1. There has been work on using Riemannian approaches before about 10 or so years ago. I will see if I can track it down. It was by two Chinese researchers and I do not think it was successful in getting published for quite some time since no one thought Riemannian stuff could ever be efficient and in fact I think that was true with their approach in solving the two Sylvester equation. I vaguely recall the iteration being very much like IRKA.
2. There was a Newton-based IRKA that Pierre and I discussed a long time ago by Beattie and Gugercin I think. It had an assumption that the Hessian was diagonal or was approximated as such for efficiency. I will try to find both of these papers.
3. The first task is too implement manifold version of IRKA and the version that they have modified to use a Riemannian step size line search and check if we can replicate the results in their paper. Complexity measures at first should probably concentrate on function calls, etc. and translate them into total operations. Space should be carefully compared to later versions. I think using Roptlib will simplify the eventual shift to timing comparably optimized and compiled versions of all the codes but certainly initial stuff can avoid this as mentioned above.
4. I would start with LRBFGS on their manifold then move to LRTRSR1. I am not sure that we can justify any storage above the simple descent that they seem to be using but Limited Memory versions might be justifiable if we can show like with Xinru's stuff that there is a sweet spot of accelerated convergence compared to their first order method.
5. It may be worth then looking at a LM Broyden with the adaptive phi\_k from Shuguang's work.
6. convergence results from our papers and comments on transport, line search termination and other parameters and techniques could be compared to what they assume and know from a systems point of view. I do not recall IRKA or its accelerated version having rigorous convergence without serious simplifying assumptions such as the diagonal Hessian.
7. The earlier work was on stiefel or grassmann. Their paper seems novel in that regard but I have looked at it carefully. There is the possibility that some of Zhifeng's work might become relevant although we would have to look at orthogonal projectors first as some subset of his fixed rank spsd set.
8. Tangential interpolation from the old work of paul, pierre and me might be interesting to eventually look at but I am not sure what the state of knowledge is about our stuff versus IRKA and more sophisticated stuff by Antoulas that uses Loewner matrices and I/O based system identification and has very little to do with projection iterations directly as far as I can tell.