Publication III

**K. Baarman, T. Eirola, and V. Havu. Direct minimization of electronic struc-
ture calculations with Householder reflections.** *arXiv:1204.1204*, **16 pages,
April 2012.**

Direct minimization of electronic structure calculations with Householder reflections

K. Baarman,[*] T. Eirola[*], and V. Havu[†]

July 11, 2012

Abstract

We consider a minimization scheme based on the Householder transport operator for the Grassman manifold, where a point on the manifold is represented by an $m \times n$ matrix with orthonormal columns. In particular, we consider the case where $m \gg n$ and present a method with asymptotic complexity mn^2. To avoid explicit parametrization of the manifold we use Householder transforms to move on the manifold, and present a formulation for simultaneous Householder reflections for $\mathbf{S}$-orthonormal columns. We compare a quasi-Newton and nonlinear conjugate gradient implementation adapted to the manifold with a projected nonlinear conjugate gradient method, and demonstrate that the convergence rate is significantly improved if the manifold is taken into account when designing the optimization procedure.

1 Introduction

We consider the optimization problem

$$\min_{\mathbf{X}^T \mathbf{X} = \mathbf{I}} f(\mathbf{X}), \tag{1}$$

that is, we attempt to minimize the real valued function f of $\mathbf{X} \in \mathbb{R}^{m \times n}$ where $m \gg n$, subject to the constraint $\mathbf{X}^T \mathbf{X} = \mathbf{I}$, and with the computable derivative $df(\mathbf{X})$. The constraint on $\mathbf{X}$ ensures that f has a minimum, but this minimum is not necessarily unique. The method we present requires that $m \geq 2n$. Extending the method to cover cases where just $m > n$ is, however, possible.

A special property we assume of f is the homogeneity condition: $f(\mathbf{X}) = f(\mathbf{XQ})$, where $\mathbf{Q}$ is any $n \times n$ orthonormal matrix. This property means that f only depends on the span of the columns of $\mathbf{X}$. The set of subspaces that are spanned by the columns of orthonormal $m \times n$ matrices is called the Grassman

---
[*]Department of Mathematics and Systems Analysis, Aalto University School of Science, Espoo, Finland, e-mail: `kurt.baarman@aalto.fi`

[†]Department of Applied Physics, Aalto University School of Science, Espoo, Finland

manifold, $\mathcal{M}$. While the solution to (1) is an equivalence class, we choose an arbitrary representative of the class since we are interested in the value of f. The closely related Stiefel manifold consists of the same problem without the homogeneity condition. While the Householder transformation is suitable for both the Grassman and Stiefel manifolds, the optimization method presented does not optimize with respect to the basis and is therefore suitable only for the Grassman manifold.

In the applications we have in mind the evaluation of f and df is expensive. For this reason we cannot employ a high quality line search to decide the step length, and the optimization method must be robust. Furthermore, we will measure the number of evaluations of f and df necessary to obtain a solution. One iteration of the optimization procedure requires the two evaluations, once to evaluate the solution candidate and once to construct a quadratic approximation along the search direction.

For computational reasons we choose an $m \times n$ matrix, $\mathbf{X}$, as a representative of a point on $\mathcal{M}$, and enforce the requirement

$$\mathbf{X}^T\mathbf{X} = \mathbf{I}. \tag{2}$$

While this representation includes more degrees of freedom than strictly necessary, the approach is suitable for use in practice [10].

We use the inner product for matrices

$$(\mathbf{A}, \mathbf{B}) = \mathrm{trace}(\mathbf{A}^T\mathbf{B}), \tag{3}$$

and the tangent spaces at $\mathbf{X}$ satisfying $\mathbf{X}^T\mathbf{X} = \mathbf{I}$

$$\{\mathbf{Z} = \mathbf{X}\mathbf{A} + \mathbf{Y} \,|\, \mathbf{Y}^T\mathbf{X} = \mathbf{0} \text{ and } \mathbf{A}^T = -\mathbf{A}\}. \tag{4}$$

On the Grassman manifold the value of f depends only on the space spanned by the columns of $\mathbf{X}$. We can therefore ignore the $\mathbf{X}\mathbf{A}$ component of the tangent space, and denote

$$\mathcal{T}_{\mathbf{X}}\mathcal{M} = \{\mathbf{Y} \,|\, \mathbf{Y}^T\mathbf{X} = \mathbf{0}\}. \tag{5}$$

On the Grassman manifold it is possible to substitute equivalence classes for the representatives we have chosen, but practical computations require us to always use a specific matrix.

We also assume that we are given a direction $\mathbf{W}$ by the minimization method in which we want to move on the manifold. We project the direction on to $\mathcal{T}_{\mathbf{X}}\mathcal{M}$ by

$$\mathbf{Y} = (\mathbf{I} - \mathbf{X}\mathbf{X}^T)\mathbf{W}. \tag{6}$$

The motivation for the problem under consideration comes from density functional theory (DFT) electronic structure calculations [17, 21]. We construct a simultaneous Householder operator that can be used to ensure that the optimization method naturally enforces the orthogonality constraint. This approach differs from several other approaches in that it does not solve the canonical electron orbitals [13, 18, 22, 4, 23, 21], instead we only solve the electron density
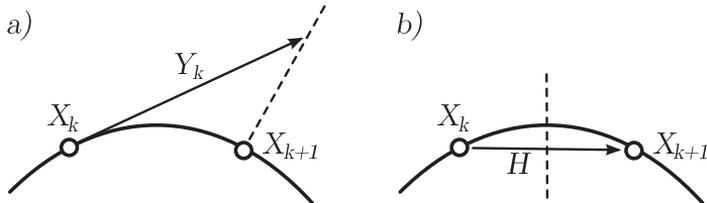
Figure 1: Conceptual difference between reorthogonalization and Householder approach. In Subfigure a) the step is taken without regard to the manifold, and after the step is taken the new solution candidate $\mathbf{X}_{k+1}$ is constructed by reorthogonalizing $\mathbf{X}_k + \mathbf{Y}_k$. Subfigure b) illustrates the case where the Householder operator, $\mathbf{H}$, constructs an update $\mathbf{X}_{k+1}$ that immediately satisfies the orthogonality condition.

that would be given by the orbitals. To obtain the canonical orbitals from the electron density a linear eigenvalue problem must then be solved in the space spanned by the columns of $\mathbf{X}$. A similar approach using polynomial filtering can be found in [24, 3]. It is also possible to solve a nonlinear eigenvalue problem instead of the minimization problem [20, 13, 16, 21].

In [10] a framework for optimization methods on the Stiefel and Grassmann manifolds is presented, while [7] discusses a Newton-like iteration scheme on a more general manifold. Univariate optimization methods for the Stiefel manifold is presented in [5], where identity plus rank one Householder transforms are given as one possible choice for moving on the manifold. The choice of coordinates can also be based on a QR factorization and polar decompositions [6, 9] or Lie groups [14]. An overview of geometric numerical integration techniques can be found in [15].

First, we present a simultaneous Householder transformation in Section 2 that we can use to move on both the Stiefel and Grassman manifolds. Then in Section 3 we recall the method of steepest descent, the quasi-Newton (QN), and the nonlinear conjugate gradient (NLCG) methods adapted for use with the Householder operator. In Section 4 we numerically demonstrate the method on a model problem that includes nonlinearities similar to a DFT problem. Finally, Section 5 presents the conclusion.

2 Householder operator

We ensure that $\mathbf{X} \in \mathcal{M}$ during the solution process by using the Householder transformation to move from one solution candidate to the next.

To do this we need to find an operator

$$\mathbf{H}(\tau) = \mathbf{I} - 2\mathbf{Q}(\tau)\mathbf{Q}(\tau)^T, \tag{7}$$

3

where
$$\mathbf{Q}(\tau)^T \mathbf{Q}(\tau) = \mathbf{I} \quad \forall \tau, \tag{8}$$
and τ is a parametrization of $\mathbf{H}$ such that
$$\mathbf{H}(0)\mathbf{X} = \mathbf{X}. \tag{9}$$
This requirement leads to
$$\mathbf{Q}(0)^T \mathbf{X} = \mathbf{0}. \tag{10}$$

$\mathbf{H}$ is unitary, and if we let
$$\mathbf{X}_{k+1} = \mathbf{H}(\tau)\mathbf{X}_k, \tag{11}$$
we obtain a sequence $\{\mathbf{X}_k\}$ that satisfies
$$\mathbf{X}_{k+1}^T \mathbf{X}_{k+1} = \mathbf{X}_k^T \mathbf{X}_k \quad \forall \, k, \tau. \tag{12}$$

The orthonormality requirement (8) on $\mathbf{Q}(\tau)$ leads to the constraint
$$\frac{\partial \mathbf{Q}}{\partial \tau}^T \mathbf{Q} + \mathbf{Q}^T \frac{\partial \mathbf{Q}}{\partial \tau} = \mathbf{0}. \tag{13}$$

We set the initial condition for $\frac{\partial \mathbf{Q}}{\partial \tau}$ by requiring that
$$\frac{\partial}{\partial \tau}\left(\mathbf{H}(\tau)\mathbf{X}\right)\Big|_{\tau=0} = \mathbf{Y}, \tag{14}$$
where $\mathbf{Y} \in \mathcal{T}_\mathbf{X}\mathcal{M}$ is a projected direction given by the minimization method we choose to employ. Differentiating with respect to τ, and using property (10), we obtain from (14)
$$-2\mathbf{Q}(0)\frac{\partial \mathbf{Q}}{\partial \tau}(0)^T \mathbf{X} = \mathbf{Y}. \tag{15}$$

We set
$$\mathbf{Q}(0) = \mathbf{V}, \tag{16}$$
where $\mathbf{VR}$ is the compact QR decomposition of $\mathbf{Y}$, and choose
$$\frac{\partial \mathbf{Q}}{\partial \tau}(0) = -\tfrac{1}{2}\mathbf{X}\mathbf{R}^T. \tag{17}$$

A solution satisfying Equation (13) and conditions (9) and (14) is
$$\tilde{\mathbf{Q}}(\tau) = \tilde{\mathbf{Q}}_0 \exp\left(\tau \begin{bmatrix} \mathbf{0} & \tfrac{1}{2}\mathbf{R} \\ -\tfrac{1}{2}\mathbf{R}^T & \mathbf{0} \end{bmatrix}\right). \tag{18}$$

Here $\tilde{\mathbf{Q}}_0 = \begin{bmatrix} \mathbf{V} & \mathbf{X} \end{bmatrix}$ and $\mathbf{Q}(\tau)$ corresponds to the first n columns of $\tilde{\mathbf{Q}}(\tau)$. Given any $\mathbf{Q}$ with orthonormal columns $\mathbf{H}$ constructed by (7) ensures that $\mathbf{X}_{k+1}^T \mathbf{X}_{k+1} = \mathbf{I}$. For this reason we also consider a Householder operator constructed from a second order expansion of the matrix exponential that has

subsequently been orthonormalized by the QR method to ensure that the orthogonality constraint is satisfied. This approach is similar to the orbital transformation but includes orthogonalization after every evaluation of the matrix exponential function [23]. To distinguish these from the basic algorithms we use AEQN and AENLCG to denote the approximate exponential versions.

Remark: If $m < 2n$ the requirement $\mathbf{Y}^T\mathbf{X} = \mathbf{0}$ restricts the number of columns in $\mathbf{Y}$ to below n. In this case the size of the first block in Equation (18) should be reduced accordingly. A similar modification must be made if $\mathbf{Y}$ is not full column rank. For the Householder transformation to work we must still have $m > n$.

3 Descent methods with orthogonality constraints

In this section we consider the method of steepest descent, a quasi-Newton method, and a nonlinear conjugate gradient method for minimization with orthogonality constraints. We also present the Householder operator for an $\mathbf{S}$-orthonormal basis, and combine this with the optimization methods.

3.1 The method of steepest descent

The method of steepest descent for the Stiefel manifold is also known as the projected gradient method [8]. At each step we simply set

$$\mathbf{Y}_k = -\sigma(\mathbf{I} - \mathbf{X}_k\mathbf{X}_k^T)\nabla f(\mathbf{X}_k). \tag{19}$$

The parameter $\sigma > 0$ is almost redundant for the method of steepest descent, but will become important for the quasi-Newton methods presented later.

To decide the step length we evaluate $f(\mathbf{H}(\tau_k^e)\mathbf{X}_k)$, where τ_k^e is an estimate step length, and construct the quadratic approximation $p(\tau)$ of $f(\mathbf{H}(\tau)\mathbf{X}_k)$. We then solve $\tau_{\min} = \operatorname{argmin} p(\tau)$ from the system

$$p(0) = f(\mathbf{H}(0)\mathbf{X}), \tag{20}$$
$$p(\tau_k^e) = f(\mathbf{H}(\tau_k^e)\mathbf{X}), \tag{21}$$
$$p'(0) = (\nabla f(\mathbf{H}(0)\mathbf{X}), \mathbf{Y}). \tag{22}$$

This permits us to compute $\mathbf{H}(\beta\tau_{\min})$ and evaluate $f(\mathbf{H}(\beta\tau_{\min})\mathbf{X})$ as well as $\nabla f(\mathbf{H}(\beta\tau_{\min})\mathbf{X})$, where β is an underrelaxation parameter. To ensure that we obtain a non-increasing iteration we choose the step length τ_k from the set $\{0, \beta\tau_{\min}, \tau_k^e\}$, such that we obtain the lowest value of f evaluated so far. If $\tau_k = 0$ we set $\tau_{k+1}^e = 0.25 \times \tau_k^e$ and otherwise we set $\tau_{k+1}^e = \min(|\tau_{\min}|, 2\,\tau_k^e)$.

Remark: It turns out that τ_k^e often is an acceptable choice for step length, and we believe that it is possible to construct a completely line search free minimization method with adaptive step length [2]. However, it must be tuned for a real-world problem, and we will not explore this option here.
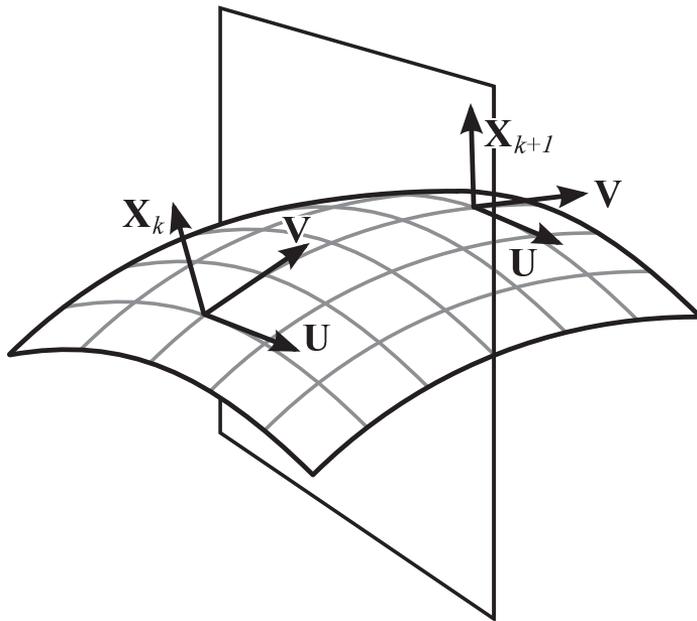
Figure 2: The curvature of the manifold must be taken into account when the trial solution is updated. The update operator, $\mathbf{H}$, corresponds to a reflection in the illustrated plane.

3.2 Vector transport on the manifold

To reduce the number of iterations we use information from previous evaluations to improve the search direction. To this end we construct a transport operator, $\mathbf{T} : \mathcal{T}_{\mathbf{X}_k}\mathcal{M} \to \mathcal{T}_{\mathbf{X}_{k+1}}\mathcal{M}$, that moves tangent vectors at $\mathbf{X}_k$ to the tangent space at $\mathbf{X}_{k+1}$.

Any given vector, $\mathbf{Z}_k \in \mathbb{R}^{m \times n}$, associated with a candidate solution, $\mathbf{X}_k$, and search direction, $\mathbf{V}$, can be decomposed into

$$\mathbf{Z}_k = \mathbf{X}_k\mathbf{A} + \mathbf{V}\mathbf{B} + \mathbf{U}\mathbf{C}, \tag{23}$$

and we demand that $\mathbf{X}_k^T\mathbf{U} = \mathbf{V}^T\mathbf{U} = \mathbf{0}$, and $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ in addition to $\mathbf{X}_k^T\mathbf{V} = \mathbf{0}$ which is satisfied by construction. For simplicity we assume that $\mathbf{Z}_k$ is such that $\mathbf{U} \in \mathbb{R}^{m \times n}$ and $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{n \times n}$. $\mathbf{A}$ and $\mathbf{B}$ can be computed by projecting $\mathbf{Z}_k$ onto $\mathbf{X}_k$ and $\mathbf{V}$ respectively, while $\mathbf{U}\mathbf{C}$ is for example the QR-decomposition of the remainder. In practice the decomposition is not explicitly constructed.

The different parts of the decompostion (23) must be transported separately when the position is updated to $\mathbf{X}_{k+1} = \mathbf{H}\mathbf{X}_k$. The component spanned by $\mathbf{X}_k$ is reflected correctly by $\mathbf{H}$, however a reflection gives the wrong sign to the $\mathbf{V}$ component of $\mathbf{Z}_k$. The final component $\mathbf{U}$ is orthogonal to both $\mathbf{X}_k$ and $\mathbf{V}$, and should not change when $\mathbf{H}$ is applied. The transport operation is shown in Figure 2. We construct the transport operator by separating the components of $\mathbf{Z}_k$ by projection and subsequent application of $\mathbf{H}$. The transport operator therefore becomes

$$\mathbf{H}\mathbf{X}_k\mathbf{X}_k^T - \mathbf{H}\mathbf{V}\mathbf{V}^T + (\mathbf{I} - \mathbf{X}_k\mathbf{X}_k^T - \mathbf{V}\mathbf{V}^T). \tag{24}$$

In practice we apply this for vectors $\mathbf{Z}_k \in \mathcal{T}_{\mathbf{X}_k}\mathcal{M}$ which satisfy $\mathbf{Z}_k^T\mathbf{X}_k = \mathbf{0}$, and we can use the simplified transport operator

$$\mathbf{T}(\tau) = -\mathbf{H}(\tau)\mathbf{V}\mathbf{V}^T + (\mathbf{I} - \mathbf{V}\mathbf{V}^T), \tag{25}$$

and the tangent vector corresponding to $\mathbf{Z}_k$ at $\mathbf{X}_{k+1}$ is

$$\mathbf{Z}_{k+1} = \mathbf{T}(\tau_k)\mathbf{Z}_k \in \mathcal{T}_{\mathbf{X}_{k+1}}\mathcal{M}. \tag{26}$$

3.3 S-orthonormal Householder and transport operator

In practice $\mathbf{X}$ is often represented by a discretization that requires a generalization of the orthonormality constraint (2). The generalized constraint is

$$\mathbf{X}^T\mathbf{S}\mathbf{X} = \mathbf{I}, \tag{27}$$

where $\mathbf{S}$ is symmetric and positive definite. This constraint arises for example as an overlap matrix in DFT or a mass matrix in finite element calculations. If $\mathbf{S}$ is sparse or has other structure that can be exploited it can be preferable not to change basis for the optimization procedure. We therefore also present the Householder transform and optimization procedure for the $\mathbf{S}$-orthonormal case.

We can construct both the Householder transform and the transport operator using the same argument as for the regular orthonormal case, as long as we account for $\mathbf{S}$ orthonormality (27). However, if $\mathbf{S}$ is full, and lacks exploitable structure, the operation $\mathbf{SX}$ has asymptotic complexity $m^2 n$ and overshadows the rest of the procedure.

The projection of the direction of steepest descent onto the $\mathbf{S}$-orthogonal manifold is

$$\mathbf{Y} = -\sigma(\mathbf{I} - \mathbf{XX}^T\mathbf{S})\nabla f(\mathbf{X}) \tag{28}$$

instead of (19) and satisfies $\mathbf{Y}^T\mathbf{SX} = \mathbf{0}$. We must also use the $\mathbf{S}$ weighted compact QR decomposition to compute a factorization $\mathbf{Y} = \mathbf{VR}$ where $\mathbf{V}^T\mathbf{SV} = \mathbf{I}$, and $\mathbf{R}$ is upper triangular.

With these modifications, the Householder operator in (7) becomes

$$\mathbf{H_S}(\tau) = \mathbf{I} - 2\mathbf{Q}(\tau)\mathbf{Q}(\tau)^T\mathbf{S}, \tag{29}$$

where $\mathbf{Q}(\tau)$ is as in Equation (18), and remains unchanged. The $\mathbf{S}$-orthonormal transport operator corresponding to (25) is

$$\mathbf{T_S}(\tau) = -\mathbf{H_S}(\tau)\mathbf{VV}^T\mathbf{S} + (\mathbf{I} - \mathbf{VV}^T\mathbf{S}). \tag{30}$$

3.4 The quasi-Newton method based on Householder transforms

The method of steepest descent generally performs poorly if the minimum of the target function is at the bottom of a narrow valley. Newton's method solves this problem, but requires that the Hessian of the function is available to determine the search direction. When the Hessian is not available we can replace it with an approximation of the true inverse Hessian of the system to obtain a quasi-Newton method.

We base our method on Broyden's second or *bad* update to construct the approximate inverse Hessian, $\mathbf{G}_k$, of f at $\mathbf{X}_k$. While Broyden's second update does not construct a symmetric approximation, or ensure that the approximation is positive definite it is a robust choice for electronic structure calculations [16, 1, 2]. However, we must take into account that our vectors are actually $\mathbb{R}^{m \times n}$ matrices, which will lead to a method identical to the generalized Broyden update. The secant condition is then

$$\mathbf{G}_{k+1}\Delta\mathbf{F}_k = \Delta\mathbf{X}_k, \tag{31}$$

where we project the orbital differences

$$\Delta\mathbf{X}_k = (\mathbf{I} - \mathbf{X}_{k+1}\mathbf{X}_{k+1}^T\mathbf{S})(\mathbf{X}_{k+1} - \mathbf{X}_k), \tag{32}$$

and gradient differences

$$\Delta\mathbf{F}_k = (\mathbf{I} - \mathbf{X}_{k+1}\mathbf{X}_{k+1}^T\mathbf{S})\nabla f(\mathbf{X}_{k+1}) - \mathbf{T}(\tau_k)(\mathbf{I} - \mathbf{X}_k\mathbf{X}_k^T\mathbf{S})\nabla f(\mathbf{X}_k) \tag{33}$$

onto $\mathcal{T}_{\mathbf{X}_{k+1}}\mathcal{M}$. The no change condition is now

$$\mathbf{G}_k\mathbf{Z} = \mathbf{G}_{k+1}\mathbf{Z} \quad \forall \mathbf{Z} \, : \, \mathbf{Z}^T\Delta\mathbf{F}_k = \mathbf{0}. \tag{34}$$

These conditions corresponds to the generalized Broyden's second update for groups of size n. We can therefore use the generalized update formula [11]

$$\mathbf{G}_{k+1} = \mathbf{G}_k + (\Delta\mathbf{X}_k - \mathbf{G}_k\Delta\mathbf{F}_k)(\Delta\mathbf{F}_k^T\mathbf{S}\Delta\mathbf{F}_k)^{-1}\Delta\mathbf{F}_k^T\mathbf{S}. \tag{35}$$

As initial guess we use $\mathbf{G}_0 = \sigma\mathbf{I}$. With this choice, the quasi-Newton method is identical with the method of steepest descent if we do not enforce any secant conditions (31). When secant conditions are enforced we can use σ to control the influence of $\mathbf{G}_0$ compared to the information gained from the secant conditions. In general, the information gained from these is reliable, and therefore σ should be small [1, 2].

To construct the search direction we use

$$\mathbf{Y}_k = -\mathbf{G}_k(\mathbf{I} - \mathbf{X}_k\mathbf{X}_k^T\mathbf{S})\nabla f(\mathbf{X}_k). \tag{36}$$

In practice, we do not store $\mathbf{G}_k$ as a full matrix. Instead we represent it as a low rank update. Details on recursive or low rank implementation of $\mathbf{G}_k$ can be found in [12, 19, 1, 2].

We also limit the number of secant conditions used to construct $\mathbf{G}_k$. Each condition requires storage of two $m \times n$ matrices, $\Delta\mathbf{X}$ and $\Delta\mathbf{F}$, and these matrices must be transported to $\mathcal{T}_{\mathbf{X}_k}\mathcal{M}$ after each step. This is done with the transport operator, $\mathbf{T}_{\mathbf{S}}(\tau)$, defined in Equation (30). As we demonstrate in Section 4 the first few secant conditions offer dramatic improvement over the method of steepest descent, but further secant conditions do not give the same benefit. For this reason we limit the secant conditions by a pre-determined history length, and simply discard older conditions.

We use the same line search as the one presented in Section 3.1. If, however

$$(\mathbf{Y}_k, \mathbf{S}(\mathbf{I} - \mathbf{X}_k\mathbf{X}_k^T\mathbf{S})\nabla f(\mathbf{X}_k)) \geq 0 \tag{37}$$

then the proposed direction is not a descent direction. In this case we restart the optimization method and forget the secant history. If the line search returns the current point $\mathbf{X}_k$, we update $\mathbf{G}_k$ but stay at $\mathbf{X}_k$.

3.5 Nonlinear conjugate gradients

The linear conjugate gradient (CG) method can be viewed as a optimization method for a quadratic problem. Several generalizations of the CG method have been presented to solve optimization problems that are not of quadratic form [19]. Below, we review a nonlinear CG method adapted to account for the curvature of the manifold [10].

Given $\mathbf{X}_0$ which satisfies $\mathbf{X}_0^T\mathbf{X}_0 = \mathbf{I}$, the gradient projected onto $\mathcal{T}_{\mathbf{X}_0}\mathcal{M}$ is

$$\mathbf{Y}_0 = (\mathbf{I} - \mathbf{X}_0\mathbf{X}_0^T\mathbf{S})\nabla f(\mathbf{X}_0), \tag{38}$$

and the initial search direction is the direction of steepest descent

$$\mathbf{P}_0 = -\mathbf{Y}_0. \tag{39}$$

On the manifold the NLCG method then proceeds by minimizing f along the path defined by the search direction $\mathbf{P}_k$. In practice we evaluate f once along the search direction and minimize the quadratic approximation as in Section 3.1. The next candidate is chosen as the best evaluated step, τ_k,

$$\mathbf{X}_{k+1} = \mathbf{H}(\tau_k)\mathbf{X}_k, \tag{40}$$

and the gradient and conjugate directions are transported to $\mathcal{T}_{\mathbf{X}_{k+1}}\mathcal{M}$ by $\mathbf{T}(\tau_k)$ in Equation (30). The new gradient

$$\mathbf{Y}_{k+1} = (\mathbf{I} - \mathbf{X}_{k+1}\mathbf{X}_{k+1}^T\mathbf{S})\nabla f(\mathbf{X}_{k+1}), \tag{41}$$

and conjugate direction

$$\mathbf{P}_{k+1} = -\mathbf{Y}_{k+1} + \gamma_k \mathbf{T}_{\mathbf{S}}(\tau_k)\mathbf{P}_k, \tag{42}$$

are then computed where

$$\gamma_k = \frac{(\mathbf{Y}_{k+1} - \mathbf{T}_{\mathbf{S}}(\tau_k)\mathbf{Y}_k, \mathbf{Y}_{k+1})}{(\mathbf{Y}_k, \mathbf{Y}_k)}. \tag{43}$$

For comparison we also implement a projected NLCG (PNLCG) method. Instead of ensuring that orbital updates satisfy $\mathbf{X}_{k+1}^T\mathbf{X}_{k+1} = \mathbf{I}$ we orthogonalize $\mathbf{X}_{k+1}$ after every update with the QR method. The conjugate directions and gradient are not transported to $\mathcal{T}_{\mathbf{X}_{k+1}}\mathcal{M}$, instead they are updated with the $\mathbf{I} - \mathbf{X}_{k+1}\mathbf{X}_{k+1}^T\mathbf{S}$ projector onto $\mathcal{T}_{\mathbf{X}_{k+1}}\mathcal{M}$.

4   Numerical experiments

We use a two dimensional model problem with the condition $\mathbf{X}^T\mathbf{S}\mathbf{X} = \mathbf{I}$ to compare the projected NLCG method with the NLCG and QN methods where satisfaction the orthonormality condition is ensured by the update operator. The model problem is inspired by electronic structure theory, and corresponds to a three dimensional system constrained to two dimensions without exchange-correlation terms.

The target function is [17]

$$f(\mathbf{X}) = -\tfrac{1}{2}\mathrm{tr}((\mathbf{S}^{1/2}\mathbf{X})^T\mathbf{L}\mathbf{S}^{1/2}\mathbf{X}) + \mathbf{v}^T\mathbf{n} + \tfrac{1}{2}\mathbf{n}^T\mathbf{P}\mathbf{n}, \tag{44}$$

where $\mathbf{L} \in \mathbb{R}^{m \times m}$ is the discretized Laplace operator, $\mathbf{v} \in \mathbb{R}^m$ the external potential, $\mathbf{n} \in \mathbb{R}^m$ the electron density, and $\mathbf{P}\mathbf{n}$ the Hartree potential. The electron density is

$$\mathbf{n}_i = \sum_{j=1}^{n}((\mathbf{S}^{1/2}\mathbf{X}) \circ (\mathbf{S}^{1/2}\mathbf{X}))_{ij}, \tag{45}$$

where ∘ is the entrywise, or Hadamard, product. We use the overlap matrix

$$\mathbf{S} = \frac{1}{9h^2}\begin{bmatrix} \mathbf{M} & \frac{\mathbf{M}}{4} & 0 & \cdots \\ \frac{\mathbf{M}}{4} & \mathbf{M} & \frac{\mathbf{M}}{4} & \ddots \\ 0 & \frac{\mathbf{M}}{4} & \mathbf{M} & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}, \tag{46}$$

where h is the one dimensional grid size and

$$\mathbf{M} = \begin{bmatrix} 4 & 1 & 0 & \cdots \\ 1 & 4 & 1 & \ddots \\ 0 & 1 & 4 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{47}$$

This corresponds to the mass matrix of a finite element discretization with bilinear quadratic element. The calculations have also been performed with a symmetric and positive definite random matrix. It turns out that as long as $\mathbf{S}$ is well conditioned, it has only a small effect on the rate of convergence.

To calculate the potentials we use

$$\mathbf{v}_i = -\sum_{j=1}^{N} \frac{Z_j}{||\mathbf{r}_i - \mathbf{R}_j|| + \alpha}, \tag{48}$$

where the sum is over the nuclei with charge Z_j and position $\mathbf{R}_j$. The position corresponding to the discretization point i is $\mathbf{r}_i$, and the parameter α is used to regularize the potential. $\mathbf{P} \in \mathbb{R}^{m \times m}$ is similarly given by

$$\mathbf{P}_{ij} = \frac{1}{||\mathbf{r}_i - \mathbf{r}_j|| + \alpha}. \tag{49}$$

We solve the problem in the unit square with zero boundary conditions corresponding to an infinite potential well. We use a uniform finite difference discretization with m inner points to obtain a system where $\mathbf{X} \in \mathbb{R}^{m \times n}$. Here n corresponds to the number of electrons. As initial guess we use the solution of the quadratic problem using the first two terms of (44), and choose $\tau_0^e = 1.0$ to initialize the minimization procedure, cf. Equation (20).

As generators for the external potential we use two nuclei, where one is placed at the grid point closest to $(\frac{1}{3}, \frac{1}{3})$, and the other at the grid point closest to $(\frac{2}{3}, \frac{13}{24})$. The off diagonal placement is chosen to break the symmetry of the system. We use

$$\epsilon_\mathbf{X} = ||(\mathbf{I} - \mathbf{X}\mathbf{X}^T)\nabla f(\mathbf{X})||/\sqrt{mn}, \tag{50}$$

to measure convergence, and consider the system converged when
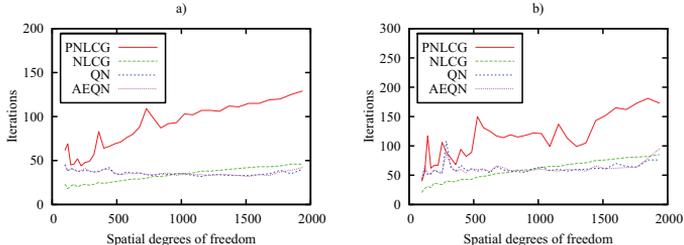
$$\epsilon_\mathbf{X} < 10^{-2}. \tag{51}$$

11

Figure 3: Iterations required for convergence as a function of the spatial degrees of freedom, m. For subfigure a) the external potential is generated by two nuclei, $Z_1 = 3$, $Z_2 = 3$, and the parameters of the calculation are $n = 6$, $\beta = 0.5$, $\sigma = 10^{-4}$, $\alpha = 2 \times 10^{-2}$, and history length is 6. For subfigure b) $Z_1 = 4$, $Z_2 = 3$, and $n = 7$. Here NLCG corresponds to the nonlinear conjugate gradient method, PNLCG to the projected NLCG, QN to the quasi-Newton method and AEQN to the approximate exponent QN method. The number of iterations required for convergence is identical for the AENLCG and NLCG methods, where AENLCG is the approximate exponent NLCG method.

At this point
$$|f(\mathbf{X}_{\mathrm{Ref}}) - f(\mathbf{X})| \approx 10^{-4}, \tag{52}$$
where the reference solution has been calculated such that
$$\epsilon_{\mathbf{X}_{\mathrm{Ref}}} < 10^{-5}. \tag{53}$$

Figure 3 presents the iterations necessary for convergence for a six and seven electron system. These iterations roughly grows as the square root of the degrees of freedom for the NLCG methods. However, the projected NLCG method performs significantly worse than the NLCG and QN methods adapted for the manifold. It turns out that the AENLCG method requires the same number of iterations to converge as the NLCG method, while a small difference is visible between QN and AEQN methods.

In Figure 4 the effect of weight, σ, of the initial approximation of the inverse Hessian and the underrelaxation, β are presented. From Figure 4 b) it is clear that σ is particularly important for fast convergence of the quasi-Newton method. The effect of β is much smaller, and while β can in some cases improve convergence the effect of σ is significantly more important. A low σ results in slower convergence as long as the more aggressive parameter choice converges well. However, when the rate of convergence begins to suffer from the more aggressive parameter choice the rate of convergence can be improved by a more conservative choice. These results agree with earlier work [1], which indicate that the secant conditions offer reliable information of the electronic structure problem, while the initial approximation of $\mathbf{G}$ is less reliable. The history length
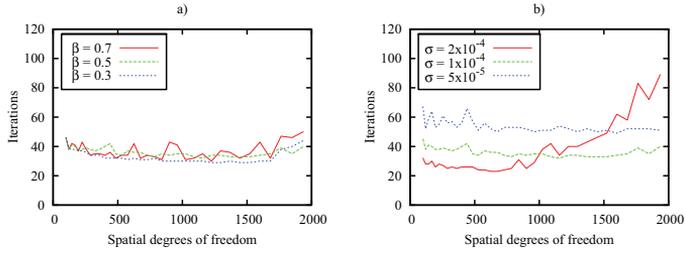
12

Figure 4: Iterations required for convergence of the QN method as a function of the spatial degrees of freedom, m. For the both figures the external potential is generated by two nuclei, $Z_1 = 3$, $Z_2 = 3$, and the parameters of the calculation are $n = 6$, $\alpha = 2 \times 10^{-2}$, and history length is 6. Unless otherwise indicated in the figure $\beta = 0.5$ and $\sigma = 10^{-4}$.
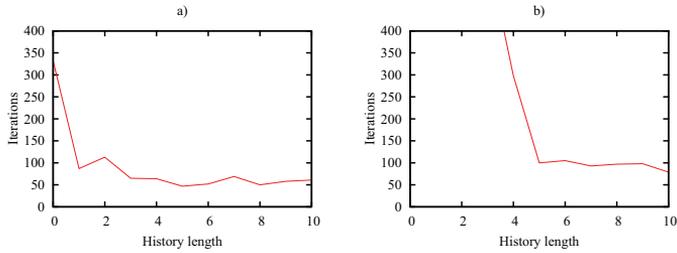

Figure 5: Iterations required for convergence as a function of history length for the QN method for a six and seven electron system for subfigure a) and b) respectively. Spatial degrees of freedom is 2500 and history length varies, parameters are otherwise identical to Figure 3.

13

of the QN method must also be sufficient for the method to perform well. This is illustrated in Figure 5.

5 Conclusion

We have presented a Householder update scheme which ensures that the columns remains orthogonal, and is suitable for both NLCG and QN methods. Furthermore, the operator allows us to transport gradient information and construct secant conditions from previous evaluations of f to the tangent space of the best candidate solution. This approach eliminates the need to parametrize the manifold, and permits us to use standard linear algebra routines to update the solution candidate.

We have demonstrated the methods numerically on a model problem inspired by the electronic structure problem, and compared them to a projected NLCG method. Taking the underlying manifold into account significantly improves convergence rate of the optimization methods, and using a second order orthonormal approximate matrix exponent does not decrease performance of the QN or NLCG methods.

The QN method is significantly improved by taking the first few secant conditions into account when constructing the approximation of the Hessian of f. However, for the secant condition history to improve convergence speed of the QN method the manifold must be taken into account. The update of the secant conditions is also based on the Householder operator that is used to update the solution candidate. While the performance of the QN method depends on the weight of the initial approximate Hessian the QN method performs well once the weight is correctly set.

While the QN method is sensitive to the correct choice of σ, the performance of the NLCG method does not depend on parameter choice. Furthermore, the NLCG method is relatively simple to implement, and only requires a one step history. For these reasons we believe that the NLCG method is a good general purpose optimization method for electronic structure problems if the method is adapted to the manifold.

6 Acknowledgments

We are grateful towards Dr. Mika Juntunen for suggestions and comments on the manuscript.

References

[1] K. Baarman, T. Eirola, and V. Havu. Robust acceleration of self consistent field calculations in density functional theory. *J. Chem. Phys.*, 134:134109, 2011.

[2] K. Baarman and J. VandeVondele. A comparison of accelerators for direct energy minimization in electronic structure calculations. *J. Chem. Phys.*, 134:244104, 2011.

[3] C. Bekas, E. Kokiopoulou, and Y. Saad. Computation of large invariant subspaces using polynomial filtered lanczos iterations with applications in density functional theory. *SIAM J. Matrix Anal. Appl.*, 30:397, 2008.

[4] Eric Cancès. Self-consistent field algorithms for kohn-sham models with fractional occupation numbers. *J. Chem. Phys.*, 114:10616, 2001.

[5] E. Celledoni and S. Fiori. Descent methods for optimization on homogeneous manifolds. *Mathematics and Computers in Simulation*, 79:1298, 2008.

[6] E. Celledoni and B. Owren. A class of intrinsic schemes for orthogonal integration. *SIAM J. Numer. Anal.*, 40:2069, 2002.

[7] M. T. Chu. On a numerical treatment for the curve-tracing of the homotopy method. *Numer. Math*, 42:323, 1983.

[8] M. T. Chu and K. R. Drissel. The projected gradient method for least squares matrix approximations with spectral constraints. *SIAM J. Numer. Anal.*, 27:1050, 1990.

[9] L. Dieci and E. S. Van Vleck. Orthonormal integrators based on Householder and Givens transformations. *Future Gener. Comp. Sy.*, 19:363, 2003.

[10] Alan Edelman, Tomás A. Arias, and Steven T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, 20:303, 1998.

[11] H.-r. Fang and Y. Saad. Two classes of multisecant methods for nonlinear acceleration. *Numer. Linear Algebra Appl.*, 16:197, 2009.

[12] M. Kawata, C. M. Cortis, and R. A. Friesner. Efficient recursive implementation of the modified broyden method and the direct inversion in the iterative subspace method: Acceleration of self-consistent calculations. *J. Chem. Phys.*, 108:4426, 1998.

[13] G. Kresse and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors usign a plan-wave basis set. *Comp. Mat. Sci.*, 6:15, 1996.

[14] S. Krogstad. A low complexity lie group method on the Stiefel manifold. *BIT*, 43:107, 2003.

[15] C. Lubich, E. Hairer, and G. Wanner. *Geometric Numerical Integration*. Springer, 2006.

[16] L. D. Marks and D. R. Luke. Robust mixing for ab initio quantum mechanical calculations. *Phys. Rev. B*, 78:075114, 2008.

[17] Dominik Marx and Jürg Hutter. *Ab Initio Molecular Dynamics*. Cambridge University Press, 2009.

[18] Nicola Marzari, David Vanderbilt, and M. C. Payne. Ensemble density-functional theory for ab initio molecular dynamics of metals and finite-temperature insulators. *Phys. Rev. Lett.*, 79:1337, 1997.

[19] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer, 2006.

[20] P. Pulay. Convergence acceleration in iterative sequences: The case of SCF iteration. *Chem. Phys. Lett.*, 73:393, 1980.

[21] Y. Saad, J. R. Chelikowsky, and S. M. Shontz. Numerical methods for electronic structure calculations of materials. *SIAM Review*, 52:3, 2010.

[22] Troy van Voorhis and Martin Head-Gordon. A geometric approach to direct minimization. *Mol. Phys.*, 100:1713, 2002.

[23] J. VandeVondele and Jürg Hutter. An efficient orbital transformation method for electronic structure calculations. *J. Chem. Phys.*, 118:4365, 2003.

[24] Y. Zhou, Y. Saad, M. L. Tiago, and J. R. Chelikowsky. Parallel self-consistent-field calculations via chebyshev-filtered subspace acceleration. *Phys. Rev. E*, 74:066704, 2006.