Mesh Shape-Quality Optimization Using the Inverse Mean-Ratio Metric: Tetrahedral Proofs*

Todd Munson†

April 20, 2004

1 Introduction

This technical report is a companion to [5] that proves the diagonal blocks of the Hessian matrix for the inverse mean-ratio metric for tetrahedral elements are positive definite. Thus, the block Jacobi preconditioner used in the inexact Newton method to solve the mesh shape-quality optimization problem using the average inverse mean-ratio metric for the objective function is positive definite. Note that [5] only proves these results for triangular elements. We first recall the proposition proved in [5] used to show convexity for fractional functions.

Definition 1.1 (Uniform Convexity [6]) *Let $f : \Re^n \to \Re$, and let $\Omega \subseteq \Re^n$ be a convex set. The function f is uniformly convex on Ω with constant κ if there exists a constant $\kappa > 0$ such that for all $x \in \Omega$, $y \in \Omega$, and $\lambda \in [0, 1]$,*

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y) - \kappa\lambda(1-\lambda)\|y-x\|_2^2.$$

Proposition 1.2 *Let $f : \Re^n \to \Re$ and $g : \Re^n \to \Re$, and let $\Omega \subseteq \Re^n$ be a convex set. Assume the following properties are satisfied:*

1. *g is a positive, concave function on Ω.*
2. *f is a nonnegative, uniformly convex function with constant κ on Ω.*
3. *For all $(x, y) \in \Theta := \left\{(x, y) \in \Omega \times \Omega \mid \frac{f(y)}{g(y)} \geq \frac{f(x)}{g(x)} \text{ and } g(y) \geq g(x)\right\}$,*

$$\left(\frac{f(y)}{g(y)} - \frac{f(x)}{g(x)}\right)(g(y) - g(x)) \leq \kappa\|y-x\|_2^2.$$

Then, $\frac{f}{g}$ is a nonnegative, convex function on Ω.

Section 2 describes the inverse mean-ratio metric for tetrahedral elements, while Section 3 proves that this metric is invariant to even permutations of the input data. These permutations reduce the number of cases that need to be considered to only one case. We then show that the metric is a convex function of each coordinate in Section 4 by establishing that the conditions needed by Proposition 1.2 are satisfied. Finally, Section 5 proves that the block Jacobi preconditioner is positive definite by showing that the Hessian matrix is invertible and assembling all the results.

*This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy, under Contract W-31-109-Eng-38.

†Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL 60439; e-mail: `tmunson@mcs.anl.gov`

2 Mean-Ratio Description

The description in this section of the inverse mean-ratio metric [4] referenced to an ideal element follows that of Knupp [2, 3] and Freitag and Knupp [1].

Let (a, b, c, d) be the coordinates for the four vertices in a tetrahedral element, where each vertex is an element of $\Re^3$. We define the incidence matrix $A \in \Re^{3\times 3}$ as

$$A := \left[\begin{array}{ccc} b-a & c-a & d-a \end{array} \right],$$

and assume that $\det(A) > 0$. That is, the incidence matrix is obtained by computing the edges of the element emanating from the first vertex in the coordinate list and concatenating them into a square matrix. Since $\det(A) > 0$ by assumption, the volume of the element at the given coordinates is nonzero.

Furthermore, let (w_a, w_b, w_c, w_d) denote the four vertices for an ideal tetrahedral element, with W denoting the incidence matrix for this element. Moreover, assume that $\det(W) > 0$. Therefore, the ideal element has a nonzero volume and W^{-1} is guaranteed to exist. Throughout this technical report, we will denote

$$W^{-1} = \left[\begin{array}{ccc} \bar{w}_{1,1} & \bar{w}_{1,2} & \bar{w}_{1,3} \\ \bar{w}_{2,1} & \bar{w}_{2,2} & \bar{w}_{2,3} \\ \bar{w}_{3,1} & \bar{w}_{3,2} & \bar{w}_{3,3} \end{array} \right],$$

where $\bar{w}_{i,j}$ is the value for the (i, j) element of W^{-1}.

The quantity AW^{-1} is the identity matrix when the trial element and the ideal element have the same shape and size. If the trial element and the ideal element have the same shape but different sizes, then AW^{-1} is a positive multiple of the identity matrix, where the multiple is the scaling factor.

The inverse mean ratio is then defined as

$$\frac{\|AW^{-1}\|_F^2}{3|\det(AW^{-1})|^{\alpha}},$$

where $\alpha = \frac{2}{3}$. When the trial element and the ideal element have the same shape with a scaling factor of $\sigma > 0$, then the numerator has a value of $3\sigma^2$. This quantity is divided by a term related to the volume of the element in order to make the entire measure independent of scaling. Furthermore, the denominator has a value of $3\sigma^2$ when the trial and ideal elements have the same shape. The resulting quantity is a dimensionless measure of the shape of the trial element with respect to the ideal element. The range of the inverse mean ratio is between one and infinity, where a value greater than one means that the trial element and the ideal element have different shapes. This metric is invariant to scaling, translating, and rotating the input values. Our proofs will be for the general case, which is not always scale invariant, where $0 \le \alpha \le 1$ is arbitrary.

A mesh M is defined by a set of vertices V and the elements E that connect these vertices, where each element is an ordered set of four vertices. The set of vertices on the boundary of the mesh is denoted by ∂M; these vertices are fixed for the duration of the computation. We let $x \in \Re^{3\times|V|}$, where $|V|$ is the number of vertices in the mesh, and define

$$A_e(x) = \left[\begin{array}{ccc} x_{e_2} - x_{e_1} & x_{e_3} - x_{e_1} & x_{e_4} - x_{e_1} \end{array} \right] W^{-1},$$

where $e \in E$ with e_j denoting the jth vertex of element e, and x_i denotes the ith column of the coordinate matrix x. That is, $A_e(x)$ is the incidence matrix for element e times the inverse incidence matrix for the ideal element.

An optimization problem to minimize the average inverse mean ratio over the entire mesh is then

$$\begin{array}{lll} \min_{x\in\Re^{3\times|V|}} & \sum_{e\in E} \frac{\|A_e(x)\|_F^2}{3\det(A_e(x))^{\alpha}} & \\ \text{subject to} & \det(A_e(x)) > 0 & \forall e \in E \\ & x_i = \bar{x}_i & \forall i \in \partial M, \end{array}$$

where $\bar{x}_i$ denotes the fixed location of the ith boundary vertex. The constraints that $\det(A_e(x)) > 0$ for all $e \in E$ ensure that the elements in the resulting mesh have a consistent orientation. The absolute value in the denominator of the mean-ratio metric has been dropped because the consistent orientation constraints ensure that this quantity is positive. The consistent orientation conditions can be dropped from the problem if we start from a feasible point because the objective function approaches infinity as the volume of any element approaches zero. Moreover, the fixed variables can be removed. This reduction leads to an unconstrained optimization problem where the objective function is twice continuously differentiable on an open set containing the level set. An inexact Newton method can be applied to solve the resulting problem.

However, the objective function is not convex. Therefore, we need to show that the preconditioner used in the inexact Newton method is positive definite. In particular, given a feasible point for the optimization problem, we obtain a block Jacobi preconditioner by taking the Hessian of the objective function, $F(x)$, with respect to each of the vertices. That is,

$$M = \begin{bmatrix} \nabla^2_{x_1,x_1} F(x) & & & & \\ & \ddots & & & \\ & & \nabla^2_{x_i,x_i} F(x) & & \\ & & & \ddots & \\ & & & & \nabla^2_{x_{|V|},x_{|V|}} F(x) \end{bmatrix},$$

where $\nabla^2_{x_i,x_i} F(x) \in \Re^{3\times 3}$. To establish that this matrix is positive definite, we prove that $\nabla^2_{x_i,x_i} F(x)$ is positive definite for each $i = \{1, \ldots, |V|\}$.

To fix notation, we define the following functions:

$$\begin{aligned} A_a(x) &= \begin{bmatrix} b-x & c-x & d-x \end{bmatrix} W^{-1} \\ A_b(x) &= \begin{bmatrix} x-a & c-a & d-a \end{bmatrix} W^{-1} \\ A_c(x) &= \begin{bmatrix} b-a & x-a & d-a \end{bmatrix} W^{-1} \\ A_d(x) &= \begin{bmatrix} b-a & c-a & x-a \end{bmatrix} W^{-1} \end{aligned}$$

$$\begin{aligned} m_a(x) &= \frac{\|A_a(x)\|_F^2}{\det(A_a(x))^\alpha} \\ m_b(x) &= \frac{\|A_b(x)\|_F^2}{\det(A_b(x))^\alpha} \\ m_c(x) &= \frac{\|A_c(x)\|_F^2}{\det(A_c(x))^\alpha} \\ m_d(x) &= \frac{\|A_d(x)\|_F^2}{\det(A_d(x))^\alpha}. \end{aligned}$$

That is, $A_a(x)$ is the incidence matrix times the inverse incidence matrix for the ideal element as a function of the first vertex position, while $A_b(x)$, $A_c(x)$, and $A_d(x)$ are the corresponding functions for the second, third, and fourth vertex positions, respectively, while $m_a(x)$, $m_b(x)$, $m_c(x)$, and $m_d(x)$ are the resulting inverse mean-ratio functions. We also define the following sets:

$$\begin{aligned} \Omega_a &= \left\{x \in \Re^3 \mid \det(A_a(x)) > 0\right\} \\ \Omega_b &= \left\{x \in \Re^3 \mid \det(A_b(x)) > 0\right\} \\ \Omega_c &= \left\{x \in \Re^3 \mid \det(A_c(x)) > 0\right\} \\ \Omega_d &= \left\{x \in \Re^3 \mid \det(A_d(x)) > 0\right\}. \end{aligned}$$

3 Permutation Properties

The first step in proving that the block Jacobi preconditioner is positive definite is to show that the mean ratio metric for tetrahedral elements is invariant to an even permutation applied to the vertices for both the trial and ideal elements. The permutation needs to be even so that we do not change the sign of $\det(W)$. This invariance means that we need to prove convexity and positive definiteness for only one function since the others can be obtained by applying permutations.

Lemma 3.1 *Let* $(w^a, w^b, w^c, w^d) \in \Re^{3\times 4}$ *be given such that* $\det(W) > 0$*, and let* $(a, b, c, d) \in \Re^{3\times 4}$ *be arbitrary. Then, the following are equivalent:*

1. $\left[\begin{array}{ccc} b-a & c-a & d-a \end{array}\right] \left[\begin{array}{ccc} w^b-w^a & w^c-w^a & w^d-w^a \end{array}\right]^{-1}$
2. $\left[\begin{array}{ccc} c-b & a-b & d-b \end{array}\right] \left[\begin{array}{ccc} w^c-w^b & w^a-w^b & w^d-w^b \end{array}\right]^{-1}$
3. $\left[\begin{array}{ccc} c-a & d-a & b-a \end{array}\right] \left[\begin{array}{ccc} w^c-w^a & w^d-w^a & w^b-w^a \end{array}\right]^{-1}$

Proof:

$$\begin{aligned}
&\left[\begin{array}{ccc} b-a & c-a & d-a \end{array}\right] \left[\begin{array}{ccc} w^b-w^a & w^c-w^a & w^d-w^a \end{array}\right]^{-1} \\
&\quad = \left[\begin{array}{ccc} c-b & a-b & d-b \end{array}\right] \begin{bmatrix} 0 & 1 & 0 \\ -1 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \left(\left[\begin{array}{ccc} w^c-w^b & w^a-w^b & w^d-w^b \end{array}\right] \begin{bmatrix} 0 & 1 & 0 \\ -1 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \right)^{-1} \\
&\quad = \left[\begin{array}{ccc} c-b & a-b & d-b \end{array}\right] \begin{bmatrix} 0 & 1 & 0 \\ -1 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ -1 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \left[\begin{array}{ccc} w^c-w^b & w^a-w^b & w^d-w^b \end{array}\right]^{-1} \\
&\quad = \left[\begin{array}{ccc} c-b & a-b & d-a \end{array}\right] \left[\begin{array}{ccc} w^c-w^b & w^a-w^b & w^d-w^a \end{array}\right]^{-1}
\end{aligned}$$

$$\begin{aligned}
&\left[\begin{array}{ccc} b-a & c-a & d-a \end{array}\right] \left[\begin{array}{ccc} w^b-w^a & w^c-w^a & w^d-w^a \end{array}\right]^{-1} \\
&\quad = \left[\begin{array}{ccc} c-a & d-a & b-a \end{array}\right] \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \left(\left[\begin{array}{ccc} w^c-w^a & w^d-w^a & w^b-w^a \end{array}\right] \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \right)^{-1} \\
&\quad = \left[\begin{array}{ccc} c-a & d-a & b-a \end{array}\right] \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}^{-1} \left[\begin{array}{ccc} w^c-w^a & w^d-w^a & w^b-w^a \end{array}\right]^{-1} \\
&\quad = \left[\begin{array}{ccc} c-a & d-a & b-a \end{array}\right] \left[\begin{array}{ccc} w^c-w^a & w^d-w^a & w^b-w^a \end{array}\right]^{-1}.
\end{aligned}$$

□

4 Convexity Properties

This section proves that $m_d(x)$ is a convex function of x on Ω_d by establishing that the assumptions for Proposition 1.2 are satisfied.

Lemma 4.1 *For any weight matrix* W^{-1}*,*

1. $\det(A_d(x))$ *is a linear function of* x*.*
2. Ω_d *is a convex set.*

Proof: From the properties of the determinant,

$$\begin{aligned}
\det(A_d(x)) &= \det\left(\begin{bmatrix} b_1-a_1 & c_1-a_1 & x_1-a_1 \\ b_2-a_2 & c_2-a_2 & x_2-a_2 \\ b_3-a_3 & c_3-a_3 & x_3-a_3 \end{bmatrix} W^{-1} \right) \\
&= \det\left(\begin{bmatrix} b_1-a_1 & c_1-a_1 & x_1-a_1 \\ b_2-a_2 & c_2-a_2 & x_2-a_2 \\ b_3-a_3 & c_3-a_3 & x_3-a_3 \end{bmatrix} \right) \det\left(W^{-1}\right) \\
&= \left(\begin{array}{l} (x_1-a_1)((b_2-a_2)(c_3-a_3)-(b_3-a_3)(c_2-a_2)) + \\ (a_2-x_2)((b_1-a_1)(c_3-a_3)-(b_3-a_3)(c_1-a_1)) + \\ (x_3-a_3)((b_1-a_1)(c_2-a_2)-(b_2-a_2)(c_1-a_1)) \end{array} \right) \det\left(W^{-1}\right)
\end{aligned}$$

This calculation shows that $\det(A_d(x))$ is a linear function of x. Furthermore, Ω_d consists of a strict linear inequality, which forms a convex set. $\Box$

Lemma 4.2 *For any weight matrix W^{-1} and for any $0 \leq \alpha \leq 1$, $\det(A_d(x))^\alpha$ is a concave function of x on Ω_d.*

Proof: Since $\det(A_d(x))$ is a linear function of x by Lemma 4.1,

$$\det(A_d((1-\lambda)x + \lambda y)) = (1-\lambda)\det(A_d(x)) + \lambda \det(A_d(y))$$

for any $\lambda \in [0,1]$. If $x \in \Omega_d$ and $y \in \Omega_d$, then $(1-\lambda)x + \lambda y \in \Omega_d$ because Ω_d is a convex set by Lemma 4.1. Therefore, $\det(A_d((1-\lambda)x + \lambda y)) > 0$. The power is applied to both sides of the equation to obtain

$$\begin{aligned} \det(A_d((1-\lambda)x + \lambda y))^\alpha &= ((1-\lambda)\det(A_d(x)) + \lambda \det(A_d(y)))^\alpha \\ &\geq (1-\lambda)\det(A_d(x))^\alpha + \lambda \det(A_d(y))^\alpha, \end{aligned}$$

where the last inequality holds because ψ^α is a concave function on the region $\psi \geq 0$ for any $0 \leq \alpha \leq 1$ [7]. $\Box$

Lemma 4.3 *For any weight matrix W^{-1} with $\det(W^{-1}) > 0$, $\|A_d(x)\|_F^2$ is a uniformly convex function of x with constant $\kappa = \bar{w}_{3,1}^2 + \bar{w}_{3,2}^2 + \bar{w}_{3,3}^2 > 0$.*

Proof: The Hessian matrix for $\|A_d(x)\|_F$ is $2(\bar{w}_{3,1}^2 + \bar{w}_{3,2}^2 + \bar{w}_{3,3}^2)I$, where I is the identity matrix. Therefore, this matrix is uniformly positive definite with constant $2(\bar{w}_{3,1}^2 + \bar{w}_{3,2}^2 + \bar{w}_{3,3}^2)$. The relationship between equivalent definitions of uniform convexity then imply that

$$\|A_d((1-\lambda)x + \lambda y)\|_F^2 \leq (1-\lambda)\|A_d(x)\|_F^2 + \lambda\|A_d(y)\|_F^2 - \left(\bar{w}_{3,1}^2 + \bar{w}_{3,2}^2 + \bar{w}_{3,3}^2\right)\lambda(1-\lambda)\|y-x\|_2^2.$$

Since $\det(W^{-1}) > 0$, either $\bar{w}_{3,1} \neq 0$, $\bar{w}_{3,2} \neq 0$, or $\bar{w}_{3,3} \neq 0$. Hence, $\|A_d(x)\|_F^2$ is a uniformly convex function with $\kappa = \bar{w}_{3,1}^2 + \bar{w}_{3,2}^2 + \bar{w}_{3,3}^2 > 0$. $\Box$

Lemma 4.4 *Let W^{-1} be any weight matrix with $\det(W^{-1}) > 0$, and let $0 \leq \alpha \leq 1$. Moreover, define $\Omega = \Omega_d$, $f(x) = \|A_d(x)\|_F^2$, and $g(x) = \det(A_d(x))^\alpha$. Then for any $(x,y) \in \Theta$, where Θ is defined in Proposition 1.2,*

$$\left(\frac{f(y)}{g(y)} - \frac{f(x)}{g(x)}\right)(g(y) - g(x)) \leq (\bar{w}_{3,1}^2 + \bar{w}_{3,2}^2 + \bar{w}_{3,3}^2)\|y-x\|_2^2.$$

Proof: This lemma is proved by showing that

$$\sup_{(x,y)\in\Theta} \left(\frac{f(y)}{g(y)} - \frac{f(x)}{g(x)}\right)(g(y) - g(x)) - (\bar{w}_{3,1}^2 + \bar{w}_{3,2}^2 + \bar{w}_{3,3}^2)\|y-x\|_2^2 \leq 0.$$

If Θ is the empty set, then there is nothing to prove. Therefore, assume Θ is nonempty, and let $(x,y) \in \Theta$. We then make the change of variables $x = R\bar{x} + a$ and $y = R\bar{y} + a$, where R is an orthogonal matrix with $\det(R) = 1$. In particular, R is defined so that

$$R^T \left[\begin{array}{cc} b-a & c-a \end{array}\right] = \left[\begin{array}{cc} \bar{d}_{1,1} & \bar{d}_{1,2} \\ 0 & \bar{d}_{2,2} \\ 0 & 0 \end{array}\right],$$

where $\bar{d}_{1,1}$ and $\bar{d}_{2,2}$ are positive and $\bar{d}_{1,2}$ is unrestricted. Such a matrix can be constructed by multiplying three rotation matrices together and using the fact that $\det(A_d(x)) > 0$.

After making this substitution, we have the following expressions for $\|A_d(x)\|_F^2$, $\det(A_d(x))$, and $\|y - x\|_2^2$:

$$
\begin{aligned}
\|A_d(x)\|_F^2 &= \|A_d(R\bar{x} + a)W^{-1}\|_F^2 \\
&= \left\| \begin{bmatrix} b-a & c-a & R\bar{x}+a-a \end{bmatrix} W^{-1} \right\|_F^2 \\
&= \left\| R \begin{bmatrix} R^T(b-a) & R^T(c-a) & \bar{x} \end{bmatrix} W^{-1} \right\|_F^2 \\
&= \left\| \begin{bmatrix} \bar{d}_{1,1} & \bar{d}_{1,2} & \bar{x}_1 \\ 0 & \bar{d}_{2,2} & \bar{x}_2 \\ 0 & 0 & \bar{x}_3 \end{bmatrix} \begin{bmatrix} \bar{w}_{1,1} & \bar{w}_{1,2} & \bar{w}_{1,3} \\ \bar{w}_{2,1} & \bar{w}_{2,2} & \bar{w}_{2,3} \\ \bar{w}_{3,1} & \bar{w}_{3,2} & \bar{w}_{3,3} \end{bmatrix} \right\|_F^2
\end{aligned}
$$

$$
\begin{aligned}
\det(A_a(x)) &= \det\left(R \begin{bmatrix} \bar{d}_{1,1} & \bar{d}_{1,2} & \bar{x}_1 \\ 0 & \bar{d}_{2,2} & \bar{x}_2 \\ 0 & 0 & \bar{x}_3 \end{bmatrix} W^{-1} \right) \\
&= \bar{d}_{1,1}\bar{d}_{2,2}\det(W^{-1})\bar{x}_3
\end{aligned}
$$

$$
\begin{aligned}
\|y - x\|_2^2 &= \|R\bar{y} + a - R\bar{x} - a\|_2^2 \\
&= \|R(\bar{y} - \bar{x})\|_2^2 \\
&= \|\bar{y} - \bar{x}\|_2^2,
\end{aligned}
$$

where the orthogonality of R is used in the norm calculations and $\det(R) = 1$ is used in the determinant.

We use the following definitions throughout the remainder of this section.

$$
\begin{aligned}
\nu &= \bar{d}_{1,1}\bar{d}_{2,2}\det(W^{-1}) \\
\Delta &= \bar{w}_{3,1}^2 + \bar{w}_{3,2}^2 + \bar{w}_{3,3}^2 \\
\delta_1(\xi) &= \left(\begin{array}{l} (\bar{d}_{1,1}\bar{w}_{1,1} + \bar{d}_{1,2}\bar{w}_{2,1} + \bar{w}_{3,1}\xi)^2 + \\ (\bar{d}_{1,1}\bar{w}_{1,2} + \bar{d}_{1,2}\bar{w}_{2,2} + \bar{w}_{3,2}\xi)^2 + \\ (\bar{d}_{1,1}\bar{w}_{1,3} + \bar{d}_{1,2}\bar{w}_{2,3} + \bar{w}_{3,3}\xi)^2 \end{array} \right) \\
\delta_2(\xi) &= (\bar{d}_{2,2}\bar{w}_{2,1} + \bar{w}_{3,1}\xi)^2 + (\bar{d}_{2,2}\bar{w}_{2,2} + \bar{w}_{3,2}\xi)^2 + (\bar{d}_{2,2}\bar{w}_{2,3} + \bar{w}_{3,3}\xi)^2
\end{aligned}
$$

The optimization problem we want to solve is then

$$
\begin{array}{ll}
\sup_{\bar{x}\in\Re^3, \bar{y}\in\Re^3} & \left(\frac{\delta_1(\bar{y}_1)+\delta_2(\bar{y}_2)+\Delta\bar{y}_3^2}{(\nu\bar{y}_3)^\alpha} - \frac{\delta_1(\bar{x}_1)+\delta_2(\bar{x}_2)+\Delta\bar{x}_3^2}{(\nu\bar{x}_3)^\alpha} \right) ((\nu\bar{y}_3)^\alpha - (\nu\bar{x}_3)^\alpha) - \Delta\|\bar{y} - \bar{x}\|_2^2 \\
\text{subject to} & \nu\bar{y}_3 \geq \nu\bar{x}_3 > 0 \\
& (\nu\bar{y}_3)^\alpha \geq (\nu\bar{x}_3)^\alpha \\
& \frac{\delta_1(\bar{y}_1)+\delta_2(\bar{y}_2)+\Delta\bar{y}_3^2}{(\nu\bar{y}_3)^\alpha} \geq \frac{\delta_1(\bar{x}_1)+\delta_2(\bar{x}_2)+\Delta\bar{x}_3^2}{(\nu\bar{x}_3)^\alpha}.
\end{array}
$$

Eliminating ν from the problem because it is a positive constant, and dropping the last two constraints, we obtain the following optimization problem, which provides an upper bound on the supremum:

$$
\begin{array}{ll}
\sup_{\bar{x}\in\Re^3, \bar{y}\in\Re^3} & \left(\frac{\delta_1(\bar{y}_1)+\delta_2(\bar{y}_2)+\Delta\bar{y}_3^2}{\bar{y}_3^\alpha} - \frac{\delta_1(\bar{x}_1)+\delta_2(\bar{x}_2)+\Delta\bar{x}_3^2}{\bar{x}_3^\alpha} \right) (\bar{y}_3^\alpha - \bar{x}_3^\alpha) - \Delta\|\bar{y} - \bar{x}\|_2^2 \\
\text{subject to} & \bar{y}_3 \geq \bar{x}_3 > 0
\end{array}
$$

Examining those terms involving $\bar{y}_3^2$ and $\bar{x}_3^2$, we obtain

$$
\begin{aligned}
\Delta((\bar{y}_3^{2-\alpha} - \bar{x}_3^{2-\alpha})(\bar{y}_3^\alpha - \bar{x}_3^\alpha) - (\bar{y}_3 - \bar{x}_3)^2) &= \Delta(\bar{y}_3^2 - \bar{y}_3^{2-\alpha}\bar{x}_3^\alpha - \bar{y}_3^\alpha\bar{x}_3^{2-\alpha} + \bar{x}_3^2 - (\bar{y}_3 - \bar{x}_3)^2) \\
&= \Delta(2\bar{y}_3\bar{x}_3 - \bar{y}_3^{2-\alpha}\bar{x}_3^\alpha - \bar{y}_3^\alpha\bar{x}_3^{2-\alpha}) \\
&= \Delta\bar{y}_3\bar{x}_3\left(2 - \frac{\bar{y}_3^{1-\alpha}}{\bar{x}_3^{1-\alpha}} - \frac{\bar{x}_3^{1-\alpha}}{\bar{y}_3^{1-\alpha}}\right) \\
&\leq 0,
\end{aligned}
$$

where the last inequality is obtained from the arithmetic-geometric mean inequality. After removing these terms, we are left with the following optimization problem, which provides an upper bound on the supremum:

$$\begin{array}{ll} \sup_{\bar{x}\in\Re^3,\bar{y}\in\Re^3} & \left(\frac{\delta_1(\bar{y}_1)+\delta_2(\bar{y}_2)}{\bar{y}_3^\alpha} - \frac{\delta_1(\bar{x}_1)+\delta_2(\bar{x}_2)}{\bar{x}_3^\alpha}\right)(\bar{y}_3^\alpha - \bar{x}_3^\alpha) - \Delta((\bar{y}_1-\bar{x}_1)^2 + (\bar{y}_2-\bar{x}_2)^2) \\ \text{subject to} & \bar{y}_3 \geq \bar{x}_3 > 0. \end{array}$$

We now write $\bar{y}_3^\alpha = \beta\bar{x}_3^\alpha$ for $\beta \geq 1$ because $\bar{x}_3^\alpha > 0$, eliminate $\bar{y}_3$ and $\bar{x}_3$, and rearrange the terms to obtain the equivalent optimization problem:

$$\sup_{\bar{y}_1\in\Re,\bar{y}_2\in\Re,\beta\geq 1}\sup_{\bar{x}_1\in\Re,\bar{x}_2\in\Re} \quad \frac{1}{\beta}\left(\begin{array}{l} (\beta-1)(\delta_1(\bar{y}_1)+\delta_2(\bar{y}_2)) - \\ \beta(\beta-1)(\delta_1(\bar{x}_1)+\delta_2(\bar{x}_2)) - \\ \beta\Delta((\bar{y}_1-\bar{x}_1)^2 + (\bar{y}_2-\bar{x}_2)^2) \end{array}\right).$$

Note that the objective function is strongly concave in the $\bar{x}_1$ and $\bar{x}_2$ variables. Therefore, we can set to zero the gradient of the objective function with respect to $\bar{x}_1$ and $\bar{x}_2$ to derive the optimal solution for $\bar{x}_1$ and $\bar{x}_2$ given $\bar{y}_1$, $\bar{y}_2$, and β.

$$\begin{aligned}
\nabla_{\bar{x}_1}\mathrm{obj}(\bar{x}_1,\bar{x}_2,\bar{y}_1,\bar{y}_2,\beta) &= -2\left((\beta-1)\left(\begin{array}{l} (\bar{d}_{1,1}\bar{w}_{1,1}+\bar{d}_{1,2}\bar{w}_{2,1}+\bar{w}_{3,1}\bar{x}_1)\bar{w}_{3,1} + \\ (\bar{d}_{1,1}\bar{w}_{1,2}+\bar{d}_{1,2}\bar{w}_{2,2}+\bar{w}_{3,2}\bar{x}_1)\bar{w}_{3,2} + \\ (\bar{d}_{1,1}\bar{w}_{1,3}+\bar{d}_{1,2}\bar{w}_{2,3}+\bar{w}_{3,3}\bar{x}_1)\bar{w}_{3,3} \end{array}\right) + \Delta(\bar{x}_1-\bar{y}_1)\right) \\
&= -2\left((\beta-1)\left(\Delta\bar{x}_1 + \left(\begin{array}{l} (\bar{d}_{1,1}\bar{w}_{1,1}+\bar{d}_{1,2}\bar{w}_{2,1})\bar{w}_{3,1} + \\ (\bar{d}_{1,1}\bar{w}_{1,2}+\bar{d}_{1,2}\bar{w}_{2,2})\bar{w}_{3,2} + \\ (\bar{d}_{1,1}\bar{w}_{1,3}+\bar{d}_{1,2}\bar{w}_{2,3})\bar{w}_{3,3} \end{array}\right)\right) + \Delta(\bar{x}_1-\bar{y}_1)\right) \\
&= -2\left(\beta\Delta\bar{x}_1 + (\beta-1)\left(\begin{array}{l} (\bar{d}_{1,1}\bar{w}_{1,1}+\bar{d}_{1,2}\bar{w}_{2,1})\bar{w}_{3,1} + \\ (\bar{d}_{1,1}\bar{w}_{1,2}+\bar{d}_{1,2}\bar{w}_{2,2})\bar{w}_{3,2} + \\ (\bar{d}_{1,1}\bar{w}_{1,3}+\bar{d}_{1,2}\bar{w}_{2,3})\bar{w}_{3,3} \end{array}\right) - \Delta\bar{y}_1\right) \\
&= -2\beta\Delta\left(\bar{x}_1 + \frac{\beta-1}{\beta\Delta}\left(\begin{array}{l} (\bar{d}_{1,1}\bar{w}_{1,1}+\bar{d}_{1,2}\bar{w}_{2,1})\bar{w}_{3,1} + \\ (\bar{d}_{1,1}\bar{w}_{1,2}+\bar{d}_{1,2}\bar{w}_{2,2})\bar{w}_{3,2} + \\ (\bar{d}_{1,1}\bar{w}_{1,3}+\bar{d}_{1,2}\bar{w}_{2,3})\bar{w}_{3,3} \end{array}\right) - \frac{\bar{y}_1}{\beta}\right)
\end{aligned}$$

$$\begin{aligned}
\nabla_{\bar{x}_2}\mathrm{obj}(\bar{x}_1,\bar{x}_2,\bar{y}_1,\bar{y}_2,\beta) &= -2\left((\beta-1)\left(\begin{array}{l} (\bar{d}_{2,2}\bar{w}_{2,1}+\bar{w}_{3,1}\bar{x}_2)\bar{w}_{3,1} + \\ (\bar{d}_{2,2}\bar{w}_{2,2}+\bar{w}_{3,2}\bar{x}_2)\bar{w}_{3,2} + \\ (\bar{d}_{2,2}\bar{w}_{2,3}+\bar{w}_{3,3}\bar{x}_2)\bar{w}_{3,3} \end{array}\right) + \Delta(\bar{x}_2-\bar{y}_2)\right) \\
&= -2\left((\beta-1)\left(\Delta\bar{x}_2 + \left(\begin{array}{l} (\bar{d}_{2,2}\bar{w}_{2,1})\bar{w}_{3,1} + \\ (\bar{d}_{2,2}\bar{w}_{2,2})\bar{w}_{3,2} + \\ (\bar{d}_{2,2}\bar{w}_{2,3})\bar{w}_{3,3} \end{array}\right)\right) + \Delta(\bar{x}_2-\bar{y}_2)\right) \\
&= -2\left(\beta\Delta\bar{x}_2 + (\beta-1)\left(\begin{array}{l} (\bar{d}_{2,2}\bar{w}_{2,1})\bar{w}_{3,1} + \\ (\bar{d}_{2,2}\bar{w}_{2,2})\bar{w}_{3,2} + \\ (\bar{d}_{2,2}\bar{w}_{2,3})\bar{w}_{3,3} \end{array}\right) - \Delta\bar{y}_2\right) \\
&= -2\beta\Delta\left(\bar{x}_2 + \frac{\beta-1}{\beta\Delta}\left(\begin{array}{l} (\bar{d}_{2,2}\bar{w}_{2,1})\bar{w}_{3,1} + \\ (\bar{d}_{2,2}\bar{w}_{2,2})\bar{w}_{3,2} + \\ (\bar{d}_{2,2}\bar{w}_{2,3})\bar{w}_{3,3} \end{array}\right) - \frac{\bar{y}_2}{\beta}\right).
\end{aligned}$$

Let

$$\begin{aligned}
\Lambda_1 &= (\bar{d}_{1,1}\bar{w}_{1,1}+\bar{d}_{1,2}\bar{w}_{2,1})\bar{w}_{3,1} + (\bar{d}_{1,1}\bar{w}_{1,2}+\bar{d}_{1,2}\bar{w}_{2,2})\bar{w}_{3,2} + (\bar{d}_{1,1}\bar{w}_{1,3}+\bar{d}_{1,2}\bar{w}_{2,3})\bar{w}_{3,3} \\
\Lambda_2 &= (\bar{d}_{2,2}\bar{w}_{2,1})\bar{w}_{3,1} + (\bar{d}_{2,2}\bar{w}_{2,2})\bar{w}_{3,2} + (\bar{d}_{2,2}\bar{w}_{2,3})\bar{w}_{3,3}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\bar{x}_1 &= \frac{\bar{y}_1}{\beta} - \frac{\beta-1}{\beta}\frac{\Lambda_1}{\Delta} \\
\bar{x}_2 &= \frac{\bar{y}_2}{\beta} - \frac{\beta-1}{\beta}\frac{\Lambda_2}{\Delta}.
\end{aligned}$$

Substituting these quantities into the objective function gives an equivalent optimization problem

$$\sup_{\bar y_1\in\Re,\bar y_2\in\Re,\beta\ge 1}\quad \frac{1}{\beta}\left(\begin{array}{l}(\beta-1)(\delta_1(\bar y_1)+\delta_2(\bar y_2))-\\ \beta(\beta-1)\left(\delta_1\left(\frac{\bar y_1}{\beta}-\frac{\beta-1}{\beta}\frac{\Lambda_1}{\Delta}\right)+\delta_2\left(\frac{\bar y_2}{\beta}-\frac{\beta-1}{\beta}\frac{\Lambda_2}{\Delta}\right)\right)-\\ \beta\Delta\left(\left(\bar y_1-\frac{\bar y_1}{\beta}-\frac{\beta-1}{\beta}\frac{\Lambda_1}{\Delta}\right)^2+\left(\bar y_2-\frac{\bar y_2}{\beta}-\frac{\beta-1}{\beta}\frac{\Lambda_2}{\Delta}\right)^2\right)\end{array}\right).$$

Examining those terms involving $\bar y_1^2$ and β, we obtain

$$\begin{aligned}
\frac{1}{\beta}\left(\begin{array}{l}(\beta-1)\delta_1(\bar y_1)-\\ \beta(\beta-1)\left(\delta_1\left(\frac{\bar y_1}{\beta}-\frac{\beta-1}{\beta}\frac{\Lambda_1}{\Delta}\right)\right)-\\ \beta\Delta\left(\left(\bar y_1-\frac{\bar y_1}{\beta}-\frac{\beta-1}{\beta}\frac{\Lambda_1}{\Delta}\right)^2\right)\end{array}\right)
&= -\frac{(\beta-1)^2}{\beta\Delta}\left(\begin{array}{l}((\bar d_{1,1}\bar w_{1,1}+\bar d_{1,2}\bar w_{2,1})\bar w_{3,2}-(\bar d_{1,1}\bar w_{1,2}+\bar d_{1,2}\bar w_{2,2})\bar w_{3,1})^2+\\ ((\bar d_{1,1}\bar w_{1,1}+\bar d_{1,2}\bar w_{2,1})\bar w_{3,3}-(\bar d_{1,1}\bar w_{1,3}+\bar d_{1,2}\bar w_{2,3})\bar w_{3,1})^2+\\ ((\bar d_{1,1}\bar w_{1,2}+\bar d_{1,2}\bar w_{2,2})\bar w_{3,3}-(\bar d_{1,1}\bar w_{1,3}+\bar d_{1,2}\bar w_{2,3})\bar w_{3,2})^2\end{array}\right)\\
&= -\frac{(\beta-1)^2}{\beta\Delta}\left(\begin{array}{l}\det\left(\begin{bmatrix}\bar w_{1,1}&\bar w_{1,2}&-\bar d_{1,2}\\ \bar w_{2,1}&\bar w_{2,2}&\bar d_{1,1}\\ \bar w_{3,1}&\bar w_{3,2}&0\end{bmatrix}\right)+\\ \det\left(\begin{bmatrix}\bar w_{1,1}&-\bar d_{1,2}&\bar w_{1,3}\\ \bar w_{2,1}&\bar d_{1,1}&\bar w_{2,3}\\ \bar w_{3,1}&0&\bar w_{3,3}\end{bmatrix}\right)+\\ \det\left(\begin{bmatrix}-\bar d_{1,2}&\bar w_{1,2}&\bar w_{1,3}\\ \bar d_{1,1}&\bar w_{2,2}&\bar w_{2,3}\\ 0&\bar w_{3,2}&\bar w_{3,3}\end{bmatrix}\right)\end{array}\right)\\
&= -\frac{(\beta-1)^2}{\beta}\frac{\det(W^{-1})^2}{\Delta}\left\|W\begin{bmatrix}-\bar d_{1,2}\\ \bar d_{1,1}\\ 0\end{bmatrix}\right\|_2^2,
\end{aligned}$$

where the last equation is derived from Cramer's rule.

Examining those terms involving $\bar y_2^2$ and β, we obtain

$$\begin{aligned}
\frac{1}{\beta}\left(\begin{array}{l}(\beta-1)\delta_2(\bar y_2)-\\ \beta(\beta-1)\left(\delta_2\left(\frac{\bar y_2}{\beta}-\frac{\beta-1}{\beta}\frac{\Lambda_2}{\Delta}\right)\right)-\\ \beta\Delta\left(\left(\bar y_2-\frac{\bar y_2}{\beta}-\frac{\beta-1}{\beta}\frac{\Lambda_2}{\Delta}\right)^2\right)\end{array}\right)
&= -\frac{(\beta-1)^2}{\beta\Delta}\left(\begin{array}{l}(\bar d_{2,2}\bar w_{2,1}\bar w_{3,2}-\bar d_{2,2}\bar w_{3,1}\bar w_{2,2})^2+\\ (\bar d_{2,2}\bar w_{2,1}\bar w_{3,3}-\bar d_{2,2}\bar w_{3,1}\bar w_{2,3})^2+\\ (\bar d_{2,2}\bar w_{2,2}\bar w_{3,3}-\bar d_{2,2}\bar w_{3,2}\bar w_{2,3})^2\end{array}\right)\\
&= -\frac{(\beta-1)^2}{\beta\Delta}\left(\begin{array}{l}\det\left(\begin{bmatrix}\bar w_{1,1}&\bar w_{1,2}&\bar d_{2,2}\\ \bar w_{2,1}&\bar w_{2,2}&0\\ \bar w_{3,1}&\bar w_{3,2}&0\end{bmatrix}\right)+\\ \det\left(\begin{bmatrix}\bar w_{1,1}&\bar d_{2,2}&\bar w_{1,3}\\ \bar w_{2,1}&0&\bar w_{2,3}\\ \bar w_{3,1}&0&\bar w_{3,3}\end{bmatrix}\right)+\\ \det\left(\begin{bmatrix}\bar d_{2,2}&\bar w_{1,2}&\bar w_{1,3}\\ 0&\bar w_{2,2}&\bar w_{2,3}\\ 0&\bar w_{3,2}&\bar w_{3,3}\end{bmatrix}\right)\end{array}\right)\\
&= -\frac{(\beta-1)^2}{\beta}\frac{\det(W^{-1})^2}{\Delta}\left\|W\begin{bmatrix}\bar d_{2,2}\\ 0\\ 0\end{bmatrix}\right\|_2^2,
\end{aligned}$$

where the last equation is derived from Cramer's rule.

Therefore, the optimization problem reduces to

$$\sup_{\beta\ge 1}\quad -\frac{(\beta-1)^2}{\beta}\frac{\det(W^{-1})^2}{\Delta}\left(\left\|W\begin{bmatrix}-\bar d_{1,2}\\ \bar d_{1,1}\\ 0\end{bmatrix}\right\|_2^2+\left\|W\begin{bmatrix}\bar d_{2,2}\\ 0\\ 0\end{bmatrix}\right\|_2^2\right).$$

The constant in this optimization problem is positive since the null space of W is the zero vector, $\det(W^{-1}) > 0$, and $\Delta > 0$. Hence, the superemum is zero. $\Box$

Theorem 4.5 *Let W^{-1} be any weight matrix with $\det(W^{-1}) > 0$, and let $0 \leq \alpha \leq 1$. Then, $m_d(x)$ is a nonnegative, convex function of x on Ω_d.*

Proof: Lemma 4.1 shows that Ω_d is a convex set, Lemma 4.2 demonstrates that $\det(A_d(x))^\alpha$ is a concave function on Ω_d, and $\det(A_d(x))^\alpha > 0$ for any $x \in \Omega_d$. Therefore, Property 1 of Proposition 1.2 is satisfied. Furthermore, $\|A_d(x)\|_F^2$ is a nonnegative function of x and Lemma 4.3 shows that $\|A_d(x)\|_F^2$ is a uniformly convex function of x with constant $\bar{w}_{3,1}^2 + \bar{w}_{3,2}^2 + \bar{w}_{3,3}^2 > 0$. Therefore, Property 2 of Proposition 1.2 holds. Lemma 4.4 shows that Property 3 is satisfied. Therefore, by Proposition 1.2, $m_d(x)$ is a nonnegative, convex function of x on Ω_d. $\Box$

5 Preconditioner Properties

Lemma 5.1 *Let W^{-1} be any weight matrix with $\det(W^{-1}) > 0$, and let $0 \leq \alpha \leq 1$. Then for any $x \in \Omega_d$, $\nabla^2_{x,x} m_d(x)$ is invertible.*

Proof: If Ω_d is empty, then there is nothing to prove. Therefore, let Ω_d be nonempty, and let $x \in \Omega_d$. Define R to be as in the proof of Lemma 4.4. Then we have

$$m_d(x) = m_d(R\bar{x} + a)$$

where $\bar{x} = R^T(x - a)$. By the chain rule,

$$\nabla_x m_d(x) = \left[\nabla_{\bar{x}} m_d(R\bar{x} + a)\right] R^T$$

and

$$\nabla^2_{x,x} m_d(x) = R\left[\nabla^2_{\bar{x},\bar{x}} m_d(R\bar{x} + a)\right] R^T.$$

We can ignore the terms involving R, since R is an orthogonal matrix with $\det(R) = 1$.

Using the definitions from Lemma 4.4, we recall that

$$\begin{aligned} m_d(R\bar{x}+a) &= \frac{\left\| \begin{bmatrix} \bar{d}_{1,1} & \bar{d}_{1,2} & \bar{x}_1 \\ 0 & \bar{d}_{2,2} & \bar{x}_2 \\ 0 & 0 & \bar{x}_3 \end{bmatrix} \begin{bmatrix} \bar{w}_{1,1} & \bar{w}_{1,2} & \bar{w}_{1,3} \\ \bar{w}_{2,1} & \bar{w}_{2,2} & \bar{w}_{2,3} \\ \bar{w}_{3,1} & \bar{w}_{3,2} & \bar{w}_{3,3} \end{bmatrix} \right\|_F^2}{\det\left(\begin{bmatrix} \bar{d}_{1,1} & \bar{d}_{1,2} & \bar{x}_1 \\ 0 & \bar{d}_{2,2} & \bar{x}_2 \\ 0 & 0 & \bar{x}_3 \end{bmatrix} W^{-1} \right)^\alpha} \\ &= \frac{\left\| \begin{bmatrix} \bar{d}_{1,1} & \bar{d}_{1,2} & \bar{x}_1 \\ 0 & \bar{d}_{2,2} & \bar{x}_2 \\ 0 & 0 & \bar{x}_3 \end{bmatrix} \begin{bmatrix} \hat{w}_{1,1} & \hat{w}_{1,2} & \hat{w}_{1,3} \\ 0 & \hat{w}_{2,2} & \hat{w}_{2,3} \\ 0 & 0 & \hat{w}_{3,3} \end{bmatrix} \right\|_F^2}{(\nu\bar{x}_3)^\alpha}, \end{aligned}$$

where we have applied an orthogonal matrix S with $\det(S) = 1$ to the right-hand side of the numerator to eliminate constants. Moreover, $\bar{d}_{1,1} > 0$, $\bar{d}_{2,2} > 0$, and $\bar{x}_3 > 0$ because $x \in \Omega_d$, and $\hat{w}_{1,1}\hat{w}_{2,2}\hat{w}_{3,3} > 0$ because $\det(W^{-1}) > 0$.

Let

$$\begin{aligned} \phi_1 &= \bar{d}_{1,1}\hat{w}_{1,1} \\ \phi_2 &= \bar{d}_{2,2}\hat{w}_{2,2} \\ \rho &= \bar{d}_{1,1}\hat{w}_{1,2} + \bar{d}_{1,2}\hat{w}_{2,2} \\ \\ \omega_1(\xi) &= \bar{d}_{1,1}\hat{w}_{1,3} + \bar{d}_{1,2}\hat{w}_{2,3} + \hat{w}_{3,3}\xi \\ \omega_2(\xi) &= \bar{d}_{2,2}\hat{w}_{2,3} + \hat{w}_{3,3}\xi. \end{aligned}$$

We then have by direct computation that the Hessian matrix is

$$\begin{bmatrix} \frac{2\hat{w}_{3,3}^2}{\nu^\alpha \bar{x}_3^\alpha} & 0 & -\frac{2\alpha\hat{w}_{3,3}\omega_1(\bar{x}_1)}{\nu^\alpha \bar{x}_3^{\alpha+1}} \\ 0 & \frac{2\hat{w}_{3,3}^2}{\nu^\alpha \bar{x}_3^\alpha} & -\frac{2\alpha\hat{w}_{3,3}\omega_2(\bar{x}_2)}{\nu^\alpha \bar{x}_3^{\alpha+1}} \\ -\frac{2\alpha\hat{w}_{3,3}\omega_1(\bar{x}_1)}{\nu^\alpha \bar{x}_3^{\alpha+1}} & -\frac{2\alpha\hat{w}_{3,3}\omega_2(\bar{x}_2)}{\nu^\alpha \bar{x}_3^{\alpha+1}} & \frac{\alpha(\alpha+1)(\phi_1^2+\phi_2^2+\rho^2+\omega_1(\bar{x}_1)^2+\omega_2(\bar{x}_2)^2)}{\nu^\alpha \bar{x}_3^{\alpha+2}} + \frac{(1-\alpha)(2-\alpha)\hat{w}_{3,3}^2}{\nu^\alpha \bar{x}_3^\alpha} \end{bmatrix}$$

We now compute the null space of this matrix by solving the system of equations:

$$\begin{bmatrix} \frac{2\hat{w}_{3,3}^2}{\nu^\alpha \bar{x}_3^\alpha} & 0 & -\frac{2\alpha\hat{w}_{3,3}\omega_1(\bar{x}_1)}{\nu^\alpha \bar{x}_3^{\alpha+1}} \\ 0 & \frac{2\hat{w}_{3,3}^2}{\nu^\alpha \bar{x}_3^\alpha} & -\frac{2\alpha\hat{w}_{3,3}\omega_2(\bar{x}_2)}{\nu^\alpha \bar{x}_3^{\alpha+1}} \\ -\frac{2\alpha\hat{w}_{3,3}\omega_1(\bar{x}_1)}{\nu^\alpha \bar{x}_3^{\alpha+1}} & -\frac{2\alpha\hat{w}_{3,3}\omega_2(\bar{x}_2)}{\nu^\alpha \bar{x}_3^{\alpha+1}} & \frac{\alpha(\alpha+1)(\phi_1^2+\phi_2^2+\rho^2+\omega_1(\bar{x}_1)^2+\omega_2(\bar{x}_2)^2)}{\nu^\alpha \bar{x}_3^{\alpha+2}} + \frac{(1-\alpha)(2-\alpha)\hat{w}_{3,3}^2}{\nu^\alpha \bar{x}_3^\alpha} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = 0.$$

The first and second constraints imply

$$\begin{aligned} z_1 &= \frac{\alpha\omega_1(\bar{x}_1)}{\hat{w}_{3,3}\bar{x}_3} z_3 \\ z_2 &= \frac{\alpha\omega_2(\bar{x}_1)}{\hat{w}_{3,3}\bar{x}_3} z_3. \end{aligned}$$

Substituting these into the third equation gives a coefficient on z_3 of

$$\begin{aligned} &\frac{\alpha(\alpha+1)(\phi_1^2+\phi_2^2+\rho^2+\omega_1(\bar{x}_1)^2+\omega_2(\bar{x}_2)^2)}{\nu^\alpha \bar{x}_3^{\alpha+2}} + \frac{(1-\alpha)(2-\alpha)\hat{w}_{3,3}^2}{\nu^\alpha \bar{x}_3^\alpha} - \frac{2\alpha\hat{w}_{3,3}\omega_1(\bar{x}_1)}{\nu^\alpha \bar{x}_3^{\alpha+1}}\frac{\alpha\omega_1(\bar{x}_1)}{\hat{w}_{3,3}\bar{x}_3} - \frac{2\alpha\hat{w}_{3,3}\omega_2(\bar{x}_1)}{\nu^\alpha \bar{x}_3^{\alpha+1}}\frac{\alpha\omega_2(\bar{x}_2)}{\hat{w}_{3,3}\bar{x}_3} \\ &= \frac{\alpha(\alpha+1)(\phi_1^2+\phi_2^2+\rho^2+\omega_1(\bar{x}_1)^2+\omega_2(\bar{x}_2)^2)}{\nu^\alpha \bar{x}_3^{\alpha+2}} + \frac{(1-\alpha)(2-\alpha)\hat{w}_{3,3}^2}{\nu^\alpha \bar{x}_3^\alpha} - \frac{2\alpha^2\omega_1(\bar{x}_1)^2}{\nu^\alpha \bar{x}_3^{\alpha+2}} - \frac{2\alpha^2\omega_2(\bar{x}_1)^2}{\nu^\alpha \bar{x}_3^{\alpha+2}} \\ &= \frac{\alpha(1+\alpha)(\phi_1^2+\phi_2^2+\rho^2)}{\nu^\alpha \bar{x}_3^{\alpha+2}} + \frac{\alpha(1-\alpha)(\omega_1(\bar{x}_1)^2+\omega_2(\bar{x}_2)^2)}{\nu^\alpha \bar{x}_3^{\alpha+2}} + \frac{(1-\alpha)(2-\alpha)\hat{w}_{3,3}^2}{\nu^\alpha \bar{x}_3^\alpha}. \end{aligned}$$

The first term is positive whenever $0 < \alpha \leq 1$ since $\phi_1 \neq 0$, $\phi_2 \neq 0$, $\nu > 0$, and $\bar{x}_3 > 0$. The second term is nonnegative whenever $0 \leq \alpha \leq 1$. The third term is positive whenever $0 \leq \alpha < 1$ since $\hat{w}_{3,3} \neq 0$, $\nu > 0$, and $\bar{x}_3 > 0$. Therefore, this coefficient is positive and $z_3 = 0$, which implies that $z_1 = 0$ and $z_2 = 0$. Since the null space is the zero vector, the matrix is invertible. □

Corollary 5.2 *Let W^{-1} be any weight matrix with* $\det(W^{-1}) > 0$, *and let $0 \leq \alpha \leq 1$. Then for any $x \in \Omega_d$, $\nabla^2_{x,x} m_d(x)$ is positive definite.*

Proof: Since $m_d(x)$ is convex on Ω_d by Theorem 4.5 and twice continuously differentiable on this set, $\nabla^2_{x,x} m_d(x)$ is positive semidefinite for any $x \in \Omega_d$. However, Lemma 5.1 shows that the Hessian matrix is also invertible for any $x \in \Omega_d$. Therefore, we conclude that $\nabla^2_{x,x} m_d(x)$ is positive definite for any $x \in \Omega_d$. □

Corollary 5.3 *Let W^{-1} be any weight matrix with* $\det(W^{-1}) > 0$, *and let $0 \leq \alpha \leq 1$. Then,*

1. *$\nabla^2_{x,x} m_a(x)$ is positive definite for any $x \in \Omega_a$.*
2. *$\nabla^2_{x,x} m_b(x)$ is positive definite for any $x \in \Omega_b$.*
3. *$\nabla^2_{x,x} m_c(x)$ is positive definite for any $x \in \Omega_c$.*

Proof: By Lemma 3.1

$$\begin{aligned} A_a(x) &= \begin{bmatrix} b-x & c-x & d-x \end{bmatrix}\begin{bmatrix} w^b-w^a & w^c-w^a & w^d-w^a \end{bmatrix}^{-1} \\ &= \begin{bmatrix} c-b & x-b & d-b \end{bmatrix}\begin{bmatrix} w^c-w^b & w^a-w^b & w^d-w^b \end{bmatrix}^{-1} \\ &= \begin{bmatrix} d-b & c-b & x-b \end{bmatrix}\begin{bmatrix} w^d-w^b & w^c-w^b & w^a-w^b \end{bmatrix}^{-1}. \end{aligned}$$

We now apply Corollary 5.2 to $m_a(x)$ using this equivalent redefinition. Therefore, $\nabla^2_{x,x} m_a(x)$ is positive definite for any $x \in \Omega_a$. A similar argument using applications of Lemma 3.1 shows that $\nabla^2_{x,x} m_b(x)$ is positive definite for any $x \in \Omega_b$ and $\nabla^2_{x,x} m_c(x)$ is positive definite for any $x \in \Omega_c$. $\square$

Corollary 5.4 *Let W^{-1} be any weight matrix with* $\det(W^{-1}) > 0$*, and let $x \in \Re^{3\times|V|}$ be given such that* $\det(A_e(x)) > 0$ *for all $e \in E$. Then, the block Jacobi preconditioner for the shape-quality optimization problem using the mean-ratio metric is positive definite.*

Proof: The objective function, $F(x)$, for the shape-quality optimization problem consists of the sum of the mean-ratio metric for each element. Therefore,

$$\nabla^2_{x_i,x_i} F(x) = \left(\begin{array}{c} \sum_{\{e\in E|e_1=i\}} \nabla^2_{x_i,x_i} m_a(x_i) + \sum_{\{e\in E|e_2=i\}} \nabla^2_{x_i,x_i} m_b(x_i) + \\ \sum_{\{e\in E|e_3=i\}} \nabla^2_{x_i,x_i} m_c(x_i) + \sum_{\{e\in E|e_4=i\}} \nabla^2_{x_i,x_i} m_d(x_i) \end{array} \right),$$

where $a = x_{e_1}$, $b = x_{e_2}$, $c = x_{e_3}$, and $d = x_{e_4}$ in the mean ratio metric for each $e \in E$. Corollary 5.2 and Corollary 5.3 now imply that $\nabla^2_{x_i,x_i} F(x)$ is positive definite because the sum of positive definite matrices is also positive definite. Therefore, the block Jacobi preconditioner is positive definite. $\square$

References

[1] L. Freitag and P. Knupp. Tetrahedral mesh improvement via optimization of the element condition number. *International Journal for Numerical Methods in Engineering*, 53:1377–1391, 2002.

[2] P. Knupp. Achieving finite element mesh quality via optimization of the Jacobian matrix norm and associated quantities, Part I – A framework for surface mesh optimization. *International Journal for Numerical Methods in Engineering*, 48:401–420, 2000.

[3] P. Knupp. Achieving finite element mesh quality via optimization of the Jacobian matrix norm and associated quantities, Part II – A framework for volume mesh optimization and the condition number of the Jacobian matrix. *International Journal for Numerical Methods in Engineering*, 48:1165–1185, 2000.

[4] A. Liu and B. Joe. Relationship between tetrahedron quality measures. *BIT*, 34:268–287, 1994.

[5] T. S. Munson. Mesh shape-quality optimization using the inverse mean-ratio metric. Preprint ANL/MCS-P1136-0304, Argonne National Laboratory, Argonne, Illinois, 2004.

[6] J. M. Ortega and W. C. Rheinboldt. *Iterative Solution of Nonlinear Equations in Several Variables.* Academic Press, San Diego, California, 1970.

[7] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, New Jersey, 1970.