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 areas of application, finite correlations lead to matrices with decay properties. By decay, we mean an approximate (perhaps bounded []) inverse relationship between matrix elements and an associated distance; this may be a simple inverse exponential relationship between elements and the Cartesian distance between support functions, or it may involve a generalized distance, *e.g.* a statistical measure between strings. In electronic structure, correlations manifest in decay properties of the gap shifted matrix sign function, as projector of the effective Hamiltonian (Fig. I). More broadly, matrix decay properties may coorespond to statistical matrices [1–5], including learned correlations in a generalized, non-orthogonal metric []. More broadly still, problems with local, non-orothogonal support are often solved with congruential transformations of the matrix inverse square root [6, 7] or a related factorization [5]; these transformations correlate local support with a representation independent form, *eg.* of the eigenproblem. Interestingly, the matrix sign function and the matrix inverse square root function are related by Higham's identity:

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

A complete overivew of matrix function theory and computation is given in Higham's enjoyable reference [8].

A well conditioned matrix $\boldsymbol{s}$ may often 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. Supporting this approximation are usefull bounds to matrix function elements [**? ?**]. The criteria τ might be a droptolerence, $\epsilon^{\boldsymbol{s}}_\tau = \{-s_{ij} * \hat{\boldsymbol{e}}_i \,|\, |s_{ij}| < \tau\}$, 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 general matrix-matrix multiply (`SpGEMM`) [9–12], 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. Comprehensive surveys of these methods in the numerical linear algebra are given by Benzi [13, 14], and by Bowler [15] and Benzi [16] for electronic structure.

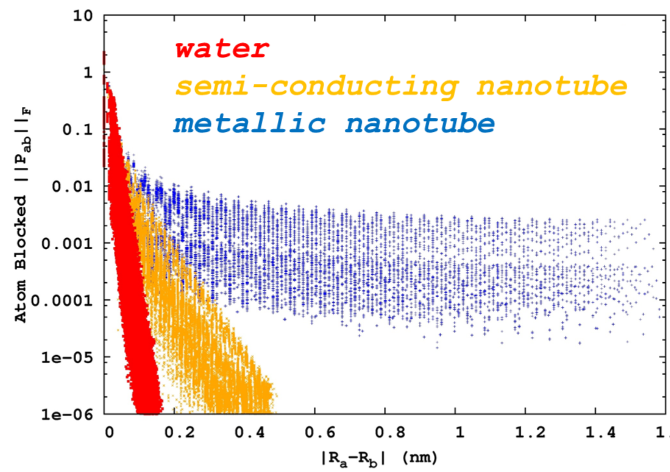


FIG. 1: Examples from electronic structure of decay for the spectral projector (gap shifted sign function) with respect to local (atomic) support. Shown is decay for systems with correlations that are short (insulating water), medium (semi-conducting 4,3 nanotube), and long (metalic 3,3 nanotube) ranged,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

Because the truncated multiplication is controled only by absolute, addititve errors in the product,

$$\overline{\boldsymbol{a}\cdot\boldsymbol{b}} \;=\; \boldsymbol{a}\cdot\boldsymbol{b} \;+\; \boldsymbol{\epsilon}^{\boldsymbol{a}}_\tau\cdot\boldsymbol{b} \;+\; \boldsymbol{a}\cdot\boldsymbol{\epsilon}^{\boldsymbol{b}}_\tau \;+\; \mathcal{O}(\tau^2) \tag{2}$$

achieving sparse, stable and rapidly convergent iteration for ill-conditioned problems can be challenging []. In cases of extreme degeneracy, hierarchical semi-seperable (reduced rank) algorithms can offer effective complexity reduction []. However, many pratical cases are somewhere in-between sparse and meaningfully degenerate regimes; effectively dense but without an exploitable reduction in rank. This is the case in electronic structure for strong but non-metalic correlation, *e.g.* towards the Mott transition [], and also in the case of local atomic support towards completeness [**? ? ?**].

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

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

In this contribution, we consider an N-body approach to the approximation of matrix functions with decay, based on the quadtree data structure [? ?]

$$\boldsymbol{a}^i = \begin{bmatrix} \boldsymbol{a}_{00}^{i+1} & \boldsymbol{a}_{01}^{i+1} \\ \boldsymbol{a}_{10}^{i+1} & \boldsymbol{a}_{11}^{i+1} \end{bmatrix} , \tag{3}$$

and orderings that are locality preserving []. Orderings that preserve data locality are well developed in the database theory [], providing fast spatial and metric querries. Locality enabled, fast data access is central to the N-Body approximation [], and an important problem for enterprise [] and runtime systems [], with memory hierarchies becoming increasingly assynchronous and decentralized [?]. For matrices with decay, orderings that preserve locality lead to block-by-magnitude matrix structures with well segregated neighborhoods, inhabited by matrix elements of like size, and efficiently resolved by the quadtree data structure [].

With block-by-magnitude ordering of matrices $\boldsymbol{a}$ and $\boldsymbol{b}$, the Sparse Approximate Matrix Multiplication (`SpAMM`) kernel, $\otimes_\tau$, carries out fast occlusion culling of insignifcant volumes in the product octree:

$$\boldsymbol{a}^i \otimes_\tau \boldsymbol{b}^i = \begin{cases} \emptyset \quad \texttt{if} \quad \|\boldsymbol{a}^{\mathtt{i}}\| \|\boldsymbol{b}^{\mathtt{i}}\| < \tau \\ \boldsymbol{a}^i \cdot \boldsymbol{b}^i \quad \texttt{if}(\mathtt{i} = \mathtt{leaf}) \\ \begin{bmatrix} \boldsymbol{a}_{00}^{i+1} \otimes_\tau \boldsymbol{b}_{00}^{i+1} + \boldsymbol{a}_{01}^{i+1} \otimes_\tau \boldsymbol{b}_{10}^{i+1} , & \boldsymbol{a}_{00}^{i+1} \otimes_\tau \boldsymbol{b}_{01}^{i+1} + \boldsymbol{a}_{01}^{i+1} \otimes_\tau \boldsymbol{b}_{11}^{i+1} \\ \boldsymbol{a}_{00}^{i+1} \otimes_\tau \boldsymbol{b}_{01}^{i+1} + \boldsymbol{a}_{01}^{i+1} \otimes_\tau \boldsymbol{b}_{11}^{i+1} , & \boldsymbol{a}_{00}^{i+1} \otimes_\tau \boldsymbol{b}_{01}^{i+1} + \boldsymbol{a}_{01}^{i+1} \otimes_\tau \boldsymbol{b}_{11}^{i+1} \end{bmatrix} \quad \texttt{else} \end{cases} , \tag{4}$$

with errors linear in τ bounded by the sub-multiplicative norms $\|\cdot\| \equiv \|\cdot\|_F$ and the Cauchy-Schwarz inequality [? ?]. In Ref.[?], we generalize this recursive task occlusion to the problem of Fock exchange.

The approximate `SpAMM` product is

$$\widetilde{\boldsymbol{a} \cdot \boldsymbol{b}} \equiv \boldsymbol{a} \otimes_\tau \boldsymbol{b} = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{\Delta}_\tau^{a \cdot b} + \mathcal{O}\left(\tau^2\right) , \tag{5}$$

with the culled contractions $\boldsymbol{\Delta}_\tau^{a \cdot b}$ obeying the `SpAMM` bound

$$\|\boldsymbol{\Delta}_\tau^{a \cdot b}\| \leq \tau \, \|\boldsymbol{a}\| \, \|\boldsymbol{b}\| , \tag{6}$$

at each level of recursion. This makes $\otimes_\tau$ *stable*, as defined by Demmel, Dumitriu and Holz (see Eq.(1), Ref. [?]). However, instead of the roundoff error, we are concerned with the deterministic `SpAMM` error, which leads to a non-associative algebra and error flows with properties of the Lie bracket

$$\widetilde{[\boldsymbol{a}, \boldsymbol{b}]} \equiv \boldsymbol{a} \otimes_\tau \boldsymbol{b} - \boldsymbol{b} \otimes_\tau \boldsymbol{a} = [\boldsymbol{a}, \boldsymbol{b}] + \boldsymbol{\Delta}_\tau^{a \cdot b} - \boldsymbol{\Delta}_\tau^{b \cdot a} . \tag{7}$$

The interesting group theory associated with the construction of matrix functions is developed by Higham, Mackey, Mackey and T in Ref.[].

`SpAMM` is similar to compressed kernels for sketching the matrix product [17, 18]. spatial join! Paugh. MAD. etc Instead of the FFT however, compression is achieved through a recursive two-sided metric querry on the `SpAMM` bound, Eq. (6), which may exprerience acceleration through localiztion effects (lensing) in the ijk (octree) space, as demonstrated shortly. These localization techniques are, in most cases, complimentary with other compresive technologies, including methods based on hierarchical semi-seperability and also methods for fast kernel summation []. In addition to compression by occlusion, octree locality is important for the communication optimality of n-body methods [? ?] that may be achieved by minimal locally essential trees []. Finally, its interesting to note that this data-centric, *locality as compression* approach is incompatible with randomization methods that employ homogenization to well condition matrices [? ?], and also to achieve domain decomposition of the `SpGEMM`, *e.g.* using the conventional SUMMA [].

II. FIRST ORDER NEWTON-SHULZ ITERATION

There are two common, first order NS iterations; the sign iteration and the square root iteration, related by the square, $\boldsymbol{I}(\cdot) = \text{sign}^2(\cdot)$. These equivalent iterations converge linearly at first, then enters a basin of stability marked by super-linear convergence. Our interest is to access this basin with the most permisive τ possible, building a foundation for future refinement at a reduced cost and with a higher precision ($\tau \to 0$) [?].

A. Sign iteration

For the NS sign iteration, this basin is marked by a behavioral change in the difference $\delta \boldsymbol{X}_k = \widetilde{\boldsymbol{X}}_k - \boldsymbol{X}_k = \text{sign}\left(\boldsymbol{X}_{k-1} + \delta \boldsymbol{X}_{k-1}\right) - \text{sign}\left(\boldsymbol{X}_{k-1}\right)$, where $\delta \boldsymbol{X}_{k-1}$ is some previous error. The change in behavior is associated with the onset of idempotence and the bounded eigenvalues of $\mathrm{sign}'(\cdot)$, leading to stable iteration when $\mathrm{sign}'(\boldsymbol{X}_{k-1})\,\delta\boldsymbol{X}_{k-1} < 1$. Global perturbative bounds on this iteration have been derived by Bai and Demmel [19], while Byers, He and Mehrmann [] developed assymptotic bounds. The automatic stability of sign iteration is a well developed theme in Ref.[8].

B. Square root iteration

In this work, we are concerned with resolution of the identity []

$$\boldsymbol{I}(\boldsymbol{s}) = \boldsymbol{s}^{1/2} \cdot \boldsymbol{s}^{-1/2}\,, \tag{8}$$

and the cooresponding cannonical (dual) square root iteration [];

$$\begin{aligned} \boldsymbol{y}_k &\leftarrow h_\alpha\left[\boldsymbol{y}_{k-1} \cdot \boldsymbol{z}_{k-1}\right] \cdot \boldsymbol{y}_{k-1} \\ \boldsymbol{z}_k &\leftarrow \boldsymbol{z}_{k-1} \cdot h_\alpha\left[\boldsymbol{y}_{k-1} \cdot \boldsymbol{z}_{k-1}\right]\ , \end{aligned} \tag{9}$$

with eigenvalues in the proper domain aggregated towards 0 or 1 by the NS map $h_\alpha[\boldsymbol{x}] = \frac{\sqrt{\alpha}}{2}(3 - \alpha\boldsymbol{x})$ []. Then, starting with $\boldsymbol{z}_0 = \boldsymbol{I}$ and $\boldsymbol{x}_0 = \boldsymbol{y}_0 = \boldsymbol{s}$, $\boldsymbol{y}_k \to \boldsymbol{s}^{1/2}$, $\boldsymbol{z}_k \to \boldsymbol{s}^{-1/2}$ and $\boldsymbol{x}_k \to \boldsymbol{I}$. As in the case of sign iteration, this dual iteration was shown by Higham, Mackey, Mackey and Tisseur [20] to remain bounded in the superlinear regime, by idempotent Frechet derivatives about the fixed point $\left(\boldsymbol{s}^{1/2}, \boldsymbol{s}^{-1/2}\right)$, in the direction $\left(\delta\boldsymbol{y}_{k-1}, \delta\boldsymbol{z}_{k-1}\right)$:

$$\delta\boldsymbol{y}_k = \frac{1}{2}\delta\boldsymbol{y}_{k-1} - \frac{1}{2}\boldsymbol{s}^{1/2} \cdot \delta\boldsymbol{z}_{k-1} \cdot \boldsymbol{s}^{1/2} \tag{10}$$

$$\delta\boldsymbol{z}_k = \frac{1}{2}\delta\boldsymbol{z}_{k-1} - \frac{1}{2}\boldsymbol{s}^{-1/2} \cdot \delta\boldsymbol{y}_{k-1} \cdot \boldsymbol{s}^{-1/2}\,. \tag{11}$$

In this contribution, we consider another aspect of convergence, namely the (hopefully) linear approach towards stability of the iteration

$$\widetilde{\boldsymbol{x}}_k \leftarrow \widetilde{\boldsymbol{y}}_k(\widetilde{\boldsymbol{x}}_{k-1}) \otimes_\tau \widetilde{\boldsymbol{z}}_k(\widetilde{\boldsymbol{x}}_{k-1})\ , \tag{12}$$

made difficult by ill-conditioning and a sketchy $\otimes_\tau$.

1. the NS map

Initially, h'_α at the smallest eigenvalue x_0 controls the rate of progress towards idempotence. As recently shown by Jie and Chen [21], for very ill-conditioned problems, a factor of two reduction in the number of NS steps can be achieved by chosing $\alpha \sim 2.85$, which is at the edge of stability. As argued by Pan and Schreiber [22], Jie and Chen [21], switching or damping the scaling factor towards $\alpha = 1$ at convergence is important, shifting emphasis away from the behavior of x_0 towards *e.g.* $x_i \in [0.01, 1]$, emphasizing overall convergence of the broad distribution [**?**]. In an approximate algebra like `SpAMM`, the potential for eigenvalues to fluctuate out of the domain of convergence must be considered. This is addressed in Section. **??**.

2. stability and ill-conditioning

Agressive scaling can reduce stability of the iteration due to the larger derivative h'_α. Also, the previous error, $\delta\boldsymbol{x}_{k-1}$, may be to large, *e.g.* due to a too large value of τ, leading to the unbounded (exponential) accumulation of errors $\delta\boldsymbol{x}_k > 1$. To first order, stability of the iteration is controled by the Frechet derivatives contributing to $\delta\boldsymbol{x}_k$. For ill-conditioned problems and the $\otimes_\tau$ kernel, these derivatives can behave differently amongst nominally equivalent implementations, for example formulations based on the assumption of a commuting algebra. In this contribution, we are interested in the cannonical (dual) square root iteration, Eq. 9, as well as the "stabilized" version with $\boldsymbol{y}_k^{\mathrm{stab}} = \boldsymbol{z}_k^\dagger \cdot \boldsymbol{s}$.

3. lensing

A feature of square root iteration with the $\otimes_\tau$ kernel is localization of the culled octree with convergence towards identity, $\widetilde{\boldsymbol{x}}_k \leftarrow \boldsymbol{I}(\widetilde{\boldsymbol{x}}_{k-1})$. In the product $\widetilde{\boldsymbol{y}}_k \otimes_\tau \widetilde{\boldsymbol{z}}_k$, large and small eigenvalues, as well as norms, are recursively brought to unity along the diagonal $i = k$, and within the cube ijk. Because the `SpAMM` error obeys the multiplicative Cauchy-Schwarz bound, Eq. (), the cooresponding culled-octree follows the $i = k$ plane[1], resolving the *relative* error in identity to within τ. This effect is shown in Figure **??**. We call this identity related, plane-wise concentration of the culled octree *lensing*. Ideally, lensing cooresponds to the formation of two-dimensional surfaces, crossing along the cube diagonal, with attenuation off the diagonal in the case of decay. This is a complexity reduction relative to the naive (full) volume of the cube, and also relative to sparsification strategies that preserve only absolute errors, as in Eq. **??**.

III. MODELING THE `SpAMM` NS ITERATION

There are multiple effects we wish to model in the $\otimes_\tau$-approximate square root iteration. Primarily, classical first order errors in $\delta\boldsymbol{x}_k$ are controled by the Fréchet derivatives [] $\boldsymbol{x}_{\delta\boldsymbol{y}_{k-1}}$ and $\boldsymbol{x}_{\delta\boldsymbol{z}_{k-1}}$ at $\boldsymbol{x}_{k-1}$ and along the unit directions $\delta\widehat{\boldsymbol{y}}_{k-1}$ and $\delta\widehat{\boldsymbol{z}}_{k-1}$;

$$\delta\boldsymbol{x}_k = \boldsymbol{x}_{\delta\widehat{\boldsymbol{y}}_{k-1}} \times \delta y_{k-1} + \boldsymbol{x}_{\delta\widehat{\boldsymbol{z}}_{k-1}} \times \delta z_{k-1}\,, \tag{13}$$

with previous displacements $\delta y_{k-1} = \|\delta\boldsymbol{y}_{k-1}\|$ and $\delta z_{k-1} = \|\delta\boldsymbol{z}_{k-1}\|$. However, these displacements are augmented by the `SpAMM` error at each step. Also, second order effects can develop (due to ill-conditioning), which invalidate $\boldsymbol{x}_{k-1}$ as reference for first order expansion.

[1] as well as the $i = j$ plane.

A. Fréchet Analyses

Our model is based on classical Fréchet analyses [] of the matrix function

$$\boldsymbol{x}\left(\boldsymbol{y}_{k-1},\boldsymbol{z}_{k-1}\right)=h_\alpha\left[\boldsymbol{y}_{k-1}\cdot\boldsymbol{z}_{k-1}\right]\cdot\boldsymbol{y}_{k-1}\cdot\boldsymbol{z}_{k-1}\cdot h_\alpha\left[\boldsymbol{y}_{k-1}\cdot\boldsymbol{z}_{k-1}\right]\;, \tag{14}$$

with all previous errors mapped to $\delta\boldsymbol{y}_{k-1}$ and $\delta\boldsymbol{z}_{k-1}$ by the complicated logistics of h_α and $\otimes_\tau$. Later, we will consider the also the first `SpAMM` error at step k. At present however, we turn to the Fréchet derivatives []:

$$\begin{aligned}\boldsymbol{x}_{\delta\widehat{\boldsymbol{y}}_{k-1}} &= \lim_{\tau\to 0}\frac{\boldsymbol{x}\left(\boldsymbol{y}_{k-1}+\tau\delta\widehat{\boldsymbol{y}}_{k-1},\boldsymbol{z}_{k-1}\right)-\boldsymbol{x}\left(\boldsymbol{y}_{k-1},\boldsymbol{z}_{k-1}\right)}{\tau}\\ &= \boldsymbol{y}_{\delta\widehat{\boldsymbol{y}}_{k-1}}\cdot\boldsymbol{z}_k+\boldsymbol{y}_k\cdot\boldsymbol{z}_{\delta\widehat{\boldsymbol{y}}_{k-1}}\end{aligned} \tag{15}$$

and

$$\begin{aligned}\boldsymbol{x}_{\delta\widehat{\boldsymbol{z}}_{k-1}} &= \lim_{\tau\to 0}\frac{\boldsymbol{x}\left(\boldsymbol{y}_{k-1},\boldsymbol{z}_{k-1}+\tau\delta\widehat{\boldsymbol{z}}_{k-1}\right)-\boldsymbol{x}\left(\boldsymbol{y}_{k-1},\boldsymbol{z}_{k-1}\right)}{\tau}\\ &= \boldsymbol{y}_{\delta\widehat{\boldsymbol{z}}_{k-1}}\cdot\boldsymbol{z}_k+\boldsymbol{y}_k\cdot\boldsymbol{z}_{\delta\widehat{\boldsymbol{z}}_{k-1}}\,.\end{aligned} \tag{16}$$

1. $\boldsymbol{x}_{\delta\widehat{\boldsymbol{y}}_{k-1}}$

For $\boldsymbol{x}_{\delta\widehat{\boldsymbol{y}}_{k-1}}$ we have

$$\boldsymbol{y}_{\delta\widehat{\boldsymbol{y}}_{k-1}}=h_\alpha\left[\boldsymbol{x}_{k-1}\right]\cdot\delta\widehat{\boldsymbol{y}}_{k-1}+h'_\alpha\left[\boldsymbol{x}_{k-1}\right]\cdot\delta\widehat{\boldsymbol{y}}_{k-1}\cdot\boldsymbol{z}_{k-1}\cdot\boldsymbol{y}_{k-1} \tag{17}$$

and

$$\boldsymbol{z}_{\delta\widehat{\boldsymbol{y}}_{k-1}}=\boldsymbol{z}_{k-1}\cdot h'_\alpha\left[\boldsymbol{x}_{k-1}\right]\cdot\delta\widehat{\boldsymbol{y}}_{k-1}\cdot\boldsymbol{z}_{k-1} \tag{18}$$

yeilding

$$\begin{aligned}\boldsymbol{x}_{\delta\widehat{\boldsymbol{y}}_{k-1}}=h_\alpha\left[\boldsymbol{x}_{k-1}\right]\cdot\delta\widehat{\boldsymbol{y}}_{k-1}\cdot\boldsymbol{z}_k &+h'_\alpha\delta\widehat{\boldsymbol{y}}_{k-1}\cdot\boldsymbol{z}_{k-1}\cdot\boldsymbol{y}_{k-1}\cdot\boldsymbol{z}_k\\ &+\boldsymbol{y}_k\cdot\boldsymbol{z}_{k-1}\cdot h'_\alpha\delta\widehat{\boldsymbol{y}}_{k-1}\cdot\boldsymbol{z}_{k-1}\end{aligned} \tag{19}$$

Closer to convergence,

$$\boldsymbol{x}_{\delta\widehat{\boldsymbol{y}}_{k-1}}\to\delta\widehat{\boldsymbol{y}}_{k-1}\cdot\boldsymbol{z}_k+2\delta\widehat{\boldsymbol{y}}_{k-1}\cdot\boldsymbol{z}_{k-1}\,. \tag{20}$$

2. $\boldsymbol{x}_{\delta\widehat{\boldsymbol{z}}_{k-1}}$

For $\boldsymbol{x}_{\delta\widehat{\boldsymbol{z}}_{k-1}}$ we have

$$\boldsymbol{y}_{\delta\widehat{\boldsymbol{z}}_{k-1}}=\boldsymbol{y}_{k-1}\cdot h'_\alpha\delta\widehat{\boldsymbol{z}}_{k-1}\cdot\boldsymbol{y}_{k-1} \tag{21}$$

and

$$\boldsymbol{z}_{\delta\widehat{\boldsymbol{z}}_{k-1}}=\delta\widehat{\boldsymbol{z}}_{k-1}\cdot h_\alpha\left[\boldsymbol{x}_{k-1}\right]+\boldsymbol{z}_{k-1}\cdot\boldsymbol{y}_{k-1}\cdot h'_\alpha\delta\widehat{\boldsymbol{z}}_{k-1}\;, \tag{22}$$

yeilding

$$\begin{aligned}\boldsymbol{x}_{\delta\widehat{\boldsymbol{z}}_{k-1}}=\boldsymbol{y}_{k-1}\cdot h'_\alpha\delta\widehat{\boldsymbol{z}}_{k-1}\cdot\boldsymbol{y}_{k-1}\cdot\boldsymbol{z}_k &+\boldsymbol{y}_k\cdot\delta\widehat{\boldsymbol{z}}_{k-1}\cdot h_\alpha\left[\boldsymbol{x}_{k-1}\right]\\ &+\boldsymbol{y}_k\cdot\boldsymbol{z}_{k-1}\cdot\boldsymbol{y}_{k-1}\cdot h'_\alpha\delta\widehat{\boldsymbol{z}}_{k-1}\end{aligned} \tag{23}$$

Closer to convergence,

$$\boldsymbol{x}_{\delta\widehat{\boldsymbol{y}}_{k-1}}\to\delta\widehat{\boldsymbol{y}}_{k-1}\cdot\boldsymbol{z}_k+2\delta\widehat{\boldsymbol{y}}_{k-1}\cdot\boldsymbol{z}_{k-1}\,. \tag{24}$$

3. $\boldsymbol{x}_{\delta\widehat{\boldsymbol{x}}_{k-1}}$

$$\boldsymbol{x}_{\delta\widehat{\boldsymbol{x}}_{k-1}}=h'_\alpha\,\delta\widehat{\boldsymbol{x}}_{k-1}\cdot\boldsymbol{y}_{k-1}\cdot\widetilde{\boldsymbol{z}}_k+\boldsymbol{y}_k\cdot\widetilde{\boldsymbol{z}}_{k-1}\cdot\delta\widehat{\boldsymbol{x}}_{k-1}\,h'_\alpha\,. \tag{25}$$

$$\left\{\boldsymbol{s}\cdot\delta\widehat{\boldsymbol{z}}_{k-1}\cdot\mathtt{m}\left[\boldsymbol{x}_{\mathtt{k-1}}\right]\right\}\to\boldsymbol{n}\left[\boldsymbol{s}\right] \tag{26}$$

Ideally, a τ can be found yeilding fast computation with precision sufficient to gain the basin of stability. From the preconditioned state then, additional corrections can be made to the residual at little additional cost. [2]

IV. 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

V. ILL-CONDITIONED SUPPORT

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

double exponential (Fig.)

[2] The ability to correct back to the argumental basis is sometimes refered to as the "variational" property of early "spectral projection as optimization" techniques, with gradients retaining proximity to the basis [].

B. Water with triple zeta and double polarization

Here's looking at you Jurg...

VI. EXPERIMENTS

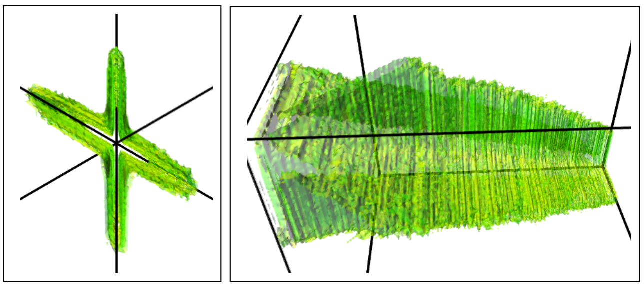


FIG. 2: Views of the $\tau = 0.03$ sign occlusion surface, for the 128x u.c. nanotube, at $\sim 14k \times 14k$ and $\kappa(\boldsymbol{s}) = 10^6$. This surface envelopes the ijk volume of the $\otimes_\tau$ kernel, cooresponding to the unscaled dual iteration step $\widetilde{\boldsymbol{x}}_{19} \leftarrow \widetilde{\boldsymbol{y}}_{19} \otimes_\tau \widetilde{\boldsymbol{z}}_{19}$ at $b = 64$, $\tau = 0.03$ and $\tau_y = 10^{-3}\,\tau$. The first pannel looks straight down the cube-diagonal $i = j = k$, from the upper bound towards (1,1,1). Remarkably, this surface forms an elongated $\times$, closely following intersection of the $i = j$ and $i = k$ planes along the cube-diagonal. The second pannel looks along the cube-diagonal, with the upper bound at upper left, and (1,1,1) at lower right.

In this section, we present numerical experiments that highlight the effects of ill-conditioning, dimensionality, and the stability of different first order NS approaches to iteration with `SpAMM`. We turn first to complexity reduction for $\otimes_\tau$ in the basin of stability, where we find a novel, compressive effect in the product octree. This effect is shown in Fig. VI, for unscaled, inverse square root duals iteration, Eqs. (**??**), on the 3,3 carbon nanotube metric at $\kappa = 10^6$.

In this example, the `SpAMM` octree culled from the ijk-cube is outlined by its occlusion surface, enclosing a volume that closely follows the $i = j$ and $i = k$ planes, forming an $\times$. The banded distribution of large norms along matrix diagonals leads to cube-diagonal dominance, with plane-following a consequence of moderate ill-conditioning, large norms along the plane-diagonals and their overlap in ijk via the multiplicative bound, Eq. (6). The tightness of this bound, and the compression gained relative to methods that control only the absolute error, *e.g.* as given by Eq. (2), will hopefully be quantified in future work.

FIG. 3: equation...

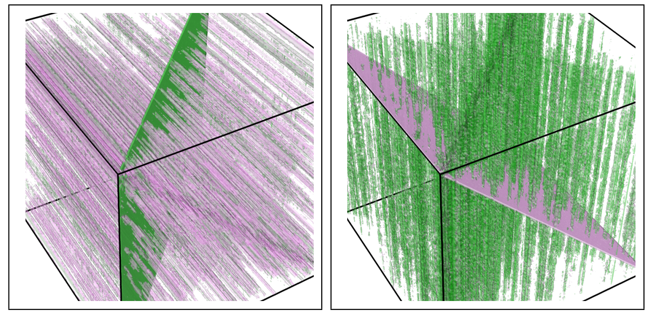


FIG. 4: equation...

`z_water_to_duals_scn1.png`

A. Occlusion Error Flows

B. Comments

$$\delta \boldsymbol{z}_{k-1} \approx \Delta_\tau^{\widetilde{\boldsymbol{z}}_{k-2}\cdot \mathtt{m}[\widetilde{\boldsymbol{x}}_{\mathtt{k}-2}]} + \boldsymbol{z}_{k-2}\cdot \mathtt{m}'\left[\widetilde{\boldsymbol{x}}_{\mathtt{k}-2}\right]\cdot \delta\boldsymbol{x}_{\mathtt{k}-2} + \delta\boldsymbol{z}_{k-2}\cdot \mathtt{m}\left[\widetilde{\boldsymbol{x}}_{\mathtt{k}-2}\right] \quad (27)$$

$$\|\delta \boldsymbol{z}_{k-1}\| \lesssim \|\boldsymbol{z}_{k-2}\|\left(\,\tau\,\|\mathtt{m}\left[\widetilde{\boldsymbol{x}}_{\mathtt{k}-2}\right]\| + \|\delta \boldsymbol{x}_{k-2}\|\|\mathtt{m}'\left[\widetilde{\boldsymbol{x}}_{\mathtt{k}-2}\right]\|\,\right) \quad (28)$$

$$\|\boldsymbol{z}_k\| \to \sqrt{\kappa\left(\boldsymbol{s}\right)} \quad (29)$$

C. Comments

D. Found Contraction

E. Comments

Pictures of the spamm structure

VII. CONCLUSION

[1] O. Penrose and J. Lebowitz, Commun. Math. Phys. **184** (1974).

[2] J. Voit, *The Statistical Mechanics of Financial Markets*, Theoretical and Mathematical Physics (Springer Berlin Heidelberg, 2006), ISBN 9783540262893, URL `https://books.google.com/books?id=6zUlh\_TkWSwC`.

[3] L. Anselin, Int. Reg. Sci. Rev. **26**, 153 (2003), ISSN 01600176, URL `http://irx.sagepub.com/cgi/doi/10.`

FIG. 5: equation...

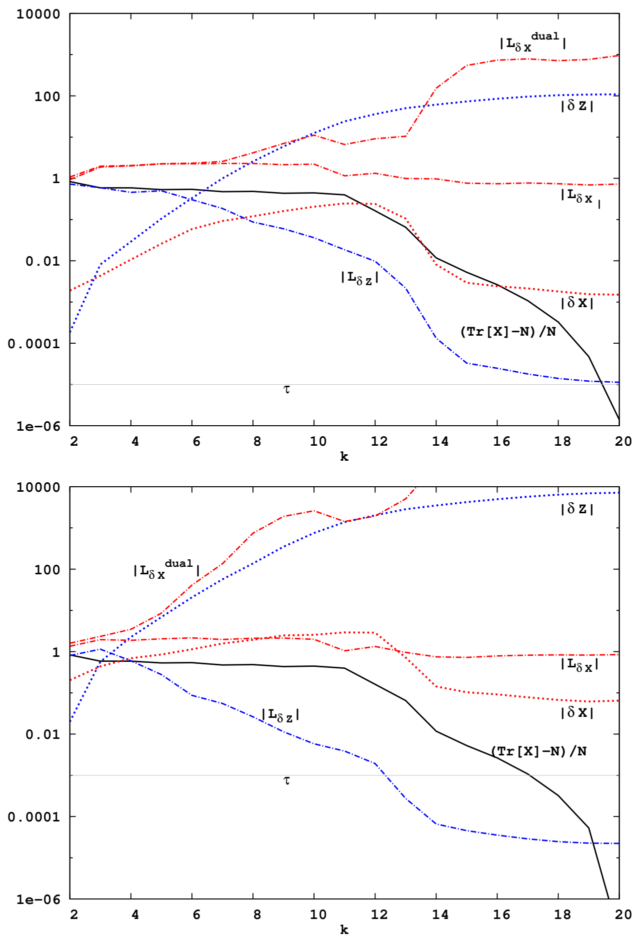
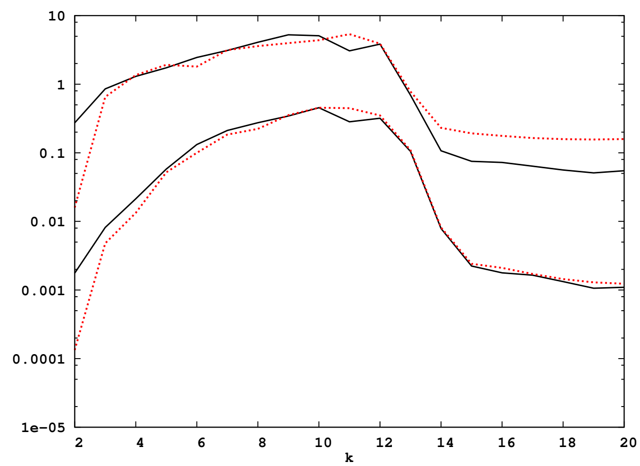


FIG. 6: equation...



1177/0160017602250972.

[4] J. Hardin, S. R. Garcia, and D. Golan, Ann. Appl. Stat. **7**, 1733 (2013), ISSN 1932-6157, arXiv:1106.5834v4.

[5] I. Krishtal, T. Strohmer, and T. Wertz, Found. Comput. . . . (2013), ISSN 1615-3375.

[6] P. O. Lowdin, Advances in Physics **5**, 1 (1956).

[7] A. R. Naidu, p. 1 (2011), 1105.3571.

[8] N. J. Higham, *Functions of Matrices* (Society for Industrial & Applied Mathematics,, 2008).

[9] F. G. Gustavson, ACM Transactions on Mathematical Software (TOMS) **4**, 250 (1978).

[10] S. Toledo, IBM J. Res. Dev. **41**, 711 (1997).

[11] M. Challacombe, Comput. Phys. Commun. **128**, 93 (2000).

[12] D. R. Bowler and T. M. andM. J. Gillan, Comp. Phys. Comm. **137**, 255 (2000).

[13] M. Benzi and M. Tuma, Appl. Numer. Math. **30**, 305 (1999).

[14] M. Benzi, J. Comput. Phys. **182**, 418 (2002).

[15] D. R. Bowler and T. Miyazaki, Reports Prog. Phys. **75**, 36503 (2012).

[16] M. Benzi, P. Boito, and N. Razouk, SIAM Rev. **55**, 3 (2013).

[17] K. Kutzkov, pp. 1–15 (2012), 1209.4508.

[18] R. Pagh, ACM Trans. Comput. Theory **5**, 1 (2013).

[19] Z. Bai and J. Demmel, SIAM J. Matrix Anal. Appl **19**, 205 (1998).

[20] N. J. . Higham, N. M. D . Steven Mackey, and F. . coise Tisseur, SIAM J. Matrix Anal. Appl. **26**, 849 (2005).

[21] J. Chen and E. Chow (preprint), URL `http://www.mcs.anl.gov/papers/P5059-0114.pdf`.

[22] V. Pan and R. Schreiber, SIAM J. Sci. Stat. Comput. (1991).