Proof that the Global Optimum of the Relaxed Constrained Maximum Likelihood Detection Approaches a Point Mass at the MLSD as $c \to 0$

John MacLaren Walsh

Technical Note ASPITRG.TN.12.09.1

Enumerate the range set in which $\mathbf{Z}$ lies as

$$\mathcal{Z} \triangleq \times_{a=1}^{A} \times_{i \in \mathcal{N}(a)} \mathcal{X}_i \triangleq \{\mathbf{z}(0), \mathbf{z}(1), \ldots, \mathbf{z}(K-1)\}$$

where we have used the typical notation of a capital letter for a random vector and a lower case variable for one of its possible realizations. Note that $\mathbf{Y}$ has the same range set as $\mathbf{Z}$. Given a joint distribution $\mathsf{q} \in \mathcal{C}(c)$ on $\mathbf{Z}$ and $\mathbf{Y}$, break it apart into its independent distributions for $\mathbf{Y}$ and $\mathbf{Z}$ as

$$\mathsf{q}(\mathbf{y}, \mathbf{z}) = q_{\mathbf{Y}}(\mathbf{y}) q_{\mathbf{Z}}(\mathbf{z})$$

It will be convenient to parameterize the distribution for $\mathbf{Z}$ and for $\mathbf{Y}$ under q as

$$\mathbf{p} := [\mathsf{q}_{\mathbf{Y}}(\mathbf{z}(0)), \mathsf{q}_{\mathbf{Y}}(\mathbf{z}(1)), \mathsf{q}_{\mathbf{Y}}(\mathbf{z}(2)), \ldots, \mathsf{q}_{\mathbf{Y}}(\mathbf{z}(K-1))]^T \quad (1)$$

$$\mathbf{q} := [\mathsf{q}_{\mathbf{Z}}(\mathbf{z}(0)), \mathsf{q}_{\mathbf{Z}}(\mathbf{z}(1)), \mathsf{q}_{\mathbf{Z}}(\mathbf{z}(2)), \ldots, \mathsf{q}_{\mathbf{Z}}(\mathbf{z}(K-1))]^T \quad (2)$$

(Note that these vectors are for notational convenience in the following development defined as length K rather than as length $K-1$ as in the paper.) Let $\hat{\mathcal{C}}(c)$ be the set of all $(\mathbf{p}, \mathbf{q})$ such that the associated q is in $\mathcal{C}(c)$, and similarly $\hat{\mathcal{P}}$ be the set of $\mathbf{p}, \mathbf{q}$ such that $\mathsf{q} \in \mathcal{P}$. Then, since

$$\mathcal{C}(c) = \left\{ \mathsf{q} \middle| \log \left(\sum_{\mathbf{z} \in \mathcal{Z}} \mathbb{P}_{\mathsf{q}}[\mathbf{Z} = \mathbf{z}] \mathbb{P}_{\mathsf{q}}[\mathbf{Z} = \mathbf{z}] \right) = c \right\} \cap \mathcal{P}$$

The constraint set in this parametrization then becomes

$$\hat{\mathcal{C}}(c) = \underbrace{\{\mathbf{p}, \mathbf{q} \,|\log(\mathbf{p} \cdot \mathbf{q}) = c\}}_{\mathcal{G}(c)} \cap \hat{\mathcal{P}}$$

Additionally, we can rewrite the objective function using this parametrization. In particular, define the length K vectors $\mathbf{u}$ and $\mathbf{v}$ whose kth elements are

$$u_k := \prod_{a=1}^{A} f_a(\mathbf{z}_a(k)), \quad v_k := \mathbf{1}\{z_{a,i}(k) = z_{b,i}(k) \ \forall a, b \in \mathcal{M}(i)\} \quad k \in \{0, 1, \ldots, K-1\} \quad (3)$$

Then the objective function becomes

$$J(\mathbf{q}, \mathbf{p}) := \log(\mathbf{u} \cdot \mathbf{q}) + \log(\mathbf{v} \cdot \mathbf{p}).$$

Hence the constrained relaxed maximum likelihood sequence detection optimization takes the form in this parametrization

$$(\mathbf{q}^*(c), \mathbf{p}^*(c)) \in \arg \max_{(\mathbf{q}, \mathbf{p}) \in \mathcal{G}(c) \cap \hat{\mathcal{P}}} \log(\mathbf{u} \cdot \mathbf{q}) + \log(\mathbf{v} \cdot \mathbf{p}) \quad (4)$$

What we would like to prove is that for $c < 0$

$$\lim_{c \to 0} \mathbf{q}^*(c) = \mathbf{q}^*(0) \quad \lim_{c \to 0} \mathbf{p}^*(c) = \mathbf{p}^*(0)$$

That is, the global optimum for $c = 0$ is continuous at the boundary from the left. This is turn will imply that $\mathbf{q}^*(c)$ and $\mathbf{p}^*(c)$ approach a point mass at the MLSD, $\mathbf{e}_{\mathbf{z}_{MLSD}}$ as $c \to 0$ because Lemma 1 from the paper proves that $\mathbf{q}^*(0) = \mathbf{p}^*(0) = \mathbf{e}_{\mathbf{z}_{MLSD}}$.

In this regard, the following lemma concerning the geometry of the constraint set will be useful. The lemma shows that the constraint set $\hat{\mathcal{C}}(c)$ for c reasonably close to 0 is contained within the union of ∞ norm balls of radius $1 - \exp(c)$ centered at point mass distributions for which $\mathbb{P}[\mathbf{Z} = \mathbf{Y} = \mathbf{z}] = 1$ for some $\mathbf{z} \in \mathcal{Z}$.

Lemma: Let $\gamma := \exp(c)$. For any c such that $\gamma > \frac{2}{3}$ we have the containment

$$\hat{C}(c) \subset \bigcup_{\mathbf{z} \in \mathcal{Z}} \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}}) \tag{5}$$

where $\mathbf{e}_{\mathbf{z}}$ is the vector whose kth element is 1 if $\mathbf{z}(k) = \mathbf{z}$ and 0 otherwise (i.e. the coordinates for the point mass distribution at $\mathbf{z}$), and $\mathcal{B}^{\infty}_{r}(\mathbf{w})$ is the K dimensional ball

$$\mathcal{B}^{\infty}_{\gamma}(\mathbf{w}) := \{\mathbf{a} \,|\, \|\mathbf{a} - \mathbf{w}\|_{\infty} \leq \gamma\}$$

Proof: Begin by defining

$$i := \arg\max_k p_k, \quad j = \arg\max_k q_k$$

and

$$p_{max} = \max_k p_k, \quad q_{max} = \max_k q_k$$

Observe then that

$$\mathbf{p}^T\mathbf{q} \leq \max_{\mathbf{q}' | \mathbf{q}' \geq \mathbf{0},\ \mathbf{1}^T\mathbf{q}' = 1} \mathbf{p}^T\mathbf{q}' = \max_{\mathbf{q}' | \mathbf{q}' \geq \mathbf{0},\ \mathbf{1}^T\mathbf{q}' = 1} \sum_k p_k q'_k \leq \max_{\mathbf{q}' | \mathbf{q}' \geq \mathbf{0},\ \mathbf{1}^T\mathbf{q}' = 1} \sum_k p_{max} q'_k = p_{max}$$

Similarly $\mathbf{p}^T\mathbf{q} \leq q_{max}$. Hence, $(\mathbf{p}, \mathbf{q}) \in \hat{\mathcal{C}}(c)$ implies that

$$q_{max} \geq \gamma, \quad p_{max} \geq \gamma \tag{6}$$

Furthermore, since

$$q_{max} = 1 - \sum_{k \neq j} q_k, \quad p_{max} = 1 - \sum_{k \neq i} p_k$$

We have, after combining with (8), that $(\mathbf{p}, \mathbf{q}) \in \hat{\mathcal{C}}(c)$ implies that

$$\sum_{k \neq i} p_k \leq 1 - \gamma, \quad \sum_{k \neq j} q_k \leq 1 - \gamma \tag{7}$$

which in turn implies that

$$p_k \leq 1 - \gamma \quad \forall k \neq i \quad \text{and} \quad q_k \leq 1 - \gamma \quad \forall k \neq j \tag{8}$$

Furthermore, (8) implies

$$1 - q_{max} \leq 1 - \gamma, \quad 1 - p_{max} \leq 1 - \gamma \tag{9}$$

But, for probability mass function vectors $\mathbf{p}$ and $\mathbf{q}$, the requirements (10) and (11) are equivalent to

$$\|\mathbf{e}_j - \mathbf{q}\|_{\infty} \leq 1 - \gamma, \quad \|\mathbf{e}_i - \mathbf{p}\|_{\infty} \leq 1 - \gamma$$

Summarizing what we have learned so far, we have

$$(\mathbf{p}, \mathbf{q}) \in \mathcal{C}(c) \Rightarrow \mathbf{p} \in \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_i) \text{ and } \mathbf{q} \in \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_j)$$

for some i and j. Finally, we must show that it suffices to take $i = j$ when $\gamma \geq \frac{2}{3}$. Suppose $i \neq j$, then (10) and (9) imply that

$$\mathbf{p}^T\mathbf{q} = p_i q_i + \sum_{k \neq i} p_k q_k \leq p_i q_i + q_j \sum_{k \neq i} p_k \leq 1(1 - \gamma) + 1(1 - \gamma) = 2(1 - \gamma) \tag{10}$$

But, since $\mathbf{p}, \mathbf{q} \in \mathcal{C}(c)$ we must then also have

$$2(1-\gamma) \geq \mathbf{p}^T \mathbf{q} = \gamma$$

which is not possible when $\gamma > \frac{2}{3}$. Hence we must have $i = j$ for $\gamma > \frac{2}{3}$. This proves the proposition. ■

Note that for $\gamma > \frac{1}{2}$, $\mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e_z}) \cap \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e_{z'}}) = \emptyset$ for any $\mathbf{z} \neq \mathbf{z'}$ with $\mathbf{z}, \mathbf{z'} \in \mathcal{Z}$, so that the union (5) is a disjoint union for $\gamma > \frac{1}{2}$. Also note that for any $\mathbf{z} \in \mathcal{Z}$, the constraint set $\hat{\mathcal{C}}(c)$ intersects the ball $\mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e_z}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e_z})$, so that $\hat{\mathcal{C}}(c) \cap \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e_z}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e_z}) \neq \emptyset$. Thus, when $\gamma > \frac{1}{2}$ we can express the optimum objective value for our original constrained optimization (4) as

$$\max_{\mathbf{z} \in \mathcal{Z}} \max \left\{ J(\mathbf{q}, \mathbf{p}) \,\middle|\, (\mathbf{q}, \mathbf{p}) \in \hat{\mathcal{C}}(c) \cap \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e_z}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e_z}) \right\} \tag{11}$$

Now, consider $\mathbf{x}_{MLSD} := \arg\max_{\mathbf{x} \in \mathcal{X}} \prod_{a=1}^{A} f_a(\mathbf{x}_a)$, the most likely variable node vector realization, and let $\mathbf{z}_{MLSD}$ be its associated replication in $\mathcal{Z}$ with $z^a_{i,MLSD} = x_{i,MLSD}$ for all $a \in \mathcal{M}(i)$ for all $i \in \{1, \ldots, M\}$. By virtue of its selection so that the edge copies of the same variable node are equal, $\mathbf{v} \cdot \mathbf{e}_{\mathbf{z}_{MLSD}} = 1$, and also by virtue of its selection

$$\mathbf{u} \cdot \mathbf{e}_{\mathbf{z}_{MLSD}} = \prod_{a=1}^{A} f_a(\mathbf{x}_{a,MLSD}) = f(\mathbf{x}_{MLSD})$$

Hence

$$\begin{aligned}
&\max \left\{ J(\mathbf{q}, \mathbf{p}) \,\middle|\, (\mathbf{q}, \mathbf{p}) \in \hat{\mathcal{C}}(c) \cap \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_{MLSD}}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_{MLSD}}) \right\} && (12)\\
&\quad \geq \min \left\{ J(\mathbf{q}, \mathbf{p}) \,\middle|\, (\mathbf{q}, \mathbf{p}) \in \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_{MLSD}}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_{MLSD}}) \right\}\\
&\quad \geq \log\left(f(\mathbf{x}_{MLSD}) - \mathbf{u} \cdot \mathbf{1}(1-\gamma)\right) + \log\left(1 - \mathbf{v} \cdot \mathbf{1}(1-\gamma)\right) && (13)
\end{aligned}$$

where the latter inequality (15) follows from the general principle that for any $\mathbf{a} \geq 0$, $\min_{\mathbf{w} \in \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{0})} \mathbf{a} \cdot \mathbf{w} = -\mathbf{a} \cdot \mathbf{1}(1-\gamma)$ (and hence $\mathbf{a} \cdot \mathbf{1}(1-\gamma) \geq \mathbf{a} \cdot \mathbf{w} \geq -\mathbf{a} \cdot \mathbf{1}(1-\gamma)$).

Next, consider any other $\mathbf{x'} \in \mathcal{X}$ such that $f(\mathbf{x'}) < f(\mathbf{x}_{MLSD})$, and let $\mathbf{z'}$ be its edge variable replication, so that $z^a_i = x_i$ for all $a \in \mathcal{M}(i)$ for all $i \in \{1, \ldots, M\}$. For it, again $\mathbf{v} \cdot \mathbf{e_{z'}} = 1$, while $\mathbf{u} \cdot \mathbf{e_{z'}} = f(\mathbf{x'})$. The maximum of the objective within this ball intersected with the constraint set is then

$$\begin{aligned}
&\max \left\{ J(\mathbf{q}, \mathbf{p}) \,\middle|\, (\mathbf{q}, \mathbf{p}) \in \hat{\mathcal{C}}(c) \cap \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e_{z'}}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e_{z'}}) \right\}\\
&\quad \leq \max \left\{ J(\mathbf{q}, \mathbf{p}) \,\middle|\, (\mathbf{q}, \mathbf{p}) \in \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e_{z'}}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e_{z'}}) \right\}\\
&\quad \leq \log\left(f(\mathbf{x'}) + \mathbf{u} \cdot \mathbf{1}(1-\gamma)\right) + \log\left(1 + \mathbf{v} \cdot \mathbf{1}(1-\gamma)\right) && (14)
\end{aligned}$$

Next, note that (16) will be < (15) whenever

$$1 - \gamma < \frac{f(\mathbf{x}_{MLSD}) - f(\mathbf{x'})}{2\mathbf{u} \cdot \mathbf{1} + \mathbf{v} \cdot \mathbf{1}(f(\mathbf{x}_{MLSD}) + f(\mathbf{x'}))}$$

(where the right side is < 1) thus we will have

$$\begin{aligned}
&\max \left\{ J(\mathbf{q}, \mathbf{p}) \,\middle|\, (\mathbf{q}, \mathbf{p}) \in \hat{\mathcal{C}}(c) \cap \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_{MLSD}}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_{MLSD}}) \right\}\\
&\quad \geq \max \left\{ J(\mathbf{q}, \mathbf{p}) \,\middle|\, (\mathbf{q}, \mathbf{p}) \in \hat{\mathcal{C}}(c) \cap \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e_{z'}}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e_{z'}}) \right\}
\end{aligned}$$

for c sufficiently close to 0 (i.e. γ sufficiently close to 1).

Finally, the last possibility for $\mathbf{z}$ is to consider $\mathbf{z}_*$ for which some edge variables incident on the same variable node are not equal (i.e. an inconsistent $\mathbf{z}_* \in \mathcal{Z}$). For such a $\mathbf{z}_*$

$$\begin{aligned}
&\max \left\{ J(\mathbf{q}, \mathbf{p}) \,\middle|\, (\mathbf{q}, \mathbf{p}) \in \hat{\mathcal{C}}(c) \cap \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_*}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_*}) \right\}\\
&\quad \leq \max \left\{ J(\mathbf{q}, \mathbf{p}) \,\middle|\, (\mathbf{q}, \mathbf{p}) \in \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_*}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_*}) \right\}\\
&\quad \leq \log\left(\mathbf{u} \cdot \mathbf{e}_{\mathbf{z}_*} + \mathbf{u} \cdot \mathbf{1}(1-\gamma)\right) + \log\left(\mathbf{v} \cdot \mathbf{1}(1-\gamma)\right) && (15)
\end{aligned}$$

This allows us to observe that (17) will be < (15) whenever

$$1 - \gamma < \frac{f(\mathbf{x}_{MLSD})}{\mathbf{u} \cdot \mathbf{1} + \mathbf{v} \cdot \mathbf{1}(f(\mathbf{x}_{MLSD}) + \mathbf{u} \cdot \mathbf{e}_{\mathbf{z}_*})}$$

where again the right hand side is always < 1. Thus, for γ close enough to 1 (and thus c close enough to 0) we also have

$$\max\left\{ J(\mathbf{q},\mathbf{p}) \,\middle|\, (\mathbf{q},\mathbf{p}) \in \hat{\mathcal{C}}(c) \cap \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_{MLSD}}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_{MLSD}}) \right\}$$
$$\geq \max\left\{ J(\mathbf{q},\mathbf{p}) \,\middle|\, (\mathbf{q},\mathbf{p}) \in \hat{\mathcal{C}}(c) \cap \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_*}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_*}) \right\}$$

But this implies that the outer maximum over $\mathbf{z} \in \mathcal{Z}$ in (13) must, for $c < 0$ sufficiently close to 0, be attained by $\mathbf{z}_{MLSD}$. This then implies that $\mathbf{q}^*(c), \mathbf{p}^*(c)$ must lie in the ball $\mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_{MLSD}}) \times \mathcal{B}^{\infty}_{1-\gamma}(\mathbf{e}_{\mathbf{z}_{MLSD}})$ for this $c < 0$ sufficiently close to zero. Since $\gamma := \exp(c)$, and the balls that $\mathbf{q}^*(c), \mathbf{p}^*(c)$ must lie in have radius $1 - \gamma = 1 - \exp(c)$, we finally must have that as $c \to 0$, $\mathbf{q}^*(c), \mathbf{p}^*(c)$ approach the center of these balls, which is nothing other than $\mathbf{e}_{\mathbf{z}_{MLSD}}, \mathbf{e}_{\mathbf{z}_{MLSD}}$ which is $\mathbf{q}^*(0), \mathbf{p}^*(0)$. Hence the desired continuity is established.