

A matrix algebra approach to approximate Hessians

WARREN HARE AND GABRIEL JARRY-BOLDUC

*Department of Mathematics, University of British Columbia, Okanagan Campus, Kelowna,
B.C. V1V 1V7, Canada*

AND

CHAYNE PLANIDEN*

*School of Mathematics and Applied Statistics, University of Wollongong, Wollongong,
NSW, 2500, Australia*

*Corresponding author: chayne@uow.edu.au

[Received on 4 July 2022; revised on 7 June 2023]

This work presents a novel matrix-based method for constructing an approximation Hessian using only function evaluations. The method requires less computational power than interpolation-based methods and is easy to implement in matrix-based programming languages such as MATLAB. As only function evaluations are required, the method is suitable for use in derivative-free algorithms. For reasonably structured sample sets, the method is proven to create an order-1 accurate approximation of the full Hessian. Under more specialized structures, the method is proved to yield order-2 accuracy. The underdetermined case, where the number of sample points is fewer than required for full interpolation, is studied and error bounds are developed for the resulting partial Hessians.

Keywords: Hessian approximation; derivative-free optimization; order- N accuracy; generalized simplex gradient; generalized simplex Hessian; generalized centered simplex Hessian.

1. Introduction

Derivative-free optimization (DFO) is the study of finding the minimum value and minimizers of a function without using gradients or higher-order derivative information (Audet & Hare, 2017). DFO is gaining in popularity in recent years and is useful in cases where gradients are unavailable, computationally expensive or simply difficult to obtain (Conn *et al.*, 2009; Audet & Hare, 2017). DFO methods have been used in both smooth (Powell, 2003; Berghen & Bersini, 2005; Powell, 2009; Wild & Shoemaker, 2013; Cocchi *et al.*, 2018; Shashaani *et al.*, 2018; Liuzzi *et al.*, 2019; Gratton *et al.*, 2020) and nonsmooth (Bagirov *et al.*, 2008; Hare & Nutini, 2013; Hare *et al.*, 2016; Larson *et al.*, 2016; Audet & Hare, 2018; Menickelly & Wild, 2019) optimization, and most commonly take the form of either direct-search (Audet *et al.*, 2008; Amaïoua *et al.*, 2018; Audet & Tribes, 2018; Berahas *et al.*, 2019; Gratton *et al.*, 2019) or model-based (Wild & Shoemaker, 2013; Verdério *et al.*, 2017; Maggiar *et al.*, 2018; Liuzzi *et al.*, 2019; Hare *et al.*, 2020b) methods, while some use a blend of both (Custódio *et al.*, 2008; Audet *et al.*, 2014). Model-based DFO methods use numerical analysis techniques to approximate gradients and Hessians in a manner that has controllable error bounds. For example, one technique for approximating gradients is the simplex gradient. This method consists of taking a set of $n + 1$ sample points in \mathbb{R}^n that are appropriately spaced, called a simplex, and using them to build a linear interpolation function

that approximates the objective function locally, then calculating the gradient of the affine function. The simplex gradient has been shown to have a nicely bounded error, for functions whose gradients are locally Lipschitz continuous (Bortz & Kelley, 1998).

One recent line of research explores methods of improving or generalizing the simplex gradient for its use in DFO (Powell, 2001; Custódio & Vicente, 2007; Conn *et al.*, 2008a,b; Regis, 2015; Coope & Tappenden, 2019; Hare & Jarry-Bolduc, 2020; Hare *et al.*, 2020a). Working from ideas in Conn *et al.* (2008b), in Hare & Jarry-Bolduc (2020) the simplex gradient was generalized so as not to require exactly $n + 1$ points; an error-controlled approximation can now be found using any finite number of properly-spaced points. In Hare *et al.* (2020a), a similar approach was used to examine the generalized centered simplex gradient, an approximation that uses twice as many points, but results in an improvement on the error control from order-1 to order-2 (Hare, 2020).

In a similar vein, researchers have also explored methods to approximate full Hessians or partial Hessians. In Custódio & Vicente (2007), the authors outline an idea for a ‘simplex Hessian’ that is constructed via quadratic interpolation through $(n + 1)(n + 2)/2$ well-poised sample points. They further suggested that if only a portion of the Hessian were desired (say the diagonal component), then fewer points could be used. These ideas were formalized in Conn *et al.* (2008b) through quadratic interpolation and analyzed through the use of Lagrange interpolating polynomials. Obtaining an approximation of the diagonal component of a Hessian is also discussed in Coope & Tappenden (2021); Jarry-Bolduc (2022).

In this work, we continue the development of these tools by introducing the *generalized simplex Hessian (GSH)* and the *generalized centered simplex Hessian (GCSH)*. The GSH is closely linked to the Hessian of a quadratic interpolation function; we will see that in certain situations, both approaches yield the same result.

This paper can be viewed as an extension of the work related to simplex Hessians introduced in Custódio & Vicente (2007); Conn *et al.* (2009). We introduce an explicit formula based on matrix algebra concepts to approximate the Hessian of a function. This formula provides several advantages. First, it provides an approach that is well-defined regardless of the number of sample points utilized. Indeed, as long as the matrices used to build the sample set of points are nonempty, the GSH and the GCSH are well-defined. In particular, the GSH provides an explicit formula to approximate the Hessian even when the quadratic interpolation function does not exist or is not unique. However, we note that when the matrices used in the computation of the GSH have a specific structure, the GSH is equivalent to a *forward-finite-difference approximation* of the Hessian.

The GSH provides an accurate approximation of the Hessian of an objective function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ under reasonable assumptions. The technique uses a point of interest x^0 and sets of directions, stored in a set of matrices $\{S, T_1, \dots, T_m\}$ whose columns are vectors that are added to x^0 to build the sample points. Defining Δ_S as the radius of S and Δ_T as the maximum radius of the T_i , we prove that if S and all T_i are full row rank and the Hessian $\nabla^2 f$ exists and is Lipschitz continuous, then the GSH is an accurate estimate of $\nabla^2 f$ to within a multiple of $\Delta_u := \max\{\Delta_S, \Delta_T\}$. In terms of order- N accuracy Hare (2020), the GSH provides order-1 accurate approximations of the full Hessian (see Definition 2.6 herein). Furthermore, the GCSH provides order-2 accurate approximations of the full Hessian. Error bounds for the *underdetermined case* are presented showing that the GSH and its centered version, the GCSH, are order-1 and order-2 accurate approximations of the appropriate partial Hessian, respectively.

The error bounds presented in this paper share similarities to those presented in Conn *et al.* (2008a,b). In those papers, the authors provide error bounds for the Hessian of a fully quadratic model (Conn *et al.*, 2008a, Theorem 3), the Hessian of a quadratic regression model (Conn *et al.*, 2008b, Theorem 3.2) and the Hessian of a quadratic underdetermined model (Conn *et al.*, 2008b, Theorem 5.12).

A second benefit of the GSH formula is that by employing a matrix structure in the definition, the Hessian approximation is extremely easy to implement in any matrix-based programming language. Moreover, after function evaluations are complete, the matrix structure requires less computational effort (in flops) than quadratic interpolation. Indeed, when S and all T_i are square matrices, then the GSH requires order $O(n^4)$ flops to compute. This is an improvement to the order $O(n^6)$ flops required to find the Hessian of the quadratic interpolation function using a set of sample points that is poised for quadratic interpolation (Conn *et al.*, 2008a) (see also Conn *et al.*, 2009, §6.2).

The remainder of this paper is organized as follows. Section 2 contains a description of notation and some needed definitions. Section 3 introduces the GSH and the GCSH. The relation between these two approximation techniques is clarified. Section 4 presents several error bounds for the GSH and the GCSH. The error bounds are divided into two categories: the case in which all matrices of directions T_i are equal, and the case in which not all matrices T_i are necessarily equal. For both cases, the error bound is valid for the underdetermined setting. In Section 5, we establish the minimal poised set for GSH computation and show that it is well-poised for quadratic interpolation. Finally, Section 6 contains concluding remarks and recommends areas of future research in this vein.

2. Preliminaries

Throughout this work, we use the standard notation found in Rockafellar & Wets (1998). The domain of a function f is denoted by $\text{dom } f$. The transpose of a matrix A is denoted by A^\top . We work in finite-dimensional space \mathbb{R}^n with inner product $x^\top y = \sum_{i=1}^n x_i y_i$ and induced norm $\|x\| = \sqrt{x^\top x}$. The identity matrix in $\mathbb{R}^{n \times n}$ is denoted by Id_n . We use $e_n^i \in \mathbb{R}^n$ for $i \in \{1, \dots, n\}$ to denote the standard unit basis vectors in \mathbb{R}^n , i.e., the i^{th} column of Id_n . The matrix $\text{Diag}(v) = \text{Diag}(v_1, \dots, v_n) =: M \in \mathbb{R}^{n \times n}$, where $v \in \mathbb{R}^n$ represents the diagonal matrix with diagonal entries $M_{i,i} = v_i$ for all $i \in \{1, \dots, n\}$. The zero vector in \mathbb{R}^n is denoted $\mathbf{0}$ and the zero matrix in $\mathbb{R}^{n \times m}$ is denoted $\mathbf{0}_{n \times m}$. The entry in the i^{th} row and j^{th} column of a matrix A is denoted $A_{i,j}$.

Given a matrix $A \in \mathbb{R}^{n \times m}$, we use the induced matrix norm

$$\|A\| = \|A\|_2 = \max\{\|Ax\|_2 : \|x\|_2 = 1\}$$

and the Frobenius norm

$$\|A\|_F = \left(\sum_{i=1}^n \sum_{j=1}^m A_{i,j}^2 \right)^{\frac{1}{2}}.$$

We denote by $B(x^0, \Delta)$ and $\bar{B}(x^0, \Delta)$ the open and closed balls, respectively, centered about x^0 with radius Δ . The Minkowski sum of two sets of vectors A and B is denoted by $A \oplus B$. That is

$$A \oplus B = \{a + b : a \in A, b \in B\}.$$

We define a quadratic function $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ to be a function of the form $Q(x) = \alpha_0 + \alpha^\top x + \frac{1}{2}x^\top \mathbf{H}x$ where $\alpha_0 \in \mathbb{R}$, $\alpha \in \mathbb{R}^n$ and $\mathbf{H} = \mathbf{H}^\top \in \mathbb{R}^{n \times n}$. An affine function $\mathcal{L} : \mathbb{R}^n \rightarrow \mathbb{R}$ is defined to be any function that can be written in the form $\mathcal{L}(x) = \alpha_0 + \alpha^\top x$. Note, affine functions and constant functions $C(x) = \alpha_0$ are also considered quadratic functions, with $\mathbf{H} = \mathbf{0}_{n \times n}$.

Next, we introduce fundamental definitions and notation that will be used throughout this paper.

DEFINITION 2.1 (Poised for quadratic interpolation). [Conn et al. \(2008a,b, 2009\)](#) The set of distinct points $\mathcal{Y} = \{y^0, y^1, \dots, y^m\} \subset \mathbb{R}^n$ with $m = \frac{1}{2}(n+1)(n+2) - 1$ is poised for quadratic interpolation if the system

$$\alpha_0 + \alpha^\top y^i + \frac{1}{2}(y^i)^\top \mathbf{H} y^i = 0, \quad i \in \{0, 1, \dots, m\}, \quad (2.1)$$

has a unique solution for $\alpha_0 \in \mathbb{R}$, $\alpha \in \mathbb{R}^n$ and $\mathbf{H} = \mathbf{H}^\top \in \mathbb{R}^{n \times n}$.

When a set of sample points \mathcal{Y} is poised for quadratic interpolation, it means that there exists a unique quadratic model passing through all the sample points in \mathcal{Y} . Therefore, the Hessian of the quadratic model defined through \mathcal{Y} is unique. Note that a set of sample points containing $(n+1)(n+2)/2$ distinct sample points is not necessarily poised for quadratic interpolation. The set of sample points is poised for quadratic interpolation if and only the matrix associated to the linear system (2.1) is square and full rank. For instance, the set of sample points

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \right\} \subset \mathbb{R}^2$$

is not poised for quadratic interpolation, since the rank of the associated matrix is $5 < 6$. The set

$$\left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \right\} \subset \mathbb{R}^2$$

is poised for quadratic interpolation, since the rank of the associated matrix is 6.

When the number of distinct sample points in \mathcal{Y} is fewer than $(n+1)(n+2)/2$, then the quadratic interpolation model is no longer unique. This is referred as the underdetermined case ([Conn et al., 2009](#), Chapter 5). In this case, a quadratic model may be defined through the minimum Frobenius norm problem ([Conn et al., 2009](#), Chapter 5) (see also [Conn et al., 2008b](#), Section 5). When the number of distinct sample points in \mathcal{Y} is greater than $(n+1)(n+2)/2$ and there exists no quadratic function that passes through all points in \mathcal{Y} , the quadratic model may be determined through a least square regression problem ([Conn et al., 2009](#), Chapter 4) (see also [Conn et al., 2008b](#), Sections 2, 3 and 4).

DEFINITION 2.2 (Quadratic interpolation function). ([Audet & Hare, 2017](#), Definition 9.9) Let $f : \text{dom} f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and let $\mathcal{Y} = \{y^0, y^1, \dots, y^m\} \subset \text{dom} f$ with $m = \frac{1}{2}(n+1)(n+2) - 1$ be poised for quadratic interpolation. Then the quadratic interpolation function of f over \mathcal{Y} is

$$Q_f(\mathcal{Y})(x) = \alpha_0 + \alpha^\top x + \frac{1}{2}x^\top \mathbf{H}x,$$

where $(\alpha_0, \alpha, \mathbf{H} = \mathbf{H}^\top)$ is the unique solution to

$$\alpha_0 + \alpha^\top y^i + \frac{1}{2}(y^i)^\top \mathbf{H} y^i = f(y^i), \quad i \in \{0, 1, \dots, m\}.$$

In the next definition, we introduce key notation used in the construction of the GSH. Within, we write a set of vectors in matrix form, by which we mean that each column of the matrix is a vector in the set.

DEFINITION 2.3 (GSH notation). Let $f : \text{dom} f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and let $x^0 \in \text{dom} f$ be the *point of interest*. Let

$$S = [s^1 \ s^2 \ \dots \ s^m] \in \mathbb{R}^{n \times m} \quad \text{and} \\ T_i = [t_i^1 \ t_i^2 \ \dots \ t_i^{k_i}] \in \mathbb{R}^{n \times k_i}, i \in \{1, \dots, m\}$$

be sets of directions contained in \mathbb{R}^n , written in matrix form. Define

$$T_{1:m} = \{T_1, \dots, T_m\}.$$

Assume that $x^0 \oplus T_i, x^0 + s^i, x^0 + s^i \oplus T_i$ are contained in $\text{dom} f$ for all $i \in \{1, \dots, m\}$. We define

$$\Delta_S = \max_{i \in \{1, \dots, m\}} \|s^i\|, \quad \Delta_{T_i} = \max_{j \in \{1, \dots, k_i\}} \|t_i^j\|, \quad \Delta_T = \max_{i \in \{1, \dots, m\}} \Delta_{T_i}, \\ \widehat{S} = \frac{1}{\Delta_S} S \tag{2.2}$$

and

$$\delta_s(x^0; T_i) = \begin{bmatrix} f(x^0 + t_i^1) - f(x^0) \\ f(x^0 + t_i^2) - f(x^0) \\ \vdots \\ f(x^0 + t_i^{k_i}) - f(x^0) \end{bmatrix} \in \mathbb{R}^{k_i}.$$

In this paper, all matrices involved are assumed to have non-null rank. This ensures that the radius of the matrix is nonzero and hence, (2.2) is always well-defined. Note that in the previous definition, $m \in \mathbb{N}$ and k_i can be any positive integer for all $i \in \{1, \dots, m\}$.

Recall that for nonsquare matrices, a generalization of the matrix inverse is the pseudoinverse. The most well-known type of matrix pseudoinverse and the one we use in the sequel is the Moore–Penrose pseudoinverse.

DEFINITION 2.4 (Moore–Penrose pseudoinverse). Let $A \in \mathbb{R}^{n \times m}$. The *Moore–Penrose pseudoinverse* of A , denoted by A^\dagger , is the unique matrix in $\mathbb{R}^{m \times n}$ that satisfies the following four equations:

- (i) $AA^\dagger A = A$,
- (ii) $A^\dagger AA^\dagger = A^\dagger$,
- (iii) $(AA^\dagger)^\top = AA^\dagger$,
- (iv) $(A^\dagger A)^\top = A^\dagger A$.

Note that, given any $A \in \mathbb{R}^{n \times m}$, there exists a unique Moore–Penrose pseudoinverse $A^\dagger \in \mathbb{R}^{m \times n}$. The following two properties hold.

- (i) If A has full column rank m , then A^\dagger is a left-inverse of A , that is $A^\dagger A = \text{Id}_m$. In this case,

$$A^\dagger = (A^\top A)^{-1} A^\top. \quad (2.3)$$

- (ii) If A has full row rank n , then A^\dagger is a right-inverse of A , that is $AA^\dagger = \text{Id}_n$. In this case, $A^\dagger = A^\top (AA^\top)^{-1}$.

Before introducing the GSH, it is valuable to recall the definition of the generalized simplex gradient.

DEFINITION 2.5 (Generalized simplex gradient). (*Hare & Jarry-Bolduc, 2020, Definition 2*) Let $f : \text{dom} f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and let $x^0 \in \text{dom} f$ be the point of interest. Let $S \in \mathbb{R}^{n \times m}$ with $x^0 \oplus S \subset \text{dom} f$. The *generalized simplex gradient* of f at x^0 over S is denoted by $\nabla_S f(x^0; S)$ and defined by

$$\nabla_S f(x^0; S) = (S^\top)^\dagger \delta_S(x^0; S).$$

The Moore–Penrose pseudoinverse allows for S to contain any finite number of columns, rather than being restricted to $m = n$. When $S \in \mathbb{R}^{n \times n}$ and $\text{rank } S = n$, then the Moore–Penrose pseudoinverse of S^\top is the inverse of S^\top and we recover the definition of the simplex gradient (*Audet & Hare, 2017, Definition 9.5*). Note that the generalized simplex gradient is a generalization of the familiar *forward-finite-difference approximation* of the gradient (*Burden et al., 2016, Section 4.1*). Indeed, if $S = H \text{Id}_n$, where $H = \text{Diag}[h_1 \ h_2 \ \dots \ h_n]$, $h_i > 0$, then we recover the forward-finite-difference approximation of the gradient.

We conclude this section by defining *order- N accuracy* as it is used in this paper several times to describe the quality of the Hessian approximation techniques developed.

DEFINITION 2.6 (Order- N accuracy). Let $P \in \mathbb{R}^{n \times n}$. Given $f \in \mathcal{C}^2$, $x^0 \in \text{dom} f$ and $\bar{\Delta} > 0$, we say that $\{\tilde{f}_\Delta\}_{\Delta \in (0, \bar{\Delta}]}$ is a class of models of f at x^0 parametrized by Δ that provides order- N accuracy of P at x^0 if there exists a scalar $\kappa(x^0)$ such that, given any $\Delta \in (0, \bar{\Delta}]$, the model \tilde{f}_Δ satisfies

$$\|P - \nabla^2 \tilde{f}_\Delta(x^0)\| \leq \kappa(x^0) \Delta^N.$$

3. The GSH and the GCSH

In this section, we begin by defining the GSH and the GCSH. The relation between the GSH and the GCSH is investigated afterwards.

The GSH and the GCSH require finite sets of directions and a point of interest at which the Hessian is approximated. Similarly to the generalized simplex gradient, the GSH involves the Moore–Penrose pseudoinverse and a difference matrix. In this case, the difference matrix consists of the difference between generalized simplex gradients.

For ease of notation, we use the definition $T_{1:m} = \{T_1, \dots, T_m\}$ where convenient.

DEFINITION 3.1 (Generalized simplex Hessian). Let $f : \text{dom} f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and let $x^0 \in \text{dom} f$ be the point of interest. Let $S = [s^1 \ s^2 \ \dots \ s^m] \in \mathbb{R}^{n \times m}$ and $T_i \in \mathbb{R}^{n \times k_i}$ with $x^0 \oplus T_i, x^0 \oplus S, x^0 + s^i \oplus T_i$ contained in $\text{dom} f$ for all $i \in \{1, \dots, m\}$. The *generalized simplex Hessian* of f at x^0 over S and $T_{1:m}$ is

denoted by $\nabla_s^2 f(x^0; S, T_{1:m})$ and defined by

$$\nabla_s^2 f(x^0; S, T_{1:m}) = (S^\top)^\dagger \delta_s^2(x^0; S, T_{1:m}), \quad (3.1)$$

where

$$\delta_s^2(x^0; S, T_{1:m}) = \begin{bmatrix} (\nabla_s f(x^0 + s^1; T_1) - \nabla_s f(x^0; T_1))^\top \\ (\nabla_s f(x^0 + s^2; T_2) - \nabla_s f(x^0; T_2))^\top \\ \vdots \\ (\nabla_s f(x^0 + s^m; T_m) - \nabla_s f(x^0; T_m))^\top \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

In the case $T_1 = T_2 = \dots = T_m$, we use $\bar{T} := T_i$ to simplify notation and write $\nabla_s^2 f(x^0; S; \bar{T})$ to emphasize the special case.

We remark that the subindex s in $\nabla_s^2 f$, δ_s and δ_s^2 stands for simplex. This becomes important as we now introduce the generalized centered simplex Hessian. The ‘square’ in δ_s^2 emphasizes the fact that we are taking differences of first-order objects (hence obtaining a second-order object).

DEFINITION 3.2 (Generalized centered simplex Hessian). Let $f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and let $x^0 \in \text{dom } f$ be the point of interest. Let $S \in \mathbb{R}^{n \times m}$ and $T_i \in \mathbb{R}^{n \times k_i}$ with $x^0 \oplus S \oplus T_i, x^0 \oplus (-S) \oplus (-T_i), x^0 \oplus (\pm S)$, and $x^0 \oplus (\pm T_i)$ contained in $\text{dom } f$ for all $i \in \{1, \dots, m\}$. The generalized centered simplex Hessian of f at x^0 over S and $T_{1:m}$ is denoted by $\nabla_c^2 f(x^0; S, T_{1:m})$ and defined by

$$\nabla_c^2 f(x^0; S, T_{1:m}) = \frac{1}{2} \left(\nabla_s^2 f(x^0; S, T_{1:m}) + \nabla_s^2 f(x^0; -S, -T_{1:m}) \right). \quad (3.2)$$

Note that the subindex c in $\nabla_c^2 f$ stands for centered.

The GCSH can be viewed as a generalization of the centered-finite difference approximation of the Hessian, also called midpoint approximation (Burden *et al.*, 2016, Section 4.1). Similarly, the GSH can be viewed as a generalization of the forward-finite-difference approximation of the Hessian (Burden *et al.*, 2016, Section 4.1). Indeed, when $S = T_1 = \dots = T_m = H \text{Id}_n$, where $H = \text{Diag}[h_1 \ h_2 \ \dots \ h_n], h_i > 0$, then the GSH (GCSH) is equivalent to a forward-finite-difference approximation (centered-finite-difference approximation) of the second-order derivatives. Moreover, the GSH can be viewed as the DFO analog of the Hessian approximation given by Taylor’s Theorem in the derivative-based context, which approximates the gradient by differences of true gradients (see Nocedal & Wright, 2006, Section 8.1). An advantage of the GSH (GCSH) over forward-finite-difference approximations (centered-finite-difference approximations) is that it is less restrictive. In particular, it is well-defined as long as the matrices S and T_1, \dots, T_m are nonempty.

REMARK 3.3 Consider the case where S and T_i are square matrices, i.e., $m = n$ and $k_i = n$ for all $i \in \{1, \dots, m\}$. The computation of the GSH begins by the construction of $n + 1$ generalized simplex gradients. Setting aside function evaluations, computing a simplex gradient is dominated by computing a matrix inverse, which is $O(n^3)$ flops. Thus, the construction of $\delta_s^2(x^0; S; T_{1:m})$ requires $O(n^4)$ flops. The inverse of S is computed ($O(n^3)$ flops) and multiplied by $\delta_s^2(x^0; S; T_{1:m})$, $O(n^3)$ flops. Hence, setting aside function evaluations, computing the GSH requires $O(n^4)$ flops.

Conversely, setting aside function evaluations, quadratic interpolation requires $O(n^6)$ flops to compute.

Consider the situation where all of the matrices S and T_i are non-null rank. The possible cases for constructing the GSH (GCSH) can be sorted into 4 categories.

- Underdetermined: the GSH (GCSH) is said to be S -underdetermined if S is nonsquare and full column rank. We say that it is $T_{1:m}$ -underdetermined if all matrices in the set $T_{1:m}$ are full column rank and at least one matrix is nonsquare.
- Determined: the GSH (GCSH) is said to be S -determined if S is square and full rank. It is $T_{1:m}$ -determined if all matrices in the set $T_{1:m}$ are square and full rank.
- Overdetermined: the GSH (GCSH) is said to be S -overdetermined if S is nonsquare and full row rank. It is $T_{1:m}$ -overdetermined if all matrices in the set $T_{1:m}$ are full row rank at least one is nonsquare.
- Nondetermined: the GSH (GCSH) is said to be S -nondetermined if it is not in any of the previous three cases. It is $T_{1:m}$ -nondetermined if the set $T_{1:m}$ is not in any of the previous three cases.

Note that S -underdetermined GSH (GCSH) implies that $\text{span } S \neq \mathbb{R}^n$, which is true if and only if $SS^\dagger \neq \text{Id}_n$. Similarly, $T_{1:m}$ -underdetermined GSH (GCSH) implies that $\text{span } T_{\bar{i}} \neq \mathbb{R}^n$ for some $\bar{i} \in \{1, \dots, m\}$, which is true if and only if $T_{\bar{i}}T_{\bar{i}}^\dagger \neq \text{Id}_n$ for \bar{i} . The language above creates 16 different cases to classify a GSH (GCSH), all of which are investigated in Section 4.

Since the GSH and the GCSH involve generalized simplex gradients, we present a bound for the absolute error between the generalized simplex gradient $\nabla_s f(x^0; S)$ and the projection onto S of the true gradient $\nabla f(x^0)$. Note that the assumption that $f \in \mathcal{C}^2$ on $B(x^0, \bar{\Delta})$ implies that ∇f is Lipschitz continuous on $\bar{B}(x^0, \Delta)$ where $\Delta < \bar{\Delta}$.

We remark that results akin to Proposition 3.4 have been proven several times using a variety of different techniques (Custódio & Vicente, 2007; Regis, 2015; Hare & Jarry-Bolduc, 2020). In Conn *et al.* (2008a,b) (see also Custódio & Vicente, 2007, and Conn *et al.*, 2009, Theorem 3.1), simplex gradients were analyzed using Lagrange interpolating polynomials. In Regis (2015) (see also Audet & Hare, 2017; Hare & Jarry-Bolduc, 2020), error bounds for simplex gradients were analyzed using matrix condition numbers. In Proposition 3.4, we present the error bound in this second format, as further results will examine GSHs in the same style.

PROPOSITION 3.4 (Error bound for the generalized simplex gradient). Let $S = [s^1 \ s^2 \ \dots \ s^m] \in \mathbb{R}^{n \times m}$ with radius $\Delta_S \geq 0$. Let $\hat{S} = S/\Delta_S$ and let $f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ be \mathcal{C}^2 on $B(x^0, \bar{\Delta})$ where x^0 is the point of interest and $\bar{\Delta} > \Delta_S$. Denote by $L_{\nabla f} \geq 0$ the Lipschitz constant of ∇f on $\bar{B}(x^0, \Delta_S)$. Then

$$\begin{aligned} \left\| \text{Proj}_S \nabla_s f(x^0; S) - \text{Proj}_S \nabla f(x^0) \right\| &= \left\| \nabla_s f(x^0; S) - \text{Proj}_S \nabla f(x^0) \right\| \\ &\leq \frac{\sqrt{m}}{2} L_{\nabla f} \left\| (\hat{S}^\top)^\dagger \right\| \Delta_S, \end{aligned}$$

where $\text{Proj}_S(\cdot) = (S^\top)^\dagger S^\top(\cdot)$.

The error bound presented in Proposition 3.4 only requires that S has non-null rank. Therefore, it covers all possible cases: S -underdetermined, S -determined, S -overdetermined and S -nondetermined.

In Section 4, error bounds with a similar format than the one in Proposition 3.4 are developed for the GSH and the GCSH. To do this, we next define the projection onto a pair of subspaces.

Given matrices $S \in \mathbb{R}^{n \times m}$ and $T_i \in \mathbb{R}^{n \times k_i}$, the projection of the matrix $H \in \mathbb{R}^{n \times n}$ onto S and $T_{1:m}$ is denoted by $\text{Proj}_{S, T_{1:m}} H$ and defined by

$$\text{Proj}_{S, T_{1:m}} H = \sum_{i=1}^m (S^\top)^\dagger e_m^i (e_m^i)^\top S^\top H T_i T_i^\dagger.$$

In the case where $T_1 = T_2 = \dots = T_m =: \bar{T}$, the projection of H onto S and \bar{T} is denoted by $\text{Proj}_{S, \bar{T}} H$, and reduces to

$$\begin{aligned} \text{Proj}_{S, \bar{T}} H &= \sum_{i=1}^m (S^\top)^\dagger e_m^i (e_m^i)^\top S^\top H \bar{T} \bar{T}^\dagger \\ &= (S^\top)^\dagger \left(\sum_{i=1}^m e_m^i (e_m^i)^\top \right) S^\top H \bar{T} \bar{T}^\dagger \\ &= (S^\top)^\dagger \text{Id}_m S^\top H \bar{T} \bar{T}^\dagger \\ &= (S^\top)^\dagger S^\top H \bar{T} \bar{T}^\dagger. \end{aligned}$$

When defining error bounds in Section 4, the matrix P in Definition 2.6 will be equal to $\text{Proj}_{S, T_{1:m}} \nabla^2 f(x^0)$. We may refer to $\text{Proj}_{S, T_{1:m}} \nabla^2 f(x^0)$ as a *partial Hessian*. Note that when S is full row rank and T_i is full row rank for all i , then $\text{Proj}_{S, T_{1:m}} \nabla^2 f(x^0)$ is equal to the true Hessian $\nabla^2 f(x^0)$. We will see that we can obtain order-1 or order-2 accuracy of the partial Hessian $\text{Proj}_{S, T_{1:m}} \nabla_s^2 f(x^0; S, T_{1:m})$ under the appropriate assumptions.

Next a formula to compute the GCSH in a format similar to the formula for the GSH (3.1) is provided. In Lemma 3.5, we introduce the notation δ_c^2 . This is a centered version of δ_s^2 ; hence, the s for simplex is replaced with c for centered.

LEMMA 3.5 Let $f: \text{dom} f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and let $x^0 \in \text{dom} f$ be the point of interest. Let $S = [s^1 \ \dots \ s^m] \in \mathbb{R}^{n \times m}$ and $T_i \in \mathbb{R}^{n \times k_i}$ with $x^0 \oplus S \oplus T_i, x^0 \oplus (-S) \oplus (-T_i), x^0 \oplus (\pm S), x^0 \oplus (\pm T_i)$ contained in $\text{dom} f$ for all $i \in \{1, \dots, m\}$. Then

$$\nabla_c^2 f(x^0; S, T_{1:m}) = (S^\top)^\dagger \delta_c^2(x^0; S, T_{1:m}),$$

where

$$\delta_c^2(x^0; S, T_{1:m}) = \frac{1}{2} \begin{bmatrix} (\delta_s(x^0 + s^1; T_1) + \delta_s(x^0 - s^1; -T_1) - \delta_s(x^0; T_1) - \delta_s(x^0; -T_1))^\top T_1^\dagger \\ \vdots \\ (\delta_s(x^0 + s^m; T_m) + \delta_s(x^0 - s^m; -T_m) - \delta_s(x^0; T_m) - \delta_s(x^0; -T_m))^\top T_m^\dagger \end{bmatrix}.$$

Proof. We have

$$\begin{aligned}
& \nabla_c^2 f(x^0; S, T_{1:m}) \\
&= \frac{1}{2} \left(\nabla_{Sf}^2(x^0; S, T_{1:m}) + \nabla_{Sf}^2(x^0; -S, -T_{1:m}) \right) \\
&= \frac{1}{2} (S^\top)^\dagger \left(\delta_s^2(x^0; S, T_{1:m}) - \delta_s^2(x^0; -S, -T_{1:m}) \right) \\
&= (S^\top)^\dagger \frac{1}{2} \begin{bmatrix} (\nabla_{Sf}(x^0 + s^1; T_1) - \nabla_{Sf}(x^0; T_1))^\top - (\nabla_{Sf}(x^0 - s^1; -T_1) - \nabla_{Sf}(x^0; -T_1))^\top \\ \vdots \\ (\nabla_{Sf}(x^0 + s^m; T_m) - \nabla_{Sf}(x^0; T_m))^\top - (\nabla_{Sf}(x^0 - s^m; -T_m) - \nabla_{Sf}(x^0; -T_m))^\top \end{bmatrix} \\
&= (S^\top)^\dagger \frac{1}{2} \begin{bmatrix} (\delta_s(x^0 + s^1; T_1) - \delta_s(x^0; T_1))^\top T_1^\dagger - ((\delta_s(x^0 - s^1; -T_1) - \delta_s(x^0; -T_1))^\top (-T_1)^\dagger) \\ \vdots \\ (\delta_s(x^0 + s^m; T_m) - \delta_s(x^0; T_m))^\top T_m^\dagger - ((\delta_s(x^0 - s^m; -T_m) - \delta_s(x^0; -T_m))^\top (-T_m)^\dagger) \end{bmatrix} \\
&= (S^\top)^\dagger \frac{1}{2} \begin{bmatrix} (\delta_s(x^0 + s^1; T_1) + \delta_s(x^0 - s^1; -T_1) - \delta_s(x^0; T_1) - \delta_s(x^0; -T_1))^\top T_1^\dagger \\ \vdots \\ (\delta_s(x^0 + s^m; T_m) + \delta_s(x^0 - s^m; -T_m) - \delta_s(x^0; T_m) - \delta_s(x^0; -T_m))^\top T_m^\dagger \end{bmatrix} \\
&= (S^\top)^\dagger \delta_c^2(x^0; S, T_{1:m}).
\end{aligned}$$

□

Notice that

$$\begin{aligned}
& \left[\delta_s(x^0 + s^i; T_i) + \delta_s(x^0 - s^i; -T_i) - \delta_s(x^0; T_i) - \delta_s(x^0; -T_i) \right]_{j_i} \\
&= \frac{1}{2} \left(f(x^0 + s^i + t_i^{j_i}) - f(x^0 + s^i) + f(x^0 - s^i - t_i^{j_i}) - f(x^0 - s^i) \right. \\
&\quad \left. - f(x^0 + t_i^{j_i}) + f(x^0) - f(x^0 - t_i^{j_i}) + f(x^0) \right) \\
&= \frac{1}{2} \left(f(x^0 + s^i + t_i^{j_i}) + f(x^0 - s^i - t_i^{j_i}) \right. \\
&\quad \left. - f(x^0 + s^i) - f(x^0 - s^i) - f(x^0 + t_i^{j_i}) - f(x^0 - t_i^{j_i}) + 2f(x^0) \right) \tag{3.3}
\end{aligned}$$

for $j_i \in \{1, \dots, k_i\}$ and $i \in \{1, \dots, m\}$. We will use (3.3) in Theorem 4.3.

Next, we show that the GCSH is actually a special case of the GSH where the matrices of directions have a specific form. This result is similar to the result in Hare *et al.* (2020a, Proposition 2.10), which showed that the generalized centered gradient is equal to the generalized simplex gradient when the matrix of directions used to compute the generalized simplex gradient has the form $A = [S \ -S]$ for some matrix $S \in \mathbb{R}^{n \times m}$. We shall use the following lemma.

LEMMA 3.6 Let $A = \begin{bmatrix} S & -S \end{bmatrix} \in \mathbb{R}^{n \times 2m}$ for some matrix $S \in \mathbb{R}^{n \times m}$. Then $A^\dagger = \frac{1}{2} \begin{bmatrix} S^\dagger \\ -S^\dagger \end{bmatrix}$.

Proof. This follows from basic algebra. □

PROPOSITION 3.7 Let $f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and let $x^0 \in \text{dom } f$ be the point of interest. Let $A = \begin{bmatrix} a^1 & a^2 & \dots & a^{2m} \end{bmatrix} = \begin{bmatrix} S & -S \end{bmatrix} \in \mathbb{R}^{n \times 2m}$ for some $S = \begin{bmatrix} s^1 & \dots & s^m \end{bmatrix} \in \mathbb{R}^{n \times m}$ and let $B_i = T_i \in \mathbb{R}^{n \times k_i}$, and $B_{m+i} = -T_i$ for all $i \in \{1, \dots, m\}$. Suppose that $x^0 \oplus S \oplus T_i, x^0 \oplus (-S) \oplus (-T_i), x^0 \oplus (\pm S), x^0 \oplus (\pm T_i)$ are contained in $\text{dom } f$ for all $i \in \{1, \dots, m\}$. Then

$$\nabla_s^2 f(x^0; A, B_{1:2m}) = \nabla_c^2 f(x^0; S, T_{1:m}).$$

Proof. We have

$$\begin{aligned} & \nabla_s^2 f(x^0; A, B_{1:2m}) \\ &= \begin{bmatrix} S^\top \\ -S^\top \end{bmatrix}^\dagger \begin{bmatrix} (\nabla_s f(x^0 + a^1; B_1) - \nabla_s f(x^0; B_1))^\top \\ \vdots \\ (\nabla_s f(x^0 + a^{2m}; B_{2m}) - \nabla_s f(x^0; B_{2m}))^\top \end{bmatrix} \\ &= \frac{1}{2} [(S^\top)^\dagger \quad -(S^\top)^\dagger] \begin{bmatrix} (\nabla_s f(x^0 + a^1; B_1) - \nabla_s f(x^0; B_1))^\top \\ \vdots \\ (\nabla_s f(x^0 + a^{2m}; B_{2m}) - \nabla_s f(x^0; B_{2m}))^\top \end{bmatrix} \\ &= \frac{1}{2} (S^\top)^\dagger \begin{bmatrix} (\nabla_s f(x^0 + s^1; T_1) - \nabla_s f(x^0; T_1))^\top \\ \vdots \\ (\nabla_s f(x^0 + s^m; T_m) - \nabla_s f(x^0; T_m))^\top \end{bmatrix} \\ &\quad + \frac{1}{2} (-S^\top)^\dagger \begin{bmatrix} (\nabla_s f(x^0 - s^1; -T_1) - \nabla_s f(x^0; -T_1))^\top \\ \vdots \\ (\nabla_s f(x^0 - s^m; -T_m) - \nabla_s f(x^0; -T_m))^\top \end{bmatrix} \\ &= \frac{1}{2} \left(\nabla_s^2 f(x^0; S, T_{1:m}) + \nabla_s^2 f(x^0; -S, -T_{1:m}) \right) \\ &= \nabla_c^2 f(x^0; S, T_{1:m}). \end{aligned}$$

□

We are now ready to examine the error bounds for the GSH and the GCSH.

4. Error bounds

In this section, we provide error bounds for the GSH and the GCSH. These error bounds are separated in two categories: the general case where all matrices T_i are not necessarily equal, and the special case where

all matrices $T_i =: \bar{T}$ are equal. The error bounds for the GSH show an order-1 accurate approximation of the full Hessian $\nabla^2 f(x^0)$, or an order-1 accurate approximation of the partial Hessian $\text{Proj}_{S, T_{1:m}} \nabla^2 f(x^0)$ if the GSH is S - or $T_{1:m}$ -underdetermined. Similarly, the error bounds for the GCSH show an order-2 accurate approximation of the full Hessian $\nabla^2 f(x^0)$, or an order-2 accurate approximation of the partial Hessian $\text{Proj}_{S, T_{1:m}} \nabla^2 f(x^0)$ if the GCSH is S - or $T_{1:m}$ -underdetermined.

Before introducing the error bounds, we note that in certain situations, the projection of the GSH (GCSH) onto S and $T_{1:m}$ is equal to the GSH (GCSH).

PROPOSITION 4.1 Let $f : \text{dom} f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and $x^0 \in \text{dom} f$. Let $S = [s^1 \ \cdots \ s^m] \in \mathbb{R}^{n \times m}$ and $T_i \in \mathbb{R}^{n \times k_i}$. Assume that $x^0 \oplus (\pm S)$, $x^0 \oplus (\pm T_i)$, $x^0 \oplus S \oplus T_i$ and $x^0 \oplus -S \oplus -T_i$ are contained in $\text{dom} f$ for all i . Then the following hold.

1. If S is full column rank **or** T_i is full row rank for all $i \in \{1, \dots, m\}$, then

$$\text{Proj}_{S, T_{1:m}} \nabla_s^2 f(x^0; S, T_{1:m}) = \nabla_s^2 f(x^0; S, T_{1:m}) \quad (4.1)$$

and

$$\text{Proj}_{S, T_{1:m}} \nabla_c^2 f(x^0; S, T_{1:m}) = \nabla_c^2 f(x^0; S, T_{1:m}). \quad (4.2)$$

2. If $T_1 = T_2 = \cdots = T_m =: \bar{T}$, then

$$\text{Proj}_{S, \bar{T}} \nabla_s^2 f(x^0; S, \bar{T}) = \nabla_s^2 f(x^0; S, \bar{T}) \quad (4.3)$$

and

$$\text{Proj}_{S, \bar{T}} \nabla_c^2 f(x^0; S, \bar{T}) = \nabla_c^2 f(x^0; S, \bar{T}). \quad (4.4)$$

Proof. Suppose S is full column rank. Since $T_i^\dagger T_i T_i^\dagger = T_i^\dagger$ by Property (ii) of the Moore–Penrose pseudoinverse, we obtain

$$\begin{aligned} \text{Proj}_{S, T_{1:m}} \nabla_s^2 f(x^0; S, T_{1:m}) &= \sum_{i=1}^m (S^\top)^\dagger e_m^i (e_m^i)^\top S^\top \nabla_s^2 f(x^0; S, T_{1:m}) T_i T_i^\dagger \\ &= \sum_{i=1}^m (S^\top)^\dagger e_m^i (e_m^i)^\top S^\top (S^\top)^\dagger \delta_s^2(x^0; S, T_{1:m}) T_i T_i^\dagger \\ &= \sum_{i=1}^m (S^\top)^\dagger e_m^i (e_m^i)^\top \text{Id}_m \delta_s^2(x^0; S, T_{1:m}) T_i T_i^\dagger \\ &= (S^\top)^\dagger \sum_{i=1}^m e_m^i (e_m^i)^\top \delta_s^2(x^0; S, T_{1:m}) T_i T_i^\dagger \end{aligned}$$

$$\begin{aligned}
&= (S^\top)^\dagger \sum_{i=1}^m \text{Diag} \left(e_m^i \right) \delta_s^2 \left(x^0; S, T_{1:m} \right) T_i T_i^\dagger \\
&= (S^\top)^\dagger \delta_s^2 \left(x^0; S, T_{1:m} \right) \\
&= \nabla_s^2 f \left(x^0; S, T_{1:m} \right).
\end{aligned}$$

Equation (4.2) follows analogously.

Now suppose that T_i is full row rank for all $i \in \{1, \dots, m\}$. Then we know $T_i T_i^\dagger = \text{Id}_n$ for all $i \in \{1, \dots, m\}$. We obtain

$$\begin{aligned}
\text{Proj}_{S, T_{1:m}} \nabla_s^2 f \left(x^0; S, T_{1:m} \right) &= \sum_{i=1}^m (S^\top)^\dagger e_m^i \left(e_m^i \right)^\top S^\top \nabla_s^2 f \left(x^0; S, T_{1:m} \right) T_i T_i^\dagger \\
&= \sum_{i=1}^m (S^\top)^\dagger e_m^i \left(e_m^i \right)^\top S^\top \nabla_s^2 f \left(x^0; S, T_{1:m} \right) \\
&= (S^\top)^\dagger \left(\sum_{i=1}^m e_m^i \left(e_m^i \right)^\top \right) S^\top \nabla_s^2 f \left(x^0; S, T_{1:m} \right) \\
&= (S^\top)^\dagger S^\top (S^\top)^\dagger \delta_s^2 \left(x^0; S, T_{1:m} \right) \\
&= (S^\top)^\dagger \delta_s^2 \left(x^0; S, T_{1:m} \right) = \nabla_s^2 f \left(x^0; S, T_{1:m} \right).
\end{aligned}$$

Equation (4.2) follows analogously.

Finally, suppose that $T_1 = T_2 = \dots = T_m =: \bar{T}$. We have

$$\begin{aligned}
\text{Proj}_{S, \bar{T}} \nabla_s^2 f \left(x^0; S, \bar{T} \right) &= (S^\top)^\dagger S^\top \nabla_s^2 f \left(x^0; S, \bar{T} \right) \bar{T} \bar{T}^\dagger \\
&= (S^\top)^\dagger S^\top (S^\top)^\dagger \delta_s^2 \left(x^0; S, \bar{T} \right) \bar{T} \bar{T}^\dagger \\
&= (S^\top)^\dagger S^\top (S^\top)^\dagger \begin{bmatrix} \left(\nabla_s f(x^0 + s^1; \bar{T}) - \nabla_s f(x^0; \bar{T}) \right)^\top \\ \vdots \\ \left(\nabla_s f(x^0 + s^m; \bar{T}) - \nabla_s f(x^0; \bar{T}) \right)^\top \end{bmatrix} \bar{T} \bar{T}^\dagger \\
&= (S^\top)^\dagger S^\top (S^\top)^\dagger \begin{bmatrix} \left((\bar{T}^\top)^\dagger \delta_s(x^0 + s^1; \bar{T}) - (\bar{T}^\top)^\dagger \delta_s(x^0; \bar{T}) \right)^\top \\ \vdots \\ \left((\bar{T}^\top)^\dagger \delta_s(x^0 + s^m; \bar{T}) - (\bar{T}^\top)^\dagger \delta_s(x^0; \bar{T}) \right)^\top \end{bmatrix} \bar{T} \bar{T}^\dagger
\end{aligned}$$

$$\begin{aligned}
&= (S^\top)^\dagger S^\top (S^\top)^\dagger \begin{bmatrix} (\delta_s(x^0 + s^1; \bar{T}) - \delta_s(x^0; \bar{T}))^\top \\ \vdots \\ (\delta_s(x^0 + s^m; \bar{T}) - \delta_s(x^0; \bar{T}))^\top \end{bmatrix} \bar{T}^\dagger \bar{T} \bar{T}^\dagger \\
&= (S^\top)^\dagger \begin{bmatrix} (\delta_s(x^0 + s^1; \bar{T}) - \delta_s(x^0; \bar{T}))^\top \\ \vdots \\ (\delta_s(x^0 + s^m; \bar{T}) - \delta_s(x^0; \bar{T}))^\top \end{bmatrix} \bar{T}^\dagger
\end{aligned}$$

by Property (ii) of the Moore–Penrose pseudoinverse. Thus, $\text{Proj}_{S, \bar{T}} \nabla_s^2 f(x^0; S, \bar{T}) = \nabla_s^2 f(x^0; S, \bar{T})$. Equation (4.4) follows analogously. \square

In the following theorem and the remainder of this paper, we use the notation

$$\begin{aligned}
\hat{T}_i &= T_i / \Delta_{T_i}, \quad i \in \{1, \dots, m\}, \\
\Delta_u &= \max \{ \Delta_S, \Delta_{T_1}, \dots, \Delta_{T_m} \}, \\
\Delta_l &= \min \{ \Delta_S, \Delta_{T_1}, \dots, \Delta_{T_m} \}, \\
\Delta_T &= \max \{ \Delta_{T_1}, \dots, \Delta_{T_m} \} \\
\hat{T} &= \hat{T}_i \quad \text{such that} \quad \|\hat{T}_i^\dagger\| \quad \text{is maximal,} \quad i \in \{1, \dots, m\}, \\
k &= \max \{ k_1, \dots, k_m \}, \\
H &= \nabla^2 f(x^0).
\end{aligned}$$

THEOREM 4.2 (Error bounds for the GSH). Let $f : \text{dom} f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ be \mathcal{C}^3 on $B(x^0, \bar{\Delta})$ where $x^0 \in \text{dom} f$ is the point of interest and $\bar{\Delta} > 0$. Denote by $L_{\nabla^2 f} \geq 0$ the Lipschitz constant of $\nabla^2 f$ on $\bar{B}(x^0, \bar{\Delta})$, respectively. Let $S = [s^1 \ s^2 \ \dots \ s^m] \in \mathbb{R}^{n \times m}$ and $T_i = [t_i^1 \ t_i^2 \ \dots \ t_i^{k_i}] \in \mathbb{R}^{n \times k_i}$ for all $i \in \{1, \dots, m\}$. Assume that $B(x^0, \Delta_{T_i}) \subset B(x^0, \bar{\Delta})$ and $B(x^0 + s^i, \Delta_{T_i}) \subset B(x^0, \bar{\Delta})$ for all $i \in \{1, \dots, m\}$.

(i) If S is full column rank or T_i is full row rank for all $i \in \{1, \dots, m\}$, then

$$\begin{aligned}
&\left\| \text{Proj}_{S, T_{1:m}} \nabla_s^2 f(x^0; S, T_{1:m}) - \text{Proj}_{S, T_{1:m}} \nabla^2 f(x^0) \right\| \\
&= \left\| \nabla_s^2 f(x^0; S, T_{1:m}) - \text{Proj}_{S, T_{1:m}} \nabla^2 f(x^0) \right\| \\
&\leq 4m\sqrt{k}L_{\nabla^2 f} \|(\hat{S}^\top)^\dagger\| \|\hat{T}^\dagger\| \left(\frac{\Delta_u}{\Delta_l} \right)^2 \Delta_u.
\end{aligned} \tag{4.5}$$

(ii) If $T_1 = T_2 = \dots = T_m =: \bar{T}$, then

$$\begin{aligned} \left\| \text{Proj}_{S, \bar{T}} \nabla_s^2 f(x^0; S, \bar{T}) - \text{Proj}_{S, \bar{T}} \nabla^2 f(x^0) \right\| &= \left\| \nabla_s^2 f(x^0; S, \bar{T}) - \text{Proj}_{S, \bar{T}} \nabla^2 f(x^0) \right\| \\ &\leq 4\sqrt{mk} L_{\nabla^2 f} \frac{\Delta_u}{\Delta_l} \left\| (\widehat{S}^\top)^\dagger \right\| \left\| \widehat{T}^\dagger \right\| \Delta_u. \end{aligned} \quad (4.6)$$

Proof. The equalities in (i) and (ii) follow from (4.1) and (4.3), respectively. We have

$$\begin{aligned} &\left\| \nabla_s^2 f(x^0; S, T_{1:m}) - \text{Proj}_{S, T_{1:m}} H \right\| \\ &= \left\| (S^\top)^\dagger \delta_s^2(x^0; S, T_{1:m}) - \sum_{i=1}^m (S^\top)^\dagger e_m^i (e_m^i)^\top S^\top H T_i T_i^\dagger \right\| \\ &\leq \frac{1}{\Delta_S} \left\| (\widehat{S}^\top)^\dagger \right\| \left\| \delta_s^2(x^0; S, T_{1:m}) - \sum_{i=1}^m (s^i)^\top H T_i T_i^\dagger \right\| \\ &= \frac{1}{\Delta_S} \left\| (\widehat{S}^\top)^\dagger \right\| \left\| \sum_{i=1}^m (e_m^i)^\top \delta_s^2(x^0; S, T_{1:m}) - (s^i)^\top H T_i T_i^\dagger \right\| \\ &\leq \sum_{i=1}^m \frac{1}{\Delta_S \Delta_{T_i}} \left\| (\widehat{S}^\top)^\dagger \right\| \left\| \widehat{T}_i^\dagger \right\| \left\| \left(\delta_s(x^0 + s^i; T_i) - \delta_s(x^0; T_i) \right)^\top - (s^i)^\top H T_i \right\|. \end{aligned} \quad (4.7)$$

Let $u_i = (\delta_s(x^0 + s^i; T_i) - \delta_s(x^0; T_i))^\top \in \mathbb{R}^{k_i}$ for all $\{1, \dots, m\}$. Next we find a bound for each $[u_i - (s^i)^\top H T_i]_{j_i}$, $j_i \in \{1, \dots, k_i\}$. Note that

$$\left[u_i - (s^i)^\top H T_i \right]_{j_i} = f(x^0 + s^i + t_i^{j_i}) - f(x^0 + s^i) - f(x^0 + t_i^{j_i}) + f(x^0) - (s^i)^\top H t_i^{j_i}.$$

Each function value $f(x^0 + s^i + t_i^{j_i})$, $f(x^0 + s^i)$ and $f(x^0 + t_i^{j_i})$ can be written as a second-order Taylor expansion about x^0 plus a remainder term $R_2(\cdot)$. Hence, we obtain

$$\begin{aligned} &\left| f(x^0 + s^i + t_i^{j_i}) - f(x^0 + s^i) - f(x^0 + t_i^{j_i}) + f(x^0) - (s^i)^\top H t_i^{j_i} \right| \\ &\leq \left| R_2(x^0 + s^i + t_i^{j_i}) \right| + |R_2(x^0 + s^i)| + \left| R_2(x^0 + t_i^{j_i}) \right| \\ &\leq \frac{1}{6} L_{\nabla^2 f} \|s^i + t_i^{j_i}\|^3 + \frac{1}{6} L_{\nabla^2 f} \|s^i\|^3 + \frac{1}{6} L_{\nabla^2 f} \|t_i^{j_i}\|^3 \\ &\leq \frac{1}{2} L_{\nabla^2 f} (\Delta_S + \Delta_{T_i})^3. \end{aligned} \quad (4.8)$$

Substituting (4.8) in (4.7), we obtain

$$\begin{aligned}
 \left\| \nabla_s^2 f(x^0; S, T_{1:m}) - \text{Proj}_{S, T_{1:m}} H \right\| &\leq \sum_{i=1}^m \frac{1}{\Delta_S \Delta_{T_i}} \left\| (\widehat{S}^\top)^\dagger \right\| \left\| \widehat{T}_i^\dagger \right\| \frac{\sqrt{k_i}}{2} L_{\nabla^2 f} (\Delta_S + \Delta_{T_i})^3 \\
 &\leq \sum_{i=1}^m \frac{1}{\Delta_l^2} \left\| (\widehat{S}^\top)^\dagger \right\| \left\| \widehat{T}^\dagger \right\| \frac{\sqrt{k}}{2} L_{\nabla^2 f} (2\Delta_u)^3 \\
 &= 4m\sqrt{k} L_{\nabla^2 f} \left\| (\widehat{S}^\top)^\dagger \right\| \left\| \widehat{T}^\dagger \right\| \left(\frac{\Delta_u}{\Delta_l} \right)^2 \Delta_u.
 \end{aligned}$$

Now, suppose that $T_1 = T_2 = \dots = T_m =: \bar{T}$. We have

$$\begin{aligned}
 \left\| \nabla_s^2 f(x^0; S, \bar{T}) - \text{Proj}_{S, \bar{T}} H \right\| &= \left\| (S^\top)^\dagger \delta_s^2(x^0; S, \bar{T}) - (S^\top)^\dagger S^\top H \bar{T} \bar{T}^\dagger \right\| \\
 &\leq \frac{1}{\Delta_S} \left\| (\widehat{S}^\top)^\dagger \right\| \left\| \delta_s^2(x^0; S, \bar{T}) - S^\top H \bar{T} \bar{T}^\dagger \right\| \\
 &\leq \frac{1}{\Delta_S \Delta_T} \left\| (\widehat{S}^\top)^\dagger \right\| \left\| \widehat{T}^\dagger \right\| \left\| (D_s^2)^\top - S^\top H \bar{T} \right\|, \tag{4.9}
 \end{aligned}$$

where

$$(D_s^2)^\top = \begin{bmatrix} (\delta_s(x^0 + s^1; \bar{T}) - \delta_s(x^0; \bar{T}))^\top \\ \vdots \\ (\delta_s(x^0 + s^m; \bar{T}) - \delta_s(x^0; \bar{T}))^\top \end{bmatrix} \in \mathbb{R}^{m \times k}.$$

Each entry $[(D_s^2)^\top - S^\top H \bar{T}]_{ij}$ has the previous bound obtained in (4.8). Since

$$\left\| (D_s^2)^\top - S^\top H \bar{T} \right\| \leq \left\| (D_s^2)^\top - S^\top H \bar{T} \right\|_F,$$

using (4.8) in (4.9), we get the inequality

$$\begin{aligned}
 \left\| \nabla_s^2 f(x^0; S, \bar{T}) - \text{Proj}_{S, \bar{T}} H \right\| &\leq \frac{\sqrt{mk}}{2} L_{\nabla^2 f} \frac{(\Delta_S + \Delta_T)^3}{\Delta_S \Delta_T} \left\| (\widehat{S}^\top)^\dagger \right\| \left\| \widehat{T}^\dagger \right\| \\
 &\leq 4\sqrt{mk} L_{\nabla^2 f} \frac{\Delta_u}{\Delta_l} \left\| (\widehat{S}^\top)^\dagger \right\| \left\| \widehat{T}^\dagger \right\| \Delta_u.
 \end{aligned}$$

□

The previous error bounds show that the GSH is an order-1 accurate approximation of a partial Hessian, or an order-1 accurate approximation of the full Hessian when the GSH is S -determined/overdetermined and $T_{1:m}$ -determined/overdetermined (equivalently, $(S^\top)^\dagger S^\top = \text{Id}_n$ and $T_i^\dagger T_i = \text{Id}_n$ for all i). The ratio $\frac{\Delta_u}{\Delta_l}$ suggests that the radii of the matrices should be taken to be equal.

If one of the radii is decreased, it suggests that all radii should be decreased by the same amount. Finally, the presence of m and k in the error bounds roughly suggests that there is no advantage of having S -overdetermined or $T_{1:m}$ -overdetermined. More details on this topic can be found in [Braun et al. \(2021\)](#); [Hare et al. \(2023\)](#).

When all matrices T_i are equal, note that (ii) covers all 16 cases. When all matrices T_i are not equal, (i) covers 12 out of the 16 cases. It does not cover the case S -nondetermined/overdetermined and $T_{1:m}$ -nondetermined/underdetermined. These cases are left as a future research direction and discussed in the conclusion.

THEOREM 4.3 (Error bounds for the GCSH). Let $f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ be \mathcal{C}^4 on $B(x^0, \bar{\Delta})$ where $x^0 \in \text{dom } f$ is the point of interest and $\bar{\Delta} > 0$. Denote by $L_{\nabla^3 f}$ the Lipschitz constant of $\nabla^3 f$ on $\bar{B}(x^0, \bar{\Delta})$. Let $S = [s^1 \ s^2 \ \dots \ s^m] \in \mathbb{R}^{n \times m}$, $T_i = [t_i^1 \ t_i^2 \ \dots \ t_i^{k_i}] \in \mathbb{R}^{n \times k_i}$ with $\bar{B}(x^0 \pm s^i, \Delta_{T_i}) \subset B(x^0, \bar{\Delta})$ for all $i \in \{1, \dots, m\}$.

(i) If S is full column rank or T_i is full row rank for all $i \in \{1, \dots, m\}$, then

$$\begin{aligned} & \left\| \text{Proj}_{S, T_{1:m}} \nabla_c^2 f(x^0; S, T_{1:m}) - \text{Proj}_{S, T_{1:m}} \nabla^2 f(x^0) \right\| \\ &= \left\| \nabla_c^2 f(x^0; S, T_{1:m}) - \text{Proj}_{S, T_{1:m}} \nabla^2 f(x^0) \right\| \\ &\leq 2m\sqrt{k}L_{\nabla^3 f} \left(\frac{\Delta_u}{\Delta_l} \right)^2 \left\| (\widehat{S}^\top)^\dagger \right\| \left\| (\widehat{T})^\dagger \right\| \Delta_u^2. \end{aligned} \quad (4.10)$$

(ii) If $T_1 = T_2 = \dots = T_m =: \bar{T} \in \mathbb{R}^{n \times k}$, then

$$\begin{aligned} & \left\| \text{Proj}_{S, \bar{T}} \nabla_c^2 f(x^0; S, \bar{T}) - \text{Proj}_{S, \bar{T}} \nabla^2 f(x^0) \right\| = \left\| \nabla_c^2 f(x^0; S, \bar{T}) - \text{Proj}_{S, \bar{T}} \nabla^2 f(x^0) \right\| \\ &\leq 2\sqrt{mk}L_{\nabla^3 f} \frac{\Delta_u}{\Delta_l} \left\| (\widehat{S}^\top)^\dagger \right\| \left\| (\widehat{T})^\dagger \right\| \Delta_u^2. \end{aligned} \quad (4.11)$$

Proof. The equalities in (i) and (ii) follow from (4.2) and (4.4), respectively.

Similar to the proof of Theorem 4.2, we have

$$\begin{aligned} & \left\| \nabla_c^2 f(x^0; S, T_{1:m}) - \text{Proj}_{S, T_{1:m}} H \right\| \\ &\leq \sum_{i=1}^m \frac{1}{\Delta_S \Delta_{T_i}} \left\| (\widehat{S}^\top)^\dagger \right\| \left\| \widehat{T}_i^\dagger \right\| \left\| \frac{1}{2} (\delta_s(x^0 + s^i; T_i) + \delta_s(x^0 - s^i; -T_i) - \delta_s(x^0; T_i) - \delta_f^s(x^0; -T_i))^\top - (s^i)^\top H T_i \right\|. \end{aligned} \quad (4.12)$$

Using (3.3), we know that

$$\begin{aligned} & \left[\frac{1}{2} \left(\delta_s \left(x^0 + s^i; T_i \right) + \delta_s \left(x^0 - s^i; -T_i \right) - \delta_s \left(x^0; T_i \right) - \delta_s \left(x^0; -T_i \right) \right) - (s^i)^\top H T_i \right]_{j_i} \\ &= \frac{1}{2} \left(f \left(x^0 + s^i + t_i^{j_i} \right) + f \left(x^0 - s^i - t_i^{j_i} \right) - f \left(x^0 + s^i \right) - f \left(x^0 - s^i \right) - f \left(x^0 + t_i^{j_i} \right) - f \left(x^0 - t_i^{j_i} \right) + 2f \left(x^0 \right) \right) \\ & \quad - (s^i)^\top H t_i^{j_i}. \end{aligned} \quad (4.13)$$

Writing each of $f(x^0 + s^i + t_i^{j_i})$, $f(x^0 - s^i - t_i^{j_i})$, $f(x^0 + s^i)$, $f(x^0 - s^i)$, $f(x^0 + t_i^{j_i})$ and $f(x^0 - t_i^{j_i})$ as a third-order Taylor expansion about x^0 plus a remainder term $R_3(\cdot)$, it follows that

$$\begin{aligned} & \left| \frac{1}{2} \left(f \left(x^0 + s^i + t_i^{j_i} \right) + f \left(x^0 - s^i - t_i^{j_i} \right) - f \left(x^0 + s^i \right) - f \left(x^0 - s^i \right) - f \left(x^0 + t_i^{j_i} \right) \right. \right. \\ & \quad \left. \left. - f \left(x^0 - t_i^{j_i} \right) + 2f \left(x^0 \right) \right) - (s^i)^\top H t_i^{j_i} \right| \\ & \leq \frac{1}{8} L_{\nabla^3 f} (\Delta_S + \Delta_{T_i})^4. \end{aligned} \quad (4.14)$$

Substituting the bound from (4.14) in (4.13) leads to

$$\left\| \nabla_c^2 f \left(x^0; S, T_{1:m} \right) - \text{Proj}_{S, T_{1:m}} H \right\| \leq 2m\sqrt{k} L_{\nabla^3 f} \left(\frac{\Delta_u}{\Delta_l} \right)^2 \left\| (\widehat{S}^\top)^\dagger \right\| \left\| (\widehat{T})^\dagger \right\| \Delta_u^2.$$

Similarly, in the case where $T_1 = T_2 = \dots = T_m =: \bar{T}$, we have

$$\left\| \nabla_c^2 f \left(x^0; S, \bar{T} \right) - \text{Proj}_{S, \bar{T}} H \right\| \leq \frac{2\sqrt{mk}}{3} L_{\nabla^3 f} \frac{\Delta_u}{\Delta_l} \left\| (\widehat{S}^\top)^\dagger \right\| \left\| \widehat{T}^\dagger \right\| \Delta_u^2.$$

Full details can be found in [Jarry-Bolduc \(2023\)](#). □

The main differences between the error bounds presented for the GCSH in Theorem 4.3 and the GSH in Theorem 4.2 are

- the error bounds in Theorem 4.3 assume that $f \in \mathcal{C}^4$ and the error bounds in Theorem 4.2 assume $f \in \mathcal{C}^3$,
- the error bounds in Theorem 4.3 involve the Lipschitz constant $L_{\nabla^3 f}$ and the error bounds in Theorem 4.2 involve the Lipschitz constant $L_{\nabla^2 f}$,
- the error bounds in Theorem 4.3 are order-2 accurate and the error bounds in Theorem 4.2 are order-1 accurate.

Item (ii) of Theorems 4.2 and 4.3 is more restrictive than (i), but in return it can offer additional benefits, such as symmetry. In the following proposition, we show that the transpose of the GSH (GCSH) of f at x^0 over S and \bar{T} is the same as the GSH (GCSH) of f at x^0 over \bar{T} and S .

PROPOSITION 4.4 Let $f : \text{dom} f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and let $x^0 \in \text{dom} f$ be the point of interest. Let $S = [s^1 \ s^2 \ \dots \ s^m] \in \mathbb{R}^{n \times m}$ and $\bar{T} = [t^1 \ t^2 \ \dots \ t^k] \in \mathbb{R}^{n \times k}$ with $x^0 \oplus S \oplus T_i, x^0 \oplus (-S) \oplus (-T_i), x^0 \oplus (\pm S)$, and $x^0 \oplus (\pm T_i)$ contained in $\text{dom} f$ for all $i \in \{1, \dots, m\}$. Then

$$\left(\nabla_s^2 f(x^0; S, \bar{T}) \right)^\top = \nabla_s^2 f(x^0; \bar{T}, S),$$

and

$$\left(\nabla_c^2 f(x^0; S, \bar{T}) \right)^\top = \nabla_c^2 f(x^0; \bar{T}, S).$$

Proof. We have

$$\begin{aligned} \left(\nabla_s^2 f(x^0; S, \bar{T}) \right)^\top &= \left((S^\top)^\dagger \delta_s^2(x^0; S, \bar{T}) \right)^\top \\ &= \left(\delta_s^2(x^0; S, \bar{T}) \right)^\top S^\dagger \\ &= [\nabla_s f(x^0 + s^1; \bar{T}) - \nabla_s f(x^0; \bar{T}) \ \dots \ \nabla_s f(x^0 + s^m; \bar{T}) - \nabla_s f(x^0; \bar{T})] S^\dagger \\ &= (\bar{T}^\top)^\dagger [\delta_s(x^0 + s^1; \bar{T}) - \delta_s(x^0; \bar{T}) \ \dots \ \delta_s(x^0 + s^m; \bar{T}) - \delta_s(x^0; \bar{T})] S^\dagger. \end{aligned}$$

Let $D_s^2 = [\delta_s(x^0 + s^1; \bar{T}) - \delta_s(x^0; \bar{T}) \ \dots \ \delta_s(x^0 + s^m; \bar{T}) - \delta_s(x^0; \bar{T})] \in \mathbb{R}^{k \times m}$. Then

$$[D_s^2]_{ij} = f(x^0 + s^i + t^j) - f(x^0 + s^i) - f(x^0 + t^j) + f(x^0).$$

Note that

$$D_s^2 = \begin{bmatrix} (\delta_s(x^0 + t^1; S) - \delta_s(x^0; S))^\top \\ \vdots \\ (\delta_s(x^0 + t^k; S) - \delta_s(x^0; S))^\top \end{bmatrix}.$$

Hence,

$$\begin{aligned} \left(\nabla_s^2 f(x^0; S, \bar{T}) \right)^\top &= (\bar{T}^\top)^\dagger \begin{bmatrix} (\delta_s(x^0 + t^1; S) - \delta_s(x^0; S))^\top \\ \vdots \\ (\delta_s(x^0 + t^k; S) - \delta_s(x^0; S))^\top \end{bmatrix} S^\dagger \\ &= (\bar{T}^\top)^\dagger \begin{bmatrix} ((S^\dagger)^\top \delta_s(x^0 + t^1; S) - (S^\dagger)^\top \delta_s(x^0; S))^\top \\ \vdots \\ ((S^\dagger)^\top \delta_s(x^0 + t^k; S) - (S^\dagger)^\top \delta_s(x^0; S))^\top \end{bmatrix} \\ &= (\bar{T}^\top)^\dagger \delta_s^2(x^0; \bar{T}, S) = \nabla_s^2 f(x^0; \bar{T}, S). \end{aligned}$$

The second equality follows from the definition of the GCSH (Definition 3.2). \square

In the next section, we investigate how to obtain an order-1 accurate approximation of the full Hessian with a minimal number of sample points.

5. Minimal poised set and quadratic interpolation

In this section, we investigate how to choose the matrices of directions S and T_i to obtain an approximation of the full Hessian with a minimal number of sample points. Reducing the number of distinct points is extremely valuable, as it will decrease the number of function evaluations necessary to compute the GSH. For all of this section, we set $T_1 = \dots = T_m =: \bar{T} \in \mathbb{R}^{n \times k}$. We show that if S and \bar{T} have a specific structure, then the number of function evaluations necessary to compute the GSH is $(n+1)(n+2)/2$. We then explore some results that occur when such a structure is used. We begin by defining a *set for GSH computation*.

DEFINITION 5.1 The set of all points utilized in the computation of $\nabla_s^2 f(x^0; S, \bar{T})$ is said to be the set for *GSH computation* and is denoted $\mathcal{S}(x^0; S, \bar{T})$.

Note that $\mathcal{S}(x^0; S, \bar{T})$ contains at most $(m+1)(k+1)$ distinct points, but can contain fewer points if some of them overlap. Next, we introduce the definition of *minimal poised set* for the GSH.

DEFINITION 5.2 (Minimal poised set for the GSH). Let $x^0 \in \mathbb{R}^n$ be the point of interest. Let $S = [s^1 \ s^2 \ \dots \ s^n] \in \mathbb{R}^{n \times n}$ and $\bar{T} = [t^1 \ t^2 \ \dots \ t^n] \in \mathbb{R}^{n \times n}$. We say that $\mathcal{S}(x^0; S, \bar{T})$ is a *minimal poised set for the GSH* at x^0 if and only if S and \bar{T} are full rank and $\mathcal{S}(x^0; S, \bar{T})$ contains exactly $(n+1)(n+2)/2$ distinct points.

This definition requires S and \bar{T} to have exactly n columns and to be full rank. This is the case where the GSH is S -determined and \bar{T} -determined. This implies that S and \bar{T} are of minimal size to ensure that the GSH is an order-1 accurate approximation of the full Hessian.

Note that a minimal poised set for the GCSH could be defined in a similar manner. Obviously, such a set must contain more than $(n+1)(n+2)/2$. An investigation of this topic is left as a future research direction. We next show that it is possible to create a minimal poised set for the GSH.

PROPOSITION 5.3 Let x^0 be the point of interest. Let $S = [s^1 \ s^2 \ \dots \ s^n] \in \mathbb{R}^{n \times n}$. Define the set U_ℓ for each index $\ell \in \{0, 1, \dots, n\}$ as

$$U_0 = S$$

and

$$U_\ell = [s^1 - s^\ell \ s^2 - s^\ell \ \dots \ s^{\ell-1} - s^\ell \ -s^\ell \ s^{\ell+1} - s^\ell \ \dots \ s^n - s^\ell] \in \mathbb{R}^{n \times n}, \ell \neq 0.$$

Then for each ℓ , $|\mathcal{S}(x^0; S, U_\ell)| \leq (n+1)(n+2)/2$. Moreover, if S has full rank, then $|\mathcal{S}(x^0; S, U_\ell)| = (n+1)(n+2)/2$.

Proof. Without loss of generality, let $x^0 = \mathbf{0}$. First, suppose $\ell \in \{1, 2, \dots, n\}$. For an arbitrary function f , consider the matrix $\delta_s^2(x^0; S, \bar{T})$ where $\bar{T} = U_\ell$. The computation of $\nabla_s f(x^0; U_\ell)$ evaluates f at the points

$$\begin{aligned} & \{\mathbf{0}, (s^1 - s^\ell), \dots, (s^{\ell-1} - s^\ell), -s^\ell, (s^{\ell+1} - s^\ell), \dots, (s^n - s^\ell)\} \\ &= \{\mathbf{0}\} \cup \{-s^\ell\} \cup \bigcup_{\substack{i=1 \\ i \neq \ell}}^n \{s^i - s^\ell\}. \end{aligned}$$

For $i \neq \ell$, the computation of $\nabla_s f(x^0 + s^i; U_k)$ evaluates f at the points

$$\begin{aligned} & \{s^i, s^i + (s^1 - s^\ell), \dots, s^i + (s^{\ell-1} - s^\ell), s^i - s^\ell, s^i + (s^{\ell+1} - s^\ell), \dots, s^i + (s^n - s^\ell)\} \\ &= \{s^i\} \cup \{s^i - s^\ell\} \cup \bigcup_{\substack{j=1 \\ j \neq \ell}}^n \{s^i + s^j - s^\ell\}. \end{aligned}$$

The computation of $\nabla_s f(x^0 + s^\ell; U_\ell)$ evaluates f at the points

$$\begin{aligned} & \{s^\ell, s^\ell + (s^1 - s^\ell), \dots, s^\ell + (s^{\ell-1} - s^\ell), s^\ell - s^\ell, s^\ell + (s^{\ell+1} - s^\ell), \dots, s^\ell + (s^n - s^\ell)\} \\ &= \{s^\ell\} \cup \{\mathbf{0}\} \cup \bigcup_{\substack{i=1 \\ i \neq \ell}}^n \{s^i\} = \{\mathbf{0}\} \cup S. \end{aligned}$$

Thus, f is evaluated at the points

$$\begin{aligned} & \left(\{\mathbf{0}\} \cup \{-s^\ell\} \cup \bigcup_{\substack{i=1 \\ i \neq \ell}}^n \{s^i - s^\ell\} \right) \cup \left(\bigcup_{\substack{i=1 \\ i \neq \ell}}^n \{s^i\} \cup \bigcup_{\substack{i=1 \\ i \neq \ell}}^n \{s^i - s^\ell\} \cup \bigcup_{\substack{i=1 \\ i \neq \ell}}^n \bigcup_{\substack{j \geq i \\ j \neq \ell}}^n \{s^i + s^j - s^\ell\} \right) \cup (\{\mathbf{0}\} \cup S) \\ &= \{\mathbf{0}\} \cup S \cup \{-s^\ell\} \cup \bigcup_{\substack{i=1 \\ i \neq \ell}}^n \{s^i - s^\ell\} \cup \bigcup_{\substack{i=1 \\ i \neq \ell}}^n \bigcup_{\substack{j \geq i \\ j \neq \ell}}^n \{s^i + s^j - s^\ell\}. \end{aligned} \quad (5.1)$$

This is at most $1 + n + 1 + (n-1) + (n-1)(n-2)/2 = (n+1)(n+2)/2$ points.

Now, suppose $\ell = 0$. Using a similar process to the above, we find that f is evaluated at the points

$$\left(\{\mathbf{0}\} \cup \bigcup_{i=1}^n \{s^i\} \right) \cup \left(\bigcup_{i=1}^n \{2s^i\} \cup \bigcup_{i=1}^n \bigcup_{j>i}^n \{s^i + s^j\} \right). \quad (5.2)$$

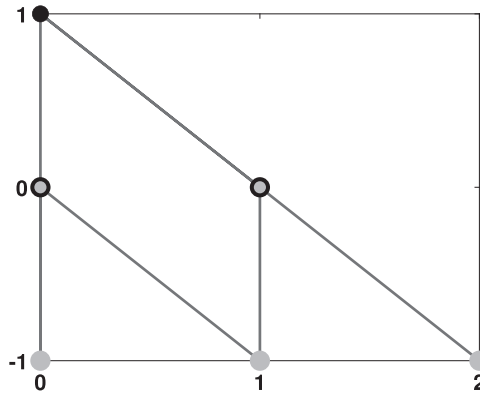
This is at most $(n+1)(n+2)/2$ points.

Finally, if S is full rank, then the four sets in (5.1) and the four sets in (5.2) are disjoint, so we have exactly $(n+1)(n+2)/2$ function evaluations. \square

Using $S = \text{Id}_n$ in Proposition 5.3, we can create $n+1$ canonical minimal poised sets for the GSH.

DEFINITION 5.4 (ℓ^{th} -canonical minimal poised set for the GSH). Let $x^0 \in \mathbb{R}^n$ be the point of interest. Let $S = \text{Id}_n$. Fix $\ell \in \{0, 1, \dots, n\}$. Let

$$E_0 = \text{Id}_n$$

FIG. 1. The 2nd-canonical minimal poised set for the GSH at x^0 in \mathbb{R}^2 .

and

$$E_\ell = [e^1 - e^\ell \ e^2 - e^\ell \dots e^{\ell-1} - e^\ell - e^\ell \ e^{\ell+1} - e^\ell \dots e^n - e^\ell], \ell \neq 0.$$

Then $\mathcal{S}(x^0; \text{Id}_n, E_\ell)$ is called the ℓ^{th} -canonical minimal poised set for the GSH at x^0 .

From Proposition 5.3 and the fact that any matrix E_k is full rank in Definition 5.4, the ℓ^{th} -canonical minimal poised set for the GSH is indeed a minimal poised set for the GSH. Henceforth, we use the notation $\mathcal{M}(x^0; S, U_k)$ to denote a minimal poised set for the GSH at x^0 that takes the form constructed in Proposition 5.3.

Note that the order of the directions in S and \bar{T} is arbitrary. Thus, it is immediately clear that if $\mathcal{S}(x^0; S, \bar{T})$ is a minimal poised set for the GSH and $P_1, P_2 \in \mathbb{R}^{n \times n}$ are permutation matrices, then $\mathcal{S}(x^0; SP_1, \bar{TP}_2)$ is also a minimal poised set for the GSH at x^0 . The next proposition expands this idea and demonstrates how to construct minimal poised sets for the GSH.

PROPOSITION 5.5 Let $x^0 \in \mathbb{R}^n$ be the point of interest. Let $S, \bar{T} \in \mathbb{R}^{n \times n}$. Let $N \in \mathbb{R}^{n \times n}$ be an invertible matrix and $P_1, P_2 \in \mathbb{R}^{n \times n}$ be permutation matrices. Then $\mathcal{S}(x^0; S, \bar{T})$ is a minimal poised set for the GSH at x^0 if and only if $\mathcal{S}(x^0; NSP_1, N\bar{TP}_2)$ is a minimal poised set for the GSH at x^0 .

Proof. The proof follows trivially from properties of matrices. \square

It follows that $\mathcal{S}(x^0; S, \bar{T})$ is a minimal poised set for the GSH at x^0 if and only if the set $\mathcal{S}(x^0; \beta SP_1, \beta \bar{TP}_2)$ is a minimal poised set for the GSH at x^0 , where β is a nonzero scalar and P_1, P_2 are permutation matrices.

EXAMPLE 5.6 Let $x^0 = (0, 0)$. The 2nd-canonical minimal poised set for the GSH in \mathbb{R}^2 contains the points $(0, -1)$, $(0, 0)$, $(0, 1)$, $(1, -1)$, $(1, 0)$ and $(2, -1)$. In this case, $S = \{e^1, e^2\}$ and $\bar{T} = \{(e^1 - e^2), -e^2\}$. Figure 1 illustrates this set.

The points in $\{(0, 0) \cup \{(0, 0) \oplus S\} = \{(0, e^1, e^2)\}$ are represented with solid black borders. These are the base points where simplex gradients will be computed. The lines represent the vectors corresponding to \bar{T} emanate from $\{(0, 0) \cup \{(0, 0) \oplus S\}$. The points in $(\{(0, 0) \cup \{(0, 0) \oplus S\}) + \bar{T}$ are represented with grey cores. These are the points used to construct the simplex gradients. Notice the points $(0, 0)$

and $(1, 0)$ have both black borders and grey cores. These are the common points that allow the number of function evaluations to be reduced to $(n + 1)(n + 2)/2 = 6$.

We next demonstrate that every minimal poised set for the GSH of the form $\mathcal{M}(x^0; S, U_\ell)$ is poised for quadratic interpolation. We then show that the converse is not true; it is possible to construct a set that is poised for quadratic interpolation, but does not take the form of $\mathcal{M}(x^0; S, U_\ell)$.

PROPOSITION 5.7 Let $S \in \mathbb{R}^{n \times n}$ be full rank. Select $\ell \in \{0, 1, \dots, n\}$ and define $\bar{T} := U_\ell$ as in Proposition 5.3. Then $\mathcal{M}(x^0; S, \bar{T})$ is poised for quadratic interpolation.

Proof. Since $\mathcal{M}(x^0; S, \bar{T})$ is poised for quadratic interpolation if and only if the set $\mathcal{M}(x^0; S, \bar{T}) \oplus -x^0$ is poised for quadratic interpolation, we assume without loss of generality that $x^0 = \mathbf{0}$.

Case I: $\bar{T} = U_\ell$ for $\ell \in \{1, 2, \dots, n\}$. Without loss of generality, by Proposition 5.5, assume that $\ell = n$. The points contained in $\mathcal{M}(x^0; S, U_n)$ are

$$\begin{aligned} & \{\mathbf{0}\} \cup S \cup \{-s^n\} \cup \bigcup_{i=1}^{n-1} \{s^i - s^n\} \cup \bigcup_{i=1}^{n-1} \bigcup_{\substack{j \geq i \\ j \neq n}} \{s^i + s^j - s^n\} \\ &= \{\mathbf{0}\} \cup S \cup \{-s^n\} \cup \bigcup_{i=1}^{n-1} \{s^i - s^n\} \cup \{2s^i - s^n\}_{i=1}^{n-1} \cup \bigcup_{i=1}^{n-1} \bigcup_{\substack{j > i \\ j \neq n}} \{s^i + s^j - s^n\}. \end{aligned}$$

We show that using this set, the only solution to (2.1) is the trivial solution. Considering the point $\mathbf{0}$, we obtain

$$\alpha_0 = 0.$$

Considering the points s^i for $i \in \{1, 2, \dots, n\}$ and noting that for all $i \in \{1, 2, \dots, n\}$, $s^i = Se^i$, we obtain

$$\bar{\alpha}^\top e^i + \frac{1}{2}(e^i)^\top \widehat{\mathbf{H}} e^i = 0, \quad (5.3)$$

where $\bar{\alpha}^\top = \alpha^\top S$ and $\widehat{\mathbf{H}} = S^\top \mathbf{H} S$. Note that $\widehat{\mathbf{H}}$ is symmetric. Considering the points $s^i - s^n$ for $i \in \{1, 2, \dots, n-1\}$, we obtain

$$\bar{\alpha}^\top e^i - \bar{\alpha}^\top e^n - (e^i)^\top \widehat{\mathbf{H}} e^n + \frac{1}{2}(e^i)^\top \widehat{\mathbf{H}} e^i + \frac{1}{2}(e^n)^\top \widehat{\mathbf{H}} e^n = 0.$$

Using (5.3), this simplifies to

$$-\bar{\alpha}^\top e^n - (e^i)^\top \widehat{\mathbf{H}} e^n + \frac{1}{2}(e^n)^\top \widehat{\mathbf{H}} e^n = 0. \quad (5.4)$$

Considering the point $-s^n$, we find

$$-\bar{\alpha}^\top e^n + \frac{1}{2}(e^n)^\top \widehat{\mathbf{H}}e^n = 0, \quad (5.5)$$

which reduces (5.4) to

$$-(e^i)^\top \widehat{\mathbf{H}}e^n = 0 \quad \text{for } i \in \{1, 2, \dots, n-1\}.$$

Thus, $(e^i)^\top \widehat{\mathbf{H}}e^n = \widehat{\mathbf{H}}_{i,n} = \widehat{\mathbf{H}}_{n,i} = 0$ for all $i \in \{1, 2, \dots, n-1\}$. Combining (5.3) at $i = n$ and (5.5) multiplied by -1 , we get

$$\bar{\alpha}^\top e^n + \frac{1}{2}(e^n)^\top \widehat{\mathbf{H}}e^n = 0 = \bar{\alpha}^\top e^n - \frac{1}{2}(e^n)^\top \widehat{\mathbf{H}}e^n,$$

which implies that $(e^n)^\top \widehat{\mathbf{H}}e^n = \widehat{\mathbf{H}}_{n,n} = 0$. Considering the points $2s^i - s^n$ for $i \in \{1, 2, \dots, n-1\}$, we get

$$2\bar{\alpha}^\top e^i - \bar{\alpha}^\top e^n + 2(e^i)^\top \widehat{\mathbf{H}}e^i + \frac{1}{2}(e^n)^\top \widehat{\mathbf{H}}e^n = 0.$$

Using (5.5), this simplifies to

$$2\bar{\alpha}^\top e^i + 2(e^i)^\top \widehat{\mathbf{H}}e^i = 0.$$

By multiplying (5.3) by 2 and substituting in the above equation, we get $(e^i)^\top \widehat{\mathbf{H}}e^i = \widehat{\mathbf{H}}_{i,i} = 0$ for all $i \in \{1, 2, \dots, n-1\}$. This now implies $\bar{\alpha}_i = \bar{\alpha}^\top e^i = 0$ for all $i \in \{1, 2, \dots, n\}$, i.e., $\bar{\alpha} = \mathbf{0}$. Lastly, consider the points $s^i + s^j - s^n$ for $i \neq j, i, j \in \{1, 2, \dots, n-1\}$. Since $\widehat{\mathbf{H}}_{i,i} = 0$ for $i \in \{1, 2, \dots, n\}$, $\widehat{\mathbf{H}}_{i,n} = \widehat{\mathbf{H}}_{n,i} = 0$, for $i \in \{1, 2, \dots, n-1\}$ and $\bar{\alpha} = \mathbf{0}$, we obtain

$$(e^j)^\top \widehat{\mathbf{H}}e^j = 0.$$

Thus $\widehat{\mathbf{H}} = \mathbf{0}_{n \times n}$. Therefore, the only solution to (2.1) is the trivial solution.

Case II: $\bar{T} = U_0$. The proof for this case is analogous to that of Case I. \square

It follows from the previous proposition that when a minimal poised set for the GSH is used, the GSH is equal to the simplex Hessian as described in Conn *et al.* (2009, Section 9.5) (see also Custódio & Vicente, 2007, Section 3). Therefore, the results obtained in Custódio & Vicente (2007); Conn *et al.* (2009) related to the simplex Hessian apply to the GSH. In this case, the GSH (and the simplex Hessian) is equal to the Hessian of the quadratic interpolation model passing through the sample points $\mathcal{M}(x^0; S, U_\ell)$. Therefore, the results developed in Custódio & Vicente (2007); Conn *et al.* (2008b, 2009) related to the simplex Hessian and the Hessian of a quadratic interpolation model are valid for the GSH. One of the advantages of computing the GSH compared to the computation of the Hessian of the quadratic interpolation model has been discussed in Remark 3.3. Another advantage of the GSH compared to the simplex Hessian described in Custódio & Vicente (2007); Conn *et al.* (2009) is that it provides a simple, explicit formula that is well-defined as long as the matrices of directions employed are nonempty.

Hence, the formula for the GSH provides an approximation of the full Hessian, or a partial Hessian, for all possible cases.

Next, we provide an example that serves to show that a set of $(n+1)(n+2)/2$ distinct points in \mathbb{R}^n that is poised for quadratic interpolation is not necessarily a minimal poised set for the GSH.

EXAMPLE 5.8 Let $x^0 = [0 \ 0]^\top$ be the point of interest. Consider $\mathcal{X} = \{x^0 = \mathbf{0}, e^1, e^2, -e^1, -e^2, -e^1 - e^2\}$. Then \mathcal{X} is poised for quadratic interpolation, but cannot be expressed as a minimal poised set for the GSH at x^0 .

Proof. Using a similar approach to the one in Proposition 5.7, one can verify that \mathcal{X} is poised for quadratic interpolation. We show that \mathcal{X} is not a minimal poised set for the GSH at x^0 using brute force. We need to build $S = \{s^1, s^2\}$ such that the matrix corresponding to S is full rank and $x^0 \oplus S \subseteq \mathcal{X}$. Hence, the possible choices for S are

$$S \in \left\{ \{e^1, e^2\}, \{e^1, -e^2\}, \{e^1, -e^1 - e^2\}, \{-e^1, e^2\}, \right. \\ \left. \{-e^1 - e^2, e^2\}, \{-e^1, -e^2\}, \{-e^1, -e^1 - e^2\}, \{-e^1 - e^2, -e^2\} \right\}.$$

Case I: $S = \{e^1, e^2\}$. In this case, we need to build $\bar{T} = \{t^1, t^2\}$ such that the matrix corresponding to \bar{T} is full rank and

$$\{t^1, e^1 + t^1, e^2 + t^1, t^2, e^1 + t^2, e^2 + t^2\} = \mathcal{X}.$$

We see that the only possible choice of t^1 such that $\{t^1, e^1 + t^1, e^2 + t^1\} \subseteq \mathcal{X}$ is $t^1 = -e^1 - e^2$ (note $t^1 \neq \mathbf{0}$ as we require full rank). However, the only possible choice of t^2 such that $\{t^2, e^1 + t^2, e^2 + t^2\} \subseteq \mathcal{X}$ is $t^2 = -e^1 - e^2$. As full rank implies t^1 cannot equal t^2 , we see $S = \{e^1, e^2\}$ cannot provide the desired properties.

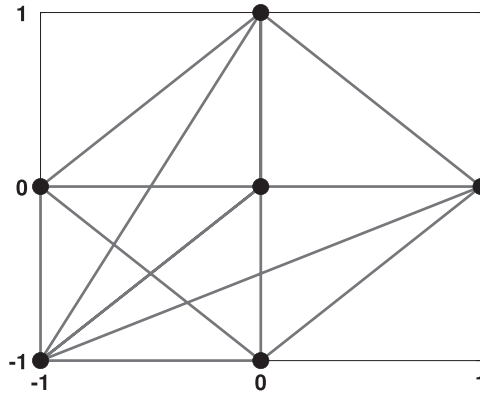
Cases II through VIII: The other options for S can be eliminated analogously.

Therefore, \mathcal{X} cannot be expressed as a minimal poised set for the GSH at x^0 . □

Figure 2 shows all possible directions connecting two points in the set \mathcal{X} from Example 5.8. If \mathcal{X} were a minimal poised set for the GSH at $x^0 = \mathbf{0}$, then it would be possible to choose two directions emerging from x^0 that connect to other points in \mathcal{Y} and these same two directions would be emerging from two other points.

We conclude this section by providing formulae to obtain all the values of the coefficients involved in the quadratic interpolation function of f over a minimal poised set for the GSH of the form $\mathcal{M}(x^0; S, U_\ell)$, which we denote by $Q_f(x^0; S, U_\ell)(x)$.

PROPOSITION 5.9 Let $f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$. Let $\mathcal{M}(x^0; S, U_\ell) \subset \text{dom } f$ be a minimal poised set for the GSH at x^0 as constructed in Proposition 5.3. If $\ell \in \{1, 2, \dots, n\}$, then the Hessian matrix \mathbf{H} of the quadratic interpolation function $Q_f(x^0; S, U_\ell)(x)$ is given by $\mathbf{H} = S^{-\top} \widehat{\mathbf{H}} S^{-1}$, where the entries of the

FIG. 2. A set poised for quadratic interpolation that is not minimal poised for the GSH at x^0 .

symmetric matrix $\widehat{\mathbf{H}} \in \mathbb{R}^{n \times n}$ are

$$\widehat{\mathbf{H}}_{i,\ell} = -f(x^0 + s^i - s^\ell) + f(x^0 + s^i) + f(x^0 - s^\ell) - f(x^0), \quad i \in \{1, 2, \dots, n\} \setminus \{\ell\},$$

$$\widehat{\mathbf{H}}_{i,i} = f(x^0 + 2s^i - s^\ell) - 2f(x^0 + s^i - s^\ell) + f(x^0 - s^\ell), \quad i \in \{1, 2, \dots, n\} \setminus \{\ell\},$$

$$\widehat{\mathbf{H}}_{\ell,\ell} = f(x^0 + s^\ell) + f(x^0 - s^\ell) - 2f(x^0),$$

$$\widehat{\mathbf{H}}_{i,j} = f(x^0 + s^i + s^j - s^\ell) - f(x^0 + s^i - s^\ell) - f(x^0 + s^j - s^\ell) + f(x^0 - s^\ell),$$

for all $i, j \in \{1, 2, \dots, n\} \setminus \{\ell\}, i \neq j$. If $\ell = 0$, then

$$\widehat{\mathbf{H}}_{i,i} = f(x^0 + 2s^i) - 2f(x^0 + s^i) + f(x^0), \quad i \in \{1, 2, \dots, n\},$$

$$\widehat{\mathbf{H}}_{i,j} = f(x^0 + s^i + s^j) - f(x^0 + s^i) - f(x^0 + s^j) + f(x^0), \quad i, j \in \{1, 2, \dots, n\}, i \neq j.$$

For all $\ell \in \{0, 1, \dots, n\}$, the vector α associated to $Q_f(x^0; S, U_\ell)(x)$ is given by $\alpha = S^{-\top} \bar{\alpha}$, where the entries of $\bar{\alpha} \in \mathbb{R}^n$ are

$$\bar{\alpha}_i = f(x^0 + s^i) - f(x^0) - \frac{1}{2} \widehat{\mathbf{H}}_{i,i} - (x^0)^\top \mathbf{H} s^i, \quad i \in \{1, 2, \dots, n\}.$$

The scalar α_0 of $Q_f(x^0; S, U_\ell)(x)$ is

$$\alpha_0 = f(x^0) - \alpha^\top x^0 - \frac{1}{2} (x^0)^\top \mathbf{H} x^0.$$

Proof. The result is obtained by using Definition 2.2. Let

$$Q_f(x^0; S, U_\ell)(x) = \alpha_0 + \alpha^\top x + \frac{1}{2} x^\top \mathbf{H} x,$$

where $a_0 \in \mathbb{R}$, $\alpha \in \mathbb{R}^n$ and $\mathbf{H} = \mathbf{H}^\top \in \mathbb{R}^{n \times n}$. Suppose $\ell \in \{1, 2, \dots, n\}$. Evaluating $Q_f(x^0; S, U_\ell)(x)$ at x^0 , we obtain

$$\alpha_0 + \alpha^\top x^0 + \frac{1}{2}(x^0)^\top \mathbf{H} x^0 = f(x^0). \quad (5.6)$$

Evaluating $Q_f(x^0; S, U_\ell)(x)$ at $x^0 + s^i$ and using (5.6), we obtain

$$f(x^0) + \bar{\alpha}^\top e^i + (x^0)^\top \bar{\mathbf{H}} e^i + \frac{1}{2} \hat{\mathbf{H}}_{i,i} = f(x^0 + s^i), \quad i \in \{1, \dots, n\}, \quad (5.7)$$

where $\bar{\alpha}^\top = \alpha^\top S$, $\bar{\mathbf{H}} = \mathbf{H}S$ and $\hat{\mathbf{H}} = S^\top \mathbf{H}S$. Evaluating $Q_f(x^0; S, U_\ell)(x)$ at $x^0 + s^i - s^\ell$ and using (5.6) and (5.7), we obtain

$$f(x^0 + s^i) - \bar{\alpha}^\top s^\ell - (x^0)^\top \bar{\mathbf{H}} e^\ell - \hat{\mathbf{H}}_{i,\ell} + \frac{1}{2} \hat{\mathbf{H}}_{\ell,\ell} = f(x^0 + s^i - s^\ell), \quad i \in \{1, \dots, n\} \setminus \{\ell\}. \quad (5.8)$$

Evaluating $Q_f(x^0; S, U_\ell)(x)$ at $x^0 - s^\ell$ and using (5.6), we find

$$-\bar{\alpha}^\top e^\ell - (x^0)^\top \bar{\mathbf{H}} e^\ell + \frac{1}{2} \hat{\mathbf{H}}_{\ell,\ell} = f(x^0 - s^\ell) - f(x^0). \quad (5.9)$$

Substituting (5.9) into (5.8), we obtain

$$\hat{\mathbf{H}}_{i,\ell} = \hat{\mathbf{H}}_{\ell,i} = -f(x^0 + s^i - s^\ell) + f(x^0 + s^i) + f(x^0 - s^\ell) - f(x^0), \quad i \in \{1, \dots, n\} \setminus \{\ell\}. \quad (5.10)$$

Evaluating $Q_f(x^0; S, U_\ell)(x)$ at $x^0 + 2s^i - s^\ell$ and using (5.6), (5.7) and (5.9), we find

$$2f(x^0 + s^i) - 2f(x^0) + f(x^0 - s^\ell) + \hat{\mathbf{H}}_{i,i} - 2\hat{\mathbf{H}}_{i,\ell} = f(x^0 + 2s^i - s^\ell), \quad i \in \{1, \dots, n\} \setminus \{\ell\}. \quad (5.11)$$

Substituting (5.10) in (5.11), we find

$$\hat{\mathbf{H}}_{i,i} = f(x^0 + 2s^i - s^\ell) + f(x^0 - s^\ell) - 2f(x^0 + s^i - s^\ell), \quad i \in \{1, 2, \dots, n\} \setminus \{\ell\}.$$

Evaluating $Q_f(x^0; S, U_\ell)(x)$ at $x^0 + s^i + s^j - s^\ell$ and using (5.6), (5.7) and (5.9), we find

$$\hat{\mathbf{H}}_{i,j} = \hat{\mathbf{H}}_{j,i} = f(x^0 + s^i + s^j - s^\ell) - f(x^0 + s^i - s^\ell) - f(x^0 + s^j - s^\ell) + f(x^0 - s^\ell),$$

for $i, j \in \{1, 2, \dots, n\} \setminus \{\ell\}$, $i \neq j$. Rearranging (5.9), we get

$$\frac{1}{2} \hat{\mathbf{H}}_{\ell,\ell} - f(x^0 - s^\ell) + f(x^0) = \bar{\alpha}^\top e^\ell + (x^0)^\top \bar{\mathbf{H}} e^\ell. \quad (5.12)$$

Substituting (5.12) into (5.7) for $i = \ell$, we obtain

$$\widehat{\mathbf{H}}_{\ell,\ell} = f(x^0 + s^\ell) + f(x^0 - s^\ell) - 2f(x^0).$$

The entries of the vector $\bar{\alpha}$ are found by isolating $\bar{\alpha}^\top e^i$ in (5.7). We obtain

$$\bar{\alpha}_i = f(x^0 + s^i) - f(x^0) - \frac{1}{2}\widehat{\mathbf{H}}_{i,i} - (x^0)^\top \mathbf{H} s^i, \quad i \in \{1, 2, \dots, n\},$$

where $\alpha_i = S^{-\top} \bar{\alpha}_i$. Lastly, the scalar α_0 is obtained from (5.6). We find

$$\alpha_0 = f(x^0) - \alpha^\top x^0 - \frac{1}{2}(x^0)^\top \mathbf{H} x^0.$$

If $\ell = 0$, a similar process can be applied to obtain \mathbf{H}, α and α_0 . □

Using $\bar{T} = U_\ell$, it is worth emphasizing that $Q_f(x^0; S, \bar{T})(x)$ can be obtained for free in terms of function evaluations whenever $\nabla_s^2 f(x^0; S, \bar{T})$ has already been computed. Indeed, all the coefficients of $Q_f(x^0; S, \bar{T})$ are computed using the same function evaluations used for finding $\nabla_s^2 f(x^0; S, \bar{T})$. Since $\mathcal{M}(x^0; S, \bar{T})$ is poised for quadratic interpolation, there exists a unique Hessian matrix \mathbf{H} . Thus, the Hessian of the quadratic interpolation function $Q_f(x^0; S, \bar{T})(x)$ must be $\nabla_s^2 f(x^0; S, \bar{T})$.

6. Conclusion

We have presented the generalized simplex Hessian (GSH), a compact and simple formula for approximating the Hessian of a function f . The GSH is well-defined as long as the matrices S and T_i are nonempty. In particular, the matrices S and T_i do not need to be full row rank nor square to compute the GSH. We have also presented a centered version of the GSH called the generalized centered simplex Hessian (GCSH) and shown its relation to the GSH (Section 3).

In Theorems 4.2 and 4.3, we developed error bounds for the GSH and GCSH. Under some assumptions, the GSH is an order-1 accurate approximation of the partial Hessian $\text{Proj}_{S, T_{1:m}} \nabla^2 f(x^0)$, and the GCSH is an order-2 accurate approximation of the partial Hessian $\text{Proj}_{S, T_{1:m}} \nabla^2 f(x^0)$. Indeed, when all matrices of directions in the set $T_{1:m}$ are equal, then we obtain error bounds that covers all possible 16 cases. When all matrices T_i are not equal, then 12 out of 16 cases are covered. Future research directions include an investigation of the cases not covered in Theorems 4.2 and 4.3: the four cases S -nondetermined/overdetermined and $T_{1:m}$ -nondetermined/underdetermined. We note that when a matrix is nondetermined or overdetermined, it is possible to remove columns (or rows) of the matrix to make it full rank. This may lead to error bounds for the nondetermined cases.

As the results in this paper are reported in terms of order- N accuracy, it is possible to apply the theory of DFO calculus developed in [Chen et al. \(2022\)](#) directly. In [Chen et al. \(2022\)](#), approximation-based product, quotient and chain rules are developed. For each rule, it is shown that when an objective function can be decomposed into a product, quotient or composition of simpler functions, then it is possible to construct approximate Hessians of the simpler functions and then use those to construct approximation Hessians of the objective function. Combining [Chen et al. \(2022\)](#) with the research herein has led to a model-based trust region algorithm using the calculus-based approximation techniques ([Hare & Jarry-Bolduc, 2023](#)).

In Section 5, we showed that if a minimal poised set for the GSH is used, the number of distinct function evaluations required to compute the GSH is $(n+1)(n+2)/2$. We investigated the relationship between a minimal poised set for the GSH and poisedness for quadratic interpolation, proving that a minimal poised set for the GSH of the form $\mathcal{M}(x^0; S, U_\ell)$ is well-poised for quadratic interpolation, but that the converse does not necessarily hold. In this case, we know that $\nabla_s^2 f(x^0; S, \bar{T})$ with $\bar{T} = U_\ell$ is equal to the Hessian of the quadratic interpolation function $Q_f(x^0; S, U_\ell)$. We also developed explicit formulae for obtaining the parameters of the quadratic interpolation function of f over a minimal poised set for the GSH of the form $\mathcal{M}(x^0; S, U_\ell)$.

Future research should explore if the sets U_0, U_1, \dots, U_n as defined in Proposition 5.3 are the only possible choices for $\bar{T} \in \mathbb{R}^{n \times n}$ such that $\mathcal{S}(x^0; S, \bar{T})$, where $S \in \mathbb{R}^{n \times n}$ is full rank and is a minimal poised set for GSH at x^0 . It can be proven that it is indeed the case by using brute force in \mathbb{R} and \mathbb{R}^2 , but it is still unclear how to generalize this claim in \mathbb{R}^n .

While it is immediately clear that a minimal poised set for GSH can be used to construct a ‘small’ set for the GCSH (using Proposition 3.7), it is not clear if such a set would be minimal. Future research will investigate the properties of a minimal poised set for the GCSH and the matrices of directions that can be used to minimize the number of sample points.

It is known that the diagonal entries of the Hessian may be obtained for free in terms of function evaluations when a centered simplex gradient is computed. Hence, a future research direction could be to verify if the main diagonal entries of $\nabla^3 f(x^0)$ may be obtained for free when a GCSH is computed.

On a final note, MATLAB implementations of the GSH and the GCSH are available upon request.

Funding

Research partially supported by NSERC of Canada Discovery Grant 2018-03865. Research partially supported by Natural Sciences and Engineering Research Council (NSERC) of Canada Discovery Grant 2018-03865. Research supported by University of Wollongong. Natural Sciences and Engineering Research Council (NSERC) of Canada (#2018-03865 to W.H. and G.J.-B.); UBC funding received through the University Graduate Fellowship award (to G.J.-B.).

Acknowledgements

The authors express their deep gratitude to the anonymous referees for the comprehensive reviews that improved the quality of this work.

REFERENCES

- AMAIQUA, N., AUDET, C., CONN, A. & LE DIGABEL, S. (2018) Efficient solution of quadratically constrained quadratic subproblems within the mesh adaptive direct search algorithm. *Eur. J. Oper. Res.*, **268**, 13–24.
- AUDET, C., BÉCHARD, V. & LE DIGABEL, S. (2008) Nonsmooth optimization through mesh adaptive direct search and variable neighborhood search. *J. Global Optim.*, **41**, 299–318.
- AUDET, C. & HARE, W. (2017) *Derivative-Free and Blackbox Optimization*. New York City: Springer.
- AUDET, C. & HARE, W. (2018) Algorithmic construction of the subdifferential from directional derivatives. *Set-Valued Var. Anal.*, **26**, 431–447.
- AUDET, C., IANNI, A., LE DIGABEL, S. & TRIBES, C. (2014) Reducing the number of function evaluations in mesh adaptive direct search algorithms. *SIAM J. Optim.*, **24**, 621–642.
- AUDET, C. & TRIBES, C. (2018) Mesh-based Nelder–Mead algorithm for inequality constrained optimization. *Comput. Optim. Appl.*, **71**, 331–352.

- BAGIROV, A., KARASÖZEN, B. & SEZER, M. (2008) Discrete gradient method: derivative-free method for nonsmooth optimization. *J. Optim. Theory Appl.*, **137**, 317–334.
- BERAHAS, A., BYRD, R. & NOCEDAL, J. (2019) Derivative-free optimization of noisy functions via quasi-Newton methods. *SIAM J. Optim.*, **29**, 965–993.
- BERGHEN, F. & BERSINI, H. (2005) CONDOR, a new parallel, constrained extension of Powell's UOBYQA algorithm: experimental results and comparison with the DFO algorithm. *J. Comput. Appl. Math.*, **181**, 157–175.
- BORTZ, D. & KELLEY, C. (1998) The simplex gradient and noisy optimization problems. *Computational Methods for Optimal Design and Control*. New York City: Springer, pp. 77–90.
- BRAUN, P., HARE, W. & JARRY-BOLDUC, G. (2021) Limiting behavior of derivative approximation techniques as the number of points tends to infinity on a fixed interval in \mathbb{R} . *J. Comput. Appl. Math.*, **386**, 113218.
- BURDEN, R., FAIRES, J. & BURDEN, A. (2016) *Numerical Analysis 10/e ie*. Boston: Brooks/Cole Cengage Learning.
- CHEN, Y., HARE, W. & JARRY-BOLDUC, G. (2022) Error analysis of surrogate models constructed through operations on submodels. *Math. Oper. Res.*
- COCCHI, G., LIUZZI, G., PAPINI, A. & SCIANDRONE, M. (2018) An implicit filtering algorithm for derivative-free multiobjective optimization with box constraints. *Comput. Optim. Appl.*, **69**, 267–296.
- CONN, A., SCHEINBERG, K. & VICENTE, L. (2008a) Geometry of interpolation sets in derivative free optimization. *Math. Programming*, **111**, 141–172.
- CONN, A., SCHEINBERG, K. & VICENTE, L. (2008b) Geometry of sample sets in derivative-free optimization: polynomial regression and underdetermined interpolation. *IMA J. Numer. Anal.*, **28**, 721–748.
- CONN, A., SCHEINBERG, K. & VICENTE, L. (2009) *Introduction to Derivative-Free Optimization*. Philadelphia: SIAM.
- COOPE, I. & TAPPENDEN, R. (2019) Efficient calculation of regular simplex gradients. *Comput. Optim. Appl.*, **72**, 561–588.
- COOPE, I. & TAPPENDEN, R. (2021) Gradient and diagonal hessian approximations using quadratic interpolation models and aligned regular bases. *Numer. Algorithms*, **88**, 767–791.
- CUSTÓDIO, A., DENNIS, J. & VICENTE, L. (2008) Using simplex gradients of nonsmooth functions in direct search methods. *IMA J. Numer. Anal.*, **28**, 770–784.
- CUSTÓDIO, A. & VICENTE, L. (2007) Using sampling and simplex derivatives in pattern search methods. *SIAM J. Optim.*, **18**, 537–555.
- GRATTON, S., ROYER, C. & VICENTE, L. (2020) A decoupled first/second-order steps technique for nonconvex nonlinear unconstrained optimization with improved complexity bounds. *Math. Programming*, **179**, 195–222.
- GRATTON, S., ROYER, C., VICENTE, L. & ZHANG, Z. (2019) Direct search based on probabilistic feasible descent for bound and linearly constrained problems. *Comput. Optim. Appl.*, **72**, 525–559.
- HARE, W. (2020) A discussion on variational analysis in derivative-free optimization. *Set-Valued Var. Anal.*, **28**, 643–659.
- HARE, W. & JARRY-BOLDUC, G. (2020) Calculus identities for generalized simplex gradients: rules and applications. *SIAM J. Optim.*, **30**, 853–884.
- HARE, W. & JARRY-BOLDUC, G. (2023) About the performance of a calculus-based approach to building model functions in a derivative-free trust-region algorithm. *Algorithms*, **16**, 84–103..
- HARE, W., JARRY-BOLDUC, G. & PLANIDEN, C. (2020a) Error bounds for overdetermined and underdetermined generalized centred simplex gradients. *IMA J. Numer. Anal.*, **42**, 744–770.
- HARE, W., PLANIDEN, C. & SAGASTIZÁBAL, C. (2020b) A derivative-free VU-algorithm for convex finite-max problems. *Optim. Methods Softw.*, **35**, 521–559.
- HARE, W., JARRY-BOLDUC, G. & PLANIDEN, C. (2023) Limiting behaviour of the generalized simplex gradient as the number of points tends to infinity on a fixed shape in \mathbb{R}^n . *Set-Valued Var. Anal.*, **31**, 1–33.
- HARE, W. & NUTINI, J. (2013) A derivative-free approximate gradient sampling algorithm for finite minimax problems. *Comput. Optim. Appl.*, **56**, 1–38.
- HARE, W., SAGASTIZÁBAL, C. & SOLODOV, M. (2016) A proximal bundle method for nonsmooth nonconvex functions with inexact information. *Comput. Optim. Appl.*, **63**, 1–28.

- JARRY-BOLDUC, G. (2022) Approximating the diagonal of a hessian: which sample set of points should be used. *Numer. Algorithms*, **91**, 1349–1361.
- JARRY-BOLDUC, G. (2023) Structures in derivative free optimization: approximating gradients/ Hessians, and positive bases. *Ph.D. Thesis*, University of British Columbia, Kelowna. <https://circle.ubc.ca/>.
- LARSON, J., MENICKELLY, M. & WILD, S. (2016) Manifold sampling for l_1 nonconvex optimization. *SIAM J. Optim.*, **26**, 2540–2563.
- LIUZZI, G., LUCIDI, S., RINALDI, F. & VICENTE, L. (2019) Trust-region methods for the derivative-free optimization of nonsmooth black-box functions. *SIAM J. Optim.*, **29**, 3012–3035.
- MAGGIAR, A., WÄCHTER, A., DOLINSKAYA, I. & STAUM, J. (2018) A derivative-free trust-region algorithm for the optimization of functions smoothed via Gaussian convolution using adaptive multiple importance sampling. *SIAM J. Optim.*, **28**, 1478–1507.
- MENICKELLY, M. & WILD, S. (2019) Derivative-free robust optimization by outer approximations. *Math. Programming* (in press).
- NOCEDAL, J. & WRIGHT, S. (2006) *Numerical Optimization*. Berlin: Springer Science & Business Media.
- POWELL, M. (2001) On the Lagrange functions of quadratic models that are defined by interpolation. *Optim. Methods Softw.*, **16**, 289–309.
- POWELL, M. (2003) On trust region methods for unconstrained minimization without derivatives. *Math. Programming*, **97**, 605–623.
- POWELL, M. (2009) The BOBYQA algorithm for bound constrained optimization without derivatives. *Cambridge NA Report NA2009/06*. Cambridge: University of Cambridge, pp. 26–46.
- REGIS, R. (2015) The calculus of simplex gradients. *Optim. Lett.*, **9**, 845–865.
- ROCKAFELLAR, R. & WETS, R. (1998) *Variational Analysis*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Berlin: Springer.
- SHASHAANI, S., HASHEMI, F. & PASUPATHY, R. (2018) ASTRO-DF: a class of adaptive sampling trust-region algorithms for derivative-free stochastic optimization. *SIAM J. Optim.*, **28**, 3145–3176.
- VERDÉRIO, A., KARAS, E., PEDROSO, L. & SCHEINBERG, K. (2017) On the construction of quadratic models for derivative-free trust-region algorithms. *EURO J. Comput. Optim.*, **5**, 501–527.
- WILD, S. & SHOEMAKER, C. (2013) Global convergence of radial basis function trust-region algorithms for derivative-free optimization. *SIAM Rev.*, **55**, 349–371.