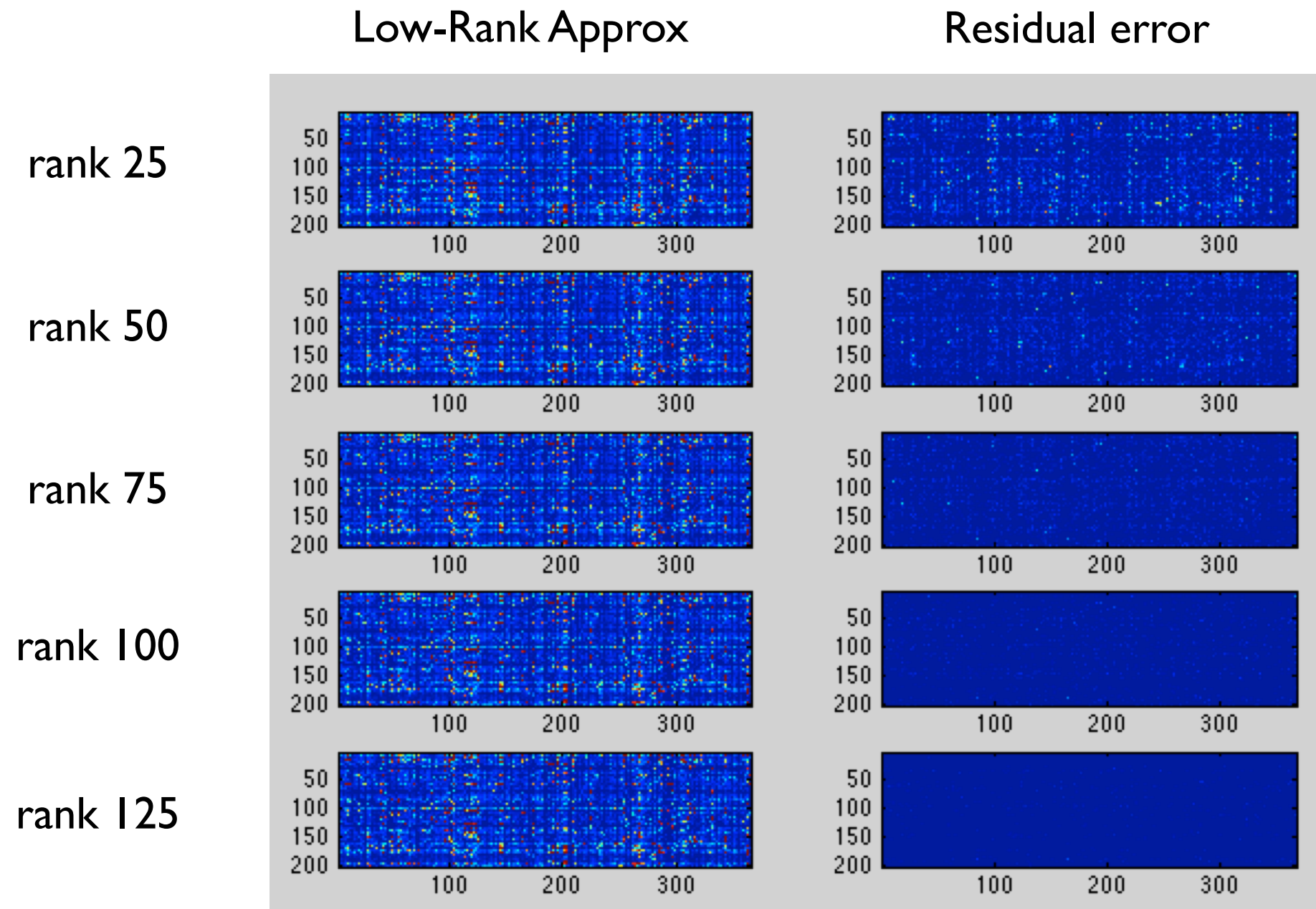


Here I took the data [GSK2013_0_1uM_matrix](#) and looked at various low-rank approximations. Each approximation is the low-rank matrix (of that rank) that is closest to the original data. The error corresponding to each case is shown in the right column. You can see that the error is negligible at rank 75. Also not that the error at rank 25 seems to have an intensity pattern similar to the original data.



Here is the same thing, except that I started with the square-root of the data, rather than the original data. I did this because the intensity of the errors on the previous page seemed to match that of the data, and taking the square-root is a “variance-stabilizing transformation” that accounts for this sort of dependence. You can see that the residuals are more noise-like (random) and less severe now. In fact, the rank 25 approximation looks pretty good. Bottom line, I think that a rank 25 approximation is pretty reasonable for these data.

