

PROOFS FOR "OFF-THE-GRID MULTI-PITCH ESTIMATION USING OPTIMAL TRANSPORT"

Anton Björkman and Filip Elvander

Department of Information and Communications Engineering, Aalto University, Finland

1. PROOF

We aim to solve the problem

$$\underset{\mathbf{M} \in \mathbb{R}_+^{F \times P}, \omega_0 \in \mathbb{R}^P}{\text{minimize}} \quad \langle \mathbf{C}_{\omega_0}, \mathbf{M} \rangle + \gamma \|\hat{\mathbf{r}} - \mathbf{A}\mathbf{M}\mathbf{1}\|_2^2 + \epsilon D(\mathbf{M}), \quad (1)$$

where, $D(\mathbf{M}) = \sum_{f,p} [\mathbf{M}]_{f,p} \log[\mathbf{M}]_{f,p} - [\mathbf{M}]_{f,p} + 1$ is the entropic regularization term. To this end, we employ a block-coordinate descent scheme, where the blocks correspond to \mathbf{M} and ω_0 . By fixating ω_0 , we get the problem

$$\underset{\mathbf{M} \in \mathbb{R}_+^{F \times P}}{\text{minimize}} \quad \langle \mathbf{C}_{\omega_0}, \mathbf{M} \rangle + \gamma \|\hat{\mathbf{r}} - \mathbf{A}\mathbf{M}\mathbf{1}\|_2^2 + \epsilon D(\mathbf{M}), \quad (2)$$

For this, the following proposition holds.

Proposition 1. *For a given \mathbf{C}_{ω_0} , the dual problem of (2) is*

$$\underset{\lambda \in \mathbb{R}^T}{\text{maximize}} \quad \lambda^T \hat{\mathbf{r}} - \epsilon \exp\left(\frac{1}{\epsilon} \lambda^T \mathbf{A}\right) (\mathbf{K}\mathbf{1}_P) - \frac{1}{4\gamma} \|\lambda\|_2^2 + \epsilon FP$$

where $\mathbf{K} = \exp(-\frac{1}{\epsilon} \mathbf{C}_{\omega_0})$. The optimal transport plan is then given by inserting the maximizing λ in

$$\mathbf{M} = \text{diag}\left(\exp\left(\frac{1}{\epsilon} \mathbf{A}^T \lambda\right)\right) \mathbf{K} \text{diag}(\mathbf{1}_P). \quad (3)$$

Furthermore, the maximizing λ is the unique root of the equation

$$\mathbf{A}\left(\exp\left(\frac{1}{\epsilon} \mathbf{A}^T \lambda\right) \odot \mathbf{K}\mathbf{1}_P\right) - \hat{\mathbf{r}} + \frac{1}{2\gamma} \lambda = 0, \quad (4)$$

where \odot denotes elementwise multiplication.

Proof. From (2) we form the equivalent problem, given by

$$\underset{\mathbf{M} \in \mathbb{R}_+^{F \times P}, \Delta \in \mathbb{R}^T}{\text{minimize}} \quad \langle \mathbf{C}_{\omega_0}, \mathbf{M} \rangle + \gamma \|\Delta\|_2^2 + \epsilon D(\mathbf{M}), \quad (5)$$

$$\text{s.t. } \hat{\mathbf{r}} = \Delta + \mathbf{A}\mathbf{M}\mathbf{1}. \quad (6)$$

The Lagrangian is then given by

$$\begin{aligned} \mathcal{L}(\mathbf{M}, \Delta, \lambda) &= \langle \mathbf{C}_{\omega_0}, \mathbf{M} \rangle + \gamma \|\Delta\|_2^2 + \epsilon D(\mathbf{M}) \\ &\quad + \lambda^T (\hat{\mathbf{r}} - \Delta - \mathbf{A}\mathbf{M}\mathbf{1}). \end{aligned} \quad (7)$$

This Lagrangian is strongly convex in \mathbf{M} and Δ with a unique minimizer

$$\mathbf{M} = \text{diag}\left(\exp\left(\frac{1}{\epsilon} \mathbf{A}^T \lambda\right)\right) \mathbf{K} \text{diag}(\mathbf{1}_P), \quad (8)$$

$$\Delta = \frac{\lambda}{2\gamma}, \quad (9)$$

where $\mathbf{K} = \exp(-\frac{1}{\epsilon} \mathbf{C}_{\omega_0})$. Inserted into the Lagrangian yields the dual problem

$$\underset{\lambda \in \mathbb{R}^T}{\text{maximize}} \quad \lambda^T \hat{\mathbf{r}} - \epsilon \exp\left(\frac{1}{\epsilon} \lambda^T \mathbf{A}\right) (\mathbf{K}\mathbf{1}_P) - \frac{1}{4\gamma} \|\lambda\|_2^2 + \epsilon FP.$$

Differentiation with respect to λ , set to 0, then yields

$$\mathbf{A}\left(\exp\left(\frac{1}{\epsilon} \mathbf{A}^T \lambda\right) \odot \mathbf{K}\mathbf{1}_P\right) - \hat{\mathbf{r}} + \frac{1}{2\gamma} \lambda = 0. \quad (10)$$

□

1.1. Solve for ω_0

We now fixate λ and solve for ω_0 . The problem in (1) then becomes

$$\underset{\omega_0 \in \mathbb{R}^P}{\text{minimize}} \quad \langle \mathbf{C}_{\omega_0}, \mathbf{M} \rangle, \quad (11)$$

where each element of \mathbf{C}_{ω_0} is given by

$$c(\omega_f, \omega_0^{(p)}) = \left(\frac{\omega_f}{\omega_0^{(p)}} - \hat{h}(\omega_f, \hat{\omega}_0^{(p)})\right)^2, \quad (12)$$

where $\hat{\omega}_0^{(p)}$ is our prior estimate for pitch p . Substituting this yields the minimization problem

$$\underset{\omega_0 \in \mathbb{R}^P}{\text{minimize}} \quad \sum_{p=1}^P \sum_{f=1}^F \left(\frac{\omega_f}{\omega_0^{(p)}} - \hat{h}(\omega_f, \hat{\omega}_0^{(p)})\right)^2 [\mathbf{M}]_{f,p},$$

where $[\mathbf{M}]_{f,p}$ is the element of \mathbf{M} that corresponds to frequency f and pitch p . Differentiation with respect to $\omega_0^{(p)}$ and setting the result to 0 yields

$$\sum_{f=1}^F -2 \left(\frac{\omega_f}{\omega_0^{(p)}} - \hat{h}(\omega_f, \hat{\omega}_0^{(p)})\right) \left(\frac{\omega_f}{(\omega_0^{(p)})^2}\right) [\mathbf{M}]_{f,p} = 0,$$

which, when solved for $\omega_0^{(p)}$, finally yields the desired result,

$$\omega_0^{(p)} = \frac{\sum_{f=1}^F \omega_f^2 [\mathbf{M}]_{f,p}}{\sum_{f=1}^F \hat{h}(\omega_f, \hat{\omega}_0^{(p)}) \omega_f [\mathbf{M}]_{f,p}}. \quad (13)$$

1.2. Interval

We want to show that for an estimate $\hat{\omega}_0$ of a harmonic signal of harmonic order H and with fundamental frequency $\omega_{0,min}$, such that

$$\frac{-1}{2H+1} \omega_{0,min} \leq \hat{\omega}_0 - \omega_{0,min} \leq \frac{1}{2H-1} \omega_{0,min}, \quad (14)$$

a local quadratic approximation of the objective function in the transport problem, minimized by $\omega_{0,min}$, is obtained, using the ground-cost function

$$c(\omega_f, \omega_0^{(p)}) = \min_{h \in \llbracket H \rrbracket} \left(\frac{\omega_f}{\omega_0^{(p)}} - h \right)^2. \quad (15)$$

Proof. Consider a harmonic signal with fundamental frequency $\omega_{0,min}$, of harmonic order H . By finding the minimizing h in (15) for each non-zero spectral component of the harmonic series for $\omega_{0,min}$, it may be noted that a local quadratic approximation of the objective function is obtained. The non-zero components of the harmonic series are all integer multiples of the fundamental frequency, given by $h\omega_{0,min}$. By letting

$$\hat{\omega}_0 = \omega_{0,min} + \Delta,$$

where $\Delta \in \mathbb{R}$, the approximation for each harmonic h of the pitch is given by

$$h\hat{\omega}_0 = h(\omega_{0,min} + \Delta).$$

We here note that the deviation from $h\omega_{0,min}$ increases linearly by the harmonic order. Thus, we only need to look at the highest harmonic, H . Inserting this into the estimate of the harmonic number grants

$$\hat{h}(H\omega_{0,min}, (\omega_{0,min} + \Delta)) = \operatorname{argmin}_{h \in \llbracket H \rrbracket} \left(\frac{H\omega_{0,min}}{(\omega_{0,min} + \Delta)} - h \right)^2.$$

What we now wish to obtain is for the optimal h to be H , which is achieved when both

$$(q - H)^2 \leq (q - (H + 1))^2, \quad (16)$$

and

$$(q - H)^2 \leq (q - (H - 1))^2, \quad (17)$$

hold, where $q = \frac{H\omega_{0,min}}{(\omega_{0,min} + \Delta)}$. We first look at (16), where we expand the terms and get

$$0 \leq -2q + 2H + 1,$$

which yields

$$q \leq H + \frac{1}{2}.$$

Substituting q gives

$$\frac{H\omega_{0,min}}{(\omega_{0,min} + \Delta)} \leq H + \frac{1}{2},$$

which when solved for Δ leads to the inequality

$$\Delta \geq \frac{-1}{2H + 1} \omega_{0,min},$$

which holds when $\omega_{0,min} + \Delta > 0$. With similar calculations, we solve (17), granting the upper bound

$$\Delta \leq \frac{1}{2H - 1} \omega_{0,min},$$

which also holds when $\omega_{0,min} + \Delta > 0$. By substituting $\Delta = \hat{\omega}_0 - \omega_{0,min}$, we finally get the desired interval,

$$\frac{-1}{2H + 1} \omega_{0,min} \leq \hat{\omega}_0 - \omega_{0,min} \leq \frac{1}{2H - 1} \omega_{0,min}. \quad (18)$$

Finally, it may be noted that for the true harmonic number of each spectral component of the signal, the ground-cost in (12) is 0 for $\omega_{0,min}$. \square