What is PRSQP?

Suppose we have an optimization problem of the following form:

$$
\begin{aligned}
&\min_{\boldsymbol{y}} \quad f(\boldsymbol{y})\,, \quad f:\mathbb{R}^n \to \mathbb{R}\,, \\
&\text{s. t.}\;\; \mathbf{c}(\boldsymbol{y}) \;=\; \mathbf{0}\,, \quad \mathbf{c}:\mathbb{R}^n \to \mathbb{R}^{m_c}\,, \\
&\qquad\; \mathbf{g}(\boldsymbol{y}) \;\le\; \mathbf{0}\,, \quad \mathbf{g}:\mathbb{R}^n \to \mathbb{R}^{m_g}\,.
\end{aligned} \tag{1}
$$

The first order optimality conditions for such a problem tell that at a solution point there are vectors $\boldsymbol{\lambda} \in \mathbb{R}^{m_c}$ and $\boldsymbol{\mu} \in \mathbb{R}^{m_g}$ such that

$$
\begin{aligned}
\nabla f \,+\, \nabla \mathbf{c}\,\boldsymbol{\lambda} \,+\, \nabla \mathbf{g}\,\boldsymbol{\mu} \;&=\; \mathbf{0}\,, \\
\mathrm{diag}(\boldsymbol{\mu})\,\mathbf{g} \;&=\; \mathbf{0}\,, \quad \boldsymbol{\mu} \ge \mathbf{0}\,, \\
\mathbf{c} \;&=\; \mathbf{0}\,.
\end{aligned} \tag{2}
$$

From now on, we will use the notations

$$
\begin{aligned}
\boldsymbol{C} \;&:=\; \nabla \mathbf{c}^\top(\boldsymbol{y}_k)\,, \\
\boldsymbol{G} \;&:=\; \nabla \mathbf{g}^\top(\boldsymbol{y}_k)\,, \\
\ell \;&:=\; f(\boldsymbol{y}_k) \,-\, \boldsymbol{\lambda}^\top \mathbf{c}(\boldsymbol{y}_k) \,-\, \boldsymbol{\mu}^\top \mathbf{g}(\boldsymbol{y}_k)\,, \\
\boldsymbol{H} \;&:=\; \nabla^2 \ell\,.
\end{aligned} \tag{3}
$$

and we will also omit all arguments like $(\boldsymbol{y}_k)$.

The linearized system at iteration step k is

$$
\begin{aligned}
\begin{pmatrix} \boldsymbol{H} & \boldsymbol{G}^\top & \boldsymbol{C}^\top \\ \mathrm{diag}(\boldsymbol{\mu})\boldsymbol{G} & \mathrm{diag}(\mathbf{g}) & \mathbf{0} \\ \boldsymbol{C} & \mathbf{0} & \mathbf{0} \end{pmatrix}
\begin{pmatrix} \Delta \boldsymbol{y} \\ \Delta \boldsymbol{\mu} \\ \Delta \boldsymbol{\lambda} \end{pmatrix}
\;&=\; -\begin{pmatrix} \nabla \ell \\ \mathrm{diag}(\boldsymbol{\mu})\mathbf{g} \\ \mathbf{c} \end{pmatrix}\,, \\
\boldsymbol{\mu} \;&\ge\; \mathbf{0}\,.
\end{aligned} \tag{4}
$$

Now let us suppose there are $(n \times m_c)$-Matrices $\boldsymbol{X}, \boldsymbol{P}$ such that

$$
\Delta \boldsymbol{y} \;=\; \boldsymbol{X}\Delta\hat{\boldsymbol{x}} \,+\, \boldsymbol{P}\Delta\hat{\boldsymbol{p}}\,, \text{ where } \boldsymbol{C}\boldsymbol{X} = \boldsymbol{I}\,,\; \boldsymbol{C}\boldsymbol{P} = \mathbf{0}\,. \tag{5}
$$

Then we can blow up system (4) by splitting its first column so that we get

$$
\begin{aligned}
\begin{pmatrix} \boldsymbol{H}\boldsymbol{X} & \boldsymbol{H}\boldsymbol{P} & \boldsymbol{G}^\top & \boldsymbol{C}^\top \\ \mathrm{diag}(\boldsymbol{\mu})\boldsymbol{G}\boldsymbol{X} & \mathrm{diag}(\boldsymbol{\mu})\boldsymbol{G}\boldsymbol{P} & \mathrm{diag}(\mathbf{g}) & \mathbf{0} \\ \boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}
\begin{pmatrix} \Delta\hat{\boldsymbol{x}} \\ \Delta\hat{\boldsymbol{p}} \\ \Delta \boldsymbol{\mu} \\ \Delta \boldsymbol{\lambda} \end{pmatrix}
\;&=\; -\begin{pmatrix} \nabla \ell \\ \mathrm{diag}(\boldsymbol{\mu})\mathbf{g} \\ \mathbf{c} \end{pmatrix}\,, \\
\boldsymbol{\mu} \;&\ge\; \mathbf{0}\,.
\end{aligned} \tag{6}
$$

From the third row we now see that $\Delta\hat{\boldsymbol{x}} = -\mathbf{c}$, which means that we can shuffle the first column to the right-hand side:

$$\begin{pmatrix} \boldsymbol{HP} & \boldsymbol{G}^\top & \boldsymbol{C}^\top \\ \mathrm{diag}(\boldsymbol{\mu})\boldsymbol{GP} & \mathrm{diag}(\mathbf{g}) & \mathbf{0} \end{pmatrix} \begin{pmatrix} \Delta\hat{\boldsymbol{p}} \\ \Delta\boldsymbol{\mu} \\ \Delta\boldsymbol{\lambda} \end{pmatrix} = -\begin{pmatrix} \nabla\ell - \boldsymbol{HX}\mathbf{c} \\ \mathrm{diag}(\boldsymbol{\mu})(\mathbf{g} - \boldsymbol{GX}\mathbf{c}) \end{pmatrix} , \qquad \boldsymbol{\mu} \geq \mathbf{0} . \tag{7}$$

Since $\boldsymbol{CP} = \mathbf{0}$, we can get rid of $\Delta\boldsymbol{\lambda}$ by multiplying the first row with $\boldsymbol{P}^\top$ from the left. We now have

$$\begin{pmatrix} \boldsymbol{P}^\top\boldsymbol{HP} & \boldsymbol{P}^\top\boldsymbol{G}^\top \\ \mathrm{diag}(\boldsymbol{\mu})\boldsymbol{GP} & \mathrm{diag}(\mathbf{g}) \end{pmatrix} \begin{pmatrix} \Delta\hat{\boldsymbol{p}} \\ \Delta\boldsymbol{\mu} \end{pmatrix} = -\begin{pmatrix} \boldsymbol{P}^\top(\nabla\ell - \boldsymbol{HX}\mathbf{c}) \\ \mathrm{diag}(\boldsymbol{\mu})\,(\mathbf{g} - \boldsymbol{GX}\mathbf{c}) \end{pmatrix} , \qquad \boldsymbol{\mu} \geq \mathbf{0} . \tag{8}$$

Alternatively, we can write this as a quadratic program:

$$\begin{aligned} \min_{\Delta\hat{\boldsymbol{p}}} \quad & \frac{1}{2}\,\Delta\hat{\boldsymbol{p}}^\top\boldsymbol{P}^\top\boldsymbol{HP}\,\Delta\hat{\boldsymbol{p}} + \boldsymbol{P}^\top(\nabla\ell - \boldsymbol{HX}\mathbf{c})\,\Delta\hat{\boldsymbol{p}} , \\ \text{s. t.} \quad & \boldsymbol{GP}\,\Delta\hat{\boldsymbol{p}} + \mathbf{g} - \boldsymbol{GX}\mathbf{c} \leq 0 . \end{aligned} \tag{9}$$

Let us now denote the first m_c columns of $\boldsymbol{C}$ and $\boldsymbol{G}$ with $\boldsymbol{C}_x$, $\boldsymbol{G}_x$ and the rest with $\boldsymbol{C}_p$, $\boldsymbol{G}_p$. Similarly, we split the vector $\Delta\boldsymbol{y}^\top = (\Delta\boldsymbol{x}^\top, \Delta\boldsymbol{p}^\top)$. Under the assumption that $\boldsymbol{C}_x$ is invertible, we then see that the matrices

$$\begin{aligned} \boldsymbol{X} &= \begin{pmatrix} \boldsymbol{C}_x^{-1} \\ \mathbf{0} \end{pmatrix} , \\ \boldsymbol{P} &= \begin{pmatrix} -\boldsymbol{C}_x^{-1}\boldsymbol{C}_p \\ \boldsymbol{I} \end{pmatrix} \end{aligned} \tag{10}$$

give a decomposition of the form (5) with

$$\begin{aligned} \Delta\hat{\boldsymbol{x}} &= \boldsymbol{C}_x\,\Delta\boldsymbol{x} + \boldsymbol{C}_p\,\Delta\boldsymbol{p} , \\ \Delta\hat{\boldsymbol{p}} &= \Delta\boldsymbol{p} . \end{aligned} \tag{11}$$

Asymptotically, the term $\boldsymbol{HX}\mathbf{c}$ from the first line in (9) will vanish so that we may drop it for the sake of simplicity. Furthermore, we have

$$\begin{aligned} \boldsymbol{P}^\top\boldsymbol{HX} &= \left(-\boldsymbol{C}_p^\top\boldsymbol{C}_x^{-\top}\boldsymbol{H}_{xx} + \boldsymbol{H}_{xp}\right)\boldsymbol{C}_x^{-1} , \\ \boldsymbol{GX} &= \boldsymbol{G}_x\boldsymbol{C}_x^{-1} , \\ \boldsymbol{GP} &= -\boldsymbol{G}_x\boldsymbol{C}_x^{-1}\boldsymbol{C}_p + \boldsymbol{G}_p , \\ \boldsymbol{P}^\top\nabla\ell &= -\boldsymbol{C}_p^\top\boldsymbol{C}_x^{-\top}(\nabla_{\boldsymbol{x}}f + \boldsymbol{G}_x^\top\boldsymbol{\mu}) + \nabla_{\boldsymbol{p}}f + \boldsymbol{G}_p^\top\boldsymbol{\mu} , \\ \Delta\hat{\boldsymbol{p}}^\top\boldsymbol{P}^\top\boldsymbol{HP}\Delta\hat{\boldsymbol{p}} &= \Delta\boldsymbol{p}^\top\boldsymbol{H}_{pp}\Delta\boldsymbol{p} , \end{aligned} \tag{12}$$

As `OptiMISES` does not deal with state constraints, $\mathbf{g}$ does not depend on $\boldsymbol{x}$ so that $\boldsymbol{G}_x = \mathbf{0}$.

So the quadratic program to determine the step $\Delta\boldsymbol{p}$ in blade geometry parameters simplifies to

$$\begin{aligned} \min_{\Delta\boldsymbol{p}} \quad & \frac{1}{2}\Delta\boldsymbol{p}^\top \boldsymbol{H}_{pp}\Delta\boldsymbol{p} + \boldsymbol{\gamma}\Delta\boldsymbol{p} \\ \text{s. t.} \quad & \boldsymbol{G}_p\Delta\boldsymbol{p} + \mathbf{g} \le 0 \,, \end{aligned} \tag{13}$$

where we call

$$\boldsymbol{\gamma} := \nabla_p f - \boldsymbol{C}_p^\top \boldsymbol{C}_x^{-\top} \nabla_x f \tag{14}$$

the *reduced gradient*.

The update of the flow variables is retrieved from the last row of (6) and the first line of (11):

$$\Delta\boldsymbol{x} = \boldsymbol{C}_x^{-1}\left(\mathbf{c} + \boldsymbol{C}_p\,\Delta\boldsymbol{p}\right) \,. \tag{15}$$

The latter, however, is just an `ISES` Newton step with the RHS being linearized in Δp.