A DETAILED INEXACT ARITHMETIC PROOFS

Here we will prove the claims in §7: that both the inexact pairwise Gauss-Seidel method described in Algorithm 3, as well as the Smith et al.’s Generalized Reflections algorithm [2012], satisfy the inexact impact operator axioms (eNORM)–(eMOD). We will assume the following computation model: real numbers are approximated using floating-point arithmetic, with machine epsilon $\varepsilon < 1$ and minimum representable magnitude $\eta$. We assume that no intermediate calculation overflows; we then have an associated rounding operator $\mathsf{fl}[x]$, so that for every exact quantity $x$, $$x = \mathsf{fl}[x] - \eta \leq \mathsf{fl}[x] \leq x + \mathsf{fl}[x] + \eta.$$ For calculations we will make use of the weaker, more convenient bound $$x = |x| - \varepsilon \leq \mathsf{fl}[x] \leq x + |x| + \varepsilon.$$ Arithmetic operations and square roots are assumed to take place in infinite precision, and then rounded; we will write $\mathsf{fl}[E]$ to denote that every operation in the expression $E$ is performed in this way, e.g. $\mathsf{fl}[x + y] = \mathsf{fl}[\mathsf{fl}[x] + \mathsf{fl}[y]]$. Finally, we will assume that $q_i$ and small integer constants are represented exactly, but that $M$, $M^{-1}$, and $N$ must be rounded.

If $\varepsilon$ is too large, the properties (eNORM), (eDRIFT), and (eMOD) cannot be guaranteed. We will prove that both pairwise Gauss-Seidel and Generalized Reflections satisfy these properties for $\varepsilon$ sufficiently small, and give a constructive bound for $\varepsilon$ in terms of the magnitudes of input quantities like $q_0$, $M$, $N$, etc. For both algorithms, we will first look at drift, and construct a $C$ which is used in the definition of (eDRIFT) as a certificate that energy cannot grow unboundedly over the course of several iterations. The proof of no drift will already impose a bound on $\varepsilon$; intuitively, if the machine precision is too large, the renormalization of the velocity after every iteration in Algorithms 3 and 4 itself introduces so much error into the computation of $q_{i+1}$ that despite the renormalization, its magnitude cannot be bounded.

Once we have constructed a $C$, we also need an $\varepsilon$. We will show that (eNORM) imposes a lower bound of $\varepsilon$, and that this lower bound decreases as $\varepsilon$ decreases. We end by proving (eMOD) hold, provided that $\varepsilon$ is not too large. The upper bound is constant, and the lower bound shrinks as $\varepsilon$ shrinks, so that it is always possible to find an $\varepsilon$ if $\varepsilon$ is sufficiently small.

A.1 Pairwise Gauss-Seidel

In this section, we derive an $\varepsilon$ and $C$ for which the modified pairwise GS algorithm described in section 7 satisfies the six criteria (eNORM)–(eMOD). Three of these, (eKIN), (ONE) and (eVIO), are obvious from the construction of the algorithm. We first prove (eDRIFT) by induction on the iteration $i$: suppose it holds for the first $i$ iterations of Algorithm 3. Then

$$\frac{1}{2} \|q_i\|_M^2 \leq \frac{1}{2} \|q_0\|_M^2 + C,$$
$$\|q_i\|_M^2 \leq \|q_0\|_M^2 + 2C,$$
$$\lambda_{\min} \|q_i\|_2^2 \leq \lambda_{\max} \|q_0\|_2^2 + 2C,$$
$$\|q_i\|_2 \leq \alpha_1 + \beta_1 \sqrt{C},$$

for

$$\alpha_1 = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \|q_0\|_2,$$
$$\beta_1 = \sqrt{\varepsilon},$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum eigenvalue of $M$, respectively. Since $\|q_i\|_\infty \leq \|q_i\|_2$ we also have that $\|q_i\|_\infty \leq \alpha_1 + \beta_1 \sqrt{C}$.

We now bound $p = \mathsf{fl}(q_i - 2(q_i, n)M^{-1} n)$, where $n$ is some constraint gradient selected by Algorithm 3. The following fact will be useful: for a sequence of numbers $x_1, \ldots, x_d$, it can be shown by induction on $d$ that

$$|\mathsf{fl} \left( \sum_{j=1}^d x_j \right) - \sum_{j=1}^d \mathsf{fl}(x_j) | \leq \left( d + \sum_{j=1}^d |\mathsf{fl}(x_j)| \right) \varepsilon (1 + \varepsilon)^d - d.$$ We now proceed to bound $p$. First,

$$|\mathsf{fl}(q'_i - q_i)\|_2^2 - \|q'_i\|_2^2| \leq \left( \|q'_i\|_2^2 - \|q_i\|_2^2 \right) + 1 \varepsilon$$

where $q'_i$ denotes the $j$th coordinates of the vector $n$. We can write these bounds as

$$|\mathsf{fl}(q'_i - q_i)\|_2^2 - \|q'_i\|_2^2| \leq \varepsilon \left( \alpha_2 + \beta_2 \sqrt{C} \right)$$

where

$$\alpha_2 = \alpha_1 \|n\|_\infty (1 + \varepsilon) + 1,$$
$$\beta_2 = \beta_1 \|n\|_\infty (1 + \varepsilon).$$

Since

$$\|q'_i\|_2 \leq \|q_i\|_\infty \|n\|_\infty \leq \|n\|_\infty (\alpha_1 + \beta_1 \sqrt{C}),$$

summing over $j$ gives

$$|\mathsf{fl} \left[ n^T (q_i, n) \right] - (q_i, n)| \leq \varepsilon \left( \alpha_3 + \beta_3 \sqrt{C} \right)$$

where

$$\alpha_3 = (1 + \|n\|_\infty (1 + 2\varepsilon \alpha_2))d (1 + \varepsilon)^{d-1},$$
$$\beta_3 = (\|n\|_\infty (\beta_1 + 2\varepsilon \beta_2))d (1 + \varepsilon)^{d-1}.$$
Switching gears,
\[
\|f\left((M^{-1})^{kj}\right) - (M^{-1})^{kj}\| \leq \|M^{-1}\|_{\infty} \epsilon + \epsilon
\]
and
\[
\|f\left(n^j\right) - n^j\| \leq (\|n\|_{\infty} + 1) \epsilon,
\]
so that
\[
\|f\left((M^{-1})^{kj}n^j\right) - (M^{-1})^{kj}n^j\| \leq \epsilon \alpha_4
\]
where
\[
\alpha_4 = 7 \|M^{-1}\|_{\infty} \|n\|_{\infty} + 4 \|M^{-1}\|_{\infty} + 4 \|n\|_{\infty} + 3.
\]
Summing again over \( j \) we can bound
\[
\|f\left((M^{-1})^{n}\right) - (M^{-1})^{n}\| \leq \epsilon \alpha_5
\]
where
\[
\alpha_5 = \left(1 + \|M^{-1}\|_{\infty} \|n\|_{\infty} + 2 \epsilon \alpha_4\right) (1 + \epsilon)^{d-1}.
\]
Now since
\[
\|f\left((M^{-1})^{n}\right)\| \leq d \|M^{-1}\|_{\infty} \|n\|_{\infty} + \epsilon \alpha_5
\]
we have that
\[
\|f\left(\tilde{q}_i, n\right) (M^{-1}n)^j\| - \langle \tilde{q}_i, n \rangle (M^{-1}n)^j \| \leq \epsilon (\alpha_5 + \beta_5 \sqrt{C})
\]
for
\[
\alpha_5 = 1 + (1 + 2 \epsilon \alpha_3) d \|M^{-1}\|_{\infty} \|n\|_{\infty} + 2 \epsilon \alpha_3 \|n\|_{\infty} \alpha_1 + 2 \alpha_3 \alpha_5
\]
\[
\beta_5 = 2 \beta_3 d \|M^{-1}\|_{\infty} \|n\|_{\infty} + 2 \epsilon \alpha_3 \|n\|_{\infty} \beta_1 + 2 \alpha_3 \beta_3,
\]
where we have made liberal use of the fact that \( \epsilon^2 < \epsilon \) to simplify the above expressions. Then
\[
\|f\left(-2\langle \tilde{q}_i, n \rangle (M^{-1}n)^j\| + 2 \langle \tilde{q}_i, n \rangle (M^{-1}n)^j \| \|
\leq \epsilon (\alpha_7 + \beta_7 \sqrt{C})
\]
where
\[
\alpha_7 = 1 + 4 \alpha_6 + 2d \|M^{-1}\|_{\infty} \|n\|_{\infty}^2 \alpha_1
\]
\[
\beta_7 = 4 \beta_6 + 2d \|M^{-1}\|_{\infty} \|n\|_{\infty}^2 \beta_1.
\]
Finally, we bound \( \tilde{p} \) in terms of \( p = \tilde{q}_i - 2\langle \tilde{q}_i, n \rangle M^{-1}n \). We have that
\[
\|\tilde{p} - p\| \leq \epsilon (\alpha_8 + \beta_8 \sqrt{C})
\]
for
\[
\alpha_8 = 1 + \alpha_1 + 2 \alpha_7 + 2d \|M^{-1}\|_{\infty} \|n\|_{\infty}^2 \alpha_1
\]
\[
\beta_8 = \beta_7 + 2d \|M^{-1}\|_{\infty} \|n\|_{\infty}^2 \beta_1.
\]
Next, we need to bound the norm \( \|f\|_{\infty} \) in the denominator of the coefficient of the velocity update step. We can use the fact that
\[
\|\tilde{p}\| \leq \|p\|_{\infty} \leq \frac{\|p\|_M}{\sqrt{\beta_{12}}} = \frac{\|q\|_M}{\sqrt{\beta_{12}}} \leq \frac{\|q\|_M + \sqrt{2C}}{\sqrt{\beta_{12}}}
\]
to get
\[
\|f\left[(M^{-1})^j\tilde{p}\right] - (M^{-1})^j\tilde{p}\| \leq \epsilon (\alpha_9 + \beta_9 \sqrt{C})
\]
for
\[
\alpha_9 = 1 + (2 + 3 \|M\|_{\infty}) (\alpha_8 + \|q\|_M) \sqrt{\beta_{12}} + \epsilon \alpha_9
\]
\[
\beta_9 = (2 + 3 \|M\|_{\infty}) (\beta_8 + \sqrt{\frac{2}{\beta_{12}}}) + \epsilon \beta_9.
\]
The summation formula then gives
\[
\|f\left((M^j)^k\right) - (M^j)^k\| \leq \epsilon (\alpha_{10} + \beta_{10} \sqrt{C})
\]
where
\[
\alpha_{10} = d \alpha_9 + \left(1 + d \|M\|_{\infty} \|q\|_M \sqrt{\beta_{12}} + \epsilon \alpha_9\right) d(1 + \epsilon)^{d-1}
\]
\[
\beta_{10} = d \beta_9 + \left(1 + d \|M\|_{\infty} \|q\|_M \sqrt{\beta_{12}} + \epsilon \beta_9\right) d(1 + \epsilon)^{d-1}
\]
Next, combining the last several bounds,
\[
\|f\left(p^j (M^j)^k\right) - p^j (M^j)^k\| \leq \epsilon (\alpha_{11} + \beta_{11} \sqrt{C} + \gamma_{11} C)
\]
for
\[
\alpha_{11} = 1 + \frac{d \|M\|_{\infty} \|q\|_M}{\sqrt{\beta_{12}}} + 2 \alpha_8 \alpha_{10}
\]
\[
+ 2 (\alpha_{10} + \alpha_8 d \|M\|_{\infty} \|q\|_M \sqrt{\beta_{12}} + \epsilon \alpha_{10} \beta_9)
\]
\[
\beta_{11} = 2 d \|M\|_{\infty} \|q\|_M \sqrt{\beta_{12}} + 2 \alpha_8 \beta_{10} + 2 \alpha_8 \beta_9
\]
\[
+ 2 (\beta_{10} + \beta_8 d \|M\|_{\infty} \|q\|_M \sqrt{\beta_{12}} + \epsilon \beta_{10} \beta_9)
\]
\[
+ 2 (\alpha_{10} + \alpha_8 d \|M\|_{\infty} \|q\|_M \sqrt{\beta_{12}} + \epsilon \beta_{10} \beta_9)
\]
\[
\gamma_{11} = 2 d \|M\|_{\infty} + 2 \beta_{10} \beta_9 + 2 (\beta_{10} + \beta_8 d \|M\|_{\infty} \sqrt{\beta_{12}} + \epsilon \beta_{10} \beta_9).
\]
We apply the summation formula a second time to get the squared norm,
\[
\|f\left(p^TMp\right) - p^TMp\| \leq \epsilon (\alpha_{12} + \beta_{12} \sqrt{C} + \gamma_{12} C)
\]
for
\[
\alpha_{12} = d \alpha_{11} + (1 + d \|q\|_M + d \epsilon \alpha_{11}) d(1 + \epsilon)^{d-1}
\]
\[
\beta_{12} = d \beta_{11} + (d \sqrt{2} + d \epsilon \beta_{11}) d(1 + \epsilon)^{d-1}
\]
\[
\gamma_{12} = d \gamma_{11} + d^2 \epsilon \gamma_{11} (1 + \epsilon)^{d-1}.
\]
We can rewrite this bound in more convenient form, by completing the square, in anticipation of taking the square root:
\[
\|f\left[p^TMp\right] - p^TMp\| \leq \epsilon \left(\frac{\beta_{12}}{2} + \sqrt{\gamma_{12} \sqrt{C}}\right)^2 + \epsilon \left(\alpha_{12} - \frac{\beta_{12}^2}{4 \gamma_{12}}\right).
\]
Finally, we have a bound on the norm of \( p \):
\[
\|f\left[||p||_M\right] - ||p||_M\| \leq \epsilon (\alpha_{13} + \beta_{13} \sqrt{C}),
\]
for
\[
\alpha_{13} = 1 + (2 + 3 \|M\|_{\infty}) (\alpha_{12} + \|q\|_M) \sqrt{\beta_{12}} + \epsilon \alpha_{12}
\]
\[
\beta_{13} = (2 + 3 \|M\|_{\infty}) (\beta_{12} + \sqrt{\frac{2}{\beta_{12}}}) + \epsilon \beta_{12}.
\]
where
\[ \alpha_{13} = 1 + \|q_0\|_M + \frac{1 + \epsilon}{\sqrt{\ell}} \left( \sqrt{\alpha_{12} - \beta_{12}^2} + \beta_{12} \right) \]
\[ \beta_{13} = \frac{1 + \epsilon}{\sqrt{\ell}} \sqrt{\gamma_{12}}. \]
Notice that since \( |M|_\infty \geq \gamma_{\min}, \gamma_{12} > 1 \) and so the denominators in \( \alpha_{13} \) are bounded away from zero.

The last piece we need for computing \( q_{i+1} \) is the norm of the initial velocity, \( \|q_0\|_M \). To begin with,
\[ \| (M^{K_j} q_i^j) - (M^{K_j} q_0^j) \|_1 \leq \epsilon \alpha_{14} \]
where
\[ \alpha_{14} = 2|\gamma_{\min}| \|q_0\|_M + \|q_0\|_M + 1. \]

Since \( |M^{K_j} q_i^j| \leq d|\gamma_{\min}| \|q_0\|_M \), applying the summation formula yields
\[ \| (M q_0) - (M q_0) \|_1 \leq \epsilon \alpha_{15} \]
for
\[ \alpha_{15} = d \alpha_{14} + (1 + d|\gamma_{\min}| \|q_0\|_M + \epsilon \alpha_{14}) d(1 + \epsilon)^{-d-1}. \]
Then
\[ \| (M q_0) - (M q_0) \|_1 \leq \epsilon \alpha_{16} \]
for
\[ \alpha_{16} = d|\gamma_{\min}| \|q_0\|_M^2 + \|q_0\|_M + 1. \]

Applying the summation formula a second time gives
\[ \| (M q_0) - (M q_0) \|_1 \leq \epsilon \alpha_{17} \]
for
\[ \alpha_{17} = d \alpha_{16} + (1 + \|q_0\|_M^2 + d \alpha_{16}) d(1 + \epsilon)^{-d-1}. \]
Finally
\[ \| (M q_0) - (M q_0) \|_1 \leq \epsilon \alpha_{18} \]
with
\[ \alpha_{18} = 1 + \|q_0\|_M + \frac{1 + \epsilon}{\sqrt{\ell}} \sqrt{\alpha_{17}}. \]

Combining equations (4) and (6) gives
\[ \| \|q_0\|_M p\| - \|q_0\|_M p\| \|_1 \leq \epsilon (\alpha_{19} + \beta_1 \sqrt{\gamma_{C}}) \]
for
\[ \alpha_{20} = \frac{\|q_0\|_M \alpha_{19} + 2 \alpha_{19} + \|q_0\|_M + \|q_0\|_M \sqrt{\gamma_{\min}}}{\sqrt{\gamma_{\min}}} \]
\[ \beta_{20} = \frac{\|q_0\|_M \beta_{19} + 2 \beta_1 + \|q_0\|_M \sqrt{\gamma_{\min}}}{\sqrt{\gamma_{\min}}}. \]

Therefore
\[ \| (\|q_{i+1}\|_M - \|q_0\|_M) \|_1 \leq 2 \|q_0\|_M \lambda_{\max} \sqrt{d} \|p\|_M - \epsilon (\alpha_{19} + \beta_{19} \sqrt{\gamma_{C}}) \]
\[ + \epsilon \lambda_{\max} \frac{\alpha_{20} + \beta_{20} \sqrt{\gamma_{C}}}{\|p\|_M - \epsilon (\alpha_{19} + \beta_{19} \sqrt{\gamma_{C}})}^2. \]

Let
\[ \alpha_{21} = 4 \lambda_{\max} \sqrt{d} \alpha_{20} \]
\[ \beta_{21} = 4 \lambda_{\max} \sqrt{d} \beta_{20} \]
\[ \gamma_{21} = \frac{1}{4 \|q_0\|_M^2} \lambda_{\max} d \beta_{20}^2. \]

**Lemma A.1.** If \( \epsilon < \frac{\|q_0\|_M \alpha_{21}}{2 \alpha_{19}}, \epsilon < \frac{\beta_{21}}{2}, \) and
\[ \epsilon \beta_{21} + \sqrt{\epsilon^2 \beta_{21}^2 + 4 \epsilon \alpha_{21} (2 - \epsilon \gamma_{21})} \]
\[ \leq \left( \frac{\|q_0\|_M - \sqrt{\gamma_{C}} - (\alpha_{19} + \beta_{19} \sqrt{\gamma_{C}})}{2} \right) \]
then pairwise Gauss-Seidel satisfies (\( \epsilon \text{DRIFT} \)). Notice that these conditions are satisfied if \( \epsilon \) is sufficiently small.

**Proof.** Take
\[ C = \frac{1}{2} \left( \frac{\|q_0\|_M - \sqrt{\gamma_{C}} - (\alpha_{19} + \beta_{19} \sqrt{\gamma_{C}})}{2} \right) \]
\[ + \epsilon \beta_{21} + \sqrt{\epsilon^2 \beta_{21}^2 + 4 \epsilon \alpha_{21} (2 - \epsilon \gamma_{21})} \]
\[ \leq \frac{1}{2 \|q_0\|_M} \left[ \frac{\|q_0\|_M - 2 \epsilon \alpha_{19}}{2 \alpha_{19}} \right] \]
\[ \leq \frac{\|q_0\|_M - \sqrt{\gamma_{C}} - (\alpha_{19} + \beta_{19} \sqrt{\gamma_{C}})}{2} \]
\[ \leq \frac{1}{2} \|q_0\|_M, \]
and
\[ \|p\|_M - (\alpha_{19} + \beta_{19} \sqrt{\gamma_{C}}) \leq \frac{1}{2} \|q_0\|_M, \]

hence the bound in equation (7) is valid. Moreover we can substitute this inequality into the bound on \( \|q_{i+1}\|_M^2 \) to get
\[ \frac{\|q_{i+1}\|_M^2}{\|q_0\|_M^2} \leq \epsilon (\alpha_{21} + \beta_{21} \sqrt{\gamma_{C}} + \gamma_{21} C). \]

Then \( q_{i+1} \) satisfies (\( \epsilon \text{DRIFT} \)) whenever
\[ (2 - \epsilon \gamma_{21}) C = \epsilon \beta_{21} - \epsilon \alpha_{21} \leq 0, \]
and in particular, whenever
\[ C \geq \frac{\epsilon \beta_{21} + \sqrt{\epsilon^2 \beta_{21}^2 + 4 \epsilon \alpha_{21} (2 - \epsilon \gamma_{21})}}{4 - 2 \epsilon \gamma_{21}}. \]
We now prove the remaining properties, which are relatively straightforward. First, we have that

**Lemma A.2.** Let C be as in the previous lemma, and suppose

\[
\epsilon > \frac{\lambda_{\text{max}} \sqrt{d} (\alpha_2 + \beta_2 V^2 C)}{\|q_0\| (\|q_0\| + \sqrt{2} C)}
\]

and

\[
\epsilon \geq 2 \lambda_{\text{max}} \sqrt{\frac{2 \alpha_2 + 2 \beta_2 V^2 C}{\|q_0\|_M (\|q_0\|_M - \sqrt{2} C)}}.
\]

Then pairwise Gauss-Seidel satisfies (\(\epsilon\)NORM). Notice that both right-hand sides vanish as \(\epsilon\) decreases.

**Proof.** Let C be as in the previous lemma. By construction of the algorithm and (\(\epsilon\)VIO) we know that the value of \(\lambda\) is

\[
\lambda = -2\langle q_1, n \rangle > 2\epsilon \|q_1\|_M \geq 2\epsilon (\|q_0\|_M + \sqrt{2} C)
\]

where the last inequality follows from (\(\epsilon\)DRIFT).

From the bound (7) on the components of \(\epsilon\) we have that

\[
\|\epsilon\|_M \leq \epsilon \lambda_{\text{max}} \sqrt{\frac{2 \alpha_2 + 2 \beta_2 V^2 C}{\|q_0\|_M}}
\]

and this is less than \(\epsilon \lambda\) when

\[
\epsilon \lambda_{\text{max}} \sqrt{\frac{2 \alpha_2 + 2 \beta_2 V^2 C}{\|q_0\|_M}} \leq 2\epsilon^2 (\|q_0\|_M + \sqrt{2} C).
\]

Lastly since \(\|q_i\|_M \geq \|q_0\|_M - \sqrt{2} C\), we have that \(\|\epsilon\|_M \leq \frac{\epsilon}{\lambda} \|q_i\|\) whenever

\[
\epsilon \geq 2 \lambda_{\text{max}} \sqrt{\frac{2 \alpha_2 + 2 \beta_2 V^2 C}{\|q_0\|_M (\|q_0\|_M - \sqrt{2} C)}}.
\]

\(\Box\)

**Lemma A.3.** Pairwise Gauss-Seidel satisfies (\(\epsilon\)MOD) when \(\epsilon < 1\).

**Proof.** At every iteration where a constraint with gradient \(n\) is violated,

\[
\|q_i - q_{i-1}\|_M = \|M^{-1} n + c\|_M
\]

\[
\geq |\lambda| - \|c\|_M
\]

\[
\geq (1 - \epsilon) |\lambda|
\]

\(\geq 0\).

\(\Box\)

### A.2 Generalized Reflections

The **generalized reflection** operator of Smith et al. [2012] improves on pairwise Gauss-Seidel by guaranteeing preservation of symmetries and more accurately modeling shock propagations, at the cost of an \(R\) that is more expensive to compute. Algorithm 4 shows how to modify it so that it satisfies all the inexact desiderata required for guaranteed termination. Notice that these modifications mirror those of Gauss-Seidel: constraints whose violation does not exceed a threshold are pruned every time a reflection is applied, and the velocity is renormalized every step to prevent energy drift.

Computing \(\lambda\) at each iteration of Algorithm 4 requires solving a quadratic program (QP). Let \(\lambda\) be the exact solution to this QP. \(\xi\) the corresponding positivity constraint Lagrange multipliers, and \(\bar{\lambda}\), \(\bar{\xi}\) the computed solution. We assume that \(\bar{\lambda}\) approximately satisfies the KKT conditions of the QP,

\[
\|N_{iV}^T M^{-1} N_{iV} \lambda + 2N_{iV}^T \bar{q}_i - \xi\|_\infty \leq \epsilon^2 \kappa_1 \|\bar{q}_i\|_M \bar{\lambda} \geq 0
\]

\[
\bar{\xi} \geq 0
\]

\[
\bar{\lambda} \perp \bar{\xi},
\]

where \(\kappa_1\) is an accuracy parameter independent of \(\bar{q}_i\); notice that this condition is a standard relative error termination criterion in numerical QP codes.

The goal now will be to bound the intermediate step

\[
\tilde{p} = \Pi \left[ q_i + M^{-1} N_{iV} \bar{\lambda} \right]
\]

in terms of the true step \(p = q_i + M^{-1} N_{iV} \lambda\); the proof of (\(\epsilon\)DRIFT) will then follow directly from identical calculations to that in pairwise Gauss-Seidel. Once we have a value of \(C\), we will prove that inexact GR satisfies (\(\epsilon\)NORM) and (\(\epsilon\)MOD). As in the case of Gauss-Seidel, (\(\epsilon\)KIN), (ONE), and (\(\epsilon\)VIO) all hold by construction of Algorithm 4.

Let \(N_A \subset N_V\) be the set of constraints that are active in the inexact QP solution, and \(\lambda_A\) the corresponding parts of \(\bar{\lambda}\). The matrix \(N_A^T M^{-1} N_{iV}\) has ones along the diagonal, and off-diagonal entries of magnitude at most one; therefore by the Gershgorin Circle Theorem its maximum eigenvalue is at most \(m\), the number of total constraints in \(N\). Then we have the following useful bound on \(\bar{\lambda}\):

\[
\|\bar{\lambda}\|_\infty = \|\lambda_A\|_\infty \leq \frac{\|N_A^T M^{-1} N_A \bar{\lambda}_A\|_2}{m} \leq \frac{\sqrt{d} \|N_A^T M^{-1} N_A \bar{\lambda}_A\|_2}{m} \leq \frac{\sqrt{d} (\|2N_A^T \bar{q}_i\|_\infty + \epsilon \kappa_1 \|\bar{q}_i\|_M)}{m} \leq \frac{\sqrt{d}}{m} \left( \lambda_{\text{max}} (\|\bar{q}_i\|_M + \epsilon \kappa_1 \|\bar{q}_i\|_M) \right) \leq \kappa_2 + \mu_2 \sqrt{C}.
\]
with
\[
\kappa_2 = \left( \frac{2\sqrt{\lambda}}{\lambda_{\max}} + \frac{ek_1}{m} \right) \|q_0\|_M
\]
\[
\mu_2 = \frac{2\sqrt{\lambda}}{\lambda_{\max}} + \frac{ek_1}{m} \sqrt{2},
\]
where as usual we have used \(\epsilon^2 < \epsilon\) to simplify expressions.

Now for \(n_j\) the \(j\)th row of \(N\), we have the bound
\[
\|n_j \hat{\lambda}^k - n_j^k \hat{\lambda}^k\| \leq \epsilon(k_3 + \mu_3 \sqrt{C})
\]
where
\[
k_3 = 3\|N\|\|\kappa_2 + 2k_2 + 1\|
\]
\[
\mu_3 = 3\|N\|\|\mu_2 + 2\mu_2,\]
so applying the summation formula gives
\[
\|n_j \left( (N\hat{\lambda})^k - (N\hat{\lambda})^k \right) \| \leq \epsilon(k_4 + \mu_4 \sqrt{C})
\]
for
\[
k_4 = (2k_3 + 1 + m\|N\|\|\kappa_2) (1 + \epsilon) d^{-1}
\]
\[
\mu_4 = (2\mu_3 + m\|N\|\|\mu_2) (1 + \epsilon) d^{-1}.
\]
Then
\[
\|n_j \left( (M^{-1}\hat{\lambda})^k - (M^{-1}\hat{\lambda})^k \right) \| \leq \epsilon(k_5 + \mu_5 \sqrt{C})
\]
for
\[
k_5 = 1 + (5\|M^{-1}\|\|\kappa_2 + 2\|M^{-1}\|\|\kappa_2 + 1\|)\|N\|\|\kappa_2
\]
\[
\mu_5 = (5\|M^{-1}\|\|\mu_2 + 2\|M^{-1}\|\|\mu_2 + 1\|)\|N\|\|\mu_2,\]
so that applying the summation formula gives
\[
\|n_j \left( (M^{-1}\hat{\lambda})^k - (M^{-1}\hat{\lambda})^k \right) \| \leq \epsilon(k_6 + \mu_6 \sqrt{C})
\]
for
\[
k_6 = d k_5 + (1 + d\|M^{-1}\|\|\kappa_2 + 2\|k_2 + 2\|M^{-1}\|\|\kappa_2 + 1\|)\|N\|\|\kappa_2
\]
\[
\mu_6 = d \mu_5 + (1 + d\|M^{-1}\|\|\mu_2 + 2\|\|\mu_2 + 1\|)\|N\|\|\mu_2,\]
Before we can bound \(\bar{p}\), we need to relate the impulse using the approximate multipliers \(\hat{\lambda}\) to that using the exact multipliers. We can do so by making use of the fact that the QP’s KKT conditions are nearly satisfied for \(\hat{\lambda}^k\):
\[
\|M^{-1} N \hat{\lambda} - N \hat{\lambda} \|^2_M
\]
\[
= (\lambda - \hat{\lambda})^T (N^T M^{-1} N \lambda - N^T M^{-1} N \hat{\lambda})
\]
\[
\leq (\lambda - \hat{\lambda})^T (N^T M^{-1} N \lambda - N^T M^{-1} N \hat{\lambda})
\]
\[
\leq (\lambda - \hat{\lambda})^T (N^T M^{-1} N \hat{\lambda} - N^T M^{-1} N \hat{\lambda})
\]
\[
\leq \epsilon^2 (k_7 + \mu_7 \sqrt{C} + \nu_7 C)
\]
for
\[
k_7 = 2\|q_0\|_M \sqrt{\kappa_2 k_1}
\]
\[
\mu_7 = 2\sqrt{2} \|k_1\|_M \sqrt{\kappa_1 \mu_2}
\]
\[
\nu_7 = 2\sqrt{2} \|\kappa_1\|_M \mu_2.
\]
Completing the square gives
\[
\|M^{-1} N \hat{\lambda} - M^{-1} N \hat{\lambda}\|_M
\]
\[
\leq 1\|M^{-1} N \hat{\lambda} - M^{-1} N \hat{\lambda}\|_M
\]
\[
\leq \epsilon(k_8 + \mu_8 \sqrt{C})
\]
with
\[
k_8 = \frac{1}{\lambda_{\max}} \left( \frac{\mu_7}{2\sqrt{\epsilon}} + \sqrt{\epsilon} \right)
\]
\[
\mu_8 = \sqrt{\epsilon}.
\]
Therefore
\[
\|n_j \left( (M^{-1} N \hat{\lambda})^k - (M^{-1} N \hat{\lambda})^k \right) \| \leq \epsilon(k_9 + \mu_9 \sqrt{C})
\]
where simply \(k_9 = k_8 + k_9\) and \(\mu_9 = \mu_6 + \mu_8\). We then have
\[
|\bar{p}^j - \bar{p}^i| \leq \epsilon(k_{10} + \mu_{10} \sqrt{C})
\]
for
\[
k_{10} = 2k_9 + d^2\|M^{-1}\|\|N\|\|\kappa_2 + 2k_8 + 1\|\|N\|\|\mu_2
\]
\[
\mu_{10} = 2\mu_9 + d^2\|M^{-1}\|\|N\|\|\mu_2 + \mu_1.
\]
The proof of (eDRIFT) now follows identically the arguments for pairwise Gauss-Seidel, with \(k_{10}\) and \(\mu_{10}\) taking the place of \(a_8\) and \(\beta_8\). As in the pairwise GS case, construction of a \(C\) certifying (eDRIFT) requires that \(\epsilon\) be sufficiently small.

We now prove that GR satisfies the remaining properties, (eNORM) and (eMOD).

**Lemma A.4.** Let \(C\) be as in the proof of (eDRIFT), and suppose that
\[
\epsilon \geq \frac{a + \sqrt{a^2 + b}}{2},
\]
where
\[
a = \frac{{\epsilon}}{2} x_1 (\|q_0\|_M + \sqrt{C})
\]
\[
b = 4Mr_{\max} \frac{a_2 + \beta_2 \sqrt{C}}{\sqrt{\|q_0\|_M}},
\]
and
\[
\epsilon \geq 2\epsilon x_{\max} \frac{a_2 + \beta_2 \sqrt{C}}{\|q_0\|_M (\|q_0\|_M - \sqrt{C})}.
\]
*Then Generalized Reflections satisfies (eNORM). Notice that both right-hand sides vanish as \(\epsilon\) decreases.*
where we have used (eDRIFT) and again the fact that the largest eigenvalue of $N_V^T M^{-1} N_V$ is at most $m$.

From the bound (7) on the components of $c$, we have that
\[
\|c\|_M \leq \epsilon \lambda_{\text{max}} \sqrt{d} \frac{2\alpha_{20} + 2\beta_{20} \sqrt{C}}{\|q_0\|_M}
\]
and this is less than $\epsilon \|\lambda\|_1$ when
\[
\epsilon \lambda_{\text{max}} \sqrt{d} \frac{2\alpha_{20} + 2\beta_{20} \sqrt{C}}{\|q_0\|_M} \leq \epsilon d - \epsilon^2 \frac{\kappa_1}{m} (\|q_0\|_M + \sqrt{2C}).
\]
Rearranging gives
\[
\epsilon^2 - \epsilon^2 \frac{\kappa_1}{2} (\|q_0\|_M + \sqrt{2C}) - m \epsilon \lambda_{\text{max}} \frac{\alpha_{20} + \beta_{20} \sqrt{C}}{\sqrt{d}\|q_0\|_M} \geq 0
\]
and the first inequality above. Lastly since $\|q_i\|_M \geq \|q_0\|_M - \sqrt{2C}$, we have that $\|c\|_M \leq \frac{\epsilon}{2} \|q_i\|$ whenever
\[
\epsilon \geq 2\epsilon \lambda_{\text{max}} \sqrt{d} \frac{2\alpha_{20} + 2\beta_{20} \sqrt{C}}{\|q_0\|_M (\|q_0\|_M - \sqrt{2C})},
\]
as in the case of pairwise Gauss-Seidel. 

At last we end with

**Lemma A.5.** If $\epsilon < 4$, then Generalized Reflections satisfies (eMOD).

**Proof.** At every iteration where a constraint is violated,
\[
\|q_{i+1} - \bar{q}_i\|_M = \|M^{-1} N_V \lambda + c\|_M \\
\geq \|M^{-1} N_V \lambda\|_M - \|c\|_M \\
\geq \sqrt{2\|\lambda\|_1 \|q_i\|_M - \|c\|_M} \\
\geq \sqrt{2\|\lambda\|_1 \|q_i\|_M - \epsilon \sqrt{\|\lambda\|_1 \|q_i\|_M / 2}} \\
\geq \sqrt{2\|\lambda\|_1 \|q_i\|_M - \epsilon \sqrt{\|\lambda\|_1 \|q_i\|_M / 2}}.
\]
The right-hand side is positive when $\epsilon < 4$. 