On Stability of Newton Schulz Iterations in an Approximate Algebra

Matt Challacombe* and Nicolas Bock†

Theoretical Division, Los Alamos National Laboratory

I. INTRODUCTION

In many disciplines, finite correlations coorespond to matrices with decay properties. Matrix decay involves an approximate (perhaps bounded) inverse relationship between matrix elements and a related distance; this may be a simple inverse exponential relationship between elements and the Cartesian distance between support functions, or it may involve a statistical distance, *e.g.* between strings. In electronic structure, correlations manifest in decay properties of the matrix sign function, as projector of the effective Hamiltonian (See Fig. **??**. More broadly, matrix decay properties may coorespond to learned correlations in a generalized, non-orthogonal metric, obtained perhaps through first order optimization involving the celebrated PLSEV line search approach to semi-definate programming based on the matrix sign function. More broadly still, problems with local, non-orothogonal support are often solved with congruential transformations based on the matrix inverse square root; these transformations determine correlations between the local support and the representation independent form [Lowdin], *eg.* of the generalized eigenproblem. Higham's identity relates the matrix sign function with the matrix inverse square root:

$$\mathrm{sign}\left(\begin{bmatrix} 0 & \boldsymbol{s} \\ \boldsymbol{I} & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & \boldsymbol{s}^{1/2} \\ \boldsymbol{s}^{-1/2} & 0 \end{bmatrix}. \tag{1}$$

A complete overivew of matrix function theory and computation is given in Higham's excellent reference [].

A well conditioned matrix $\boldsymbol{s}$ may correspond to matrix sign and inverse square root functions with rapid exponential decay, and be amenable to the sparse matrix approximation $\bar{\boldsymbol{s}} = \boldsymbol{s} + \boldsymbol{\epsilon}^{\boldsymbol{s}}_{\tau}$, where $\boldsymbol{\epsilon}^{\boldsymbol{s}}_{\tau}$ is the error introduced according to some criteria τ. This criteria might be a drop-tolerence, $\epsilon^{\boldsymbol{s}}_{\tau} = \{-s_{ij} * \hat{\boldsymbol{e}}_i \,|\, |s_{ij}| < \tau\}$, a radial cutoff, $\epsilon^{\boldsymbol{s}}_{\tau} = \{-s_{ij} * \hat{\boldsymbol{e}}_i \,|\, \|\boldsymbol{r}_i - \boldsymbol{r}_j\| > \tau\}$, or some other approach to truncation, perhaps involving a sparsity pattern chosen *a priori*. Then, conventional computational kernels may be employed, such as the Sparse Matrix Multiply (`SpMM`) [**? ? ?**], yeiding fast solutions for multiplication rich iterations and a modulated fill in. These and related incomplete/inexact approaches to the computation of sparse approximate matrix functions often lead to $\mathcal{O}(n)$ algorithms, finding wide use in technologically important preconditioning schemes, the information sciences, electronic structure and many other disciplines. A comprehensive overview of these methods in the numerical linear algebra is given by Benzi []. See also Bowler for a complete development of these methods in electronic structure [**?**], and a current summary of high performance results [].

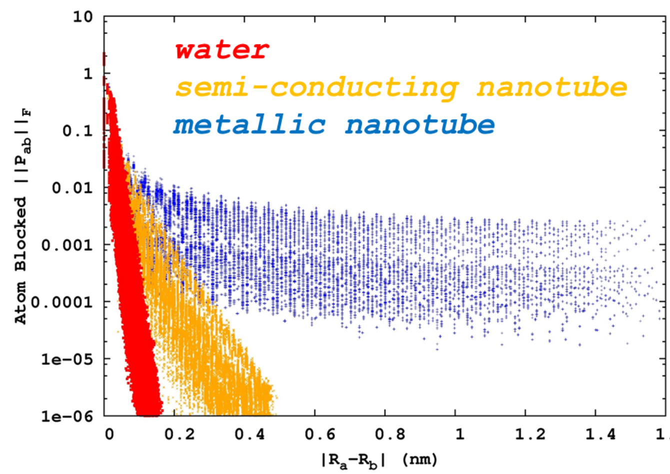


FIG. 1: Examples from electronic structure of decay for the spectral projector (gap shifted sign function) with respect to local (atomic) support. Shown is decay for systems with correlations that are short (insulating water), medium (semi-conducting 4,3 nanotube), and long (metalic 3,3 nanotube) ranged, from exponential (insulating) to algebraic (metallic).

Incompleteness -¿ sparse approximations dense problems, uses conventional sparse infrastructure, second order errors in matrix multiplication. Often adhoc.

$\tilde{\boldsymbol{a}} \cdot \tilde{\boldsymbol{b}} = \boldsymbol{a} \cdot \boldsymbol{b} + \delta\boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \delta\boldsymbol{b} + \delta\boldsymbol{a} \cdot \delta\boldsymbol{b}$

The variations do not express in the overal context of the product. Because the error in the incomplete case is additive, the

For example, $\boldsymbol{a}$ may be small, but $\delta\boldsymbol{a} \cdot \boldsymbol{b}$ large, leading extra work. Also, once a truncation error is commited, it is encounted in all subsiquent steps; it becomes difficult to impossilbe to manage error flows of differing magnitude in complex maps.

For extended quasi-degenerate correlations, these matrix functions may encounter ill-conditioning, and associated slow rates of decay. For extremely slow decay, maybe even oscillatory, low order algebraic decay, methods that compresion.... For fast decay,

correlation and the support

*Electronic address: `matt.challacombe@freeon.org`; URL: `http://www.freeon.org`

†Electronic address: `nicolasbock@freeon.org`; URL: `http://www.freeon.org`

Also, matrices with decay arise from the application of . Generally, ill-conditining is associated with slower decay,

Decay principles, often very sparse but very ill-conditioned problems.

A. Retaining the Eigenspace

Gradients lack convergence properties Iteration without orig drives away from basis NS has both. Difference between scalar iteration, Higham page 92.

B. Approximate Algebra as N-Body Problem

SpAMM is the recursive Cauchy-Schwarz occlusion product $\otimes_\tau$ on matrix quadtrees

$$\boldsymbol{a}^i = \begin{bmatrix} \boldsymbol{a}_{00}^{i+1} & \boldsymbol{a}_{01}^{i+1} \\ \boldsymbol{a}_{10}^{i+1} & \boldsymbol{a}_{11}^{i+1} \end{bmatrix} \tag{2}$$

$$\boldsymbol{a}^i \otimes_\tau \boldsymbol{b}^i = \begin{cases} \emptyset \quad \texttt{if} \quad \|\boldsymbol{a}^{\mathtt{i}}\|\|\boldsymbol{b}^{\mathtt{i}}\| < \tau \\ \boldsymbol{a}^i \cdot \boldsymbol{b}^i \quad \texttt{if}(\mathtt{i} = \mathtt{leaf}) \\ \begin{bmatrix} \boldsymbol{a}_{00}^{i+1} \otimes_\tau \boldsymbol{b}_{00}^{i+1} + \boldsymbol{a}_{01}^{i+1} \otimes_\tau \boldsymbol{b}_{10}^{i+1} \,, & \boldsymbol{a}_{00}^{i+1} \otimes_\tau \boldsymbol{b}_{01}^{i+1} + \boldsymbol{a}_{01}^{i+1} \otimes_\tau \boldsymbol{b}_{11}^{i+1} \\ \boldsymbol{a}_{00}^{i+1} \otimes_\tau \boldsymbol{b}_{01}^{i+1} + \boldsymbol{a}_{01}^{i+1} \otimes_\tau \boldsymbol{b}_{11}^{i+1} \,, & \boldsymbol{a}_{00}^{i+1} \otimes_\tau \boldsymbol{b}_{01}^{i+1} + \boldsymbol{a}_{01}^{i+1} \otimes_\tau \boldsymbol{b}_{11}^{i+1} \end{bmatrix} \quad \texttt{else} \end{cases} \tag{3}$$

database orientation, Cauchy sch Approximate Algebra, SpAMM Cauchy Schwarz occlusion, n-body approach to numerical linear algebra, first order errors in matrix multiplication. Based on Cauchy Schwarz inequality.

$$\boldsymbol{a} \otimes_\tau \boldsymbol{b} = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{\Delta}_\tau^{a \cdot b} \tag{4}$$

where $\boldsymbol{\Delta}_\tau^{a \cdot b}$ is a deterministic (assymetric) first order variation cooresponding to the branch pattern set by Cauchy-Schwarz occlusion, with length $\|\boldsymbol{\Delta}_\tau^{a \cdot b}\| \leq \tau \|\boldsymbol{a}\|\|\boldsymbol{b}\|$. The opperator $\otimes_\tau$ leads to a non-associative algebra with Lie bracket

$$[\boldsymbol{a}, \boldsymbol{b}]_\tau = \boldsymbol{a} \otimes_\tau \boldsymbol{b} - \boldsymbol{b} \otimes_\tau \boldsymbol{a} = [\boldsymbol{a}, \boldsymbol{b}] + \boldsymbol{\Delta}_\tau^{a \cdot b} - \boldsymbol{\Delta}_\tau^{b \cdot a}. \tag{5}$$

determined by the occlusion field. Our challenge is to master the error flows of these occlusion fields under iteration, for ill-conditioned problems and with permisive values of τ.

II. NEWTON SHULZ ITERATION

A. Idempotence

B. The Scaled Map

C. Alternative Formulations

dual, stabilized and naive

III. OCCLUSION FLOWS

$\delta \boldsymbol{x}_k$ and $\delta \boldsymbol{z}_k$ arrize from itteration with $\otimes_\tau$, and are deterministic flows away from the manifold of $\boldsymbol{s}$ determined by sensitivity of the NS iteration to these numerical insults.

$$\delta \boldsymbol{x}_k^{\mathrm{naiv}} = \delta \widetilde{\boldsymbol{z}}_k \cdot \boldsymbol{s} \cdot \widetilde{\boldsymbol{z}}_k + \widetilde{\boldsymbol{z}}_k \cdot \boldsymbol{s} \cdot \delta \widetilde{\boldsymbol{z}}_k \tag{6}$$

$$\delta \boldsymbol{x}_k^{\mathrm{dual}} = \delta \widetilde{\boldsymbol{y}}_k \cdot \widehat{\boldsymbol{z}}_k + \widetilde{\boldsymbol{y}}_k \cdot \delta \widetilde{\boldsymbol{z}}_k \tag{7}$$

$$\begin{aligned} \widetilde{\boldsymbol{x}}_k &= f\left[\widetilde{\boldsymbol{z}}_{k-1}, \widetilde{\boldsymbol{x}}_{k-1}\right] \\ &= \mathtt{m}\left[\widetilde{\boldsymbol{x}}_{\mathtt{k-1}}\right] \cdot \widetilde{\boldsymbol{z}}_{\mathtt{k-1}}^{\dagger} \cdot \boldsymbol{s} \cdot \widetilde{\boldsymbol{z}}_{\mathtt{k-1}} \cdot \mathtt{m}\left[\widetilde{\boldsymbol{x}}_{\mathtt{k-1}}\right] \end{aligned} \tag{8}$$

$$\delta \boldsymbol{x}_k = f_{\delta \boldsymbol{z}_{k-1}} \|\delta \boldsymbol{z}_{k-1}\| + f_{\delta \boldsymbol{x}_{k-1}} \|\delta \boldsymbol{x}_{k-1}\| + \mathcal{O}\left(\tau^2\right) \tag{9}$$

generalized Gateaux differential

$$\begin{aligned} f_{\delta \boldsymbol{z}_{k-1}} &= \lim_{\tau \to 0} \frac{f[\boldsymbol{z}_{k-1} + \tau \delta \widehat{\boldsymbol{z}}_{k-1}, \widetilde{\boldsymbol{x}}_{k-1}] - f[\, \boldsymbol{z}_{k-1}, \widetilde{\boldsymbol{x}}_{k-1}]}{\tau} \\ &= L_{\widetilde{\boldsymbol{x}}_k}\left(\widetilde{\boldsymbol{z}}_k, \delta \widehat{\boldsymbol{z}}_{k-1}\right) \end{aligned} \tag{10}$$

$$\begin{aligned} f_{\delta \boldsymbol{x}_{k-1}} &= \lim_{\tau \to 0} \frac{f[\widetilde{\boldsymbol{z}}_{k-1}, \boldsymbol{x}_{k-1} + \tau \delta \widehat{\boldsymbol{x}}_{k-1}] - f[\widetilde{\boldsymbol{z}}_{k-1}, \boldsymbol{x}_{k-1}]}{\tau} \\ &= L_{\widetilde{\boldsymbol{x}}_k}\left(\widetilde{\boldsymbol{z}}_k, \delta \widehat{\boldsymbol{x}}_{k-1}\right) \end{aligned} \tag{11}$$

$$L_{\widetilde{\boldsymbol{x}}_k}\left(\widetilde{\boldsymbol{z}}_k, \delta\widehat{\boldsymbol{x}}_{k-1}\right) = \delta\widehat{\boldsymbol{x}}^{\dagger}_{k-1}\cdot \mathtt{m}'\left[\boldsymbol{x}_{\mathtt{k-1}}\right]\cdot\left\{\widetilde{\boldsymbol{z}}^{\dagger}_{\mathtt{k-1}}\cdot \boldsymbol{s}\cdot\widetilde{\boldsymbol{z}}_{\mathtt{k}}\right\} + \left\{\widetilde{\boldsymbol{z}}^{\dagger}_{k}\cdot \boldsymbol{s}\cdot\widetilde{\boldsymbol{z}}_{k-1}\right\}\cdot \mathtt{m}'\left[\boldsymbol{x}_{\mathtt{k-1}}\right]\cdot\delta\widehat{\boldsymbol{x}}_{\mathtt{k-1}} \tag{12}$$

$$L_{\widetilde{\boldsymbol{x}}_k}\left(\widetilde{\boldsymbol{z}}_k, \delta\widehat{\boldsymbol{z}}_{k-1}\right) = \left\{\mathtt{m}\left[\boldsymbol{x}_{\mathtt{k-1}}\right]\cdot\delta\widehat{\boldsymbol{z}}^{\dagger}_{\mathtt{k-1}}\cdot\boldsymbol{s}\right\}\cdot\widetilde{\boldsymbol{z}}_{\mathtt{k}} + \widetilde{\boldsymbol{z}}^{\dagger}_{k}\cdot\left\{\boldsymbol{s}\cdot\delta\widehat{\boldsymbol{z}}_{k-1}\cdot\mathtt{m}\left[\boldsymbol{x}_{\mathtt{k-1}}\right]\right\} \tag{13}$$

$$\left\{\widetilde{\boldsymbol{z}}^{\dagger}_{k}\cdot\boldsymbol{s}\cdot\widetilde{\boldsymbol{z}}_{k-1}\right\}\rightarrow\boldsymbol{p}_{+}\left[\boldsymbol{s}\right] \tag{14}$$

$$\left\{\boldsymbol{s}\cdot\delta\widehat{\boldsymbol{z}}_{k-1}\cdot\mathtt{m}\left[\boldsymbol{x}_{\mathtt{k-1}}\right]\right\}\rightarrow\boldsymbol{n}\left[\boldsymbol{s}\right] \tag{15}$$

IV. BASIS SET ILL-CONDITIONING IN ELECTRONIC STRUCTURE

A. 3,3 carbon nanotube with diffuse *sp*-function

double exponential (Fig.)

B. Water with triple zeta and double polarization

Here's looking at you Jurg...

V. IMPLEMENTATION

A. Methods

FP, F08, OpenMP 4.0

B. A Modified NS Map

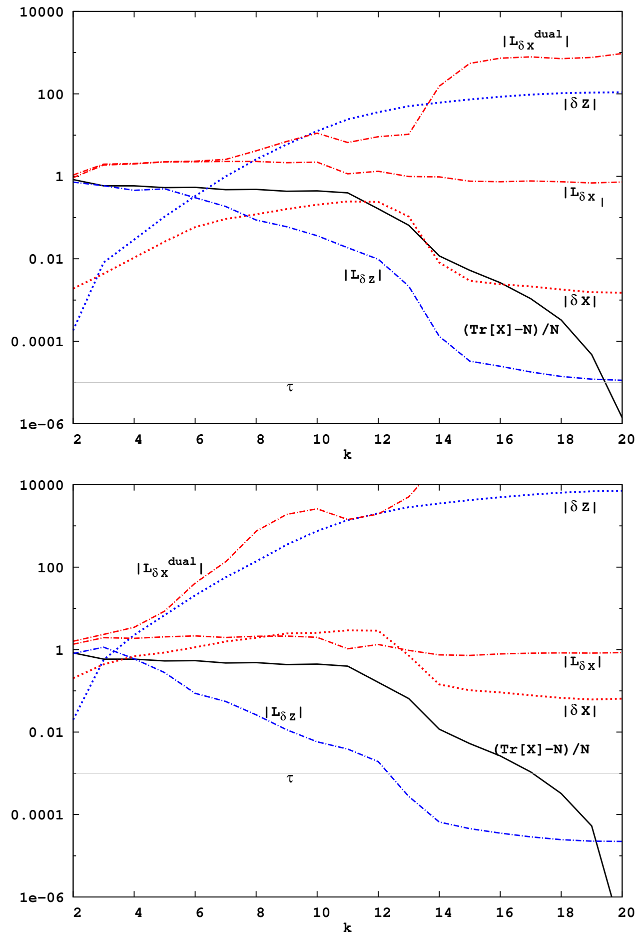
C. $\delta\boldsymbol{x}_k$ and $\delta\boldsymbol{x}_k$ channels

tau= Figure showing channels etc.

D. Convergence

Map switching and etc based on TrX

FIG. 2: equation...



VI. EXPERIMENTS

A. Occlusion Flows

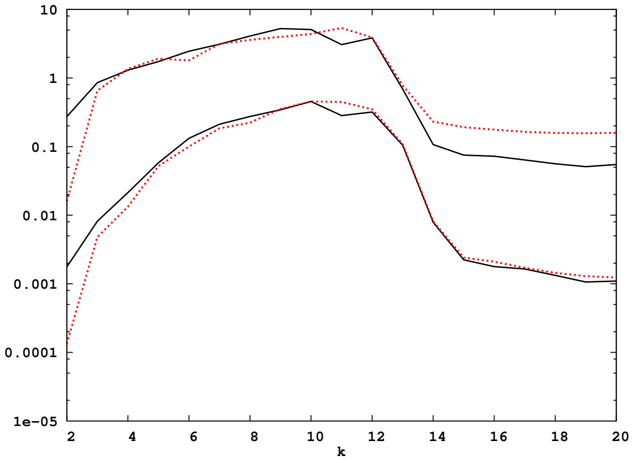
B. Comments

$$\delta\boldsymbol{z}_{k-1}\approx\Delta_{\tau}^{\widetilde{\boldsymbol{z}}_{k-2}\cdot\mathtt{m}\left[\widetilde{\boldsymbol{x}}_{\mathtt{k-2}}\right]} + \boldsymbol{z}_{k-2}\cdot\mathtt{m}'\left[\widetilde{\boldsymbol{x}}_{\mathtt{k-2}}\right]\cdot\delta\boldsymbol{x}_{\mathtt{k-2}} + \delta\boldsymbol{z}_{k-2}\cdot\mathtt{m}\left[\widetilde{\boldsymbol{x}}_{\mathtt{k-2}}\right] \tag{16}$$

$$\|\delta\boldsymbol{z}_{k-1}\|\lesssim\|\boldsymbol{z}_{k-2}\|\left(\,\tau\,\|\mathtt{m}\left[\widetilde{\boldsymbol{x}}_{\mathtt{k-2}}\right]\| + \|\delta\boldsymbol{x}_{k-2}\|\|\mathtt{m}'\left[\widetilde{\boldsymbol{x}}_{\mathtt{k-2}}\right]\|\,\right) \tag{17}$$

$$\|\boldsymbol{z}_k\|\rightarrow\sqrt{\kappa\left(\boldsymbol{s}\right)} \tag{18}$$

FIG. 3: equation...



C. Scaling

D. Comments

Pictures of the spamm structure

VII. CONCLUSION