## E WAN BENCHMARKING

In Fig. 5, we depict the benchmarking results of our $\chi^2$-test algorithms in the WAN setting and compare them to the LAN results.

We evaluated the algorithms both with the GMW and Yao's GC protocol (cf. §2.4.4 and §2.4.5, respectively). To do so, we replaced Boolean sharing with Yao sharing in all our algorithms. However, we found that GMW outperforms Yao's GC protocol despite a non-constant number of communication rounds. Thus, we report only results using Boolean sharing.

In our benchmarks, the algorithms were approximately by an order of magnitude slower in the WAN setting than in the LAN setting. This substantial difference was motivation to find a way for shifting the use of our algorithms to the LAN setting for spatially distant institutes, resulting in the outsourcing scenario discussed in §4.7 and §5.1.5. Interestingly, in the WAN setting the large-scale $\chi^2$-test algorithm performs better than the optimized version due to the by factor ~2x lower AND-depth (cf. §4.6).

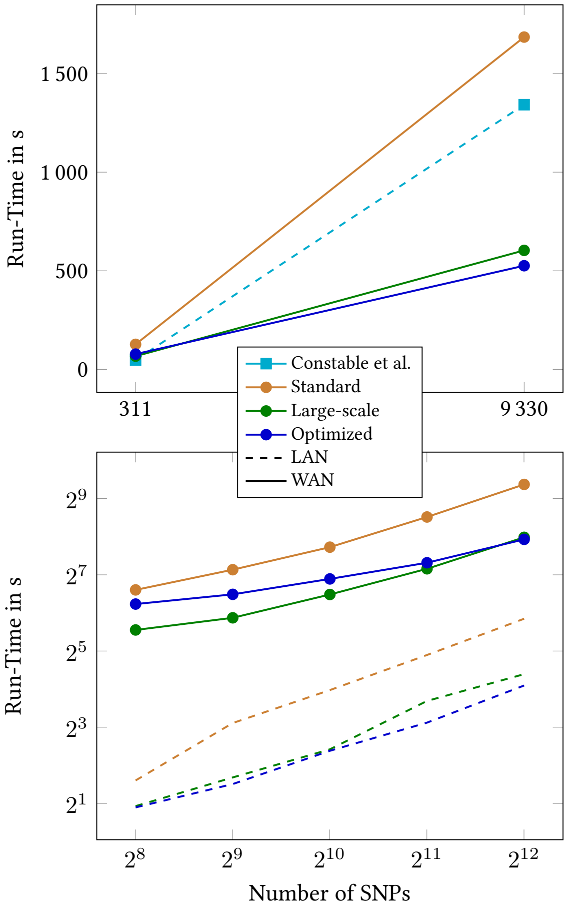

**Figure 5: Run-times of the $\chi^2$-test algorithms in the LAN and WAN setting. The upper graph shows that our optimized and large-scale algorithms are even faster in the WAN setting than the algorithm by Constable et al. [14] in the LAN setting.**

## F COMMUNICATION COSTS

The total amortized communication costs for performing a test on one SNP between two institutes or between two STTPs are shown in Tab. 10. Communication is almost equally distributed among the two parties. Hence, each party sends and receives about half the amount of data given in the table.

**Table 10: Total amortized communication in megabytes of our $\chi^2$-, G-, and P-test algorithms for one SNP.**

| $\chi^2_s$ | $\chi^2_l$ | $\chi^2_o$ | $g_s$ | $g_l$ | $g_o$ | $p$ (32-bit) | $p$ (64-bit) |
|---|---|---|---|---|---|---|---|
| 2.95 | 0.98 | 0.82 | 4.06 | 3.32 | 2.63 | 0.01 | 0.02 |

In Fig. 6 we depict the communication costs for the different versions of the $\chi^2$- and G-test algorithms. Note that the communication costs are independent of the network setting and whether the protocols are run directly between two institutes or in the outsourcing scenario.

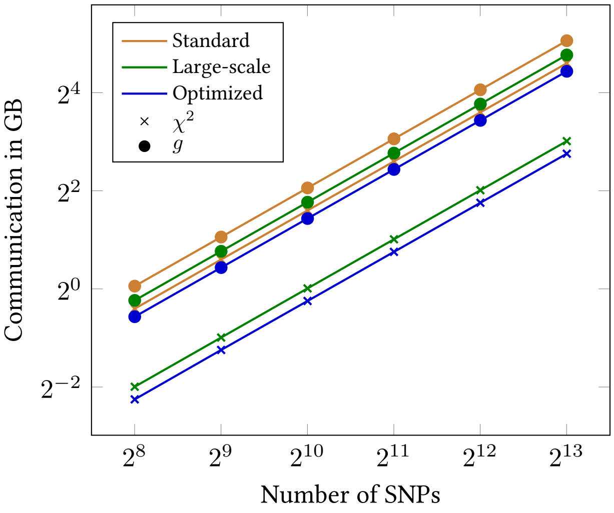

**Figure 6: Communication costs of our $\chi^2$- and G-test algorithms for different numbers of SNPs.**

In Tab. 11 we give the amortized communication costs of the extended $\chi^2$- and G-test for different codeword counts. Obviously, the algorithms scale linearly in the number of codeword counts.

**Table 11: Total amortized communication in megabytes of our extended $\chi^2$- and G-test algorithms for different codeword counts and one SNP.**

| # Codeword Counts | 4 | 8 | 16 | 32 | 64 | 128 | 256 |
|---|---|---|---|---|---|---|---|
| $\chi^2_e$ | 2.8 | 5.7 | 11 | 23 | 46 | 92 | 184 |
| $g_e$ | 3.9 | 7.9 | 15 | 31 | 63 | 127 | 255 |