Supplementary Data

I. Using QVZ with trimmed or hard-clipped reads

The current implementation of *QVZ* works with reads of equal length. However, it is possible to compress trimmed or hard-clipped reads if "padding" on those reads is performed prior to compression. For best results, we recommend padding with a fixed q-value (which must be from the same alphabet as the original reads). Note that the user also needs to store which reads are padded and by what amount in order to be able to perform the reverse operation after decompression.

II. Arithmetic Encoder

As described in the main manuscript, for the last step of the proposed lossy compressor *QVZ*, we use an arithmetic encoder. The input to the arithmetic encoder is a probability model P and a value to be compressed, denoted by X. The arithmetic encoder compresses the value X with roughly $-\log_2(P(X))$ bits, thus achieving a compression rate close to the entropy. For a more detailed explanation on how the arithmetic encoder uses the model P to compute the output bits, we refer the reader to [Amir Said, 2004].

In the proposed lossy compressor, a different probability model P is used per column and per previous quantized value. That is, there are as many models to compress the quantized values in column i as possible quantized values column $i-1$ can take. Thus, for each value to be compressed the algorithm chooses the corresponding model based on the previous quantized value and the current column. This, together with the value, are the input to the arithmetic encoder.

III. Data sets

In this section we give more details about the data sets used in this work. Specifically, we show in Table I the different data sets, and specify their size, the number of reads and the length of the reads.

Name	Size [GB]	Num. of Reads	Length of Reads	Coverage
NA12878.HiSeq.WGS.bwa.cleaned.recal.hg19.20.bam	5.7	51,585,658	101	82.7×
SRR622461chr20.bam	0.198	3,262,065	101	5.2×
SRR32209.bam	0.715	18,668,648	36	0.25×
SRR1179906.fastq	4.4	16,586,358	100	132.7×
ERR011354.fastq	2.0	11,433,255	36	3.35×

TABLE I: Data sets used in this work.

The file *NA12878.HiSeq.WGS.bwa.cleaned.recal.hg19.20.bam*, hereafter denoted as *NA12878chr20*, corresponds to the chromosome 20 of a *H. Sapiens* individual, and it was downloaded from the GATK bundle (*www.broadinstitute.org/gatk*). Also corresponding to the chromosome 20 of the same *H. Sapiens* is the data set *SRR622461chr20.bam*[1]. The data set *SRR32209.bam*[2] belongs to the *M. Musculus* species. *SRR1179906.fastq*[3] belongs to the *Saccharomyces Cerevisiae* species, and the data set *ERR011354.fastq*[4]

[1]Downloaded from *ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/data/NA12878/sequence_read/*

[2]Downloaded from *http://trace.ddbj.nig.ac.jp/DRASearch/run?acc=SRR032209*

[3]Downloaded from *http://trace.ddbj.nig.ac.jp/DRASearch/run?acc=SRR1179906*.

[4]Downloaded from *http://trace.ddbj.nig.ac.jp/DRASearch/run?acc=ERR011354*

to the *Drosophila melanogaster* species.

Regarding *SRR622461chr20.bam* and *SRR32209.bam*, we first downloaded the corresponding FASTQ files from the links provided above, and then used an alignment program to generate them. Specifically, we used Bowtie2 [Langmead *et al*., 2014]. For the *M. Musculus* data set, we modified the script *make_m_musculus_ncbi37.sh* provided by Bowtie2 to use the latest release of the *M. Musculus GRCm38*, instead of the release *ncbi37*, and use it to generate the index. To generate the index of the *H. Sapiens* we used the script provided by Bowtie2. We then run Bowtie2 with the following command:

./bowtie2 –fast –threads 2 –mm -x *index* -U *fastqFile* > *samFile*

Finally, to generate the different sorted SAM files by position, used as input by some of the considered compressor algorithms, we run the following commands with samtools [Li *et al*., 2014]:

samtools view -bS samFile — samtools sort - sortedBamFile

samtools view -h sortedBamFile > sortedSamFile

IV. Compression programs

In this section of the *Supplementary Data* we provide more details of the state of the art compression algorithms used for comparison in the main manuscript, that is, *QualComp* [Ochoa *et al*., 2014], the two algorithms *Pblock* and *Rblock* presented in [Canovas *et al*., 2014], *CRAM* [Fritz *et al*., 2011] and *DSRC2* [Roguski *et al*., 2011].

A. QualComp algorithm

We downloaded QualComp from *http://web.stanford.edu/∼iochoa/QualComp.html*. We modified the script runCompress.sh (and some of the functions contained on it) so as to take as input a SAM file instead of a FASTQ file, and called it runCompress_samFile.sh. We also modified the runDecompress.sh script (and some of the functions contained on it) so as to generate an output file containing only the reconstructed quality values, and call it runDecompress_samFile.sh. We then used the script provided in [Canovas *et al*., 2014] that takes as input a SAM file and replaces the quality values contained on it with those provided in a separate file. Note that the modifications performed on the software of QualComp are only regarding the input and output formats, the compression and decompression of the quality values remain as they were in the original publication.

We run the program with the following command:

./runCompress_samFile.sh -i <SAMfile> -c <numClusters> -r <rate>

To reconstruct the quality values, we run the program with the following command:

./runDecompress_samFile.sh -p <prefixCompressedFiles> -c <numClusters> -r <rate>

B. Pblock and Rblock algorithms

We run the algorithm presented in [Canovas *et al*., 2014] with both options Pblock and Rblock. We downloaded the program from *https://github.com/rcanovas/libCSAM*.

To compress, we run the following commands:

Pblock: ./CompressQualFile <QualityValueFile> -q 1 -m 1 -l <p>

Rblock: ./CompressQualFile <QualityValueFile> -q 2 -m 1 -l <r>

We chose *-m 1* because it is the default value for m.

To decompress, we run the following command:

./DecompressQual <CompressedFile>,

where <CompressedFile> is the compressed file generated in the compression step.

C. CRAM

We downloaded CRAM from *http://www.ebi.ac.uk/ena/software/cram-toolkit*. CRAM allows for the compression of the SAM file with or without the quality values. In order to compute the compressed size used solely by the quality values, we subtract the size of the compressed file that does not include the quality values from the size of the compressed original file.

Specifically, the command to generate the compressed file that includes the quality values is:

java -jar cramtools-2.1.jar cram –input-is-sam -R <ReferenceFastaFile> -Q < <InputSamFile> > <CompressedFile>

To compress the file without the quality values, we drop the option *-Q* in the above command. For the lossy compression of the quality values, we use the same commands as above but with the extra parameter *-L* $*8$. Finally, we use the following command for the decompression:

java -jar cramtools-2.1.jar bam -R <ReferenceFastaFile> < <InputCompressedFile> > <SamFile>

D. DSRC2

We downloaded DSRC2 from *http://sun.aei.polsl.pl/dsrc*. DSRC2 is an algorithm for the lossless compression of FASTQ files. It also allows for lossy compression of the quality values using the Illumina binning scheme, as mentioned in the main manuscript.

From the several compression options that are available we use the high compression option, -m2. The algorithm does not provide the compressed size of the different fields of the FASTQ file. Thus, in order to compute the size used by the quality values we set all the IDs and the reads in the FASTQ file to the exact same value. Since the program uses an arithmetic encoder as the entropy encoder, almost no bits are needed to compress the IDs and the reads [Amir Said, 2004]. Thus, the compressed size of the modified FASTQ file will account for the actual compressed size of the quality values.

Specifically, we use the following command for lossless compression:

dsrc c -m2 <InputFastqFile> <OutputFile>

For lossy compression we use the following command:

dsrc c -m2 -l <InputFastqFile> <OutputFile>

And for decompression:
dsrc d <InputCompressedFile> <OutputFastqFile>

V. Variant Call File Generation

In order to generate the VCF files we follow the workflow proposed in *htslib.org*. Thus, after generating the sorted SAM file using the commands mentioned above, we run the following command to generate the variants:

samtools mpileup -ugf <ReferenceFastaFile> <InputSamFile> | bcftools call -vV indels -mO v -o <OutputVCF>

VI. More results on Rate-Distortion

In this section we provide similar rate-distortion curves as the ones shown in the main manuscript, but for the data sets *SRR032209*, *SRR622461chr20*, *SRR1179906* and *ERR011354*.

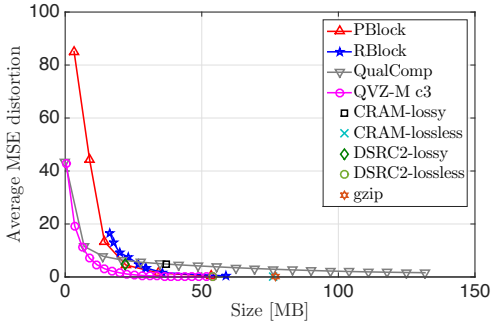
We also provide at the end of the *Supplementary Data* the exact values of the rate-distortion curves for all the data sets considered in this work. Moreover, we provide the value of all the distortions presented in [Canovas *et al*., 2014], that is, Manhattan (L1), Max:Min, MSE, Chebyshev, Soergel and Lorentzian. Note that in [Canovas *et al*., 2014] the Lorentzian distortion is defined without normalizing by the read length, and thus there is a scaling factor between the Lorentzian distortion shown in the plots and that of the tables.

A. SRR622461 (chr 20)

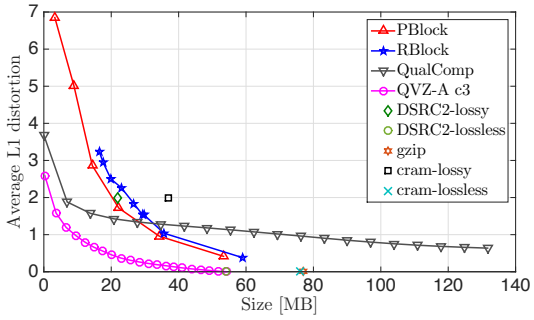
Fig. 1a, 1b and 1c show the rate-distortion curves for the MSE, L1 and Lorentzian distortions, respectively, for the different algorithms. As can be observed from the three plots, the proposed algorithm *QVZ* outperforms the previously proposed algorithms in all three distortions. Note that this is also true for the points achieved by *CRAM* and *DSRC2* with the Illumina binning. *QualComp* achieves a lower distortion than the *Pblock* and *Rblock* algorithms for smaller rates, whereas the latter two achieve a lower distortion for higher rates. This phenomenon is observed for all the three distortions.

Finally, it is worth mentioning that the proposed algorithm *QVZ* achieves the zero distortion point (lossless compression), much faster than the other lossy compression algorithms. This phenomenon is more noticeable for distortions other than MSE. Note also that it achieves a lossless rate smaller than that of *CRAM* and *gzip*, and similar to that of *DSRC2*.

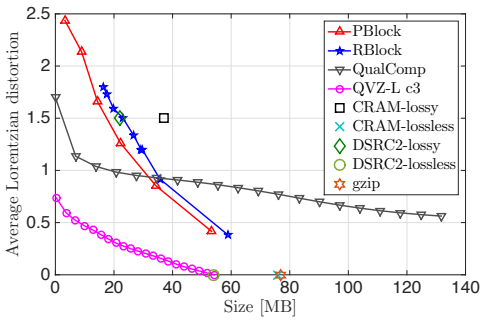
Note that the general behavior observed in this data set is similar to the one observed for the *NA12878chr20* data set, which is discussed in the main manuscript.



(a) MSE distortion.



(b) L1 distortion.

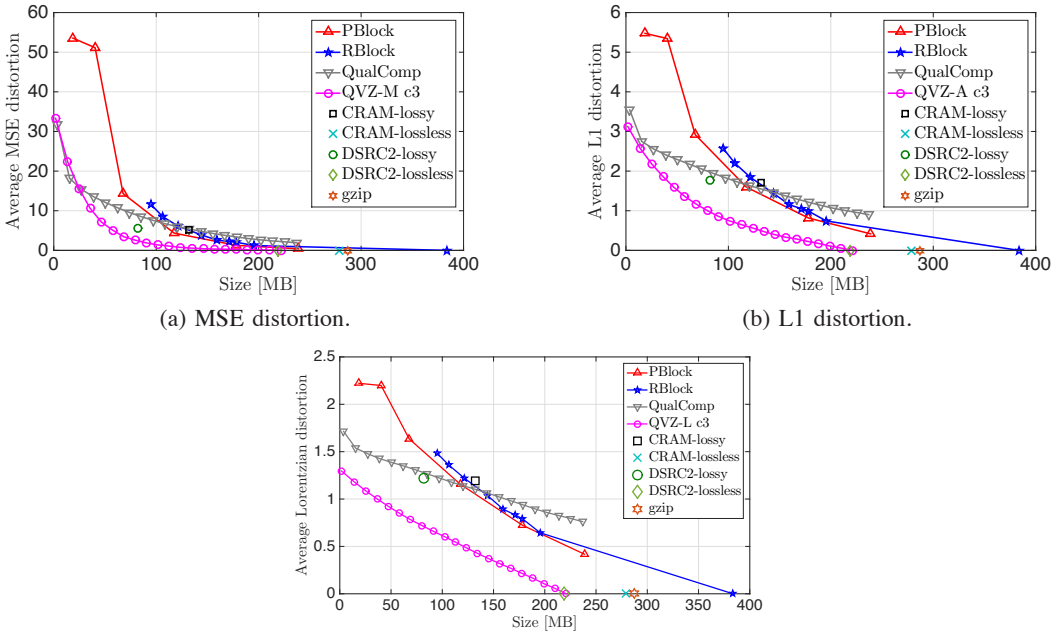


(c) Lorentzian distortion.

Fig. 1: Rate-Distortion curves for the *SRR622461chr20* data set.

B. SRR032209

The rate-distortion curves for this data set for the MSE, L1 and Lorentzian distortions are presented in Fig. 2a, 2b and 2c, respectively. The observed results are in line with what it was observed for the data sets *NA12878chr20* and *SRR622461chr20*, thus corroborating our findings. That is, the proposed algorithm *QVZ* presents a lower distortion level than *QualComp*, *Pblock* and *Rblock* for any rate, and achieves a lossless compression with a smaller rate. It also outperforms the performance obtained with the Illumina binning, and *QVZ*'s lossless compression is similar to that of *DSRC2*, and better than that of of *CRAM* and *gzip*.



(a) MSE distortion.

(b) L1 distortion.

(c) Lorentzian distortion.

Fig. 2: Rate-Distortion curves for the *SRR032209* data set.

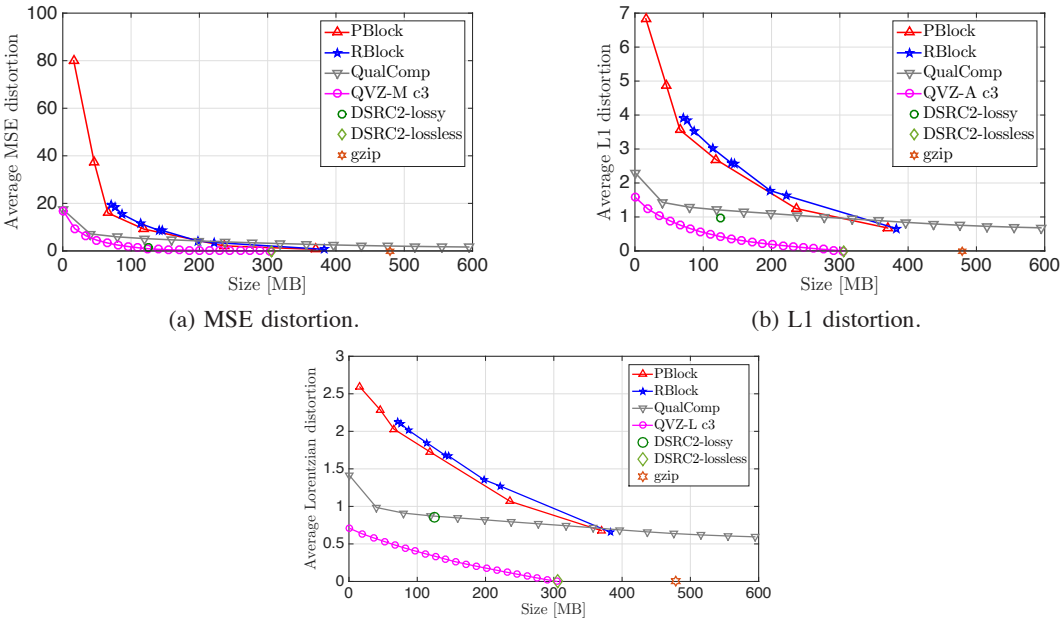
C. SRR1179906

The rate-distortion curves for this data set for the MSE, L1 and Lorentzian distortions are presented in Fig. 3a, 3b and 3c, respectively. We did not include the compression results of *CRAM*, as the algorithm does not accept FASTQ as input. Moreover, based on the previous results, we can conclude that the performance of *CRAM* is outperformed by that of *DSRC2*, both for lossy and lossless.

As can be observed from the figures, the results are in line with the results obtained for the previous data sets. The main difference is that for the MSE distortion, the rate-distortion point achieved by *DSRC2* when implementing the Illumina binning is close to that achieved by the proposed algorithm *QVZ*. Note that this is not the case for the other two distortion metrics.

D. ERR011354

The rate-distortion curves for this data set for the MSE, L1 and Lorentzian distortions are presented in Fig. 4a, 4b and 4c, respectively. Similarly to the previous data set, and for the same reasons, we did not include the compression results of *CRAM*. As can be observed from the figures, the results are in line with the results obtained for all the previous data sets. We note, however, that in this case the performance of *DSRC2* when implementing the Illumina binning is similar to that of *QVZ*, for all three distortions.

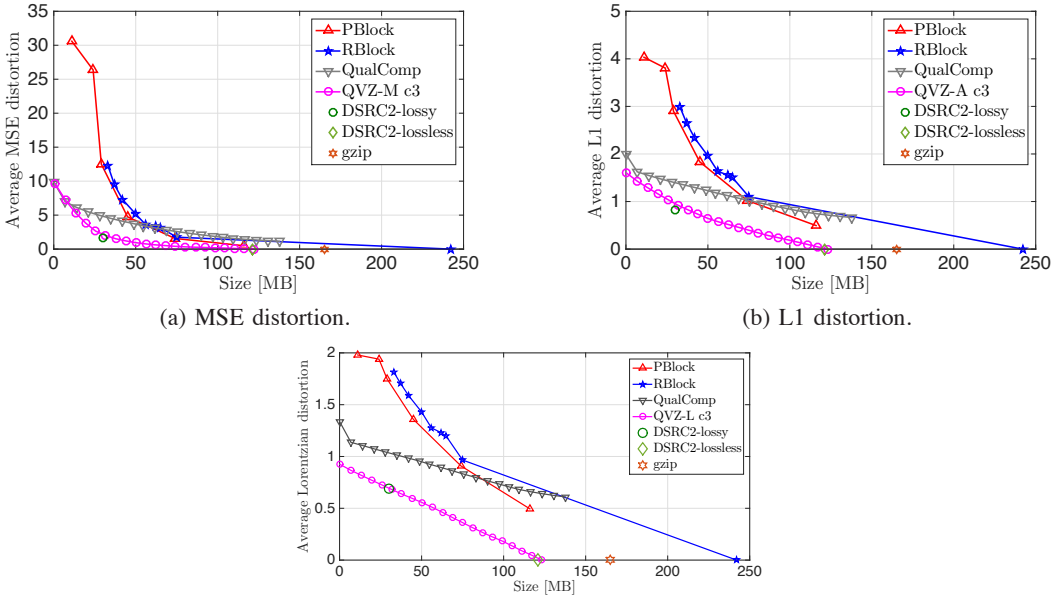


(a) MSE distortion.

(b) L1 distortion.

(c) Lorentzian distortion.

Fig. 3: Rate-Distortion curves for the *SRR1179906* data set.



(a) MSE distortion.

(b) L1 distortion.

(c) Lorentzian distortion.

Fig. 4: Rate-Distortion curves for the *ERR011354* data set.

VII. More results on genotyping

As discussed in the main manuscript, we analyze the differences in the SNP (Single Nucleotide Polymorphism) calling of the original SAM file and the SAM files where the quality scores are replaced with the reconstructed ones (after applying any lossy compression).

Recall that in the main manuscript we show the results for the *NA12878chr20* data set, for two different approaches. Specifically, the first approach aims to show the difference in the SNP calling of the SAM file with reconstructed quality scores (i.e., after the lossy compression and decompression) with that of

the SAM file containing the original quality scores. The second approach, on the contrary, assumes a "ground truth", and compares the SNP calling results of each SAM file (with or without the original quality scores) to the given "ground truth". For the latter case, in the main manuscript we assumed the "ground truth" was given by the SNP's provided by the Broad Institute Resource Bundle after following their Best Practices. Note that in the main manuscript the results presented for the *QVZ* algorithm in both experiments corresponded to the mode *QVZ-M* with three clusters. In this section we also show the results for the other modes, for both one and three clusters, and conclude that *QVZ-M* with three clusters shows the best performance.

In this section we also provide the results for the two experiments described above when the other *H. Sapiens* data set is used, that is, the *SRR622461chr20* data set. Moreover, we also analyze the SNP calling results for the case where the "ground truth" is given by the NIST proposed Standard[5], for both *H. Sapiens* data sets. These results are presented next.

A. NA12878.HiSeq.WGS.bwa.cleaned.recal.hg19.20

We first show the SNP calling results of the different modes of *QVZ*, for one and three clusters, for the two experiments analyzed in the main manuscript. Recall that in the main manuscripts we only presented the results of *QVZ-M* with three clusters, due to space constraints. Fig. 5 shows the results for the M, L and A modes. Note that the size of each point is proportional to the size of the compressed file that it represents, that is, the bigger the point, the bigger the size (and thus the smaller the distortion). All the following figures follow the same approach regarding the size of the points in order to give a sense of the size of each file. For exact values, we refer the reader to the Appendix at the end of this *Supplementary Data*[6].

As can be observed from the figure, the use of three clusters offers better genotyping results than that of one cluster, that is, for similar size, the genotyping results obtained by using three clusters is better than that of using a single cluster. We observe this phenomenon for the three modes. We also observe that *QVZ-M* converges faster to the uncompressed file than *QVZ-A* and *QVZ-L*.

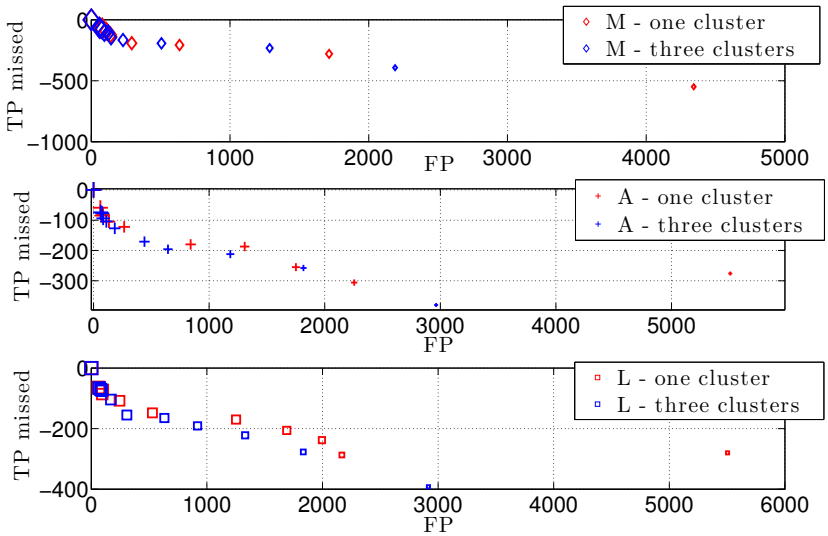


Fig. 5: SNP calling results of the original SAM file (*NA12878*), denoted as uncompressed, and those generated with the different modes of *QVZ*, for one and three clusters.

[5]Downloaded from *ftp://ftp-trace.ncbi.nih.gov/giab/ftp/data/NA12878/variant_calls/GIAB_integration/*

[6]For ease of visualization, we only plot every other point.

We show similar results in Fig. 6, but in this case assuming that the "ground truth" is given by that provided by the Broad Institute Resource Bundle. As can be observed from the figure, the shown behavior of the three modes is similar to the one observed in the previous case (Fig. 5). Specifically, the *QVZ-M* shows a faster convergence than the other two modes, being this fact more pronounced with the use of three clusters. Moreover, the number of points on the upper left corner (taking the uncompressed point as the center) shows that all three modes can offer compression while improving the genotyping, being *QVZ-M* the one that achieves more points in this quadrant.

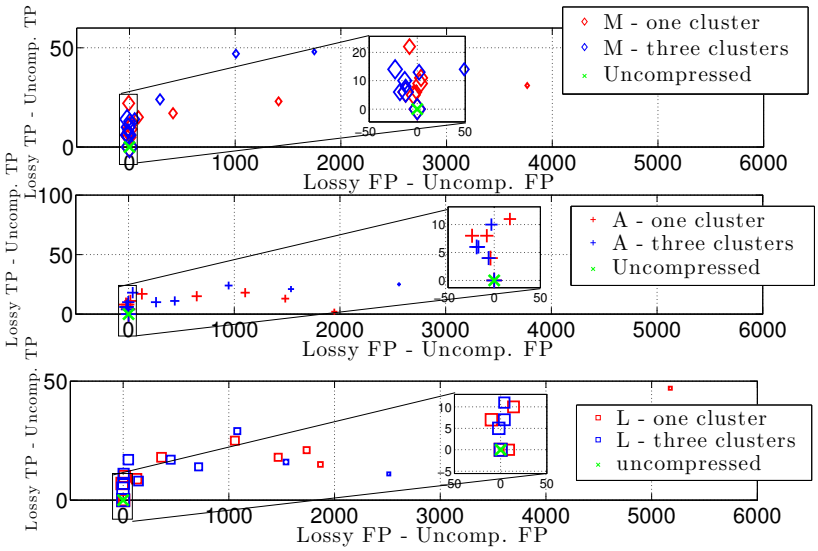


Fig. 6: SNP calling results of the original SAM file (*NA12878*), denoted as uncompressed, and those generated with the different modes of *QVZ*, for one and three clusters. We assume the "ground truth" is given by the one provided by the Broad Institute.

Finally, we show new results not presented in the main manuscript, for the case where the "ground truth" is assumed to be the one provided by the NIST proposed Standard. We first show in Fig. 7 the performance of *QVZ*, for the three modes and the options of one and three clusters. In this case we also observe that *QVZ-M* offers better performance than the other two modes. However, the number of points in the upper left quadrant decreases, having a higher concentration on the lower left quadrant. These points offer less true positives, while reducing the number of false positives that would have been called with the uncompressed version.

Fig. 8 compares the genotyping results of *QVZ* with the previously proposed algorithms, assuming as "ground truth" the SNPs provided by the NIST proposed Standard. For ease of visualization and based on the previous results shown in Fig. 7, we only show in the figure the results for *QVZ-M* with three clusters.

It is interesting to observe the difference converge of *QVZ* and *QualComp* with that of *Pblock* and *Rblock*. Whereas the first two maintain almost the same number of true positives for different rates and decrease the number of false positives as the rate (size) increases, the latter two maintain almost constant the number of false positives and increase the number of true positives as the rate increases. If we zoom in the area around the uncompressed point, we observe that the files generated with *QVZ* are concentrated around the uncompressed point, with several points in the upper right quadrant. On the contrary, the *Pblock* and *Rblock* generate more points on the lower left quadrant. Finally, *QualComp* does not show a specific behavior. Note also that the Illumina binning obtains the same number of T.P.s as the uncompressed file, but it generates more F.P.s.

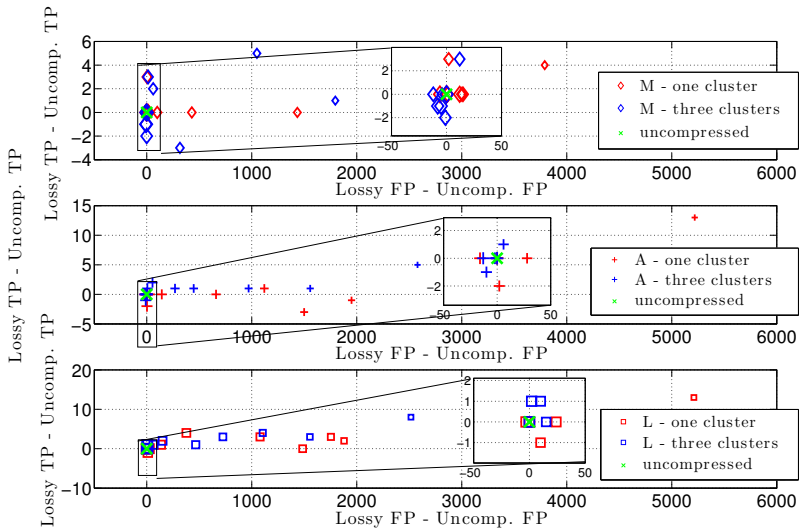


Fig. 7: SNP calling results of the original SAM file (*NA12878*), denoted as uncompressed, and those generated with the different modes of *QVZ*, for one and three clusters. We assume the "ground truth" is given by the one provided by the NIST proposed Standard.

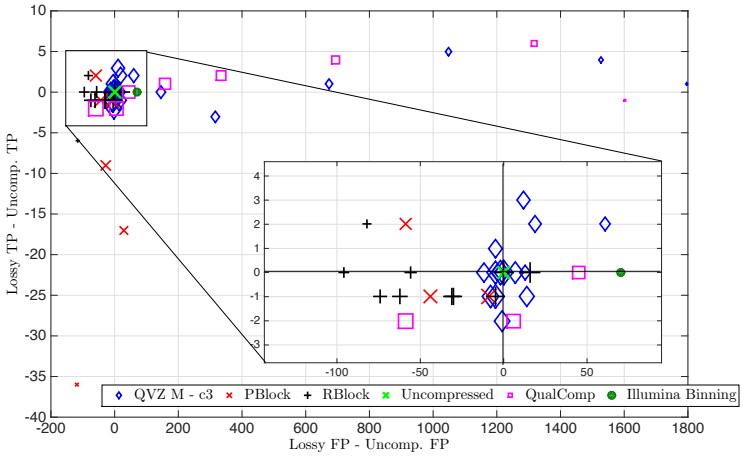


Fig. 8: SNP calling results of the original SAM file (*NA12878*), denoted as uncompressed, and those generated with the different compression algorithms. We assume the "ground truth" is given by the one provided by the NIST proposed Standard.

B. SRR622461chr20

In this section we provide extra genotyping results for the *SRR622461chr20* data set. We first show the behavior of *QVZ-M*, *QVZ-A* and *QVZ-L* with three clusters when the "ground truth" is given by the uncompressed SAM file, the Broad Institute and the NIST. The results are shown in Fig. 9. Note that the observed behavior is similarly to the one observed with the *NA12878chr20* data set, that is, the number of true positives remains similar to the uncompressed case. Regarding the number of false positives, they decrease as the size (rate) increases.

Finally, in Fig. 10 we compare the performance of *QVZ* to that of *QualComp*, *Pblock* and *Rblock*. We also compare it with the Illumina binning. The results are very similar to the ones observed in the previous results. For example, when the "ground truth" is given by the NIST Proposed Standard, *QVZ* can achieve points with more true positives than those achieved with *Pblock* and *Rblock* for the same number

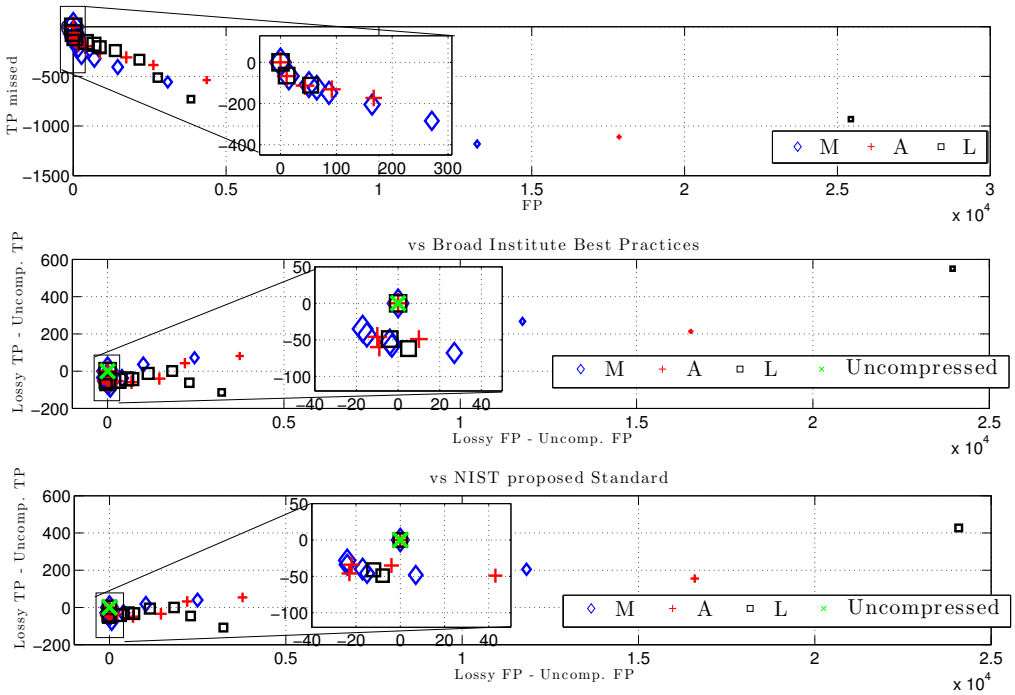


Fig. 9: SNP calling results of the original SAM file (*SRR622461*), denoted as uncompressed, and those generated with the different modes of *QVZ* with three clusters.

of false positives. Finally, we observe that for this data set, *QualComp* achieves points that increase both the number of true positives and false positives. Note also that the Illumina binning finds in general the same number of T.P.s as the uncompressed data sets, but it gets more F.P.s.

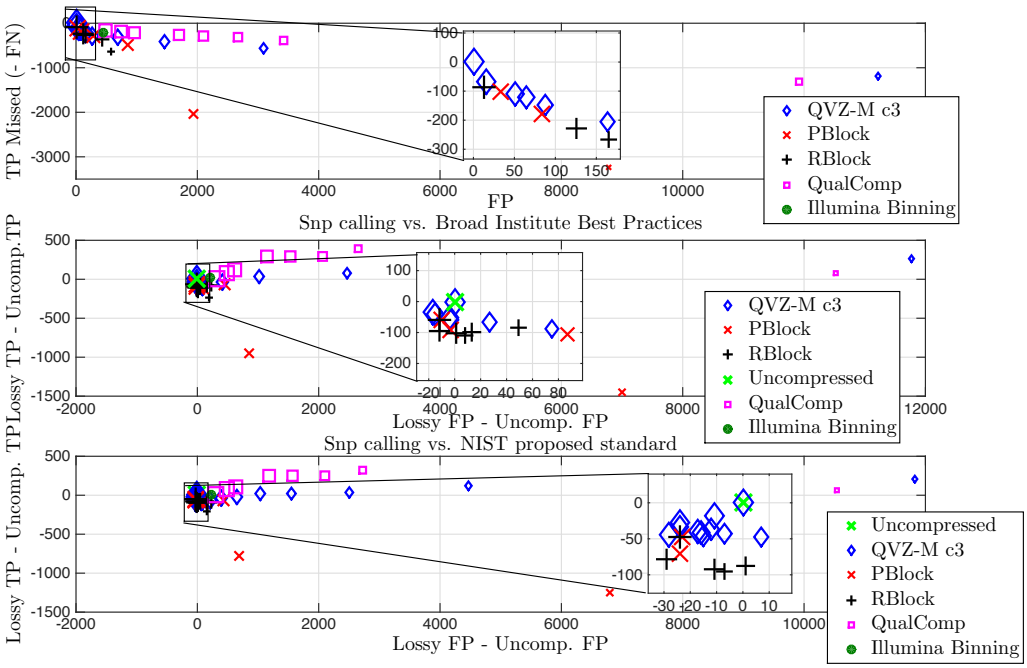


Fig. 10: SNP calling results of the original SAM file (*SRR622461*), denoted as uncompressed, and those generated with *QVZ-M* with three clusters, for the cases where the "ground truth" is given by the uncompressed file, the Broad Institute and the NIST.

REFERENCES

[Canovas *et al*., 2014] Canovas, R., Moffat, A. and Turpin, A., "Lossy compression of quality scores in genomic data", Bioinformatics, 2014.

[Ochoa *et al*., 2014] Ochoa, I., et al, "QualComp: a new lossy compressor for quality scores based on rate distortion theory", BMC bioinformatics 14.1 (2013): 187.

[Li *et al*., 2014] Li, Heng, et al, "The sequence alignment/map format and SAMtools", Bioinformatics 25.16 (2009): 2078-2079.

[Langmead *et al*., 2014] Langmead, Ben, et al., "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome", Genome Biol 10.3 (2009): R25.

[Fritz *et al*., 2011] Fritz, Markus, et al., "Efficient storage of high throughput DNA sequencing data using reference-based compression", Genome research, vol. 21, 2011.

[Roguski *et al*., 2011] Roguski, Lukasz and Deorowicz, Sebastian, "DSRC 2—Industry-oriented compression of FASTQ files", Bioinformatics, vol. 30, 2014.

[Amir Said, 2004] Said, Amir, "Introduction to arithmetic coding-theory and practice", Hewlett Packard Laboratories Report, 2004.

VIII. Appendix: Raw data of the rate-distortion

In this section we provide tables representing the rate-distortion points for each of the considered data sets and algorithms. Each table specifies the parameters used by the corresponding algorithm, the size of the generated file, and the value of the different distortions.

A. NA12878 (chr 20)

Next we show the rate-distortion results for the *NA12878chr20* data set. Tables II, III and IV contain the results for the *Pblock, Rblock* and *QualComp* algorithms. The *QVZ* results for one cluster and MSE, Lorentzian and L1 distortions are shown in Tables V, VI and VII, respectively. The same results but for three clusters are shown in Tables VIII, IX and X, and for five clusters in Tables XI, XII and XIII. Finally, the results for *CRAM, DSRC2* and *gzip* are shown in Table XIV.

B. SRR032209

We show the rate-distortion results for this data set for *Pblock, Rblock* and QualComp in Tables XV, XVI and XVII, respectively. For *QVZ* we show the results for three clusters and modes *M, L* and *A* in Tables XVIII, XIX and XX, respectively. Finally, the results for *CRAM, DSRC2* and *gzip* are shown in Table XXI.

C. SRR622461 (chr 20)

We show the rate-distortion results for this data set for *Pblock, Rblock* and QualComp in Tables XXII, XXIII and XXIV, respectively. For *QVZ* we show the results for three clusters and modes *M, L* and *A* in Tables XXV, XXVI and XXVII, respectively. Finally, the results for *CRAM, DSRC2* and *gzip* are shown in Table XXVIII.

D. SRR1179906

We show the rate-distortion results for this data set for *Pblock, Rblock* and QualComp in Tables XXIX, XXX and XXXI, respectively. For *QVZ* we show the results for three clusters and modes *M, L* and *A* in Tables XXXII, XXXIII and XXXIV, respectively. Finally, the results for *DSRC2* and *gzip* are shown in Table XXXV.

E. ERR011354

We show the rate-distortion results for this data set for *Pblock, Rblock* and QualComp in Tables XXXVI, XXXVII and XXXVIII, respectively. For *QVZ* we show the results for three clusters and modes *M, L* and *A* in Tables XXXIX, XL and XLI, respectively. Finally, the results for *DSRC2* and *gzip* are shown in Table XLII.

				P-Block			
p	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
1	1,673	0.430085	1.03832	0.430085	0.960807	0.00650891	43.4386
2	1,021	0.88712	1.07816	1.37786	1.92197	0.0134632	79.3134
4	588.5	1.83367	1.16461	5.06639	3.76427	0.0278605	132.466
8	268.4	3.26223	1.34255	15.3149	6.08972	0.0497792	185.678
16	146.7	5.22011	2.46974	43.7473	8.5919	0.081773	231.063
32	51.6	6.88308	3.49512	75.9404	9.64762	0.107954	262.299

TABLE II: Compression size and distortion results of the P-Block algorithm for the NA12878 (chr 20) data set.

				R-Block			
p	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
2.6	3,229	0	1	0	0	0	0
3.2	1,935	0.333631	1.02673	0.333631	0.909916	0.0049616	33.6968
4.2	1,695	0.420857	1.03296	0.420857	0.95943	0.00634601	42.5066
7.4	1,080	0.858535	1.06322	1.33672	1.91412	0.0129411	76.6897
8	1,044	0.887043	1.06716	1.40022	1.92164	0.0134207	78.8354
9	942.4	1.22499	1.08372	2.47516	2.76657	0.0184379	99.965
10	821.1	1.32722	1.09352	2.8321	2.86051	0.0200284	106.184
11	801.9	1.35619	1.09718	2.92918	2.87371	0.0204886	107.995
12	768.9	1.51536	1.10733	3.52301	3.2748	0.0228806	116.808
15	623.8	1.90449	1.13351	5.38085	3.83121	0.0287952	135.772
20	424.9	2.96895	1.1764	12.3523	5.40652	0.0448134	177.015
25	361.8	3.21744	1.19922	14.4739	5.88049	0.0486985	185.08
30	325.1	3.42664	1.21812	16.5345	6.31795	0.0519905	191.09

TABLE III: Compression size and distortion results of the R-Block algorithm for the NA12878 (chr 20) data set.

		QualComp (num. of clusters = 3)					
r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
0	0	3.82586	4.36983	40.1315	13.8833	0.0630282	181.303
17	109.6	2.20389	3.29692	12.1285	11.2104	0.0358412	136.624
34	219.2	1.99202	3.01702	9.643	9.95104	0.0321932	129.577
51	328.8	1.88793	2.82773	8.4207	9.0566	0.0305027	126.163
68	438.5	1.79798	2.66132	7.45408	8.30193	0.0290826	123.049
85	548.1	1.70428	2.51009	6.5668	7.61828	0.0276142	119.511
102	657.7	1.60775	2.37722	5.7616	6.99978	0.0261085	115.534
119	767.3	1.51066	2.26334	5.0398	6.44536	0.0245965	111.226
136	876.9	1.41746	2.16759	4.42083	5.96524	0.0231401	106.797
153	986.6	1.32951	2.08541	3.90176	5.55274	0.0217497	102.303
170	1,096	1.24443	2.00335	3.44142	5.15501	0.0203925	97.7293
187	1,205	1.15783	1.92097	3.00839	4.74722	0.0190007	92.874
204	1,315	1.07232	1.84038	2.61298	4.34269	0.0176207	87.8767
221	1,425	0.991383	1.76365	2.26411	3.9783	0.0163109	82.9724
238	1,534	0.922791	1.6947	1.98165	3.69122	0.0151959	78.7167
255	1,644	0.868346	1.63912	1.76351	3.47016	0.0143082	75.298
272	1,753	0.820624	1.5939	1.57967	3.26666	0.0135289	72.2531
289	1,863	0.775354	1.55362	1.41225	3.03605	0.0127868	69.3134
306	1,973	0.733789	1.51735	1.26405	2.76767	0.0121003	66.571
323	2,082	0.69734	1.48526	1.13759	2.49102	0.0114913	64.1327

TABLE IV: Compression size and distortion results of QualComp for the NA12878 (chr 20) data set.

			QVZ - M					
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
1	0	3	7.45943	5.59097	112.885	15.4656	0.112821	274.3
1	0.05	81.68	2.84912	4.97752	25.6138	15.0717	0.0470624	152.296
1	0.1	159.91	2.05068	4.40239	15.0464	13.4041	0.0337731	121.292
1	0.15	236.82	1.58657	3.78316	9.45276	11.2456	0.0260387	101.937
1	0.2	319.23	1.31531	3.18372	6.26068	9.13194	0.0214434	91.965
1	0.25	405.61	1.11945	2.80342	4.4195	7.70018	0.0181346	83.6364
1	0.3	489.61	0.956204	2.5059	3.1363	6.36219	0.0154009	75.9828
1	0.35	573.61	0.827516	2.22983	2.28526	5.33218	0.0132582	69.369
1	0.4	659.37	0.726635	2.0053	1.71122	4.45536	0.0115706	63.4774
1	0.45	740.00	0.640478	1.7804	1.29284	3.71672	0.0101386	57.9716
1	0.5	819.11	0.570146	1.56009	0.99511	3.06094	0.00898411	52.9051
1	0.55	902.86	0.502864	1.36195	0.749554	2.51953	0.00788826	47.7998
1	0.6	989.67	0.439853	1.21116	0.55725	2.08374	0.00687488	42.7559
1	0.65	1,067	0.37117	1.1238	0.419273	1.82918	0.0057869	36.7209
1	0.7	1,149	0.31338	1.09494	0.340657	1.6329	0.00488993	31.1588
1	0.75	1,232	0.260254	1.07835	0.276983	1.46648	0.00406754	25.9565
1	0.8	1,314	0.199795	1.06223	0.201224	1.01979	0.00312844	20.1526
1	0.85	1,395	0.147708	1.05678	0.147972	0.973907	0.00232356	14.9132
1	0.9	1,477	0.0968763	1.05343	0.0968792	0.95861	0.00153139	9.78445
1	0.95	1,558	0.0512244	1.04547	0.0512247	0.935011	0.000809626	5.17366
1	1	1,641	0	1	0	0	0	0

TABLE V: Compression size and distortion results of QVZ (with one cluster) when minimizing the MSE for the NA12878 (chr 20) data set.

			QVZ - L					
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
1	0	0	5.70243	6.4643	140.901	18.0555	0.0828724	159.299
1	0.05	89.81	2.4829	6.29135	35.5512	16.7989	0.0399995	107.045
1	0.1	176.32	1.98826	5.79761	23.2145	15.6243	0.0318969	97.4477
1	0.15	258.37	1.68564	5.20061	16.9062	14.4305	0.0270076	89.8209
1	0.2	338.74	1.46228	4.68376	12.968	13.3573	0.0234124	83.0876
1	0.25	420.96	1.29265	4.38551	10.2857	12.6586	0.020692	77.3473
1	0.3	502.95	1.13862	4.03193	8.08586	11.843	0.0182212	71.799
1	0.35	583.55	1.00695	3.6795	6.44276	11.0163	0.0161076	66.6385
1	0.4	662.43	0.859494	3.21351	4.62798	9.6808	0.0137396	61.1764
1	0.45	744.98	0.725287	2.74915	3.09896	7.97178	0.0115722	55.8626
1	0.5	827.86	0.617259	2.30187	2.09178	6.45925	0.00981521	50.4262
1	0.55	908.68	0.524564	1.91287	1.41008	5.21503	0.00831275	44.9879
1	0.6	991.01	0.441695	1.58365	0.938965	4.10582	0.00696912	39.6835
1	0.65	1,070	0.373087	1.38193	0.663197	3.27888	0.00586813	34.7006
1	0.7	1,152	0.309668	1.25965	0.456308	2.59573	0.00486108	29.5705
1	0.75	1,232	0.254999	1.17854	0.333026	2.16819	0.00400332	24.6376
1	0.8	1,315	0.201961	1.14387	0.251144	1.84618	0.00317493	19.6828
1	0.85	1,396	0.152001	1.12229	0.186219	1.66127	0.00239765	14.8318
1	0.9	1,478	0.100084	1.08843	0.118785	1.4216	0.00158405	9.80531
1	0.95	1,559	0.0517059	1.05485	0.0571916	1.11121	0.000818994	5.13
1	1	1,641	0	1	0	0	0	0

TABLE VI: Compression size and distortion results of QVZ (with one cluster) when minimizing the distortion function $\log_2(1 + L1)$, for the NA12878 (chr 20) data set.

			QVZ - A					
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
1	0	0	5.66824	6.41224	136.7	17.7288	0.0827925	165.376
1	0.05	86.13	2.41846	6.14301	33.7088	16.6082	0.0390532	107.621
1	0.1	165.48	1.86291	5.5359	19.7793	14.9741	0.0299802	97.5045
1	0.15	245.40	1.5269	4.91617	12.9587	13.0369	0.0245254	89.9355
1	0.2	328.01	1.30133	4.35643	9.22746	11.387	0.0208609	83.285
1	0.25	411.15	1.13555	3.93533	6.8627	10.2281	0.0181834	77.4116
1	0.3	493.05	0.994232	3.51675	5.02365	9.16145	0.0158833	71.9468
1	0.35	577.24	0.877306	3.19449	3.86592	8.25544	0.0139922	66.5727
1	0.4	658.86	0.763793	2.78711	2.84374	7.09546	0.0121539	61.095
1	0.45	740.48	0.650848	2.22689	1.76796	5.34129	0.0102986	55.7063
1	0.5	823.35	0.566793	1.86319	1.2486	4.31013	0.00894061	50.4646
1	0.55	904.69	0.492053	1.58429	0.899123	3.5185	0.00774639	45.2663
1	0.6	988.31	0.420446	1.3163	0.619925	2.7245	0.00659559	39.936
1	0.65	1,071	0.357393	1.16525	0.431627	2.06554	0.00559139	34.9332
1	0.7	1,151	0.302098	1.13077	0.34515	1.81578	0.00472963	29.8052
1	0.75	1,232	0.251495	1.10094	0.278768	1.61989	0.00394275	24.9151
1	0.8	1,314	0.199659	1.07848	0.213893	1.40558	0.00313359	19.8915
1	0.85	1,395	0.148083	1.06699	0.155496	1.23807	0.00233043	14.8108
1	0.9	1,477	0.0972047	1.05927	0.100143	1.07285	0.00153559	9.76389
1	0.95	1,559	0.050659	1.04653	0.0511411	0.957198	0.000801361	5.1072
1	1	1,641	0	1	0	0	0	0

TABLE VII: Compression size and distortion results of QVZ (with one cluster) when minimizing the L1 distortion, for the NA12878 (chr 20) data set.

			QVZ - M					
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	6.17	3.65473	4.51024	40.3378	14.2915	0.0603724	170.447
3	0.05	90.43	2.35392	4.1526	18.5538	13.3925	0.0384732	132.376
3	0.1	170.29	1.8376	3.69332	11.9401	11.8884	0.0298818	112.218
3	0.15	247.50	1.47084	3.28187	7.9224	10.0973	0.0238619	96.5692
3	0.2	326.49	1.24717	2.89821	5.50022	8.50846	0.0201519	88.2613
3	0.25	410.94	1.06946	2.55925	3.94849	7.27139	0.0172034	80.6546
3	0.3	491.52	0.922179	2.27819	2.82109	6.08432	0.0147759	73.7819
3	0.35	575.76	0.799615	2.00957	2.04284	5.01117	0.0127571	67.4041
3	0.4	657.24	0.703336	1.79682	1.51986	4.14079	0.0111593	61.8109
3	0.45	737.06	0.622728	1.59793	1.14284	3.35591	0.00983489	56.6669
3	0.5	816.86	0.555726	1.43146	0.887514	2.78619	0.00874648	51.8734
3	0.55	899.72	0.493291	1.27874	0.686891	2.29437	0.00773993	47.1558
3	0.6	982.56	0.429172	1.18396	0.529021	1.97397	0.00671897	41.8593
3	0.65	1,060	0.363268	1.11048	0.404336	1.75695	0.00567869	36.016
3	0.7	1,142	0.307881	1.08658	0.330131	1.55348	0.00481119	30.6861
3	0.75	1,222	0.256411	1.07584	0.271781	1.4318	0.00401031	25.5927
3	0.8	1,304	0.196814	1.0619	0.198485	1.03508	0.00308725	19.8458
3	0.85	1,384	0.145522	1.05641	0.14568	0.969295	0.00228936	14.6947
3	0.9	1,465	0.0957157	1.0528	0.0957248	0.95894	0.00150973	9.66712
3	0.95	1,545	0.050266	1.04477	0.0502666	0.933134	0.000792555	5.07686
3	1	1,626	0	1	0	0	0	0

TABLE VIII: Compression size and distortion results of QVZ (with three clusters) when minimizing the MSE, for the NA12878 (chr 20) data set.

			QVZ - L					
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	4.54	3.133	5.73508	54.2172	16.6495	0.0501163	118.447
3	0.05	93.89	2.1893	5.70338	27.0572	15.8098	0.0348541	102.14
3	0.1	179.34	1.83063	5.16491	19.0369	14.856	0.0290092	94.1332
3	0.15	261.49	1.58609	4.69715	14.3501	13.9071	0.0251128	87.4887
3	0.2	341.16	1.38998	4.34493	11.1384	13.1013	0.0220113	81.3766
3	0.25	421.44	1.2278	3.97561	8.7729	12.2557	0.0194439	75.9141
3	0.3	502.63	1.08669	3.5172	6.87444	11.2243	0.0171866	70.6602
3	0.35	582.29	0.958447	3.1245	5.39457	10.2336	0.0151473	65.5388
3	0.4	660.86	0.824183	2.71661	3.8977	8.89517	0.0130223	60.321
3	0.45	742.68	0.698535	2.30338	2.6237	7.24717	0.0110278	55.0734
3	0.5	824.83	0.598602	1.97369	1.82653	5.92126	0.00943132	49.7233
3	0.55	905.14	0.512552	1.71243	1.28194	4.92	0.00805914	44.4369
3	0.6	986.01	0.433875	1.45896	0.865917	3.90516	0.00680713	39.2937
3	0.65	1,066	0.36639	1.29823	0.613093	3.14569	0.00573851	34.3029
3	0.7	1,147	0.305425	1.20259	0.431279	2.49404	0.00477966	29.2834
3	0.75	1,227	0.251756	1.153	0.324596	2.11016	0.00394287	24.3817
3	0.8	1,309	0.199229	1.12407	0.242677	1.77718	0.00312406	19.4743
3	0.85	1,390	0.149865	1.10745	0.180923	1.59397	0.00235672	14.6662
3	0.9	1,471	0.0992715	1.08111	0.117567	1.39809	0.00156553	9.73162
3	0.95	1,551	0.051006	1.05139	0.0561067	1.10508	0.000805723	5.06166
3	1	1,633	0	1	0	0	0	0

TABLE IX: Compression size and distortion results of QVZ (with three clusters) when minimizing the distortion function $\log_2(1 + L1)$, for the NA12878 (chr 20) data set.

			QVZ - A					
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	4.83	3.03788	5.57218	49.6366	16.2886	0.0488057	118.969
3	0.05	90.95	2.0733	5.38509	23.3656	15.2609	0.0331741	102.437
3	0.1	169.93	1.69675	4.88057	15.5301	13.8974	0.0271023	94.4044
3	0.15	249.25	1.44163	4.45853	11.0481	12.4465	0.0230146	87.9356
3	0.2	329.81	1.24889	4.03366	8.11792	11.0933	0.0199201	81.9698
3	0.25	412.32	1.09612	3.59859	6.09748	9.98021	0.0174511	76.3235
3	0.3	495.20	0.96361	3.13344	4.47094	8.83999	0.0153003	70.9603
3	0.35	577.98	0.851039	2.76979	3.39729	7.85533	0.0134803	65.6413
3	0.4	658.08	0.743695	2.3788	2.48624	6.75606	0.0117528	60.3282
3	0.45	739.04	0.638733	1.92758	1.58653	5.11698	0.0100691	55.1305
3	0.5	821.67	0.554196	1.64095	1.11033	3.95679	0.00871804	49.8057
3	0.55	902.77	0.480426	1.44124	0.807639	3.23801	0.00754384	44.5957
3	0.6	985.08	0.413697	1.26192	0.580393	2.54487	0.00648287	39.5414
3	0.65	1,066	0.351613	1.14826	0.41957	2.02227	0.00550358	34.4333
3	0.7	1,147	0.298078	1.11222	0.33638	1.75711	0.00466551	29.4702
3	0.75	1,227	0.248585	1.09265	0.273409	1.58261	0.00389508	24.6612
3	0.8	1,308	0.197537	1.07517	0.210749	1.37631	0.00309819	19.6976
3	0.85	1,389	0.146837	1.0671	0.155463	1.26174	0.002309	14.6657
3	0.9	1,470	0.0964269	1.05812	0.0999105	1.09659	0.00152047	9.67283
3	0.95	1,551	0.0499813	1.04586	0.0506887	0.96689	0.000789301	5.03357
3	1	1,632	0	1	0	0	0	0

TABLE X: Compression size and distortion results of QVZ (with three clusters) when minimizing the L1 distortion, for the NA12878 (chr 20) data set.

			QVZ - M					
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
5	0	7.65	2.93041	3.26488	26.7796	14.3411	0.0475321	153.235
5	0.05	90.54	2.1767	3.00301	15.6346	13.1088	0.0351028	127.136
5	0.1	169.98	1.74727	2.72752	10.6362	11.6251	0.0281282	108.968
5	0.15	247.53	1.42226	2.4417	7.22598	9.81508	0.0228941	94.78
5	0.2	325.79	1.21587	2.19693	5.05316	8.16546	0.0195208	86.9919
5	0.25	410.81	1.0505	1.98787	3.65727	6.93086	0.016816	79.888
5	0.3	491.55	0.908397	1.81463	2.62661	5.76468	0.0144936	73.1615
5	0.5	817.53	0.553882	1.28247	0.851655	2.69161	0.00870668	51.824
5	0.7	1,142.64	0.307295	1.07222	0.32751	1.51573	0.00479762	30.6551
5	0.9	1,465.21	0.0956949	1.0475	0.0957061	0.958772	0.00150848	9.66499

TABLE XI: Compression size and distortion results of QVZ (with five clusters) when minimizing the MSE, for the NA12878 (chr 20) data set.

			QVZ - L					
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
5	0	6.80	2.61656	4.18365	39.2848	16.3346	0.042024	110.257
5	0.05	95.56	2.02484	4.09743	23.0631	15.6321	0.0322503	98.6656
5	0.1	180.75	1.75163	3.70145	17.2231	14.6441	0.027726	91.895
5	0.15	263.75	1.54345	3.41478	13.4643	13.7865	0.0243691	85.9015
5	0.2	344.20	1.36594	3.17853	10.6849	12.9776	0.0215557	80.2839
5	0.25	422.55	1.2	2.9238	8.32174	12.0445	0.0189524	74.8312
5	0.3	501.56	1.06282	2.66041	6.56299	11.0803	0.0167783	69.7307
5	0.5	821.02	0.590516	1.67711	1.76994	5.85972	0.00929381	49.2496
5	0.7	1,141.18	0.30335	1.16305	0.437488	2.52029	0.00474591	29.0404
5	0.9	1,463.10	0.0984504	1.07029	0.117593	1.40067	0.00155065	9.63342

TABLE XII: Compression size and distortion results of QVZ (with five clusters) when minimizing the distortion function $\log_2(1 + L1)$, for the NA12878 (chr 20) data set.

		QVZ - A						
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
5	0	7.44	2.4836	4.01538	32.7521	15.778	0.0400153	111.182
5	0.05	93.12	1.88854	3.84653	19.019	14.8114	0.030258	98.8858
5	0.1	173.28	1.5944	3.49388	13.4254	13.5301	0.0254685	91.914
5	0.15	252.53	1.37921	3.20175	9.84986	12.2191	0.0219974	86.0983
5	0.2	332.78	1.2095	2.95851	7.38151	11.0285	0.0192636	80.6352
5	0.25	413.96	1.07153	2.67307	5.67094	9.99421	0.0170308	75.2524
5	0.3	496.00	0.947702	2.37202	4.2531	8.838	0.0150258	70.0907
5	0.5	817.79	0.547014	1.44208	1.08515	3.91271	0.00860426	49.2816
5	0.7	1,140.81	0.295277	1.09135	0.333728	1.74263	0.00462229	29.1879
5	0.9	1,461.51	0.0955797	1.05314	0.100221	1.13358	0.0015054	9.56466

TABLE XIII: Compression size and distortion results of QVZ (with five clusters) when minimizing the L1 distortion, for the NA12878 (chr 20) data set.

		CRAM, DSRC2 and gzip					
Option	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
CRAM (lossless)	2,000	0	1	0	0	0	0
CRAM (lossy)	980	1.32159	1.46856	3.13707	3.22962	0.02303	105.214
DSRC2 (lossless)	1,625	0	1	0	0	0	0
DSRC2 (lossy)	646	1.32159	1.46856	3.13707	3.22962	0.02303	105.214
gzip (lossless)	1,999	0	1	0	0	0	0

TABLE XIV: Compression size and distortion results of CRAM, DSRC2 and gzip for the NA12878 (chr 20) data set.

		P-Block					
p	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
1	238.83	0.417308	1.09837	0.417308	0.884194	0.00747858	15.0231
2	178.19	0.815811	1.13353	1.25985	1.69127	0.0139317	26.052
4	117.49	1.5977	1.20827	4.35765	3.09774	0.0266624	41.6468
8	67.64	2.91892	1.39512	14.3413	5.16139	0.0483476	58.8952
16	40.34	5.34461	2.56087	51.1372	7.81506	0.0906889	79.1403
32	18.66	5.47496	2.62416	53.5475	7.85338	0.0927586	79.9788

TABLE XV: Compression size and distortion results of the P-Block algorithm, for the SRR032209 data set.

		R-Block					
p	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	383.63	0	1	0	0	0	0
8	195.65	0.741303		1.05711	1.62062	0.0118432	23.1525
10	177.81	0.994809	1.07463	2.06871	2.21176	0.0158938	28.6018
11	171.29	1.05089	1.07915	2.24309	2.32844	0.0168113	29.9042
12	159.15	1.15553	1.08614	2.60104	2.3971	0.0185233	32.1982
15	143.96	1.44064	1.105	3.90235	2.99577	0.0231362	37.394
20	121.32	1.84716	1.13351	6.13705	3.67067	0.0297399	43.9882
25	106.33	2.19959	1.15807	8.51259	4.23544	0.0354711	48.9328
30	95.09	2.56712	1.18273	11.5421	4.771	0.0414356	53.3382

TABLE XVI: Compression size and distortion results of the R-Block algorithm, for the SRR032209 data set.

QualComp (num. of clusters = 3)							
r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
0	0	3.54311	2.99438	31.7281	11.8997	0.0611354	61.7206
5	11.66	2.75072	2.51649	18.2096	9.85558	0.0492125	55.3954
10	23.33	2.5593	2.3908	15.3957	9.18237	0.0457256	53.2542
15	35.00	2.41274	2.2987	13.5558	8.65556	0.0430231	51.3856
20	46.66	2.28924	2.22481	12.0916	8.07397	0.0407869	49.9325
25	58.33	2.17283	2.15971	10.7733	7.52631	0.0387264	48.6025
30	70.00	2.05576	2.10235	9.57429	7.03105	0.0366436	47.1267
35	81.67	1.94174	2.05262	8.50641	6.579	0.0346275	45.5816
40	93.34	1.83204	2.00829	7.56067	6.16505	0.0326941	43.9985
45	105.01	1.72773	1.96796	6.71521	5.7812	0.0308813	42.4505
50	116.67	1.63452	1.93122	5.99334	5.42477	0.0292774	41.0586
55	128.34	1.54835	1.89758	5.37622	5.09676	0.027794	39.7048
60	140.01	1.46433	1.8672	4.81634	4.78749	0.026353	38.319
65	151.68	1.37749	1.83836	4.28388	4.47811	0.024864	36.8126
70	163.34	1.29276	1.81064	3.79555	4.17599	0.0234251	35.297
75	175.01	1.21234	1.78442	3.36082	3.89289	0.0220485	33.8026
80	186.68	1.1342	1.75874	2.97109	3.62743	0.0206847	32.2715
85	198.35	1.06518	1.7337	2.64167	3.39583	0.0194683	30.8858
90	210.02	1.0084	1.70936	2.38052	3.20884	0.0184592	29.7175
95	221.68	0.955332	1.68386	2.15074	3.03786	0.0175221	28.5888
100	233.35	0.901853	1.65552	1.93544	2.86893	0.0165794	27.4113

TABLE XVII: Compression size and distortion results of QualComp, for the SRR032209 data set.

QVZ - M								
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	1.78	3.70521	3.57601	33.3377	12.5371	0.0632959	65.1187
3	0.05	13.57	2.87329	3.3062	22.3807	11.5634	0.0487546	54.0962
3	0.1	24.76	2.33823	2.95106	15.5458	10.2908	0.0398404	47.5182
3	0.15	35.63	1.91755	2.60885	10.5559	8.64216	0.0328175	42.2316
3	0.2	46.47	1.62258	2.3287	7.1494	6.95223	0.0277688	38.7658
3	0.25	57.76	1.38658	2.08497	5.05323	5.8036	0.0236902	35.5328
3	0.3	68.58	1.13965	1.86044	3.41206	4.64823	0.0194787	30.9558
3	0.35	79.80	0.974813	1.72912	2.54484	3.98472	0.0166476	27.5521
3	0.4	90.34	0.840381	1.6196	1.89119	3.36397	0.0143362	24.7925
3	0.45	101.38	0.721033	1.5208	1.38139	2.80667	0.0122763	22.21
3	0.5	112.22	0.684138	1.42125	1.09625	2.35253	0.0115407	22.095
3	0.55	123.63	0.539274	1.26437	0.754142	1.844	0.00915068	17.9553
3	0.6	134.86	0.458298	1.21022	0.564681	1.5769	0.007772	15.7649
3	0.65	145.96	0.380458	1.16516	0.409775	1.16808	0.00644822	13.4891
3	0.7	156.53	0.326843	1.14715	0.336589	0.997879	0.0055325	11.6976
3	0.75	167.26	0.279551	1.13295	0.284843	0.945619	0.00472855	10.0257
3	0.8	178.04	0.227722	1.11879	0.230827	0.915718	0.00385597	8.17482
3	0.85	188.87	0.173201	1.10393	0.175236	0.895835	0.00293665	6.22003
3	0.9	199.60	0.111728	1.08331	0.112941	0.848324	0.00190094	4.01313
3	0.95	210.61	0.0588487	1.05703	0.0593538	0.728493	0.00100226	2.11478
3	1	221.65	0	1	0	0	0	0

TABLE XVIII: Compression size and distortion results of QVZ when minimizing the MSE, for the SRR032209 data set.

		QVZ - L						
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	1.71	3.3849	4.18396	48.0109	14.6952	0.0563652	46.6396
3	0.05	14.09	2.84281	3.99093	35.9821	13.853	0.0475026	42.5238
3	0.1	25.58	2.41682	3.6914	27.048	12.8904	0.0404568	39.0439
3	0.15	36.75	2.06146	3.40178	20.4186	11.8623	0.0345688	35.9678
3	0.2	47.43	1.77008	3.10273	15.4418	10.7256	0.0296907	33.2202
3	0.25	57.87	1.52056	2.84101	11.7256	9.63916	0.0255591	30.6655
3	0.3	68.65	1.31345	2.55801	8.93676	8.59948	0.0221144	28.2375
3	0.35	80.17	1.15663	2.35056	7.14964	7.85618	0.0194593	25.9229
3	0.4	91.33	1.00205	2.13526	5.42717	6.96863	0.0168642	23.7283
3	0.45	102.49	0.866927	2.00058	4.17519	6.13269	0.0146134	21.6694
3	0.5	112.97	0.775261	1.89104	3.75621	5.79329	0.0130378	19.6116
3	0.55	123.30	0.650845	1.74385	2.73682	4.91596	0.0109676	17.48
3	0.6	134.19	0.533867	1.52112	1.80256	4.00292	0.00901487	15.3975
3	0.65	145.19	0.442548	1.36548	1.33569	3.35692	0.0074751	13.3666
3	0.7	156.15	0.367389	1.28945	1.00655	2.77877	0.00620693	11.447
3	0.75	166.93	0.310236	1.27158	0.850209	2.49463	0.00524593	9.68268
3	0.8	177.47	0.2496	1.23146	0.679841	2.19322	0.00423016	7.81291
3	0.85	188.18	0.18564	1.16587	0.447653	1.74984	0.00314828	5.95901
3	0.9	199.53	0.108748	1.09915	0.143447	1.04673	0.00185258	3.76654
3	0.95	210.12	0.0580991	1.0652	0.0707855	0.823256	0.00099161	2.01991
3	1	220.94	0	1	0	0	0	0

TABLE XIX: Compression size and distortion results of QVZ when minimizing the distortion function $\log_2(1 + L1)$, for the SRR032209 data set.

		QVZ - A						
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	1.86	3.12411	3.93028	38.7661	13.4458	0.0525461	48.3446
3	0.05	13.87	2.57594	3.70052	27.0729	12.5438	0.0435234	44.2397
3	0.1	25.22	2.1784	3.38943	19.5224	11.4893	0.0368833	40.4238
3	0.15	36.30	1.85866	3.11281	14.4569	10.4068	0.0315059	36.8918
3	0.2	46.95	1.58787	2.8389	10.5715	9.10507	0.0269383	33.8267
3	0.25	57.46	1.35998	2.58202	7.68892	7.79121	0.0230903	31.1632
3	0.3	68.63	1.16728	2.29135	5.51382	6.62256	0.0198209	28.6425
3	0.35	79.54	1.00209	2.07172	3.98153	5.58773	0.017009	26.1458
3	0.4	90.26	0.859204	1.86277	2.81466	4.63171	0.0145813	23.8074
3	0.45	101.43	0.740518	1.70566	1.99345	3.85584	0.01257	21.6131
3	0.5	113.06	0.646569	1.59569	1.56962	3.40629	0.010961	19.4941
3	0.55	124.15	0.560125	1.46489	1.19887	2.9121	0.00948298	17.5037
3	0.6	134.93	0.47073	1.33381	0.824679	2.2856	0.00796846	15.4864
3	0.65	145.88	0.392789	1.22703	0.627321	1.89127	0.00664369	13.322
3	0.7	156.22	0.330236	1.18365	0.467471	1.50907	0.0055859	11.4292
3	0.75	166.50	0.286944	1.16959	0.411117	1.47649	0.00484746	9.90161
3	0.8	177.33	0.226497	1.15343	0.345705	1.39841	0.00384146	7.76925
3	0.85	188.03	0.174386	1.13203	0.272056	1.28099	0.0029601	5.97248
3	0.9	199.19	0.108971	1.08302	0.110752	0.842502	0.00185308	3.90965
3	0.95	209.96	0.0577795	1.05738	0.0585722	0.725191	0.000983184	2.07414
3	1	220.89	0	1	0	0	0	0

TABLE XX: Compression size and distortion results of QVZ when minimizing the L1 distortion, for the SRR032209 data set.

		CRAM, DSRC2 and gzip					
Option	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
CRAM (lossless)	279	0	1	0	0	0	0
CRAM (lossy)	132	1.71023	1.55322	5.18326	3.47144	0.0350165	42.8699
DSRC2 (lossless)	219	0	1	0	0	0	0
DSRC2 (lossy)	82	1.76809	1.95028	5.5304	3.64515	0.0363586	43.8444
gzip (lossless)	287	0	1	0	0	0	0

TABLE XXI: Compression size and distortion results of CRAM, DSRC2 and gzip for the SRR032209 data set.

		P-Block					
p	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
1	53.33	0.415684	1.0336	0.415684	0.951602	0.00604734	41.984
2	34.28	0.958315	1.0651	1.46311	1.82838	0.013937	86.2095
4	22.24	1.71604	1.12048	4.50992	3.08026	0.0250841	127.314
8	14.39	2.8605	1.22261	13.3331	4.81216	0.0420172	167.483
16	8.94	5.00683	1.88672	44.3076	7.15026	0.0746254	215.839
32	3.26	6.85048	3.25355	85.0713	8.72042	0.102789	246.093

TABLE XXII: Compression size and distortion results of the P-Block algorithm, for the SRR622461 (chr 20) data set.

				R-Block			
p	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	58.96	0.380134	1.02556	0.380134	0.921032	0.00543476	38.3935
8	35.68	1.03714	1.05806	1.61254	1.88233	0.0149692	92.6917
10	29.74	1.53159	1.0794	3.35199	2.66733	0.0220529	120.782
11	29.25	1.53998	1.08172	3.38048	2.67796	0.0221915	121.3
12	26.72	1.83665	1.09274	4.87256	3.18344	0.0265035	134.796
15	23.01	2.26294	1.11117	7.50162	3.74913	0.0326897	151.831
20	19.90	2.50104	1.12948	9.21274	4.20628	0.0362775	160.769
25	17.61	2.94205	1.15199	13.0804	4.74952	0.0427442	174.391
30	16.43	3.23775	1.16905	16.4523	5.18982	0.0471278	181.997

TABLE XXIII: Compression size and distortion results of the R-Block algorithm, for the SRR622461 (chr 20) data set.

				QualComp (num. of clusters = 3)			
r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
0	0	3.68393	4.21884	43.438	14.0235	0.059044	171.672
17	6.93	1.88065	3.06951	11.5605	10.0468	0.0296905	114.711
34	13.86	1.57255	2.80037	7.8183	8.75104	0.0243804	104.849
51	20.79	1.42657	2.58768	6.38594	8.04905	0.0219263	99.3024
68	27.72	1.34242	2.44947	5.57227	7.53926	0.0205813	96.0998
85	34.65	1.28284	2.34537	5.02337	7.1035	0.0196424	93.7982
102	41.58	1.23193	2.25442	4.58155	6.71682	0.018827	91.7621
119	48.52	1.18199	2.17027	4.1901	6.35499	0.018037	89.6053
136	55.45	1.12988	2.08938	3.82947	6.00699	0.0172179	87.1411
153	62.38	1.07465	2.01412	3.49333	5.66928	0.0163548	84.2898
170	69.31	1.01867	1.94748	3.18986	5.33886	0.0154816	81.1857
187	76.25	0.962676	1.88946	2.92391	5.02941	0.0146173	77.8412
204	83.18	0.904327	1.83233	2.67304	4.73057	0.0137251	74.1879
221	90.11	0.849172	1.77612	2.44631	4.46806	0.0128759	70.6661
238	97.04	0.798912	1.72244	2.24452	4.23959	0.0120944	67.4234
255	103.97	0.753892	1.67108	2.06625	4.01888	0.0113895	64.499
272	110.90	0.71552	1.62534	1.91677	3.78592	0.0107856	61.9802
289	117.84	0.684452	1.5851	1.79791	3.53926	0.0102926	59.9229
306	124.77	0.658596	1.54751	1.69736	3.30133	0.00987602	58.2242
323	131.70	0.63673	1.51223	1.61084	3.09882	0.00952053	56.7999

TABLE XXIV: Compression size and distortion results of QualComp, for the SRR622461 (chr 20) data set.

		QVZ - M						
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	0.37	3.51361	4.3059	42.9423	13.9931	0.056296	162.936
3	0.05	3.49	1.92048	4.1203	19.2066	13.2338	0.0310979	95.6051
3	0.1	6.25	1.31864	3.70917	11.3421	11.8982	0.0214217	69.8164
3	0.15	8.92	0.978445	3.25002	7.12132	10.2994	0.0158818	56.1717
3	0.2	11.62	0.772468	2.7894	4.62767	8.7006	0.0124794	48.8472
3	0.25	14.42	0.625288	2.32867	3.11122	7.42649	0.0100463	42.7944
3	0.3	17.15	0.524691	2.0263	2.22272	6.36511	0.00839019	38.4431
3	0.35	19.87	0.443157	1.79316	1.60046	5.35068	0.00705582	34.4762
3	0.4	22.55	0.367884	1.61354	1.1044	4.24897	0.00583698	30.5024
3	0.45	25.37	0.31434	1.46615	0.806319	3.48649	0.00496434	27.1934
3	0.5	28.22	0.26741	1.34194	0.579504	2.82508	0.00420645	24.0503
3	0.55	31.01	0.228073	1.25421	0.421415	2.28936	0.00357384	21.1793
3	0.6	33.72	0.194998	1.17856	0.301367	1.84786	0.0030456	18.692
3	0.65	36.34	0.167368	1.11937	0.220556	1.49408	0.00259917	16.4068
3	0.7	38.77	0.136857	1.07427	0.159903	1.24826	0.00211997	13.5854
3	0.75	41.22	0.104974	1.05083	0.112229	1.0645	0.00162747	10.5174
3	0.8	43.83	0.0786102	1.04097	0.0799614	0.938421	0.00122085	7.92388
3	0.85	46.48	0.0565586	1.0364	0.0567142	0.870532	0.000878199	5.70997
3	0.9	49.05	0.0358552	1.03187	0.0358633	0.788182	0.000557476	3.62123
3	0.95	51.65	0.0164848	1.02393	0.0164853	0.619079	0.000257609	1.66496
3	1	54.11	0	1	0	0	0	0

TABLE XXV: Compression size and distortion results of QVZ when minimizing the MSE, for the SRR622461 (chr 20) data set.

		QVZ - L						
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	0.28	2.62966	5.45917	58.8739	16.0366	0.0412496	74.5058
3	0.05	3.73	1.71205	5.76178	29.3522	15.6739	0.0274712	59.7582
3	0.1	6.87	1.32301	5.3004	18.702	14.5068	0.0212186	52.637
3	0.15	9.91	1.08875	4.6655	13.3337	13.182	0.0174583	47.4001
3	0.2	12.81	0.930696	4.20487	10.2479	11.9852	0.0149363	43.297
3	0.25	15.54	0.762502	3.55667	7.37527	10.7662	0.0121858	38.8364
3	0.3	18.16	0.605981	2.97476	4.93281	9.53684	0.00958305	34.4699
3	0.35	20.71	0.484742	2.47638	3.24857	8.14137	0.0076241	30.7239
3	0.4	23.17	0.406423	2.15628	2.25947	7.01431	0.00636777	27.856
3	0.45	25.67	0.351957	1.89236	1.71943	6.24706	0.00550149	25.3241
3	0.5	28.18	0.303251	1.67013	1.28606	5.5175	0.00472687	22.8733
3	0.55	30.78	0.263107	1.4973	0.992319	4.95653	0.00408205	20.5704
3	0.6	33.41	0.225606	1.36533	0.767315	4.41124	0.00348861	18.2541
3	0.65	36.01	0.188776	1.26895	0.565081	3.78914	0.00291051	15.8684
3	0.7	38.58	0.150451	1.19205	0.38398	3.03232	0.00231825	13.1993
3	0.75	41.17	0.109486	1.0876	0.174608	1.79336	0.00169132	10.3967
3	0.8	43.91	0.077884	1.04548	0.0828847	1.03866	0.00120522	7.78987
3	0.85	46.61	0.0559333	1.03808	0.0573665	0.903985	0.000866018	5.63036
3	0.9	49.20	0.03588	1.03209	0.0359814	0.795056	0.000555863	3.62239
3	0.95	51.81	0.0164465	1.02418	0.0164766	0.624681	0.000255876	1.66078
3	1	54.28	0	1	0	0	0	0

TABLE XXVI: Compression size and distortion results of QVZ when minimizing the distortion function $\log_2(1 + L1)$, for the SRR622461 (chr 20) data set.

QVZ - A								
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	0.30	2.5904	5.33073	53.7149	15.8607	0.0413345	76.5222
3	0.05	3.63	1.5822	5.41177	24.4272	15.0142	0.0253296	59.5968
3	0.1	6.62	1.19458	4.89521	15.2739	13.6541	0.0190562	51.9025
3	0.15	9.53	0.973053	4.46788	10.7211	12.4783	0.0155065	46.584
3	0.2	12.26	0.794386	3.92926	7.43616	11.1051	0.0126478	41.96
3	0.25	14.87	0.660733	3.34287	5.29551	9.88709	0.0105091	38.1034
3	0.3	17.51	0.555552	2.77763	3.88294	8.76163	0.00878627	34.6733
3	0.35	20.14	0.453999	2.32776	2.58716	7.42768	0.00713778	31.25
3	0.4	22.96	0.366929	1.89668	1.5874	5.53864	0.00576609	27.8125
3	0.45	25.74	0.307968	1.6193	1.0838	4.43852	0.00483841	24.7771
3	0.5	28.42	0.261794	1.42785	0.739017	3.54423	0.00410367	22.2094
3	0.55	31.08	0.222442	1.28157	0.50111	2.82513	0.00347266	19.7837
3	0.6	33.66	0.190315	1.18914	0.348975	2.32777	0.00295947	17.617
3	0.65	36.23	0.162519	1.12925	0.25302	1.97739	0.00251592	15.4448
3	0.7	38.52	0.134804	1.08518	0.176324	1.57046	0.00208318	13.1068
3	0.75	41.12	0.103973	1.05447	0.116018	1.18015	0.00160768	10.3375
3	0.8	43.88	0.0773809	1.04201	0.0796197	0.964543	0.00119836	7.7833
3	0.85	46.57	0.0559668	1.03672	0.0563294	0.879235	0.000866974	5.64651
3	0.9	49.15	0.0359006	1.03198	0.0359574	0.793297	0.000556651	3.62501
3	0.95	51.77	0.0164366	1.02413	0.0164431	0.623091	0.000255993	1.66004
3	1	54.23	0	1	0	0	0	0

TABLE XXVII: Compression size and distortion results of QVZ when minimizing the L1 distortion, for the SRR622461 (chr 20) data set.

CRAM, DSRC2 and gzip							
Option	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
CRAM (lossless)	76	0	1	0	0	0	0
CRAM (lossy)	37	1.99273	1.42671	4.84353	2.90951	0.0319889	151.451
DSRC2 (lossless)	54	0	1	0	0	0	0
DSRC2 (lossy)	22	1.99273	1.42671	4.84353	2.90951	0.0319889	151.451
gzip (lossless)	77	0	1	0	0	0	0

TABLE XXVIII: Compression size and distortion results of CRAM, DSRC2 and gzip for the SRR622461 (chr 20) data set.

P-Block							
p	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
1	370	0.675827	1.03013	0.675827	0.998772	0.00942267	67.5827
2	236	1.24724	1.06062	2.11412	1.99762	0.0174121	106.735
4	119	2.68705	1.13686	9.16467	3.9953	0.0375226	172.461
8	66	3.56066	1.43196	15.9803	5.57092	0.049982	202.857
16	46	4.87745	1.98773	37.3432	7.82002	0.0692436	228.397
32	16	6.83257	2.31688	80.0011	8.69095	0.0970828	259.083

TABLE XXIX: Compression size and distortion results of the P-Block algorithm, for the SRR1179906 data set.

				R-Block			
p	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	383	0.656682	1.02699	0.656682	0.993007	0.00913461	65.6682
8	222	1.6261	1.07233	3.55301	2.71779	0.0226816	126.915
10	198	1.7774	1.0888	4.14751	2.99418	0.0248206	135.495
11	146	2.57121	1.10452	8.56521	3.90812	0.0358082	167.128
12	141	2.58979	1.10829	8.65147	3.97675	0.0360913	168.038
15	114	3.01573	1.12874	11.4561	4.64921	0.0420766	184.665
20	87	3.51964	1.14622	15.4203	5.17995	0.0491625	201.576
25	76	3.83363	1.16095	18.5273	5.6413	0.0536552	210.213
30	71	3.90975	1.1679	19.3679	5.78838	0.0547856	212.421

TABLE XXX: Compression size and distortion results of the R-Block algorithm, for the SRR1179906 data set.

		QualComp (num. of clusters = 3)					
r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
0	0.31	2.30161	2.96579	17.4224	13.1209	0.0332686	141.707
20	40.26	1.4306	1.91067	7.15987	8.07479	0.0206523	98.5238
40	80.21	1.28702	1.75856	5.90857	7.13854	0.0186115	91.0318
60	119.21	1.21277	1.68469	5.2065	6.58116	0.0175479	87.457
80	159.01	1.15486	1.6253	4.67622	6.14071	0.0167114	84.7025
100	198.91	1.10037	1.57394	4.2142	5.74857	0.015921	82.1078
120	237.81	1.04555	1.52795	3.80004	5.36044	0.0151242	79.4035
140	277.71	0.993135	1.49208	3.43055	4.98781	0.0143604	76.7975
160	317.71	0.942722	1.45843	3.09998	4.61226	0.0136259	74.2397
180	357.21	0.892686	1.4263	2.80036	4.25544	0.0128969	71.5833
200	396.71	0.841304	1.39565	2.52018	3.95766	0.0121484	68.718
220	437.31	0.794719	1.36785	2.27745	3.72785	0.0114675	66.0477
240	475.81	0.75622	1.34458	2.07556	3.50446	0.0109016	63.8185
260	515.41	0.724625	1.32151	1.90608	3.2192	0.0104343	61.9841
280	554.91	0.699997	1.29897	1.7663	2.902	0.0100656	60.5831
300	594.41	0.679757	1.27822	1.64572	2.62994	0.00975828	59.4535

TABLE XXXI: Compression size and distortion results of QualComp, for the SRR1179906 data set.

			QVZ - M					
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	0.86	1.85036	3.00542	16.785	12.8801	0.0269202	106.101
3	0.05	17.54	1.41419	2.77706	9.36286	11.5432	0.0206023	91.8337
3	0.1	33.59	1.17822	2.5548	6.42772	10.1598	0.0171507	81.3785
3	0.15	49.69	0.973773	2.34834	4.52878	8.94525	0.0141732	70.2615
3	0.2	65.56	0.816807	2.14673	3.32184	7.88272	0.0118861	61.2816
3	0.25	80.67	0.677248	1.95999	2.46533	6.91546	0.00986304	52.2421
3	0.3	95.60	0.576201	1.80495	1.86179	6.04232	0.00839068	45.9752
3	0.35	109.96	0.487582	1.65212	1.34373	5.0057	0.00710146	40.5833
3	0.4	124.65	0.412738	1.49617	0.937004	3.90153	0.00600423	35.9647
3	0.45	139.55	0.351891	1.36989	0.673787	3.1285	0.00510827	31.6721
3	0.5	154.82	0.299497	1.26192	0.459073	2.44353	0.00433943	27.7746
3	0.55	170.35	0.259539	1.18676	0.352151	2.05759	0.00375461	24.5746
3	0.6	185.48	0.222422	1.12048	0.275727	1.74557	0.00321291	21.4345
3	0.65	200.32	0.18793	1.06724	0.215472	1.49461	0.00271186	18.3458
3	0.7	215.33	0.158287	1.0573	0.175158	1.38436	0.0022825	15.5505
3	0.75	230.16	0.128262	1.04611	0.137409	1.25023	0.00184666	12.671
3	0.8	244.92	0.100817	1.04104	0.106248	1.17177	0.00145034	9.98767
3	0.85	259.76	0.0723305	1.03613	0.0752415	1.09518	0.00104001	7.1816
3	0.9	274.32	0.046696	1.03164	0.0479076	1.01416	0.00067077	4.64749
3	0.95	288.94	0.022562	1.02508	0.0228575	0.854281	0.000323953	2.25094
3	1	302.84	0	1	0	0	0	0

TABLE XXXII: Compression size and distortion results of QVZ when minimizing the MSE, for the SRR1179906 data set.

			QVZ - L					
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	0.45	1.62473	3.41661	23.7493	14.3394	0.0232634	70.933
3	0.05	19.24	1.30311	3.34585	15.2447	13.3916	0.0187523	63.5074
3	0.1	36.32	1.11637	3.11766	11.6645	12.5393	0.0160938	57.9957
3	0.15	52.47	0.956112	2.89011	8.83581	11.6277	0.0138275	53.0118
3	0.2	67.74	0.820443	2.65983	6.73768	10.6786	0.0118946	48.397
3	0.25	82.42	0.708227	2.43755	5.14687	9.68912	0.010282	44.3055
3	0.3	97.27	0.616176	2.21051	3.96606	8.64746	0.00895381	40.4991
3	0.35	112.30	0.539652	2.00392	3.20497	7.58439	0.00784424	36.7974
3	0.4	126.62	0.464236	1.81987	2.41965	6.40359	0.00675147	33.4111
3	0.45	141.14	0.387232	1.60552	1.66053	5.17548	0.00562884	29.7461
3	0.5	156.19	0.32292	1.44179	1.12398	4.3228	0.00468307	26.2008
3	0.55	171.52	0.273285	1.32189	0.792397	3.64849	0.00395511	23.1711
3	0.6	186.62	0.229471	1.19341	0.54204	2.97265	0.00331518	20.2691
3	0.65	201.64	0.192756	1.12731	0.378228	2.48852	0.00277926	17.5639
3	0.7	216.67	0.160419	1.09918	0.290801	2.20649	0.00230892	14.8497
3	0.75	231.52	0.12957	1.07608	0.213687	1.85964	0.0018633	12.2036
3	0.8	246.27	0.101235	1.06394	0.152324	1.6033	0.00145426	9.66181
3	0.85	260.97	0.0731642	1.04634	0.0924096	1.27585	0.00105034	7.1281
3	0.9	275.60	0.0469608	1.03347	0.0513922	1.06558	0.000673607	4.64068
3	0.95	290.44	0.0226369	1.02568	0.0238246	0.871324	0.000324592	2.24864
3	1	304.43	0	1	0	0	0	0

TABLE XXXIII: Compression size and distortion results of QVZ when minimizing the distortion function $\log_2(1 + L1)$, for the SRR1179906 data set.

			QVZ - A					
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	0.51	1.57605	3.38354	20.5267	14.1169	0.0227979	72.8617
3	0.05	19.19	1.23165	3.18256	13.0382	13.0966	0.017784	63.2128
3	0.1	35.96	1.03194	2.97211	9.4568	12.1785	0.014955	57.3078
3	0.15	51.56	0.883016	2.77364	7.11588	11.2519	0.0128259	52.4542
3	0.2	66.70	0.755089	2.55121	5.23241	10.1457	0.0109835	48.1575
3	0.25	81.54	0.654151	2.32431	3.92723	9.06703	0.00952075	44.3989
3	0.3	96.02	0.56531	2.08043	2.82845	7.7714	0.00822693	40.804
3	0.35	110.88	0.490736	1.90372	2.15633	6.60496	0.00713758	37.1981
3	0.4	125.60	0.419515	1.69883	1.54179	5.20536	0.00610022	33.5416
3	0.45	140.65	0.358138	1.51302	1.0877	4.20193	0.00520108	29.936
3	0.5	155.87	0.30679	1.3724	0.770637	3.53723	0.00444484	26.4603
3	0.55	171.52	0.261114	1.25591	0.538266	2.92877	0.00377459	23.3619
3	0.6	186.74	0.221904	1.15388	0.3882	2.44983	0.00320158	20.4297
3	0.65	201.79	0.188445	1.09006	0.293359	2.11724	0.00271382	17.6841
3	0.7	216.80	0.157779	1.07523	0.23331	1.91771	0.0022688	14.9504
3	0.75	231.62	0.128159	1.06137	0.179565	1.67462	0.00184169	12.2667
3	0.8	246.34	0.0999881	1.05157	0.132477	1.45187	0.00143603	9.66961
3	0.85	260.99	0.0723261	1.03855	0.079949	1.16584	0.00103827	7.12953
3	0.9	275.66	0.0468332	1.03234	0.0491655	1.03465	0.000671869	4.64892
3	0.95	290.42	0.0225827	1.02534	0.0232251	0.863534	0.000323893	2.2487
3	1	304.43	0	1	0	0	0	0

TABLE XXXIV: Compression size and distortion results of QVZ when minimizing the L1 distortion, for the SRR1179906 data set.

			DSRC2 and gzip				
Option	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
DSRC2 (lossless)	306	0	1	0	0	0	0
DSRC2 (lossy)	125	0.966278	1.30868	1.60483	2.67583	0.013956	85.4307
gzip (lossless)	479	0	1	0	0	0	0

TABLE XXXV: Compression size and distortion results of DSRC2 and gzip for the SRR1179906 data set.

			P-Block				
p	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
1	116	0.494621	1.03592	0.494621	0.999064	0.00765235	17.8063
2	74	1.02185	1.06931	1.56608	1.92809	0.0158093	32.7206
4	45	1.83605	1.12058	4.74436	3.16897	0.0284683	48.8863
8	29	2.89916	1.21386	12.5104	4.47146	0.0451387	62.8938
16	24	3.80581	1.78984	26.3612	5.54198	0.0597094	69.8261
32	11	4.03354	1.86002	30.593	5.59874	0.0632586	71.2728

TABLE XXXVI: Compression size and distortion results of the P-Block algorithm, for the ERR011354 data set.

		R-Block					
p	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	242	0	1	0	0	0	0
8	75	1.10088	1.06663	1.75281	1.98379	0.017003	34.7613
10	65	1.50357	1.08714	3.15847	2.7178	0.0232297	43.0532
11	62	1.55685	1.09075	3.32602	2.78254	0.0240568	44.2606
12	56	1.6359	1.09575	3.60776	2.81718	0.0252918	45.9321
15	50	1.95789	1.11145	5.15092	3.35837	0.0303041	51.4532
20	42	2.33184	1.13185	7.28877	3.86274	0.03614	57.2075
25	37	2.65029	1.14773	9.48003	4.2478	0.0411172	61.5107
30	33	2.98556	1.16395	12.3115	4.61044	0.0463807	65.3647

TABLE XXXVII: Compression size and distortion results of the R-Block algorithm, for the ERR011354 data set.

		QualComp (num. of clusters = 3)					
r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
0	0.06	1.99196	2.35639	9.84632	7.97148	0.0310317	48.097
5	6.93	1.62407	2.17362	6.98131	7.03942	0.0253825	40.9238
10	14.02	1.54616	2.08264	6.12095	6.51466	0.02416	39.8053
15	21.01	1.47723	2.01276	5.48274	6.08153	0.0230951	38.6641
20	27.89	1.4145	1.95701	4.96537	5.72622	0.0221238	37.5762
25	34.88	1.35466	1.9095	4.51425	5.41415	0.0211951	36.5175
30	41.86	1.29609	1.86633	4.11008	5.12332	0.0202842	35.4542
35	48.74	1.23949	1.82588	3.74623	4.84596	0.0194038	34.4024
40	54.73	1.1839	1.78804	3.41598	4.58146	0.0185408	33.3329
45	61.71	1.12757	1.75194	3.10845	4.32403	0.017666	32.2044
50	68.60	1.0718	1.71755	2.82271	4.07505	0.0167988	31.0634
55	75.58	1.01621	1.6863	2.55956	3.83816	0.0159336	29.8924
60	83.06	0.960591	1.65699	2.31319	3.61559	0.0150669	28.6906
65	90.16	0.907903	1.62954	2.09234	3.41513	0.0142446	27.5217
70	97.26	0.859636	1.60371	1.8974	3.23126	0.01349	26.4313
75	103.36	0.8158	1.57889	1.72585	3.04999	0.0128036	25.4252
80	109.46	0.777057	1.55475	1.57518	2.86098	0.0121955	24.5386
85	116.46	0.743121	1.53096	1.44505	2.67038	0.011661	23.7557
90	123.56	0.713539	1.50815	1.33367	2.48375	0.0111938	23.0648
95	130.66	0.687907	1.4865	1.23783	2.31112	0.0107876	22.4625
100	137.76	0.665672	1.46589	1.15499	2.15785	0.0104339	21.9373

TABLE XXXVIII: Compression size and distortion results of QualComp, for the ERR011354 data set.

			QVZ - M					
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	0.79	1.70387	2.39442	9.61586	8.27705	0.0265501	39.4603
3	0.05	7.14	1.47785	2.19154	7.20942	7.40389	0.0230506	36.224
3	0.1	13.30	1.29112	1.95661	5.33077	6.38396	0.0201646	33.3627
3	0.15	19.44	1.12902	1.75123	3.8329	5.31983	0.0176493	30.8395
3	0.2	25.50	0.983131	1.53727	2.651	4.20675	0.0153824	28.3356
3	0.25	31.65	0.869683	1.41951	2.00554	3.60023	0.0136124	26.0759
3	0.3	37.86	0.771607	1.33905	1.51723	3.08543	0.0120751	24.0403
3	0.35	44.05	0.681725	1.29376	1.20336	2.74932	0.0106706	21.8164
3	0.4	50.15	0.597886	1.24973	0.941161	2.37155	0.00936256	19.634
3	0.45	56.26	0.53502	1.20962	0.760876	2.03228	0.00837798	17.9674
3	0.5	62.37	0.478945	1.16727	0.615572	1.72883	0.00750241	16.4059
3	0.55	68.44	0.427334	1.11651	0.508464	1.51052	0.00669358	14.8552
3	0.6	74.18	0.378218	1.06439	0.406014	1.24423	0.00592216	13.4295
3	0.65	80.31	0.322198	1.05108	0.329178	1.0868	0.00504404	11.5527
3	0.7	86.57	0.273641	1.04585	0.276415	1.03374	0.00428241	9.8337
3	0.75	92.55	0.228825	1.04403	0.231216	1.02517	0.00358587	8.22261
3	0.8	98.43	0.188084	1.04283	0.190209	1.01626	0.00294823	6.75744
3	0.85	104.24	0.149231	1.04078	0.150324	0.988473	0.00234142	5.36594
3	0.9	110.19	0.101829	1.03763	0.102453	0.929258	0.00160019	3.66244
3	0.95	116.18	0.0503075	1.02975	0.0503114	0.735825	0.000792985	1.81104
3	1	122.18	0	1	0	0	0	0

TABLE XXXIX: Compression size and distortion results of QVZ when minimizing the MSE, for the ERR011354 data set.

			QVZ - L					
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	0.19	1.70469	2.67278	14.9375	9.7352	0.0261031	33.2797
3	0.05	6.87	1.53928	2.53774	12.3351	9.25424	0.0236416	31.1986
3	0.1	13.28	1.41161	2.4514	10.6989	8.84187	0.0217444	29.4374
3	0.15	19.62	1.29712	2.3689	9.45689	8.43262	0.0200229	27.7441
3	0.2	25.84	1.18109	2.30646	8.19314	7.99989	0.0182817	26.1048
3	0.25	31.87	1.0749	2.24272	7.03766	7.54384	0.016679	24.5101
3	0.3	37.64	0.97962	2.13523	5.85036	7.1334	0.015227	23.0167
3	0.35	44.36	0.891869	2.05431	4.97373	6.80814	0.0138835	21.4202
3	0.4	50.39	0.801573	1.94001	4.0916	6.27171	0.0124969	19.9684
3	0.45	56.52	0.713274	1.84343	3.37192	5.75486	0.0111354	18.3315
3	0.5	62.95	0.64138	1.80567	3.07196	5.54518	0.0100247	16.5233
3	0.55	69.00	0.558788	1.69227	2.53166	5.04044	0.00874749	14.733
3	0.6	75.04	0.474592	1.5526	1.9585	4.413	0.00743797	13.0073
3	0.65	81.10	0.38914	1.35797	1.37169	3.59408	0.00610167	11.2663
3	0.7	87.32	0.311446	1.22441	0.885211	2.76183	0.00488613	9.55407
3	0.75	93.18	0.258552	1.20388	0.688757	2.39862	0.00406226	8.05605
3	0.8	99.17	0.212341	1.16472	0.56123	2.19632	0.00334063	6.6145
3	0.85	105.16	0.155937	1.1214	0.389819	1.79864	0.00245722	4.93292
3	0.9	111.26	0.101457	1.07562	0.250553	1.47546	0.00160151	3.18724
3	0.95	117.26	0.0509787	1.04718	0.106652	0.932764	0.000808194	1.65846
3	1	123.27	0	1	0	0	0	0

TABLE XL: Compression size and distortion results of QVZ when minimizing the distortion function $\log_2(1 + L1)$, for the ERR011354 data set.

			QVZ - A					
c	r	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
3	0	0.36	1.59865	2.59485	12.0674	9.20412	0.0246654	33.807
3	0.05	6.94	1.4276	2.40681	9.57657	8.58562	0.0220709	31.6036
3	0.1	13.25	1.29012	2.28338	7.88956	8.02059	0.0199937	29.7694
3	0.15	19.49	1.16101	2.17242	6.4478	7.41818	0.0180321	27.966
3	0.2	25.63	1.04036	2.10929	5.2313	6.80247	0.0161897	26.2306
3	0.25	31.75	0.922154	2.01387	4.14575	6.13951	0.0143779	24.3522
3	0.3	37.89	0.824592	1.85874	3.27871	5.44334	0.0128715	22.6948
3	0.35	43.88	0.733488	1.72855	2.5047	4.66851	0.0114619	21.1354
3	0.4	49.96	0.645413	1.56699	1.77628	3.75	0.0100932	19.6099
3	0.45	56.37	0.573596	1.41944	1.3216	3.13303	0.00897314	18.0268
3	0.5	62.71	0.5142	1.36651	1.10535	2.87352	0.00804371	16.4338
3	0.55	68.95	0.454905	1.3175	0.887289	2.53862	0.00711701	14.843
3	0.6	74.91	0.39729	1.24488	0.716263	2.26203	0.00621422	13.2094
3	0.65	80.76	0.340276	1.16427	0.572328	1.99585	0.00532195	11.4836
3	0.7	86.98	0.283198	1.13264	0.444573	1.72074	0.0044335	9.66971
3	0.75	92.92	0.239534	1.12408	0.384277	1.63768	0.0037552	8.15623
3	0.8	98.87	0.193744	1.09718	0.312878	1.51974	0.00304258	6.59558
3	0.85	104.72	0.146765	1.07445	0.223296	1.34051	0.00230836	5.02823
3	0.9	110.82	0.0976967	1.04586	0.124385	1.05564	0.00153979	3.41919
3	0.95	116.93	0.0473244	1.03117	0.0534913	0.748348	0.00074751	1.68082
3	1	122.92	0	1	0	0	0	0

TABLE XLI: Compression size and distortion results of QVZ when minimizing the L1 distortion, for the ERR011354 data set.

			DSRC2 and gzip				
Option	Size [MB]	Manhattan	Max:Min	MSE	Chebyshev	Soergel	Lorentzian
DSRC2 (lossless)	121	0	1	0	0	0	0
DSRC2 (lossy)	30	0.826888	1.31279	1.62224	2.92699	0.012929	24.79
gzip (lossless)	165	0	1	0	0	0	0

TABLE XLII: Compression size and distortion results of DSRC2 and gzip for the ERR011354 data set.