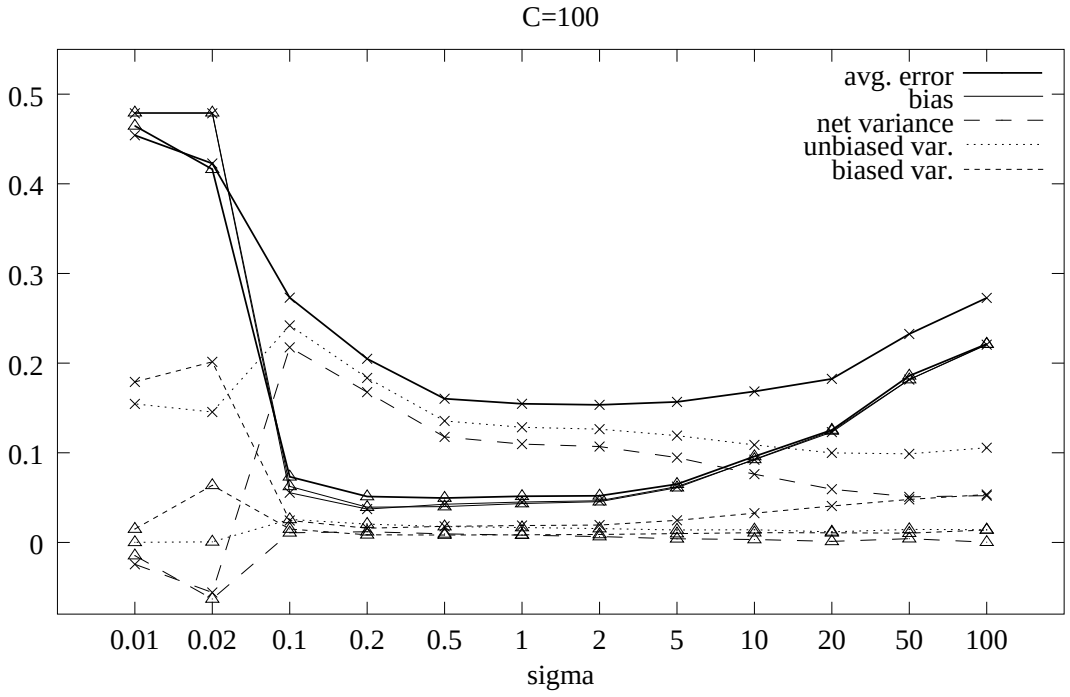
High bias region. In this region the the error of single and RA SVMs is about equal, and it is characterized by a very high bias. The net-variance is close to 0, because biased variance is about equal to the unbiased variance. In most cases they are both close to 0. In some cases they are equal but greater than 0 with significantly larger values in single that in in RA SVMs (Fig. 64).

Transition region. The bias goes down in the transition region at about the same rate in single and RA SVM ensembles. The net-variance mainitains the wave-shape also in RA SVMs, but it is lower. In some data sets (Fig. 63, 67) 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 RA 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 RA 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 RA SVM enesmbles.

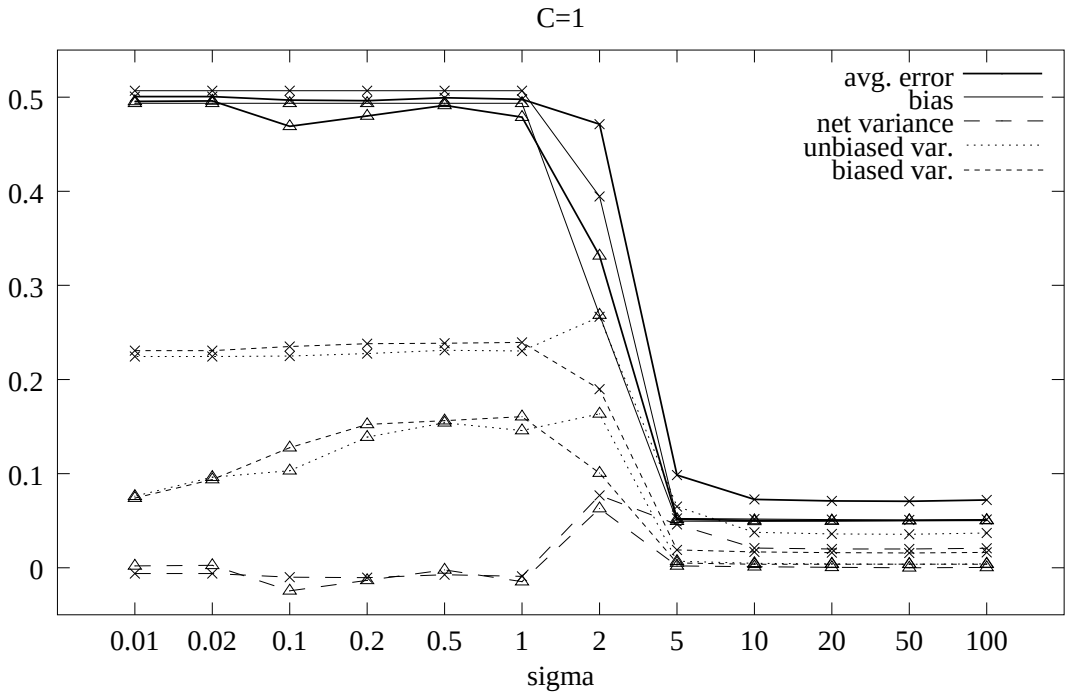


(a)

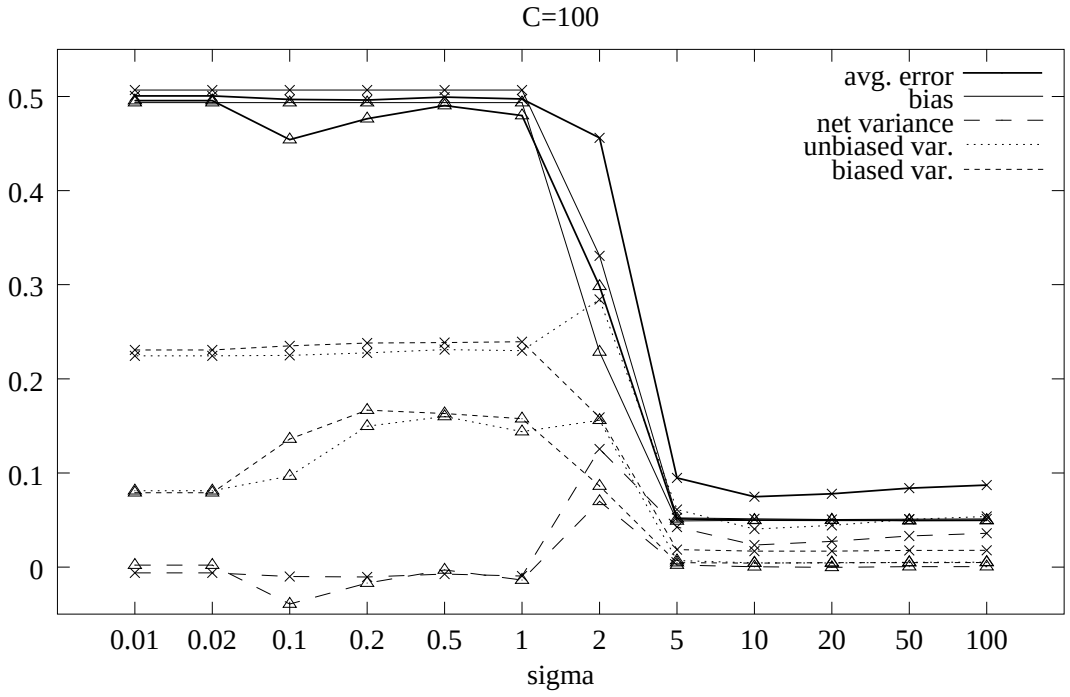


(b)

Figure 63: P2 data set. Comparison between bias-variance decomposition between single RBF-SVMs (lines labeled with crosses) and RA SVM RBF ensembles (lines labeled with triangles), while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

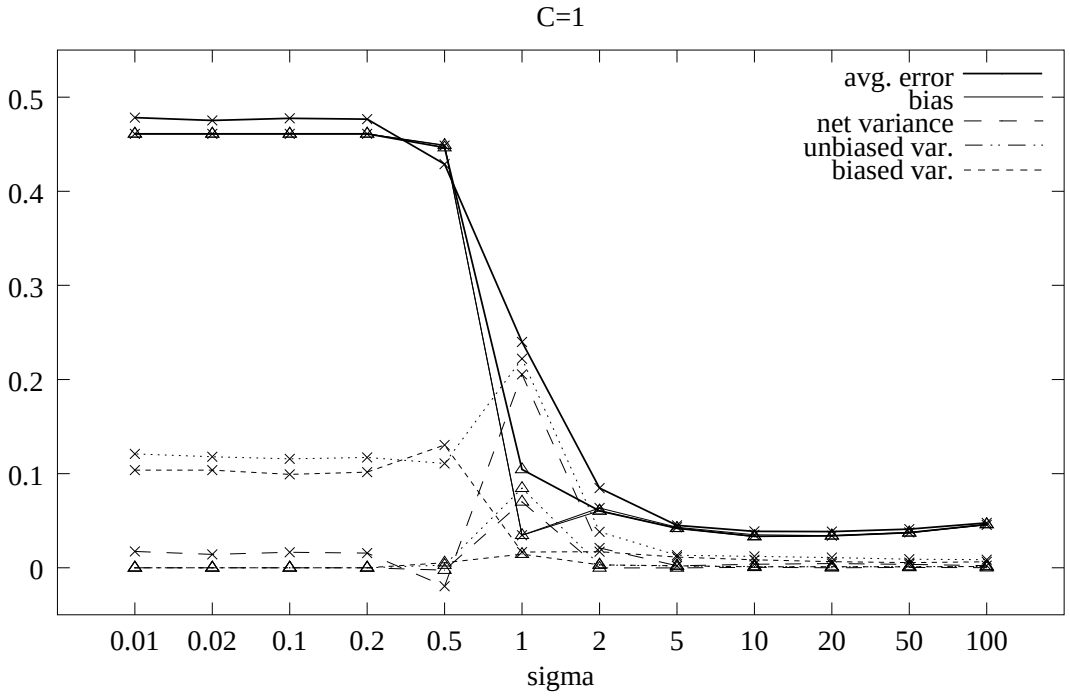


(a)

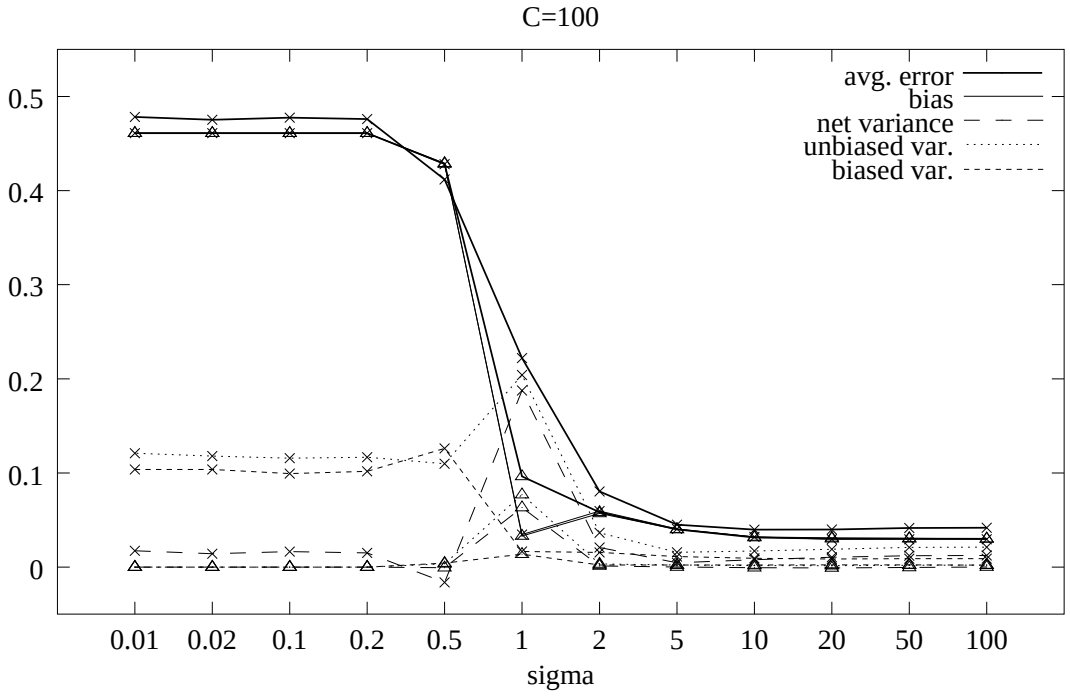


(b)

Figure 64: Waveform data set. Comparison of bias-variance decomposition between single RBF-SVMs (lines labeled with crosses) and RA SVM RBF ensembles (lines labeled with triangles), while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

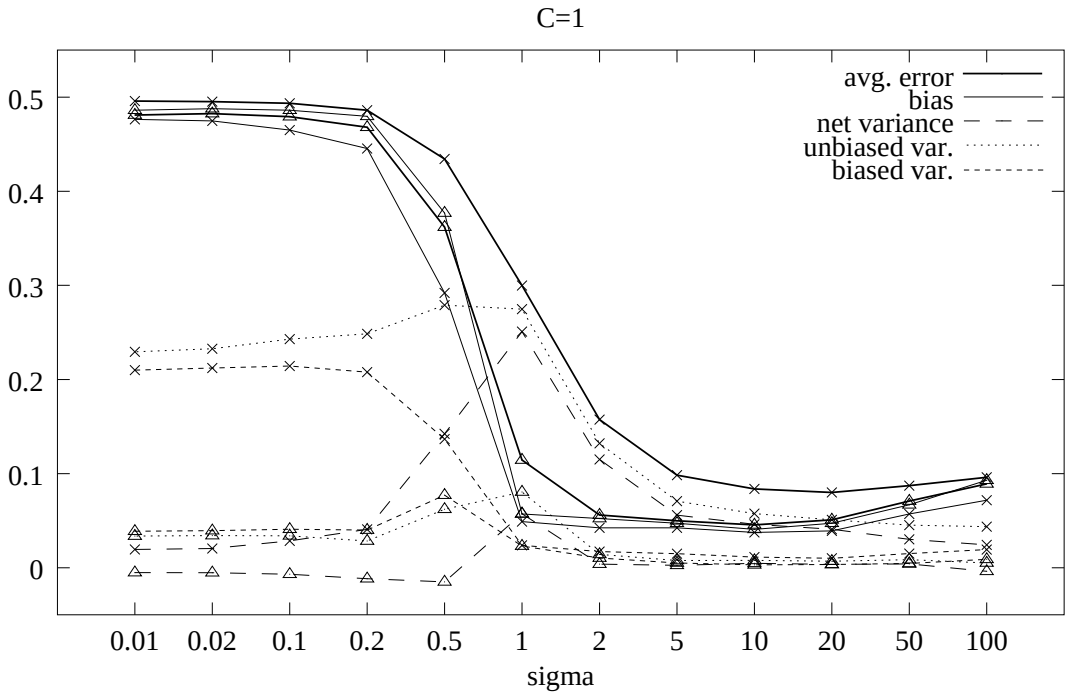


(a)

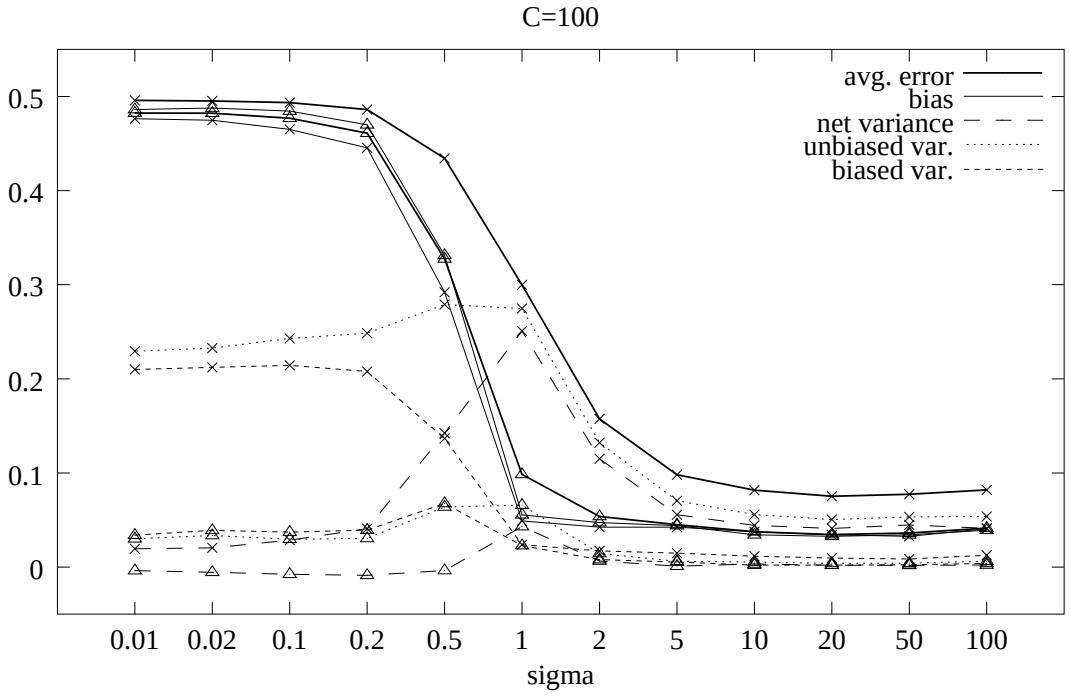


(b)

Figure 65: Grey-Landsat data set. Comparison of bias-variance decomposition between single RBF-SVMs (lines labeled with crosses) and RA SVM RBF ensembles (lines labeled with triangles), while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

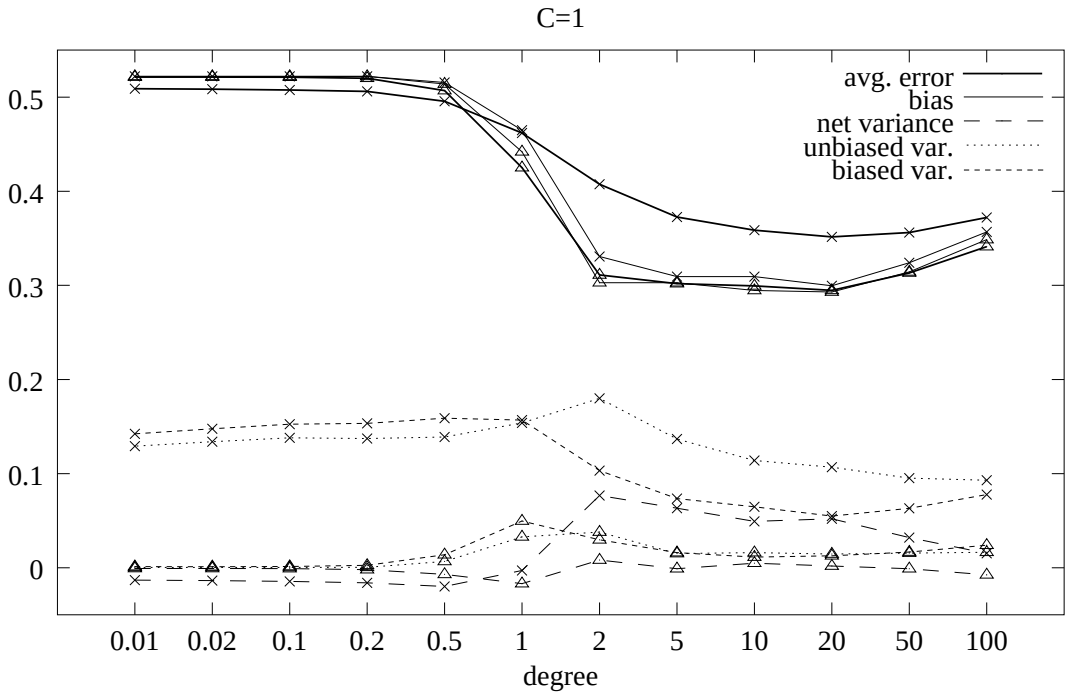


(a)

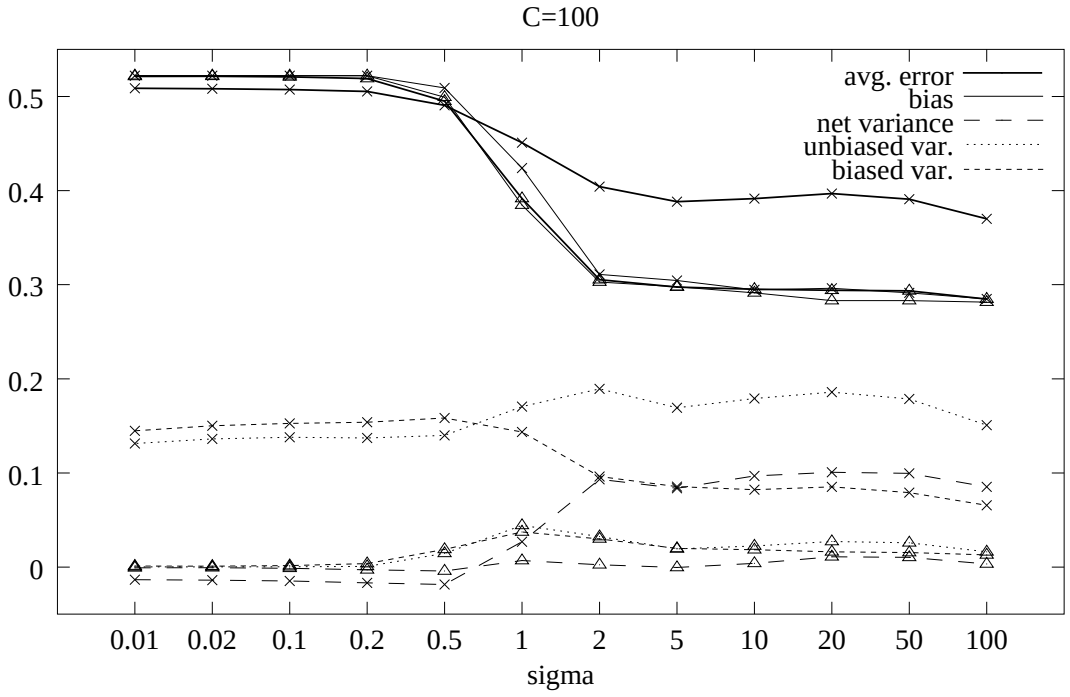


(b)

Figure 66: Letter-Two data set. Comparison of bias-variance decomposition between single RBF-SVMs (lines labeled with crosses) and RA SVM RBF ensembles (lines labeled with triangles), while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

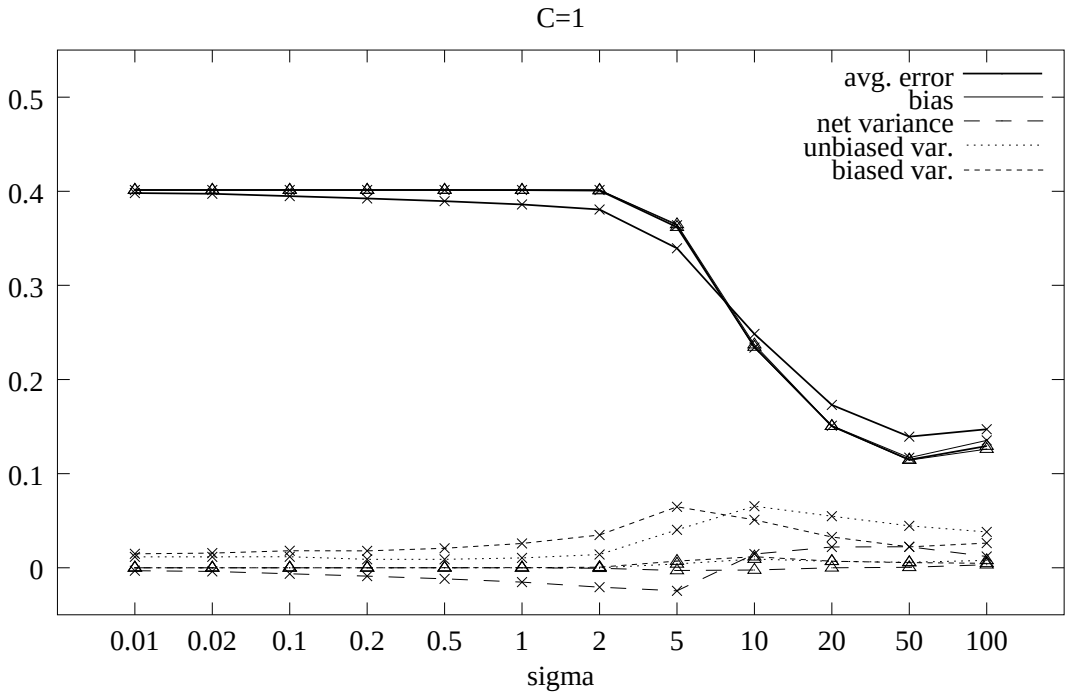


(a)

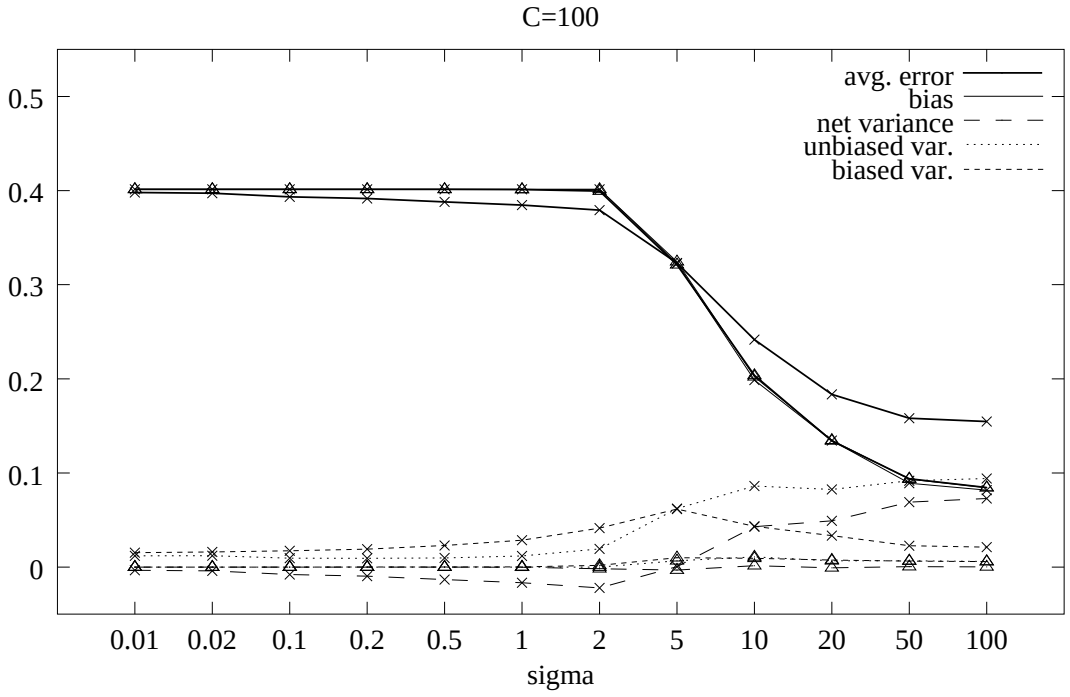


(b)

Figure 67: Letter-Two with noise data set. Comparison of bias-variance decomposition between single RBF-SVMs (lines labeled with crosses) and RA SVM RBF ensembles (lines labeled with triangles), while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

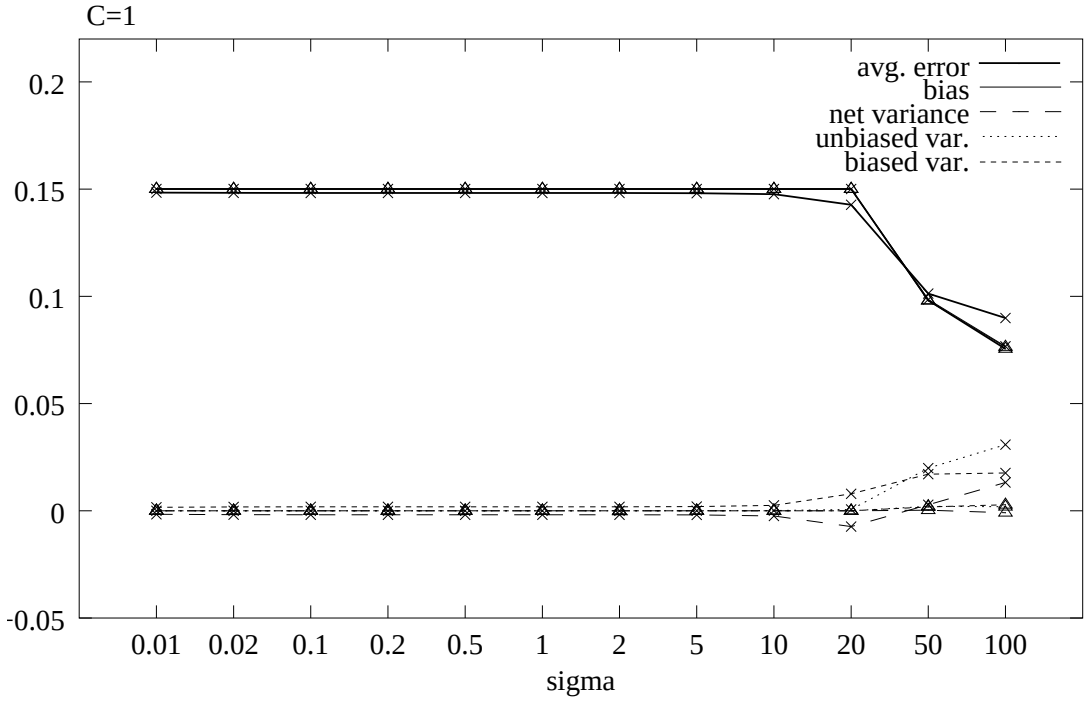


(a)

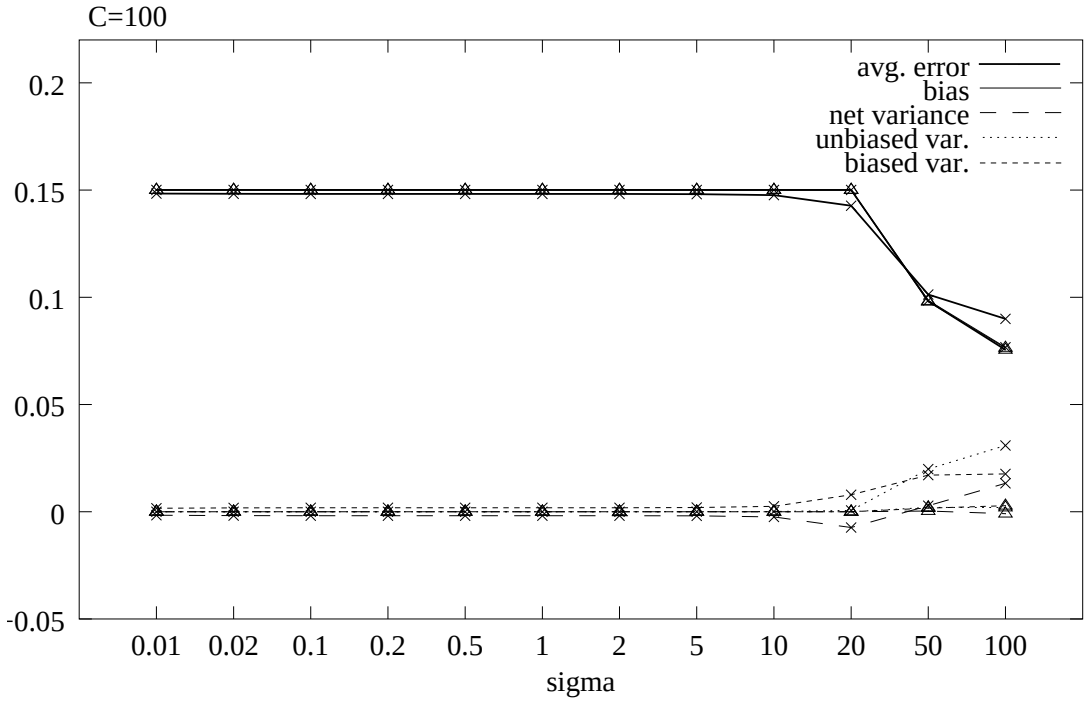


(b)

Figure 68: Spam data set. Comparison of bias-variance decomposition between single RBF-SVMs (lines labeled with crosses) and RA 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 69: Musk data set. Comparison of bias-variance decomposition between single RBF-SVMs (lines labeled with crosses) and RA SVM RBF ensembles (lines labeled with triangles), while varying σ and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

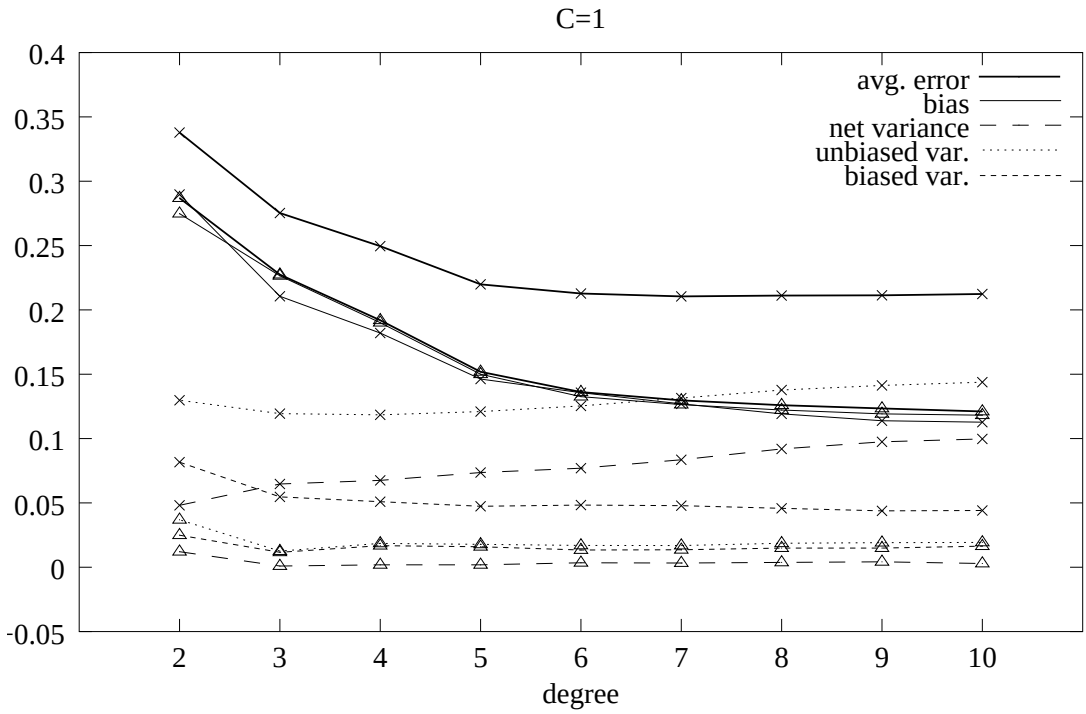
4.2 Comparison between single and RA polynomial SVM

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

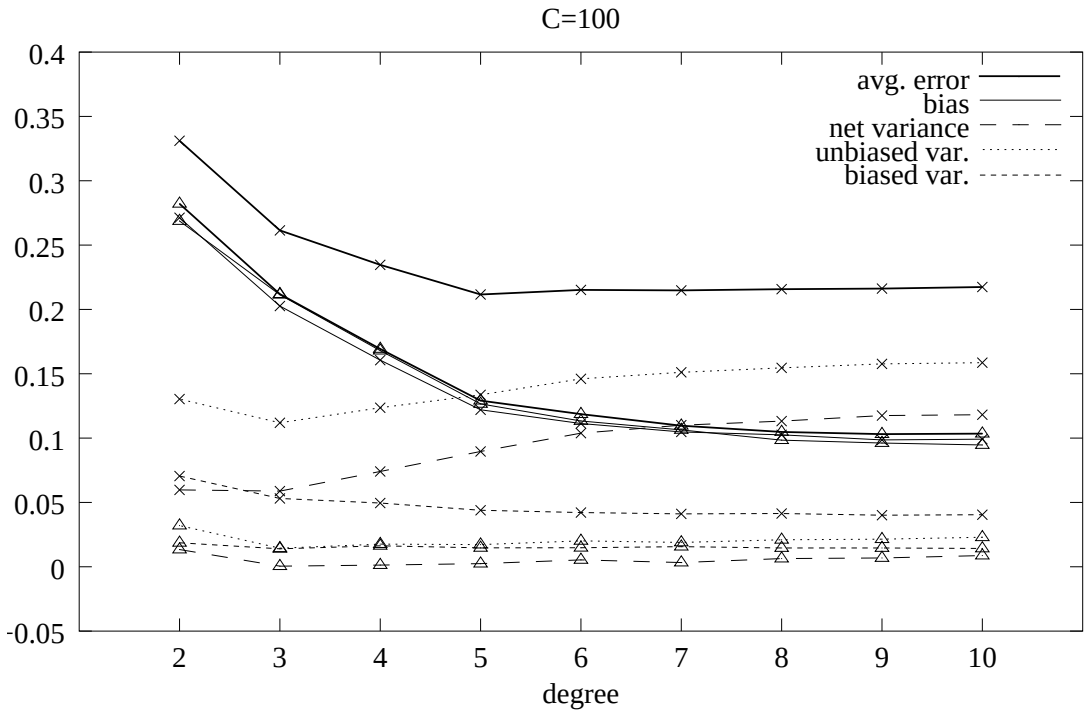
The variance component is close to 0: its contribute to the error is almost removed.

In RA 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.

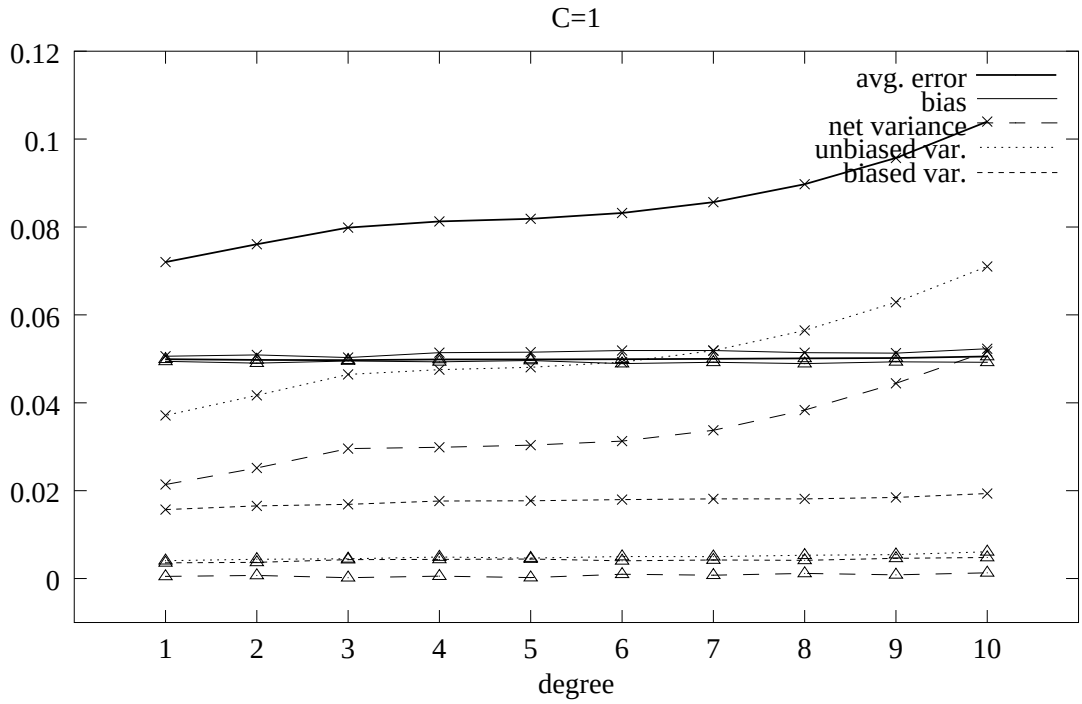


(a)

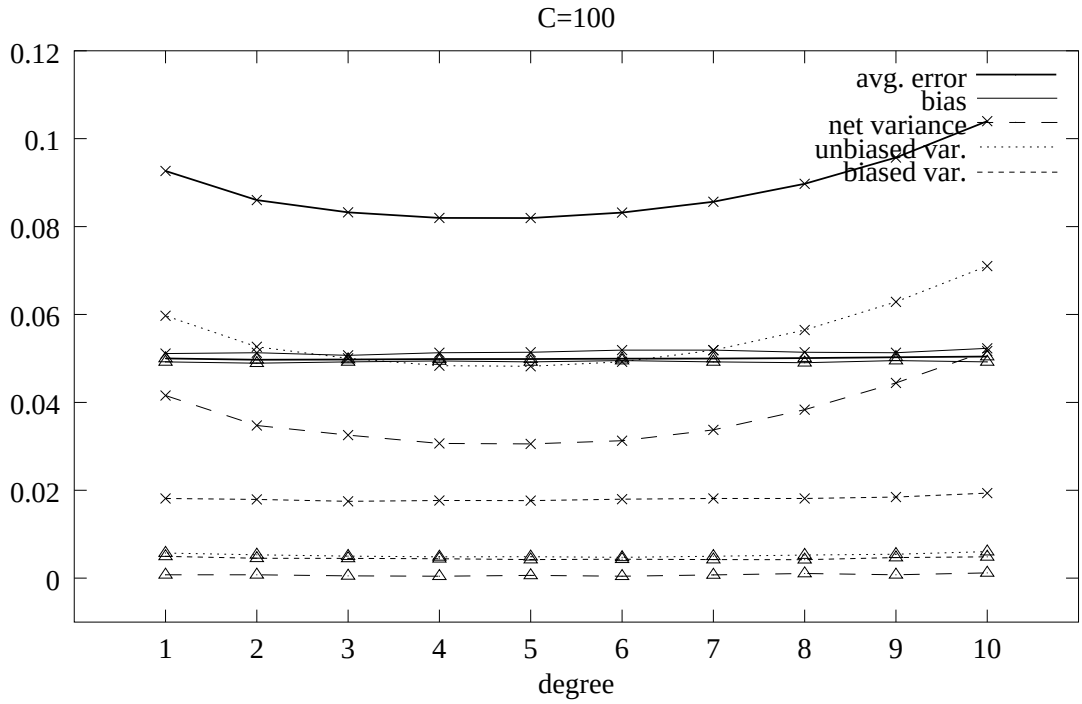


(b)

Figure 70: P2 data set. Comparison of bias-variance decomposition between single polynomial SVMs (lines labeled with crosses) and RA polynomial SVM ensembles (lines labeled with triangles), while varying the degree and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

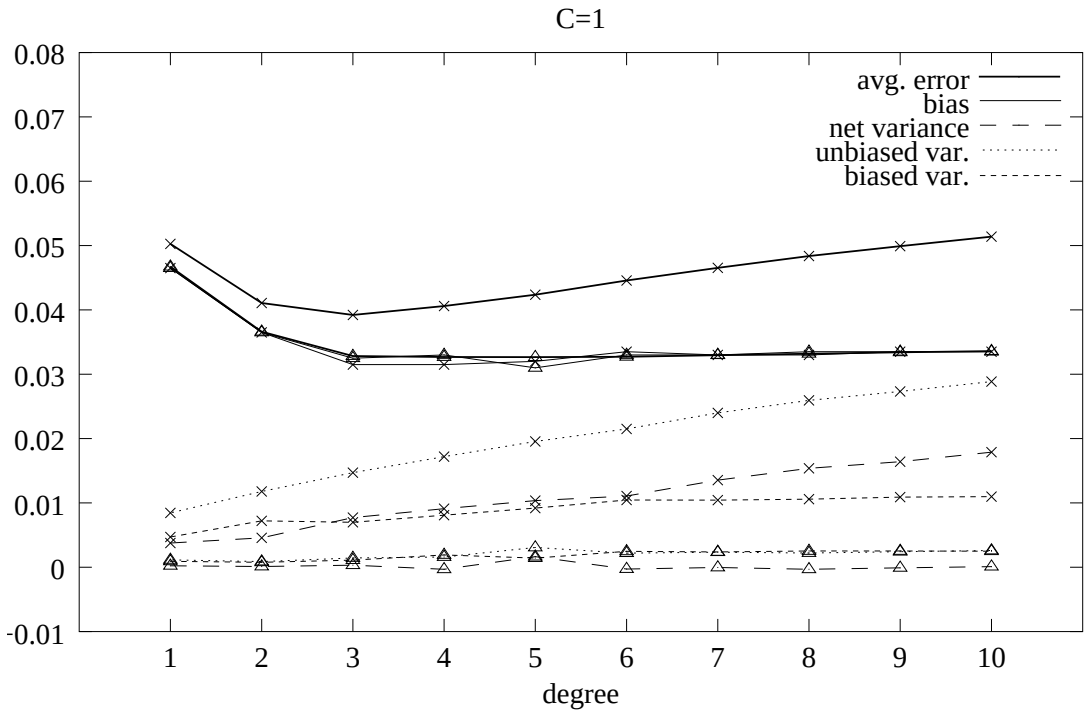


(a)

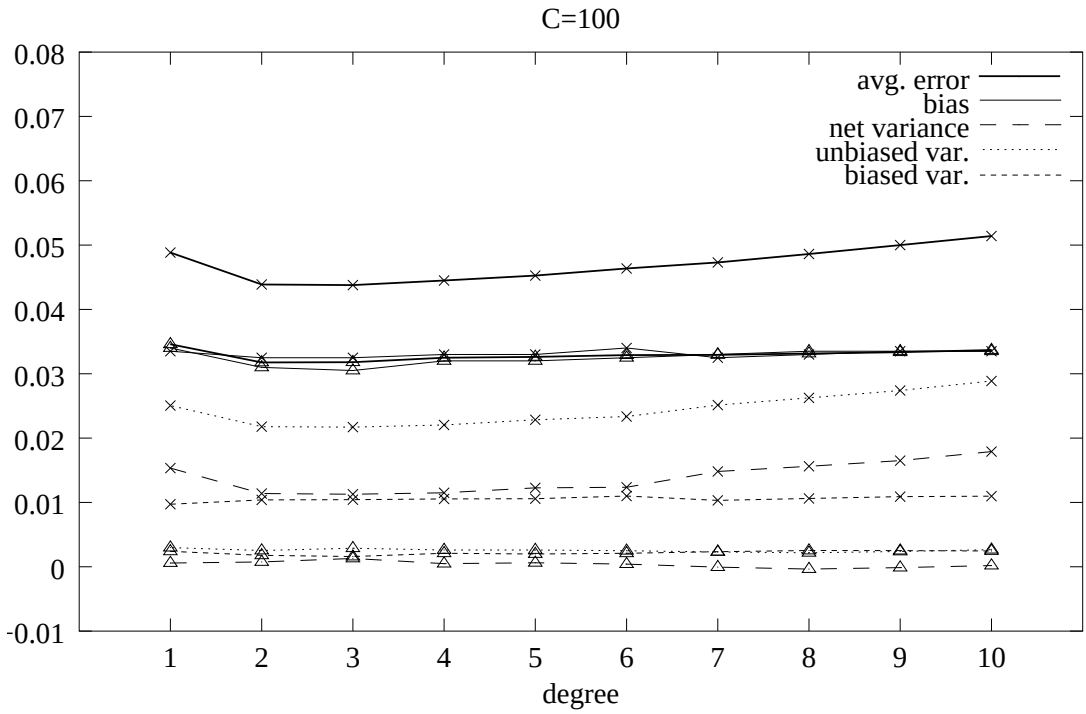


(b)

Figure 71: Waveform data set. Comparison of bias-variance decomposition between single polynomial SVMs (lines labeled with crosses) and RA polynomial SVM ensembles (lines labeled with triangles), while varying the degree and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

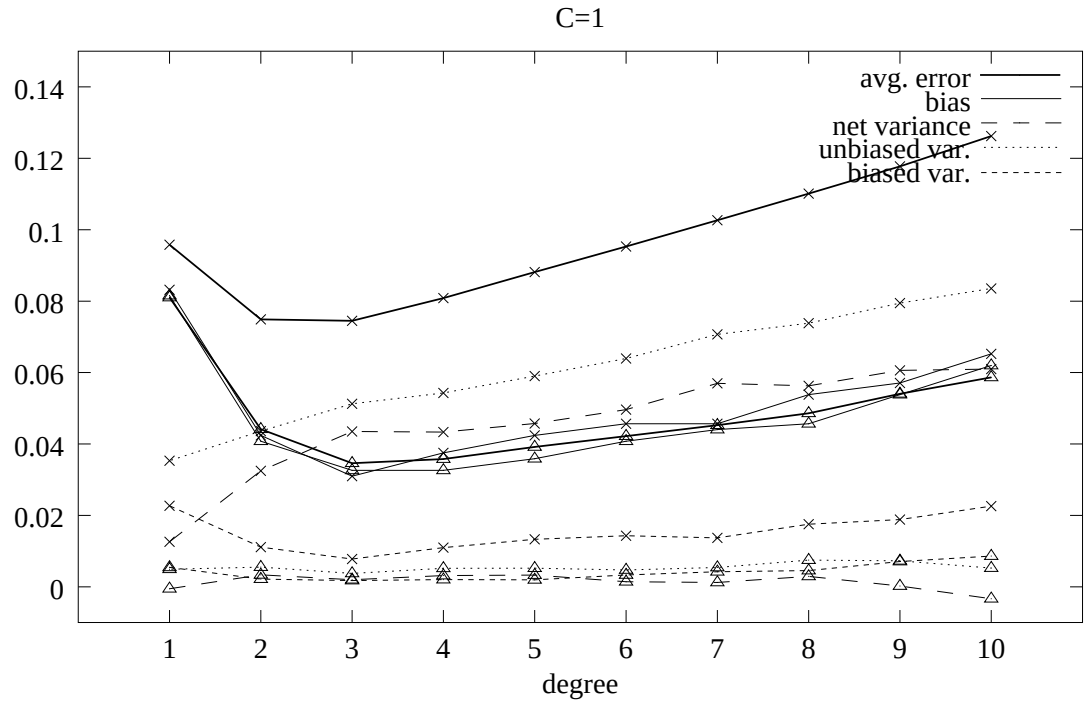


(a)

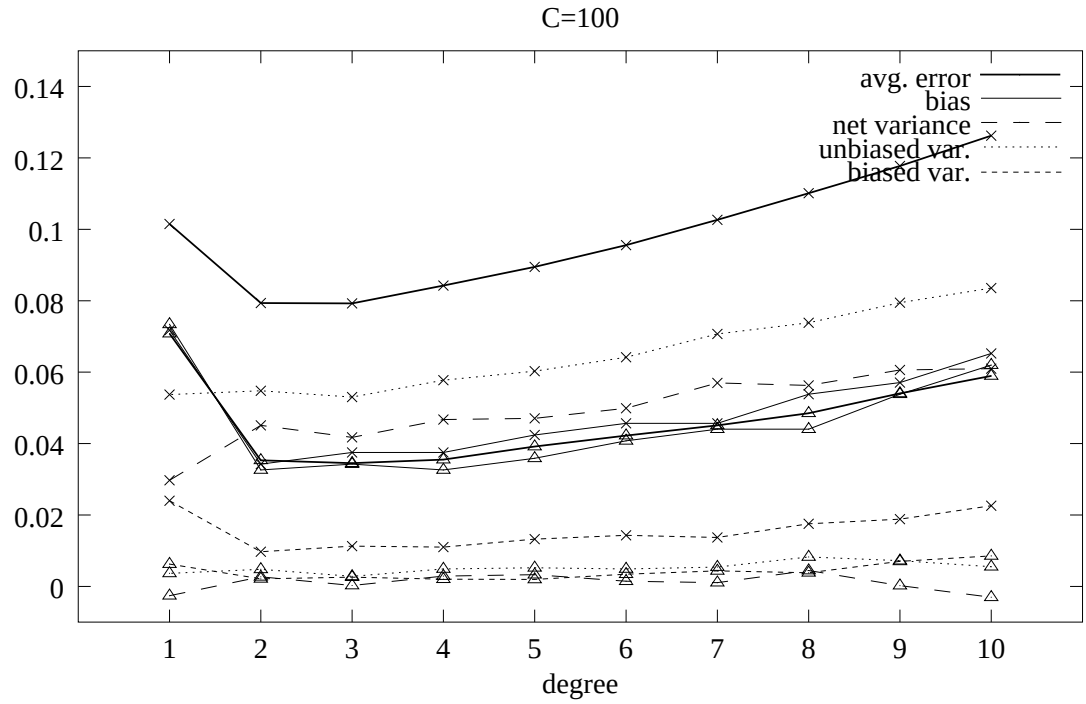


(b)

Figure 72: Grey-Landsat data set. Comparison of bias-variance decomposition between single polynomial SVMs (lines labeled with crosses) and RA polynomial SVM ensembles (lines labeled with triangles), while varying the degree and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

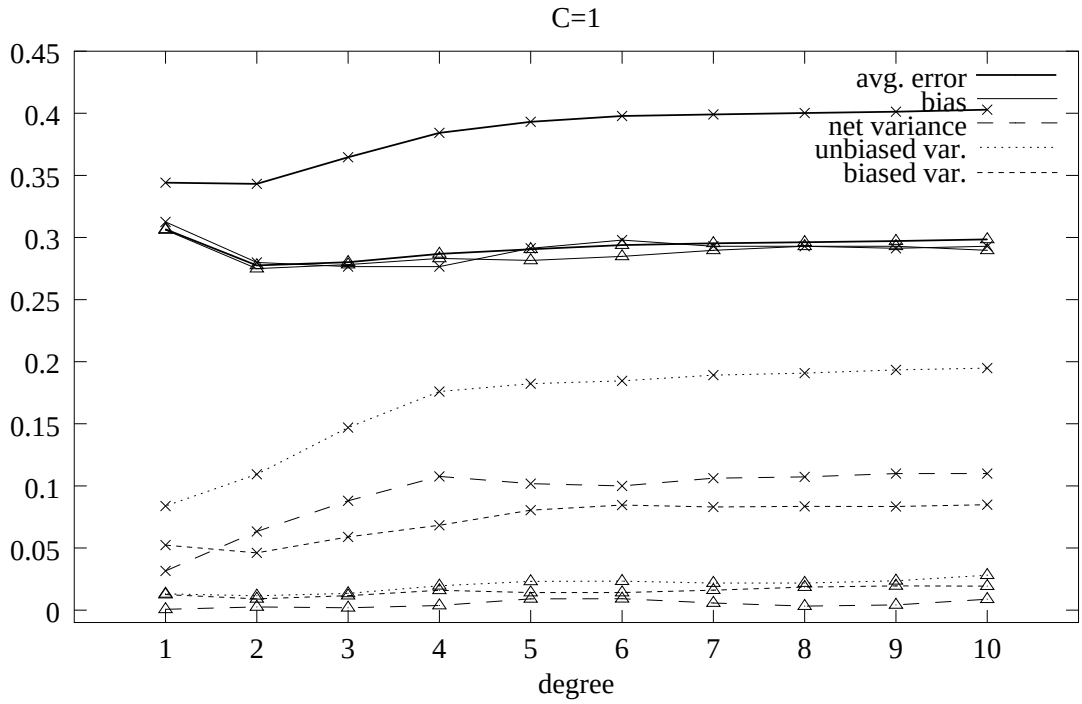


(a)

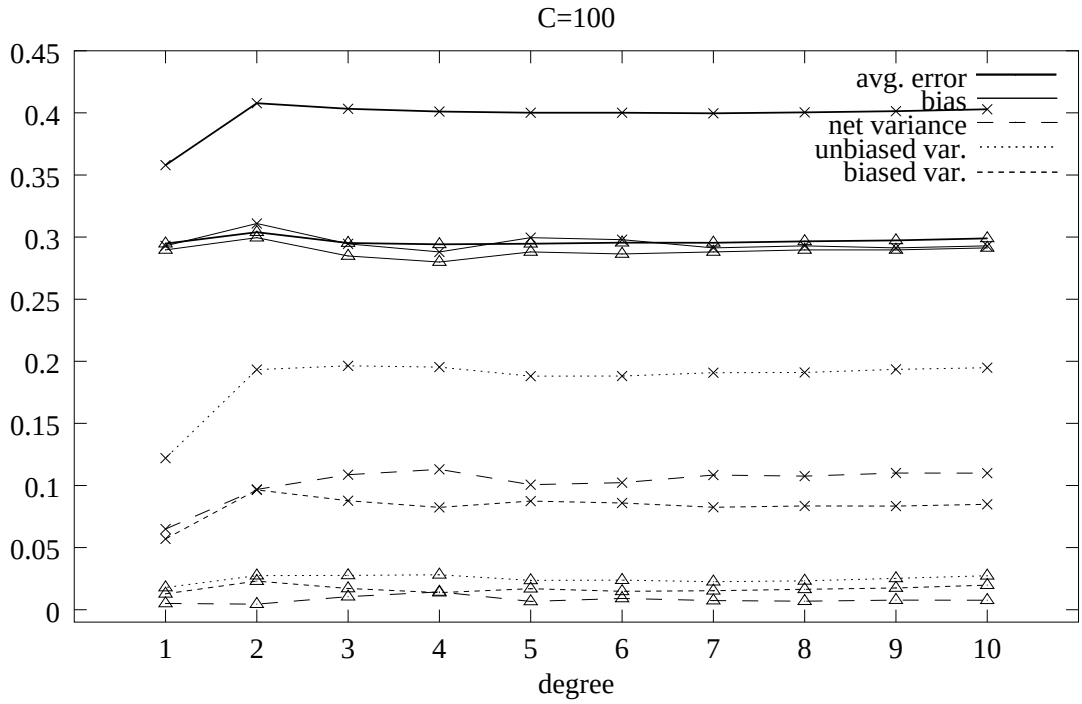


(b)

Figure 73: Letter-Two data set. Comparison of bias-variance decomposition between single polynomial SVMs (lines labeled with crosses) and RA polynomial SVM ensembles (lines labeled with triangles), while varying the degree and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

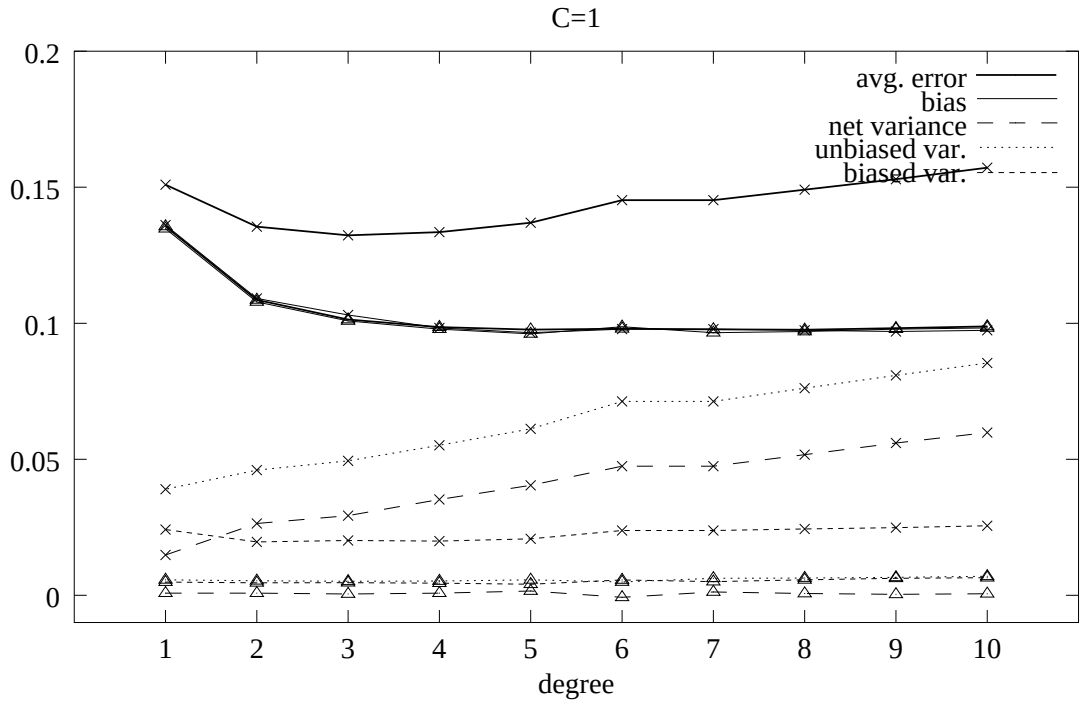


(a)

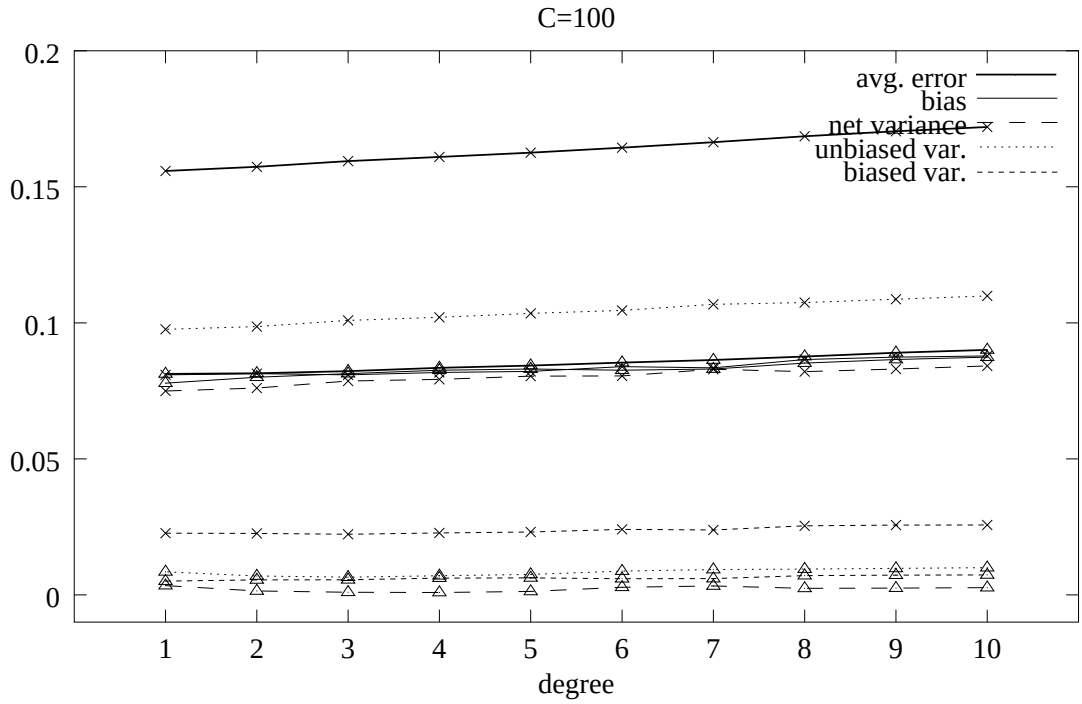


(b)

Figure 74: Letter-Two with noise data set. Comparison of bias-variance decomposition between single polynomial SVMs (lines labeled with crosses) and RA polynomial SVM ensembles (lines labeled with triangles), while varying the degree and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

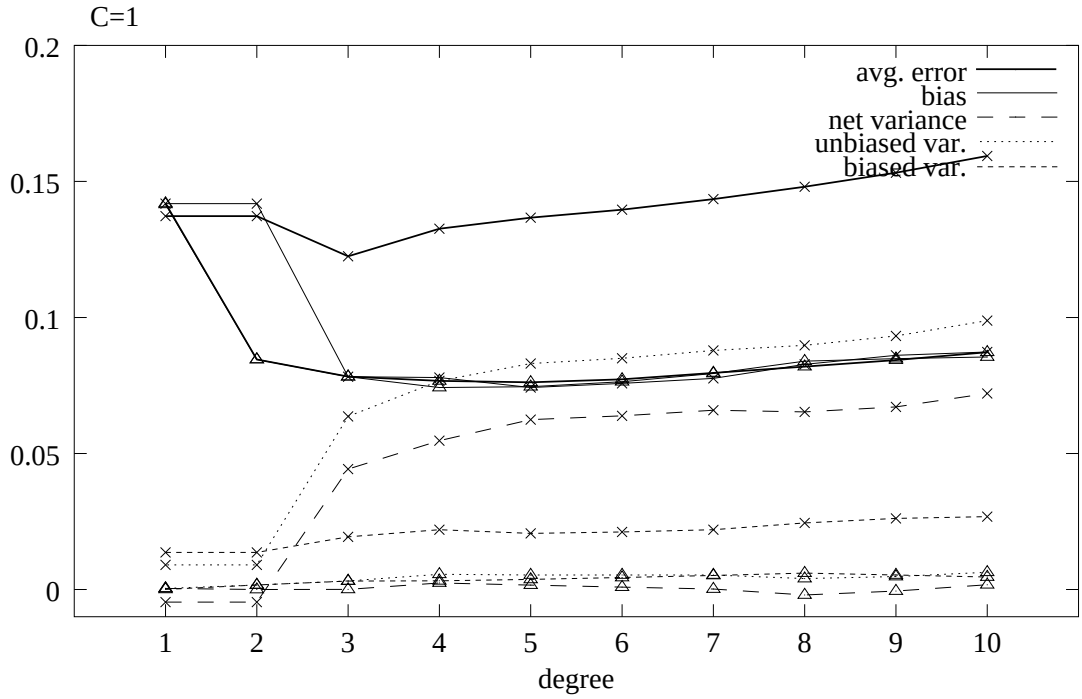


(a)

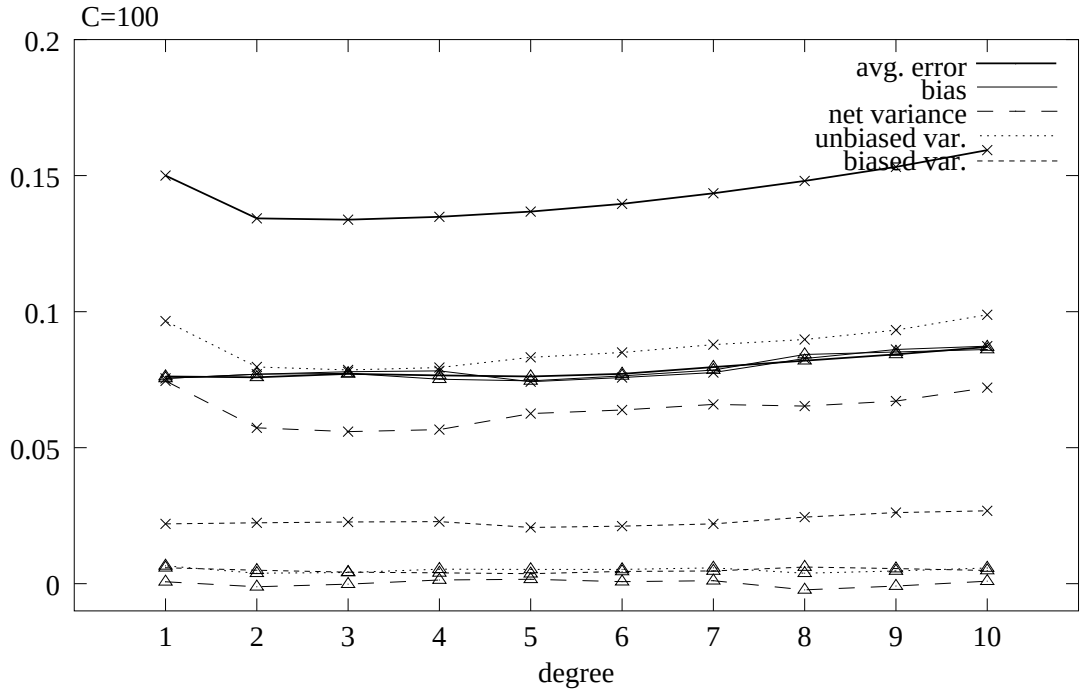


(b)

Figure 75: Spam data set. Comparison of bias-variance decomposition between single polynomial SVMs (lines labeled with crosses) and RA 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)

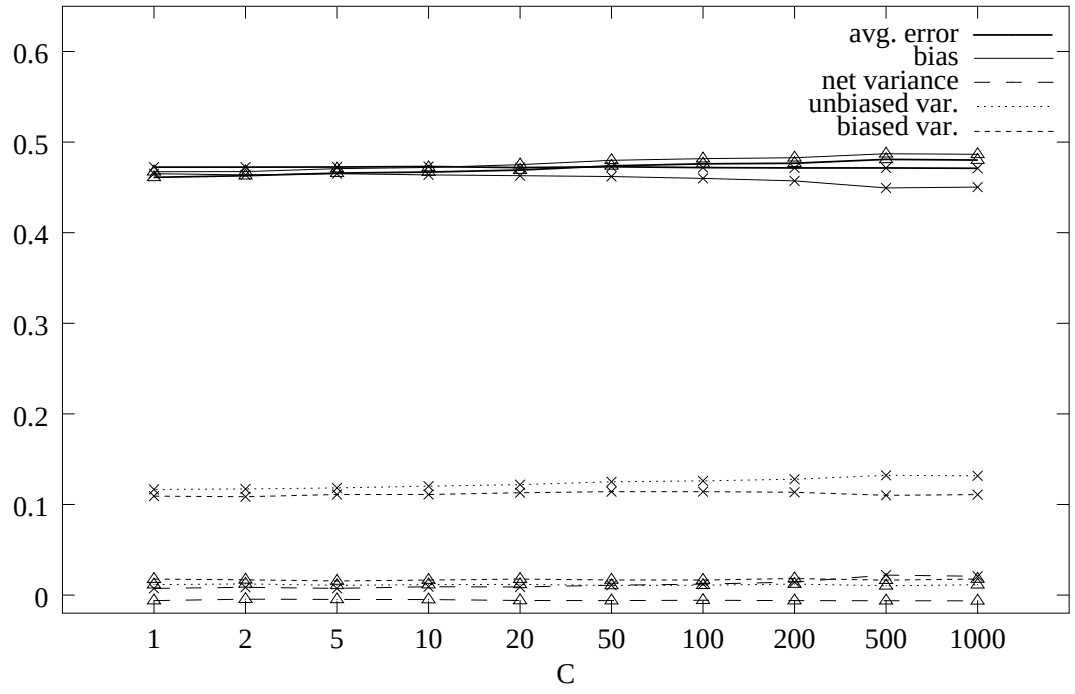


(b)

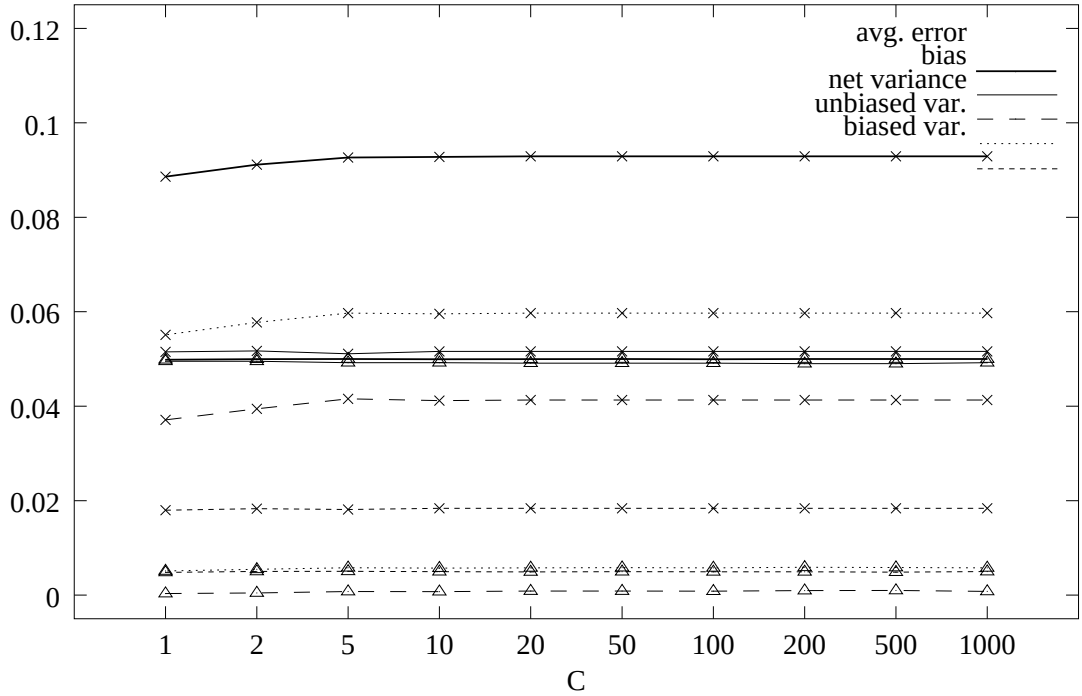
Figure 76: Musk data set. Comparison of bias-variance decomposition between single polynomial SVMs (lines labeled with crosses) and RA polynomial RBF ensembles (lines labeled with triangles), while varying the degree and for some fixed values of C: (a) $C = 1$, (b) $C = 100$.

4.3 Comparison between single and RA 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 $C \geq 1$.

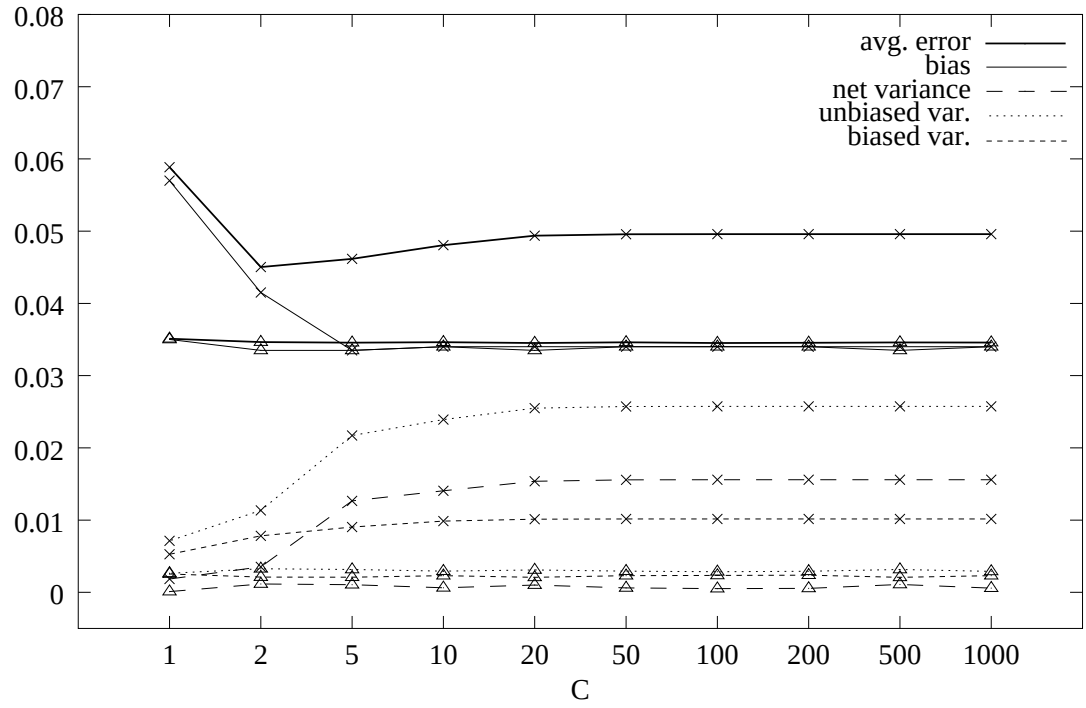


(a)

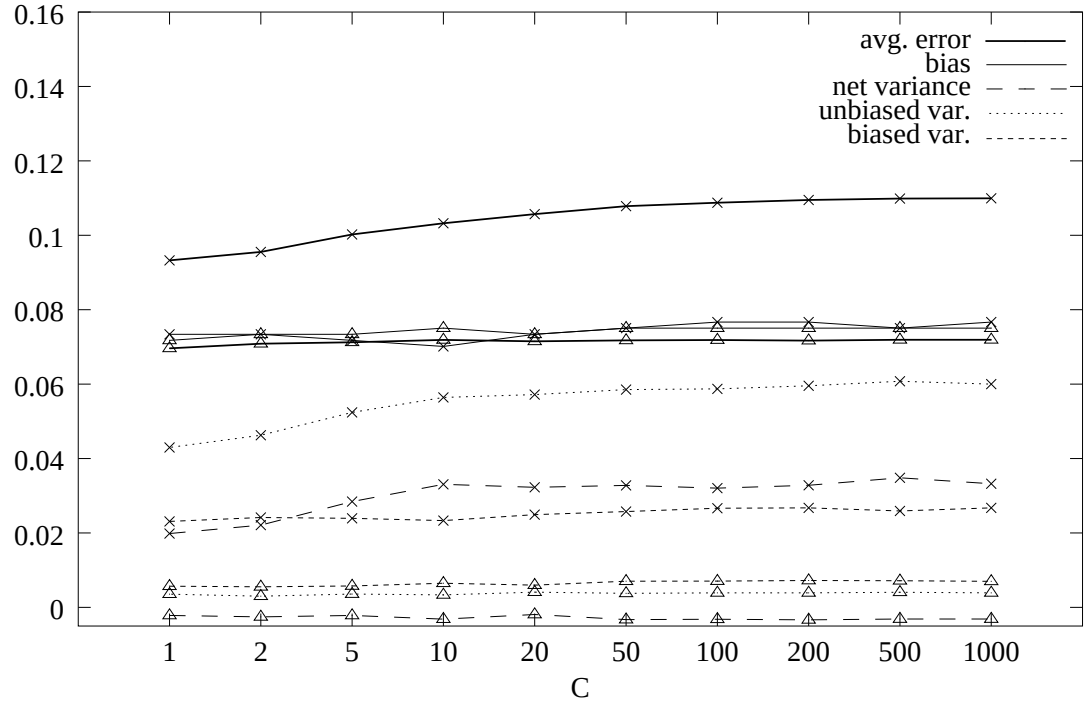


(b)

Figure 77: Comparison of bias-variance decomposition between single dot-product SVMs (lines labeled with crosses) and RA dot-product SVM ensembles (lines labeled with triangles), while varying C: (a) P2, (b) Waveform.

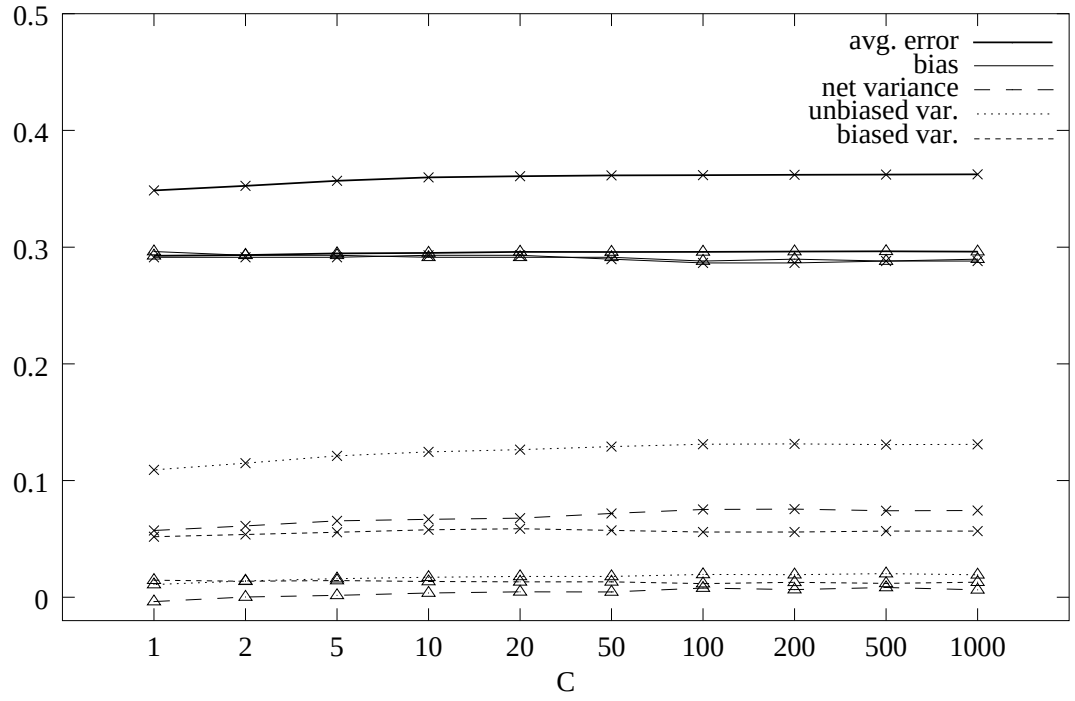


(a)

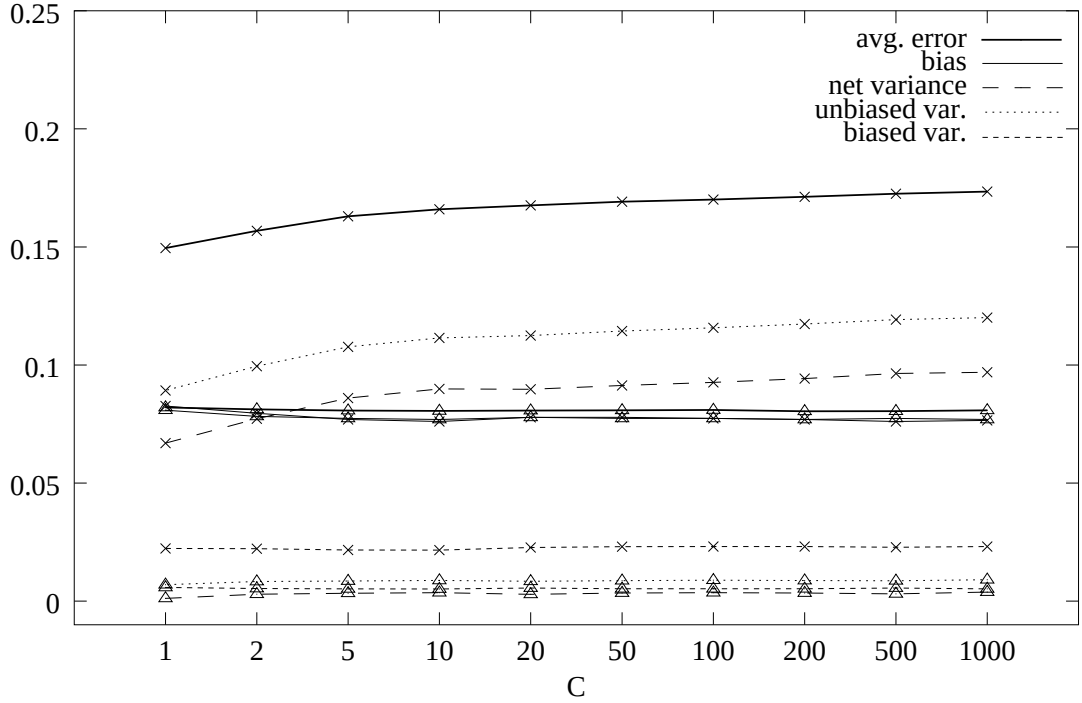


(b)

Figure 78: Comparison of bias-variance decomposition between single dot-product SVMs (lines labeled with crosses) and RA dot-product SVM ensembles (lines labeled with triangles), while varying C: (a) Grey-Landsat, (b) Letter-Two.



(a)



(b)

Figure 79: Comparison of bias-variance decomposition between single dot-product SVMs (lines labeled with crosses) and RA dot-product SVM ensembles (lines labeled with triangles), while varying C: (a) Letter-Two with noise, (b) Spam

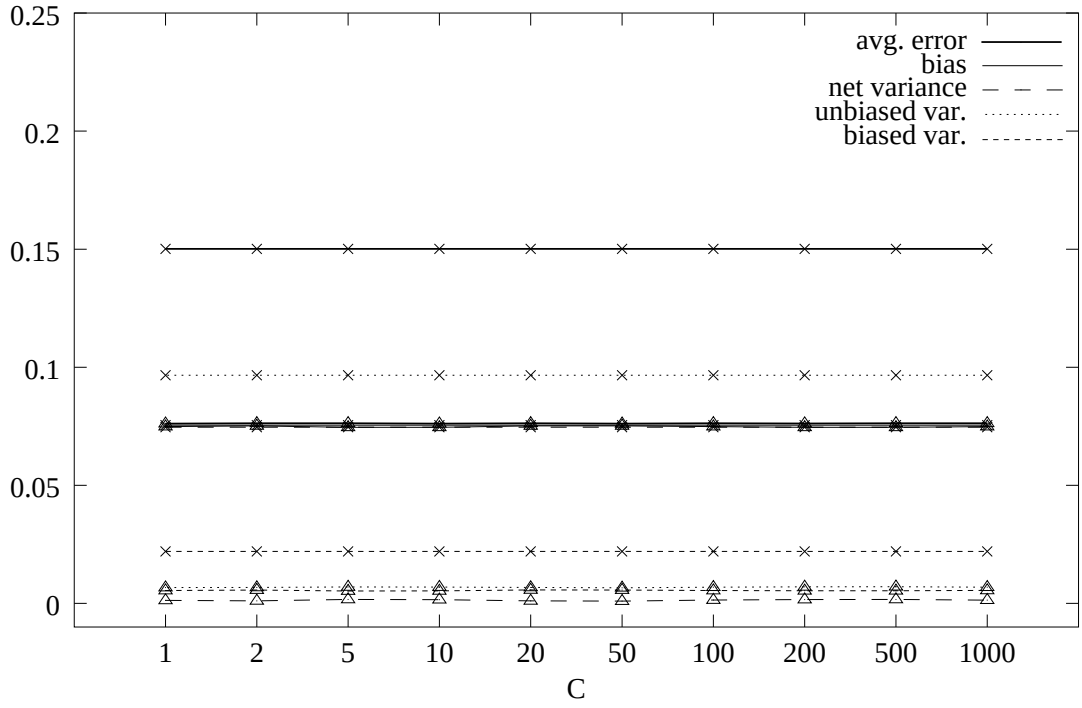


Figure 80: Comparison of bias-variance decomposition between single dot-product SVMs (lines labeled with crosses) and RA dot-product SVM ensembles (lines labeled with triangles), while varying C with Musk data set

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

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

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

$\%Bias\ reduction$, $\%NetVar\ reduction$ and $\%UnbVar\ reduction$ corresponds respectively to the percent bias, net–variance and unbiased variance reduction between single and RA ensemble of SVMs. The negative signs means that we have a larger error in the RA 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 RA RBF-SVMs

The main results of the comparison are:

- The reduction of the error varies from 30 to about 70%, depending on the data set (but for the musk dataset the reduction is limited to 15%.
- We have no error reduction in the high bias region, while in the transition and the stabilized region the reduction of the error is very significant
- The reduction is slightly larger for high values of the C parameter
- The reduction is due primarily to the reduction of the unbiased variance. Indeed in all data sets the reduction of the unbiased varianced amount to about 90%.
- Bias remains substantially unchanged (we have only small reduction or also small increments between single and RA SVMs
- For some data sets (Letter=Two, Letter-Two with noise, SPam and Musk we have more erro reduction with relatively high values of the C parameter

Table 1: Comparison of the bias variance decomposition between single RBF SVM and RA RBF SVM, with the P2 data set.

Parameters σ and $\mathbf{C}$	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
σ	$\mathbf{C = 1}$					
0.01	0.4614	0.4688	-1.62	-0.14	-39.21	99.90
0.02	0.4343	0.4390	-1.07	-0.14	-8.93	99.54
0.1	0.2891	0.0888	69.28	-9.22	93.56	87.78
0.2	0.2154	0.0558	74.05	-4.60	93.89	89.30
0.5	0.1650	0.0595	63.90	1.54	92.33	86.71
1	0.1603	0.0750	53.22	-2.75	95.50	87.73
2	0.1705	0.0962	43.57	-1.08	96.15	88.01
5	0.2008	0.1608	19.89	-3.97	97.33	84.04
10	0.2306	0.1956	15.15	0.76	95.23	85.86
20	0.2814	0.2190	22.15	-2.44	93.12	82.53
50	0.3600	0.2952	17.99	6.37	91.00	73.67
100	0.4097	0.3437	16.10	1.15	87.12	87.04
	$\mathbf{C = 10}$					
0.01	0.4542	0.4631	-1.96	-0.14	-33.69	99.88
0.02	0.4227	0.4150	1.831	-0.14	15.10	99.60
0.1	0.2732	0.0743	72.77	-13.51	94.76	89.18
0.2	0.2048	0.0512	74.97	-4.25	92.77	88.83
0.5	0.1598	0.0498	68.82	4.22	92.29	87.08
1	0.1531	0.0528	65.49	0.65	93.23	87.28
2	0.1517	0.0585	61.39	2.45	93.04	87.02
5	0.1624	0.0892	45.07	-0.93	96.08	87.63
10	0.1849	0.1270	31.29	-0.32	96.23	87.36
20	0.2172	0.1781	18.00	-0.51	97.03	85.65
50	0.2814	0.2238	20.44	0.17	100.9	86.32
100	0.3294	0.2532	23.15	-0.76	83.51	74.78
	$\mathbf{C = 100}$					
0.01	0.4542	0.4646	-2.29	-0.14	-39.66	99.95
0.02	0.4227	0.4163	1.51	-0.14	12.74	99.61
0.1	0.2730	0.0732	73.16	-12.81	95.05	89.38
0.2	0.2048	0.0512	74.98	-5.64	92.88	88.89
0.5	0.1602	0.0495	69.09	6.10	91.91	86.94
1	0.1546	0.0514	66.71	3.55	92.63	87.08
2	0.1534	0.0520	66.07	2.15	93.86	87.80
5	0.1567	0.0650	58.52	1.92	95.75	88.39
10	0.1684	0.0959	43.01	-0.43	95.70	87.05
20	0.1823	0.1253	31.25	-0.97	98.15	88.20
50	0.2324	0.1858	20.06	-0.05	91.91	85.35
100	0.2727	0.2213	18.84	-0.22	99.49	86.59

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

Parameters σ and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
σ	**C = 1**					
0.01	0.5006	0.4953	1.04	2.63	-131.8	66.13
0.02	0.5006	0.4959	0.93	2.63	-140.3	57.10
0.1	0.4967	0.4689	5.60	2.63	143.84	54.16
0.2	0.4961	0.4799	3.27	2.63	27.116	39.01
0.5	0.4993	0.4911	1.64	2.63	-69.40	33.33
1	0.4976	0.4786	3.81	2.63	61.22	36.69
2	0.4712	0.3313	29.69	31.956	18.06	38.64
5	0.0984	0.0514	47.73	5.49	95.81	89.24
10	0.0725	0.0502	30.73	4.79	95.02	87.54
20	0.0709	0.0500	29.51	2.89	98.00	88.58
50	0.0706	0.0501	29.08	1.14	100.58	89.63
100	0.0720	0.0504	29.90	2.30	98.29	89.16
	C = 10					
0.01	0.5006	0.4959	0.93	2.63	-140.3	57.10
0.02	0.5006	0.4955	1.01	2.63	-134.0	63.88
0.1	0.4967	0.4612	7.15	2.63	220.69	57.46
0.2	0.4961	0.4724	4.78	2.63	98.07	32.23
0.5	0.4993	0.4910	1.65	2.63	-68.77	37.64
1	0.4972	0.4781	3.84	2.63	60.46	40.53
2	0.4560	0.2998	34.25	30.22	44.88	46.06
5	0.0947	0.0511	46.01	5.51	95.72	88.96
10	0.0746	0.0499	33.10	3.07	98.97	89.02
20	0.0777	0.0497	36.08	0.952	101.08	90.05
50	0.0802	0.0497	38.05	3.47	99.01	89.85
100	0.0774	0.0497	35.78	3.27	99.18	90.21
	C = 100					
0.01	0.5006	0.4955	1.01	2.63	-134.0	63.88
0.02	0.5006	0.4955	1.01	2.63	-134.0	63.88
0.1	0.4967	0.4541	8.57	2.63	291.08	57.00
0.2	0.4961	0.4763	3.99	2.63	61.14	34.21
0.5	0.4993	0.4903	1.80	2.63	-58.51	30.74
1	0.4972	0.4796	3.54	2.63	44.76	37.43
2	0.4560	0.2983	34.57	30.89	44.29	45.14
5	0.0947	0.0512	45.94	6.08	94.86	88.71
10	0.0746	0.0498	33.19	3.27	98.86	89.08
20	0.0778	0.0497	36.04	0.75	101.28	90.23
50	0.0838	0.0496	40.80	3.29	98.68	89.94
100	0.0870	0.0497	42.86	3.47	98.88	90.33

Table 3: Comparison of the bias variance decomposition between single RBF SVM and RA RBF SVM, with the Grey-Landsat data set.

Parameters σ and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
σ	**C = 1**					
0.01	0.4782	0.461	3.60	0	100	100
0.02	0.4752	0.461	2.99	0	100	100
0.1	0.4774	0.461	3.44	0	100	100
0.2	0.4766	0.461	3.27	0	100	100
0.5	0.4288	0.4463	-4.09	-0.11	-86.75	97.50
1	0.2398	0.1044	56.45	0	65.93	61.98
2	0.0845	0.0603	28.62	4.72	100.7	92.44
5	0.0450	0.0418	7.09	2.32	108.4	86.21
10	0.0386	0.0333	13.66	5.71	89.62	88.89
20	0.0384	0.0339	11.69	0	101.5	91.24
50	0.0410	0.0372	9.18	1.33	92.11	86.65
100	0.0476	0.0460	3.32	-1.09	97.88	84.16
	C = 10					
0.01	0.4782	0.461	3.60	0	100	100
0.02	0.4752	0.461	2.99	0	100	100
0.1	0.4774	0.461	3.43	0	100	100
0.2	0.4759	0.461	3.15	0	100	100
0.5	0.4117	0.4283	-4.04	-0.23	-96.27	96.87
1	0.2220	0.0962	56.64	1.44	66.79	62.59
2	0.0802	0.0581	27.58	4.20	94.49	90.89
5	0.0450	0.0400	11.14	1.23	99.55	85.94
10	0.0397	0.0314	21.00	0	114.8	92.42
20	0.0393	0.0302	23.06	-1.66	102.6	90.65
50	0.0388	0.0307	20.88	3.27	85.11	85.97
100	0.0384	0.0314	18.18	0	93.39	89.75
	C = 100					
0.01	0.4782	0.461	3.60	0	100	100
0.02	0.4752	0.461	2.99	0	100	100
0.1	0.4774	0.461	3.43	0 1	100	100
0.2	0.4759	0.461	3.15	0	100	100
0.5	0.4117	0.4283	-4.02	-0.23	-95.90	96.81
1	0.2220	0.0962	56.63	4.34	66.25	62.29
2	0.0802	0.0582	27.51	4.20	94.22	90.74
5	0.0450	0.04	11.19	1.23	100	85.94
10	0.0398	0.0313	21.22	0	108.1	91.21
20	0.0398	0.0301	24.45	-5.08	108.4	92.29
50	0.0414	0.0301	27.44	-3.38	103.3	90.80
100	0.0417	0.0300	28.03	-1.69	99.71	90.17

Table 4: Comparison of the bias variance decomposition between single RBF SVM and RA RBF SVM, with the Letter-Two data set.

Parameters σ and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
σ	**C = 1**					
0.01	0.4953	0.4809	2.91	-1.36	132.85	85.07
0.02	0.4950	0.4824	2.53	-2.74	125.91	85.45
0.1	0.4942	0.4792	3.05	-3.47	128.33	85.93
0.2	0.4894	0.4679	4.38	-6.13	131.06	88.55
0.5	0.4492	0.3618	19.46	-8.45	114.74	76.53
1	0.3231	0.1144	64.58	-9.37	78.83	72.87
2	0.1686	0.0560	66.78	0	96.72	89.80
5	0.1059	0.0498	52.94	0	95.68	89.93
10	0.0946	0.0454	51.91	3.84	90.96	88.04
20	0.0925	0.0508	45.05	0	92.20	87.98
50	0.1005	0.0708	29.48	-2.5	88.66	83.62
100	0.1164	0.0891	23.46	0	116.48	90.03
	C = 10					
0.01	0.4957	0.4821	2.75	-2.39	129.13	87.41
0.02	0.4951	0.4812	2.81	-2.74	132.01	88.05
0.1	0.4935	0.4769	3.36	-4.21	126.54	87.73
0.2	0.4861	0.4616	5.04	-5.86	124.07	86.34
0.5	0.4343	0.3271	24.67	-14.52	105.11	77.88
1	0.2998	0.1007	66.38	-16.66	82.58	75.49
2	0.1573	0.0534	66.06	-11.53	94.69	89.35
5	0.0980	0.0446	54.42	-3.846	98.82	91.32
10	0.0817	0.0374	54.25	8.69	92.88	90.61
20	0.0744	0.0345	53.56	0	95.32	92.48
50	0.0746	0.0408	45.28	-4.54	91.37	88.75
100	0.0784	0.0552	29.54	-6.45	94.84	85.08
	C = 100					
0.01	0.4957	0.4823	2.71	-2.05	119.56	86.88
0.02	0.4951	0.4822	2.60	-2.74	126.98	85.60
0.1	0.4935	0.4768	3.38	-4.21 1	126.94	87.85
0.2	0.4861	0.4610	5.16	-5.49	121.59	87.74
0.5	0.4343	0.3274	24.60	-13.40	102.61	77.10
1	0.2998	0.0984	67.16	-13.33	82.865	76.06
2	0.1573	0.0537	65.86	-11.53	94.41	89.10
5	0.0980	0.0450	54.09	-3.846	98.24	91.16
10	0.0817	0.0375	54.13	8.69	92.66	90.46
20	0.0752	0.0345	54.15	4.76	95.42	92.32
50	0.0773	0.0361	53.33	-4.99	95.84	92.04
100	0.0810	0.0413	49.01	7.69	94.34	88.21

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

Parameters σ and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
σ	**C = 1**					
0.01	0.5089	0.5212	-2.42	0	-93.54	99.78
0.02	0.5083	0.5212	-2.54	0	-94.15	99.80
0.1	0.5075	0.5211	-2.67	0	-93.45	99.77
0.2	0.5060	0.5199	-2.75	0	-86.83	99.70
0.5	0.4956	0.5068	-2.27	0.31	-64.79	95.15
1	0.4618	0.4250	7.96	4.92	462.84	78.71
2	0.4073	0.3109	23.65	8.41	89.29	78.96
5	0.3725	0.3017	19.00	2.11	101.60	89.02
10	0.3585	0.2993	16.50	4.76	90.26	85.86
20	0.3515	0.2947	16.14	2.18	96.56	86.43
50	0.3561	0.3131	12.05	3.03	103.31	83.62
100	0.3720	0.3411	8.31	2.29	149.16	82.33
	C = 10					
0.01	0.5087	0.5212	-2.46	0	-93.75	99.78
0.02	0.5081	0.5213	-2.60	0	-94.61	99.81
0.1	0.5073	0.5208	-2.65	0	-91.34	99.81
0.2	0.5053	0.5189	-2.70	0	-81.46	99.53
0.5	0.4905	0.4945	-0.82	1.92	-74.95	89.37
1	0.4508	0.3923	12.98	7.72	95.63	75.76
2	0.4040	0.3058	24.30	2.10	98.43	83.33
5	0.3872	0.2972	23.23	2.70	96.84	87.71
10	0.3842	0.2943	23.38	-1.12	100.2	88.46
20	0.3728	0.2879	22.77	2.28	90.64	86.40
50	0.3456	0.2767	19.93	1.19	90.74	88.40
100	0.3361	0.2771	17.55	2.92	90.26	87.04
	C = 100					
0.01	0.5087	0.5212	-2.47	0	-94.00	99.83
0.02	0.5081	0.5213	-2.61	0	-94.96	99.83
0.1	0.5073	0.5207	-2.64	0 1	-90.90	99.79
0.2	0.5053	0.5191	-2.72	0	-82.14	99.49
0.5	0.4905	0.4949	-0.90	1.92	-77.05	89.56
1	0.4508	0.3915	13.15	9.26	74.32	74.06
2	0.4041	0.3052	24.47	2.63	97.33	82.87
5	0.3881	0.2973	23.38	2.15	100.58	88.57
10	0.3914	0.2952	24.57	1.11	95.96	87.50
20	0.3968	0.2940	25.90	4.41	89.13	85.39
50	0.3908	0.2934	24.91	2.80	89.62	85.59
100	0.3699	0.2847	23.02	1.14	96.17	89.29

Table 6: Comparison of the bias variance decomposition between single RBF SVM and RA RBF SVM, with the Spam data set.

Parameters σ and $\mathbf{C}$	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
σ	$\mathbf{C = 1}$					
0.01	0.3980	0.4013	-0.81	0	-100	100
0.02	0.3973	0.4013	-0.99	0	-100	100
0.1	0.3948	0.4013	-1.63	0	-100	100
0.2	0.3923	0.4013	-2.28	0	-100	100
0.5	0.3895	0.4012	-3.01	0	-99.70	100
1	0.3859	0.4012	-3.94	0	-99.54	100
2	0.3805	0.4006	-5.28	0	-97.06	100
5	0.3393	0.3618	-6.60	-0.23	-87.87	89.90
10	0.2484	0.2344	5.614	-1.30	117.00	85.94
20	0.1729	0.1503	13.09	0.28	100.55	87.60
50	0.1392	0.1147	17.56	2.23	98.01	86.90
100	0.1471	0.1292	12.15	6.75	73.60	79.55
	$\mathbf{C = 10}$					
0.01	0.3978	0.4013	-0.86	0	-100	100
0.02	0.3970	0.4013	-1.06	0	-100	100
0.1	0.3930	0.4013	-2.08	0	-100	100
0.2	0.3914	0.4012	-2.52	0	-99.78	100
0.5	0.3878	0.4012	-3.44	0	-99.67	100
1	0.3844	0.4010	-4.32	0	-98.68	100
2	0.3790	0.3995	-5.42	0	-92.29	100
5	0.3210	0.3207	0.11	-0.13	123.98	87.78
10	0.2374	0.2026	14.64	-1.32	91.30	85.98
20	0.1766	0.1332	24.55	0	100.46	91.06
50	0.1391	0.0939	32.48	3.70	92.30	88.35
100	0.1292	0.0899	30.39	0	101.27	90.96
	$\mathbf{C = 100}$					
0.01	0.3979	0.4013	-0.85	0	-100	100
0.02	0.3972	0.4013	-1.02	0	-100	100
0.1	0.3933	0.4013	-2.02	0	-100	100
0.2	0.3915	0.4012	-2.47	0	-99.77	100
0.5	0.3879	0.4012	-3.42	0	-99.67	100
1	0.3845	0.4010	-4.29	0	-98.75	100
2	0.3791	0.3994	-5.35	0	-91.70	100
5	0.3230	0.3211	0.57	-0.53	851.68	89.20
10	0.2415	0.2036	15.70	-1.75	96.56	87.48
20	0.1834	0.1340	26.93	-0.32	101.53	91.90
50	0.1581	0.0938	40.64	-4.87	99.43	92.74
100	0.1546	0.0844	45.37	-3.19	99.82	93.70

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

Parameters σ and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
σ	**C = 1**					
0.01	0.1495	0.1500	-0.31	0	-100	0
0.02	0.1489	0.1500	-0.72	0	-100	0
0.1	0.1482	0.1500	-1.21	0	-100	0
0.2	0.1482	0.1500	-1.23	0	-100	0
0.5	0.1482	0.1500	-1.23	0	-100	0
1	0.1482	0.1500	-1.23	0	-100	0
2	0.1482	0.1500	-1.24	0	-100	0
5	0.1481	0.1500	-1.24	0	-100	0
10	0.1481	0.1500	-1.27	0	-100	100
20	0.1472	0.1500	-1.88	0	-100	100
50	0.1138	0.1133	0.37	-1.36	67.72	76.87
100	0.0997	0.0907	9.02	0.98	106.9	92.51
	C = 10					
0.01	0.1483	0.1500	-1.11	0	-100	0
0.02	0.1482	0.1500	-1.21	0	-100	0
0.1	0.1482	0.1500	-1.23	0	-100	0
0.2	0.1482	0.1500	-1.23	0	-100	0
0.5	0.1482	0.1500	-1.24	0	-100	0
1	0.1482	0.1500	-1.24	0	-100	0
2	0.1482	0.1500	-1.24	0	-100	0
5	0.1481	0.1500	-1.30	0	-100	100
10	0.1476	0.1500	-1.63	0	-100	100
20	0.1426	0.1500	-5.14	0	-100	100
50	0.1013	0.0981	3.10	0.61	90.86	89.67
100	0.0898	0.0754	16.02	0.39	106.7	93.85
	C = 100					
0.01	0.1483	0.1500	-1.11	0	-100	0
0.02	0.1482	0.1500	-1.21	0	-100	0
0.1	0.1482	0.1500	-1.23	0	-100	0
0.2	0.1482	0.1500	-1.23	0	-100	0
0.5	0.1482	0.1500	-1.24	0	-100	0
1	0.1482	0.1500	-1.24	0	-100	0
2	0.1482	0.1500	-1.24	0	-100	0
5	0.1481	0.1500	-1.30	0	-100	100
10	0.1476	0.1500	-1.63	0	-100	100
20	0.1426	0.1500	-5.14	0	-100	100
50	0.1013	0.0980	3.24	0.61	96.19	89.95
100	0.0899	0.0755	16.04	0.39	106.6	93.72

5.2 Comparison between single and RA polynomial SVMs

The main results of the comparison are:

- The reduction of the error varies from 30 to about 55%, depending on the data set.
- The reduction is due primarily to the reduction of the unbiased variance. Indeed in all data sets the reduction of the unbiased varianced amount to about 90%.
- Bias remains substantially unchanged
- For some data sets (P2, Grey-Landsat, Spam) we have more error reduction with relatively high values of the C parameter

Table 8: Comparison of the bias variance decomposition between single polynomial SVM and RA polynomial SVM, with the P2 data set.

Parameters degree and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
degree	**C = 1**					
2	0.3378	0.2868	15.09	5.14	74.97	71.55
3	0.2752	0.2274	17.38	-7.54	98.54	89.50
4	0.2495	0.1919	23.06	-4.45	97.24	84.31
5	0.2198	0.1518	30.92	-2.52	97.46	85.27
6	0.2127	0.1360	36.01	2.28	95.45	86.49
7	0.2105	0.1296	38.40	0.47	96.08	87.15
8	0.2111	0.1259	40.35	-2.51	95.92	86.48
9	0.2113	0.1234	41.61	-4.65	95.68	86.48
10	0.2123	0.1211	42.97	-4.88	97.07	86.59
	C = 10					
2	0.3313	0.2819	14.90	0.62	79.51	76.12
3	0.2621	0.2143	18.23	-6.64	98.64	87.26
4	0.2363	0.1764	25.37	-3.94	96.10	83.76
5	0.2108	0.1338	36.54	-1.75	99.24	88.33
6	0.2087	0.1276	38.86	-5.68	98.59	87.25
7	0.2103	0.1218	42.09	-7.46	97.65	86.59
8	0.2120	0.1153	45.59	-0.09	93.74	85.39
9	0.2130	0.1107	48.00	1.67	95.01	86.16
10	0.2145	0.1078	49.71	2.87	94.20	86.02
	C = 100					
2	0.3310	0.2820	14.79	0.95	77.61	75.42
3	0.2614	0.2118	18.95	-4.34	99.19	87.13
4	0.2346	0.1693	27.83	-4.60	98.20	85.73
5	0.2116	0.1290	39.01	-3.77	97.24	87.13
6	0.2151	0.1186	44.87	-1.79	94.87	86.21
7	0.2148	0.1096	48.95	-1.52	97.04	87.47
8	0.2157	0.1047	51.43	4.00	94.37	86.43
9	0.2161	0.1030	52.31	2.43	94.16	86.39
10	0.2173	0.1034	52.40	4.53	92.59	85.52

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

Parameters degree and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
degree	**C = 1**					
1	0.0720	0.0499	30.67	2.32	97.67	89.05
2	0.0760	0.0497	34.59	3.68	97.12	89.44
3	0.0798	0.0497	37.76	1.54	99.37	90.32
4	0.0812	0.0498	38.63	4.04	98.17	89.69
5	0.0818	0.0498	39.10	3.64	99.24	90.29
6	0.0831	0.0499	39.99	5.73	96.84	89.81
7	0.0856	0.0499	41.63	5.15	97.74	90.39
8	0.0897	0.0500	44.17	4.81	96.94	90.57
9	0.0957	0.0501	47.58	3.85	98.08	91.33
10	0.1039	0.0505	51.40	5.88	97.47	91.42
	C = 10					
1	0.0841	0.0497	40.85	3.86	98.27	90.24
2	0.0859	0.0497	42.15	3.86	98.59	90.21
3	0.0832	0.0497	40.27	2.91	98.46	90.10
4	0.0819	0.0498	39.20	3.66	98.68	90.04
5	0.0819	0.0498	39.17	4.23	97.95	90.01
6	0.0831	0.0499	39.93	4.39	98.92	90.45
7	0.0856	0.0500	41.61	4.96	97.98	90.46
8	0.0897	0.0500	44.20	4.62	97.27	90.70
9	0.0957	0.0502	47.5	4.04	97.68	91.02
10	0.1039	0.0505	51.43	5.88	97.52	91.41
	C = 100					
1	0.0926	0.0499	46.04	3.67	98.14	90.41
2	0.0860	0.0496	42.23	4.63	97.76	89.92
3	0.0832	0.0497	40.25	2.91	98.41	89.97
4	0.0819	0.0498	39.15	3.66	98.53	89.96
5	0.0819	0.0498	39.15	4.23	97.90	89.87
6	0.0831	0.0499	39.94	4.58	98.62	90.35
7	0.0856	0.0499	41.65	5.15	97.79	90.40
8	0.0897	0.0500	44.19	4.62	97.24	90.69
9	0.0957	0.0502	47.47	3.46	98.30	91.35
10	0.1039	0.0504	51.50	5.88	97.66	91.47

Table 10: Comparison of the bias variance decomposition between single polynomial SVM and RA polynomial SVM, with the Grey-Landsat data set.

Parameters degree and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
degree	**C = 1**					
1	0.0502	0.0467	7.03	0	94.14	86.50
2	0.0410	0.0366	10.87	0	97.59	92.57
3	0.0392	0.0328	16.28	-3.17	95.78	90.23
4	0.0406	0.0326	19.51	-4.76	103.5	90.78
5	0.0423	0.0326	22.91	3.12	84.06	84.24
6	0.0445	0.0327	26.59	1.49	102.6	89.81
7	0.0465	0.0329	29.19	0	100.3	90.26
8	0.0483	0.0331	31.38	-1.51	102.0	91.49
9	0.0499	0.0334	33.07	0	100.6	91.12
10	0.0513	0.0336	34.60	0	99.44	91.00
	C = 10					
1	0.0488	0.0345	29.19	-1.49	96.21	88.10
2	0.0438	0.0317	27.64	4.61	93.36	88.35
3	0.0437	0.0318	27.35	6.15	88.47	86.68
4	0.0445	0.0324	26.98	3.03	95.69	88.20
5	0.0452	0.0326	27.94	3.03	94.94	88.60
6	0.0463	0.0329	29.03	4.41	96.72	89.29
7	0.0473	0.0329	30.36	-1.53	100.3	90.87
8	0.0486	0.0331	31.82	-1.51	102.2	91.69
9	0.0499	0.0333	33.27	0	100.8	91.33
10	0.0513	0.0336	34.44	0	98.91	90.80
	C = 100					
1	0.0488	0.0345	29.19	-1.49	96.21	88.10
2	0.0438	0.0317	27.64	4.61	93.36	88.35
3	0.0437	0.0318	27.35	6.15	88.47	86.68
4	0.0445	0.0324	26.98	3.03	95.69	88.20
5	0.0452	0.0326	27.94	3.03	94.94	88.60
6	0.0463	0.0329	29.03	4.41	96.72	89.29
7	0.0473	0.0329	30.36	-1.53	100.3	90.87
8	0.0486	0.0331	31.82	-1.51	102.2	91.69
9	0.0499	0.0333	33.27	0	100.8	91.33
10	0.0513	0.0336	34.44	0	98.91	90.80

Table 11: Comparison of the bias variance decomposition between single polynomial SVM and RA polynomial SVM, with the Letter-Two data set.

Parameters degree and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
degree	**C = 1**					
1	0.0957	0.0810	15.41	1.96	104.2	86.08
2	0.0748	0.0441	41.05	3.84	89.65	87.35
3	0.0744	0.0346	53.54	-5.26	95.46	92.71
4	0.0808	0.0357	55.73	13.04	92.69	90.47
5	0.0881	0.0391	55.56	15.38	92.83	91.15
6	0.0953	0.0421	55.73	10.71	97.17	92.62
7	0.1026	0.0452	55.91	3.57	97.88	92.33
8	0.1101	0.0485	55.86	15.15	94.81	89.87
9	0.1177	0.0540	54.07	5.71	99.62	90.82
10	0.1262	0.0586	53.53	5.00	105.4	93.71
	C = 10					
1	0.0911	0.0686	24.67	4.34	119.4	91.56
2	0.0782	0.0349	55.32	9.09	94.46	91.75
3	0.0792	0.0343	56.66	8.69	99.80	94.64
4	0.0842	0.0356	57.67	13.04	93.50	91.24
5	0.0894	0.0391	56.20	15.38	92.99	91.31
6	0.0955	0.0420	56.00	10.71	97.48	92.85
7	0.1026	0.0451	55.99	3.57	98.02	92.36
8	0.1101	0.0486	55.85	18.18	91.88	88.81
9	0.1177	0.0538	54.25	5.71	99.97	91.15
10	0.1262	0.0591	53.15	5.00	104.6	93.41
	C = 100					
1	0.1014	0.0707	30.24	-2.27	108.7	93.22
2	0.0793	0.0353	55.50	4.76	94.06	91.27
3	0.0792	0.0345	56.46	8.69	99.41	94.67
4	0.0842	0.0354	57.86	13.04	93.85	91.55
5	0.0894	0.0391	56.22	15.38	93.03	91.34
6	0.0955	0.0422	55.78	10.71	97.05	92.37
7	0.1026	0.0451	56.06	3.57	98.13	92.33
8	0.1101	0.0484	55.97	18.18	92.11	88.81
9	0.1177	0.0540	54.09	5.71	99.65	90.90
10	0.1262	0.0589	53.29	5.00	104.9	93.40

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

Parameters **degree** and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
degree	**C = 1**					
1	0.3441	0.3066	10.89	2.09	98.13	84.38
2	0.3431	0.2775	19.13	1.75	95.96	89.43
3	0.3646	0.2800	23.18	-0.59	97.89	90.84
4	0.3842	0.2868	25.35	-2.36	96.57	88.86
5	0.3931	0.2905	26.09	3.37	91.11	87.29
6	0.3978	0.2939	26.11	4.39	90.84	87.38
7	0.3991	0.2953	25.99	1.11	94.66	88.51
8	0.4001	0.2962	25.98	0	96.97	88.57
9	0.4012	0.2971	25.96	-0.56	96.23	87.80
10	0.4028	0.2984	25.91	1.11	91.99	85.60
	C = 10					
1	0.3453	0.2927	15.25	1.66	94.03	86.53
2	0.3859	0.2941	23.78	2.74	94.89	87.03
3	0.3995	0.2930	26.64	2.84	88.13	85.49
4	0.4005	0.2945	26.47	4.46	86.38	85.12
5	0.4000	0.2945	26.37	4.91	90.28	86.52
6	0.4001	0.2956	26.11	3.84	90.95	87.15
7	0.3996	0.2955	26.04	1.12	93.05	88.08
8	0.4004	0.2961	26.04	0.55	95.49	88.42
9	0.4013	0.2976	25.83	0.56	92.74	86.86
10	0.4028	0.2991	25.75	0	94.40	86.28
	C = 100					
1	0.3579	0.2947	17.65	1.11	92.21	85.39
2	0.4077	0.3039	25.45	3.68	95.40	85.78
3	0.4032	0.2952	26.78	3.33	90.34	85.95
4	0.4010	0.2941	26.65	2.84	87.35	85.64
5	0.4001	0.2946	26.35	3.82	93.42	87.47
6	0.4001	0.2954	26.15	3.84	91.13	87.33
7	0.3996	0.2953	26.09	1.12	93.23	88.17
8	0.4004	0.2965	25.94	1.11	93.62	87.82
9	0.4013	0.2973	25.90	0.56	93.01	86.97
10	0.4028	0.2989	25.80	0.55	93.07	85.96

Table 13: Comparison of the bias variance decomposition between single polynomial SVM and RA polynomial SVM, with the Spam data set.

Parameters degree and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
degree	**C = 1**					
1	0.1509	0.1355	10.15	0.95	94.52	85.45
2	0.1355	0.1086	19.85	1.19	97.03	88.35
3	0.1323	0.1013	23.38	2.10	98.27	89.57
4	0.1334	0.0986	26.13	0.44	97.80	90.50
5	0.1369	0.0976	28.66	0.45	96.04	90.69
6	0.1452	0.0979	32.56	-0.88	101.5	93.13
7	0.1452	0.0977	32.71	1.33	97.41	91.21
8	0.1490	0.0976	34.51	0.44	98.68	91.64
9	0.1529	0.0981	35.80	-0.89	99.37	91.82
10	0.1571	0.0988	37.13	-0.89	99.05	91.78
	C = 10					
1	0.1325	0.0919	30.64	0	100.6	92.82
2	0.1369	0.0855	37.52	-0.51	98.51	91.34
3	0.1460	0.0843	42.29	1.03	97.96	91.91
4	0.1521	0.0845	44.44	-1.04	99.11	92.61
5	0.1564	0.0853	45.47	-0.52	96.35	91.87
6	0.1593	0.0866	45.62	0.52	94.73	90.89
7	0.1614	0.0876	45.74	3.04	93.96	90.36
8	0.1636	0.0888	45.73	5.41	92.93	89.63
9	0.1659	0.0904	45.52	1.91	98.28	91.16
10	0.1678	0.0914	45.51	0.47	99.25	91.49
	C = 100					
1	0.1558	0.0812	47.86	3.76	95.45	91.30
2	0.1573	0.0814	48.24	1.60	98.11	92.94
3	0.1594	0.0822	48.42	-0.53	98.79	93.54
4	0.1609	0.0834	48.14	-1.06	98.89	93.11
5	0.1625	0.0843	48.13	-1.05	98.42	92.74
6	0.1644	0.0853	48.05	1.55	96.53	91.65
7	0.1664	0.0863	48.11	0.52	96.01	91.29
8	0.1686	0.0876	48.03	1.50	97.08	91.18
9	0.1704	0.0890	47.77	0.99	97.00	91.01
10	0.1719	0.0900	47.64	0.49	96.83	90.88

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

Parameters **degree** and **C**	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	%**Error reduction**	%**Bias reduction**	%**NetVar reduction**	%**UnbVar reduction**
degree	**C = 1**					
1	0.1372	0.1419	-3.41	0.21	-108.28	95.40
2	0.1372	0.0845	38.39	40.38	-99.73	81.70
3	0.1224	0.0782	36.09	0	99.86	95.04
4	0.1326	0.0766	42.18	4.66	95.59	92.67
5	0.1366	0.0761	44.26	-0.40	97.40	93.57
6	0.1396	0.0772	44.65	-0.80	98.61	93.74
7	0.1435	0.0795	44.53	-2.34	99.72	93.91
8	0.1480	0.0819	44.63	-1.46	103.0	95.49
9	0.1531	0.0843	44.95	1.40	100.8	94.88
10	0.1593	0.0872	45.26	2.08	97.56	93.58
	C = 10					
1	0.1209	0.0762	36.94	2.72	98.85	93.04
2	0.1339	0.0759	43.34	-0.39	101.9	95.23
3	0.1337	0.0770	42.40	0.77	100.4	94.73
4	0.1348	0.0766	43.18	3.48	97.99	93.40
5	0.1396	0.0760	45.53	2	97.20	93.70
6	0.1396	0.0772	44.66	-1.2	99.09	93.96
7	0.1435	0.0797	44.45	-0.78	97.70	93.08
8	0.1480	0.0819	44.65	-1.83	103.5	95.74
9	0.1531	0.0843	44.93	1.05	101.2	95.02
10	0.1593	0.0868	45.47	1.73	98.46	94.03
	C = 100					
1	0.1500	0.0761	49.22	0	99.07	93.21
2	0.1342	0.0758	43.50	0	101.9	95.18
3	0.1337	0.0771	42.34	0.77	100.2	94.79
4	0.1348	0.0765	43.22	3.875	97.55	93.26
5	0.1367	0.0761	44.30	-0.40	97.42	93.66
6	0.1396	0.0771	44.74	-0.80	98.80	93.81
7	0.1435	0.0796	44.52	-1.17	98.32	93.39
8	0.1480	0.0819	44.61	-1.83	103.4	95.69
9	0.1531	0.0843	44.96	1.05	101.3	94.96
10	0.1593	0.0869	45.41	1.38	98.74	94.15

5.3 Comparison between single and RA dot-product SVMs

he main results of the comparison are:

- The reduction of the error varies from 30 to about 50%, depending on the data set. In Letter-Two with noise the error reduction is limited to 18%, while in P2 we have an high error (due to a high bias) and no error reduction
- The reduction of the error is due primarily to the reduction of the unbiased variance. Indeed in all data sets the reduction of the unbiased varianced amount to about 90%.
- Bias remains substantially unchanged
- For some data sets (P2, Grey-Landsat, Spam) we have more error reduction with relatively high values of the C parameter

Table 15: Comparison of the bias variance decomposition between single dot-prod SVM and RA dot-prod SVM, with the P2 data set.

ParameterC	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
1	0.4724	0.4611	2.39	-0.47	183.36	90.15
2	0.4724	0.4628	2.02	-0.77	153.70	89.51
5	0.4724	0.4658	1.40	-1.20	165.66	90.91
10	0.4730	0.4670	1.26	-1.76	154.28	90.36
20	0.4719	0.4691	0.58	-2.61	166.56	90.37
50	0.4728	0.4737	-0.19	-3.87	156.29	91.47
100	0.4718	0.4760	-0.89	-4.76	147.55	91.39
200	0.4715	0.4766	-1.09	-5.59	142.64	90.45
500	0.4715	0.4809	-1.98	-8.38	128.49	92.26
1000	0.4711	0.4801	-1.91	-8.01	130.67	91.33

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

ParameterC	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
1	0.0886	0.0498	43.74	3.84	99.12	90.69
2	0.0911	0.0499	45.17	4.21	98.85	90.54
5	0.0926	0.0499	46.05	3.67	98.15	90.29
10	0.0927	0.0499	46.15	4.60	98.21	90.41
20	0.0929	0.0499	46.20	4.80	97.92	90.31
50	0.0929	0.0499	46.20	4.80	97.91	90.23
100	0.0929	0.0499	46.23	4.80	97.99	90.34
200	0.0929	0.0499	46.19	4.99	97.66	90.12
500	0.0929	0.0500	46.17	4.99	97.62	90.18
1000	0.0929	0.0500	46.18	4.60	98.11	90.33

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

ParameterC	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
1	0.0450	0.0351	22.03	15.66	97.16	76.80
2	0.0461	0.0346	24.93	0	90.91	84.95
5	0.0480	0.0345	28.07	1.47	92.45	86.78
10	0.0493	0.0346	29.85	0	95.87	88.48
20	0.0495	0.0345	30.38	1.47	93.48	87.99
50	0.0495	0.0346	30.19	0	96.05	88.61
100	0.0495	0.0345	30.39	0	96.69	88.92
200	0.0495	0.0345	30.34	0	96.53	88.67
500	0.0495	0.0345	30.24	1.47	93.00	87.70
1000	0.0495	0.0345	30.25	0	96.24	88.73

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

ParameterC	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
1	0.0955	0.0696	27.11	2.22	109.73	92.31
2	0.1002	0.0708	29.28	-2.27	108.88	94.24
5	0.1032	0.0712	30.99	-4.65	106.55	93.63
10	0.1056	0.0718	31.98	-2.22	109.75	94.12
20	0.1078	0.0715	33.67	2.17	105.77	92.97
50	0.1087	0.0717	33.99	2.12	110.17	93.55
100	0.1094	0.0718	34.35	2.12	109.64	93.45
200	0.1098	0.0716	34.75	0	109.59	93.58
500	0.1099	0.0719	34.55	2.12	109.32	93.20
1000	0.1099	0.0719	34.55	2.12	109.32	93.47

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

ParameterC	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
1	0.3485	0.2925	16.07	-1.68	106.4	89.97
2	0.3524	0.2932	16.81	-0.56	99.57	87.79
5	0.3568	0.2945	17.43	-0.56	97.50	86.81
10	0.3597	0.2950	17.98	0.55	94.48	86.21
20	0.3607	0.2959	17.96	0.55	93.14	85.88
50	0.3614	0.2958	18.14	-0.56	93.66	86.20
100	0.3616	0.2959	18.16	-0.57	89.47	85.04
200	0.3619	0.2962	18.16	-1.14	91.33	85.24
500	0.3621	0.2964	18.14	0	88.69	84.52
1000	0.3623	0.2960	18.29	-0.56	91.39	85.27

Table 20: Comparison of the bias variance decomposition between single dot-prod SVM and RA dot-prod SVM, with the Spam data set.

ParameterC	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
1	0.1495	0.0820	45.14	2.10	98.26	92.16
2	0.1568	0.0812	48.21	1.63	96.17	91.65
5	0.1629	0.0807	50.46	-0.56	96.11	92.07
10	0.1659	0.0805	51.44	-1.14	95.97	92.14
20	0.1675	0.0806	51.84	0	96.80	92.49
50	0.1691	0.0808	52.22	0.55	96.25	92.41
100	0.1700	0.0809	52.38	0	96.13	92.35
200	0.1712	0.0804	53.04	0	96.34	92.58
500	0.1725	0.0804	53.35	-1.71	96.81	92.78
1000	0.1734	0.0808	53.41	-0.56	96.03	92.45

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

ParameterC	$\mathbf{E_{SVM}}$	$\mathbf{E_{RA}}$	**%Error reduction**	**%Bias reduction**	**%NetVar reduction**	**%UnbVar reduction**
1	0.1501	0.0761	49.28	0.80	98.30	93.03
2	0.1501	0.0762	49.19	0.40	98.53	92.99
5	0.1501	0.0762	49.19	1.20	97.72	92.78
10	0.1501	0.0761	49.28	1.20	97.89	92.86
20	0.1501	0.0762	49.19	0.40	98.54	92.98
50	0.1501	0.0761	49.28	0.40	98.70	93.15
100	0.1501	0.0762	49.18	0.80	98.11	92.91
200	0.1501	0.0761	49.24	1.20	97.82	92.76
500	0.1501	0.0762	49.19	1.20	97.72	92.69
1000	0.1501	0.0762	49.20	0.80	98.14	92.88