Source Separation and Clustering of Phase-Locked Subspaces: Derivations and Proofs

Miguel Almeida, Jan-Hendrik Schleimer, José Bioucas-Dias, Ricardo Vigário

***Abstract*—Due to space limitations, our submission "Source Separation and Clustering of Phase-Locked Subspaces", accepted for publication on the IEEE Transactions on Neural Networks in 2011, presented some results without proof. Those proofs are provided in this paper.**

***Index Terms*—phase-locking, synchrony, source separation, clustering, subspaces**

Appendix A
Gradient of $|\varrho|^2$ in RPA

In this section we derive that the gradient of $|\varrho|^2$ is given by Eq. 6 of [1], where $|\varrho|$ is defined as in Eq. 5 of [1]. Recall that $\Delta\phi(t) = \phi(t) - \psi(t)$, where $\phi(t)$ is the phase of the estimated source $y(t) = \mathbf{w}^T\mathbf{x}(t)$ and $\psi(t)$ is the phase of the reference $u(t)$. Further, define $\varrho \equiv |\varrho|e^{\mathrm{i}\Phi}$.

We begin by noting that $|\varrho|^2 = (|\varrho|\cos(\Phi))^2 + (|\varrho|\sin(\Phi))^2$, so that

$$\begin{aligned}\nabla|\varrho|^2 &= \nabla(|\varrho|\cos(\Phi))^2 + \nabla(|\varrho|\sin(\Phi))^2 \\ &= 2|\varrho|\left[\cos(\Phi)\nabla(|\varrho|\cos(\Phi)) + \sin(\Phi)\nabla(|\varrho|\sin(\Phi))\right].\end{aligned}$$

Note that we have $\frac{1}{T}\sum_{t=1}^{T}\cos(\Delta\phi(t))) = |\varrho|\cos(\Phi)$ and $\frac{1}{T}\sum_{t=1}^{T}\sin(\Delta\phi(t))) = |\varrho|\sin(\Phi)$, so we get

$$\begin{aligned}\nabla|\varrho|^2 &= 2|\varrho|\left\{\cos(\Phi)\nabla\left[\frac{1}{T}\sum_{t=1}^{T}\cos(\Delta\phi(t)))\right] + \right. \\ &\qquad \left. + \sin(\Phi)\nabla\left[\frac{1}{T}\sum_{t=1}^{T}\sin(\Delta\phi(t)))\right]\right\} \\ &= \frac{2|\varrho|}{T}\sum_{t=1}^{T}\Big[\sin(\Phi)\cos(\Delta\phi(t)) - \cos(\Phi)\sin(\Delta\phi(t))\Big]\times \\ &\qquad \times \nabla\Delta\phi(t) \\ &= \frac{2|\varrho|}{T}\sum_{t=1}^{T}\sin[\Phi - \Delta\phi(t)]\nabla\Delta\phi(t). \qquad (1)\end{aligned}$$

Miguel Almeida is with the Institute of Telecommunications, Superior Technical Institute, Portugal. Email: malmeida@lx.it.pt

Jan-Hendrik Schleimer is with the Bernstein Center for Computational Neuroscience, Humboldt University, Berlin. Email: jan-hendrik.schleimer@bccn-berlin.de

José Bioucas-Dias is with the Institute of Telecommunications, Superior Technical Institute, Portugal. Email: bioucas@lx.it.pt

Ricardo Vigário is with the Adaptive Informatics Research Centre, Aalto University, Finland. Email: ricardo.vigario@hut.fi

Let's now take a closer look on $\nabla\Delta\phi(t)$. Note that

$$\phi(t) = \arctan\left(\frac{\mathbf{w}^\mathsf{T}\mathbf{x_h}(t)}{\mathbf{w}^\mathsf{T}\mathbf{x}(t)}\right) \quad \text{or}$$
$$\phi(t) = \arctan\left(\frac{\mathbf{w}^\mathsf{T}\mathbf{x_h}(t)}{\mathbf{w}^\mathsf{T}\mathbf{x}(t)}\right) + \pi.$$

Because of this we can say, if $\mathbf{w}^\mathsf{T}\mathbf{x}(t) \neq 0$, that $\nabla\phi(t) = \nabla\arctan\left(\frac{\mathbf{w}^\mathsf{T}\mathbf{x_h}(t)}{\mathbf{w}^\mathsf{T}\mathbf{x}(t)}\right)$. On the other hand, since $\Delta\phi(t) = \phi(t) - \psi(t)$ and $\psi(t)$ does not depend on $\mathbf{w}$, we have (we will omit the time dependence for the sake of clarity):

$$\begin{aligned}\nabla\Delta\phi &= \nabla\phi - \nabla\psi = \nabla\phi = \nabla\arctan\left(\frac{\mathbf{w}^\mathsf{T}\mathbf{x_h}}{\mathbf{w}^\mathsf{T}\mathbf{x}}\right) \\ &= \frac{\mathbf{x_h}\cdot\mathbf{w}^\mathsf{T}\mathbf{x} - \mathbf{x}\cdot\mathbf{w}^\mathsf{T}\mathbf{x_h}}{\left[1 + \left(\frac{\mathbf{w}^\mathsf{T}\mathbf{x_h}}{\mathbf{w}^\mathsf{T}\mathbf{x}}\right)^2\right]\cdot\left(\mathbf{w}^\mathsf{T}\mathbf{x}\right)^2} = \frac{\boldsymbol{\Gamma}_x(t)\cdot\mathbf{w}}{Y^2(t)},\end{aligned}$$

where $Y^2(t) = \left(\mathbf{w}^\mathsf{T}\mathbf{x}(t)\right)^2 + \left(\mathbf{w}^\mathsf{T}\mathbf{x_h}(t)\right)^2$ is the squared magnitude of the estimated source, and $\boldsymbol{\Gamma}_x(t) = \mathbf{x_h}(t)\mathbf{x}^\mathsf{T}(t) - \mathbf{x}(t)\mathbf{x_h}^\mathsf{T}(t)$, thus $\boldsymbol{\Gamma}_{x_{ij}}(t) = X_i(t)X_j(t)\sin(\phi_i(t) - \phi_j(t))$.

We can now replace $\nabla\Delta\phi(t)$ in (1) to obtain

$$\begin{aligned}\nabla|\varrho|^2 &= \frac{2|\varrho|}{T}\left[\sum_{t=1}^{T}\frac{\sin[\Phi - \Delta\phi(t)]}{Y^2(t)}\boldsymbol{\Gamma}_x(t)\right]\mathbf{w} \\ &= 2|\varrho|\left\langle\frac{\sin[\Phi - \Delta\phi(t)]}{Y^2(t)}\boldsymbol{\Gamma}_x(t)\right\rangle\mathbf{w}. \qquad (2)\end{aligned}$$

Appendix B
Gradient of J_l in IPA

In this section we show that the gradient of J_l in Eq. 7 of [1] is given by Eq. 8 of [1]. Throughout this whole section, we will omit the dependence on the subspace l, for the sake of clarity. In other words, we are assuming (with no loss of generality) that only one subspace was found. Whenever we write $\mathbf{W}$, $\mathbf{y}$, y_m or $\mathbf{z}$, we will be referring to $\mathbf{W}_l$, $\mathbf{y}_l$, $(\mathbf{y}_l)_m$ or $\mathbf{z}_l$.

The derivative of $\log|\det\mathbf{W}|$ is $\mathbf{W}^{-\mathsf{T}}$. We will therefore focus on the gradient of the first term of Eq. 7 of [1], which we will denote by P:

$$P \equiv \frac{1-\lambda}{N^2}\sum_{m,n}|\varrho_{mn}|^2.$$

Let's rewrite P as $P = \frac{1-\lambda}{N^2}\sum_{m,n}p_{mn}$ with $p_{mn} = |\varrho_{mn}|^2$. Define $\Delta\phi_{mn} = \phi_m - \phi_n$. Omitting the time dependency, we have

$$\nabla_{\mathbf{w}_j} p_{mn} = 2|\varrho_{mn}|\nabla_{\mathbf{w}_j} \left|\left\langle e^{\mathrm{i}\Delta\phi_{mn}}\right\rangle\right|$$
$$= |\varrho_{mn}| \times \left(\left\langle \cos(\Delta\phi_{mn})\right\rangle^2 + \mathrm{i}\left\langle \sin(\Delta\phi_{mn})\right\rangle^2\right)^{-1/2} \times$$
$$\times \Big[2\left\langle \cos(\Delta\phi_{mn})\right\rangle \nabla_{\mathbf{w}_j}\left\langle \cos(\Delta\phi_{mn})\right\rangle +$$
$$+ 2\left\langle \sin(\Delta\phi_{mn})\right\rangle \nabla_{\mathbf{w}_j}\left\langle \sin(\Delta\phi_{mn})\right\rangle\Big]$$
$$= 2|\varrho_{mn}| \left|\left\langle e^{\mathrm{i}\Delta\phi_{mn}}\right\rangle\right|^{-1} \times$$
$$\times \Big[- \left\langle \cos(\Delta\phi_{mn})\right\rangle \left\langle \sin(\Delta\phi_{mn})\nabla_{\mathbf{w}_j}\Delta\phi_{mn}\right\rangle +$$
$$+ \left\langle \sin(\Delta\phi_{mn})\right\rangle \left\langle \cos(\Delta\phi_{mn})\nabla_{\mathbf{w}_j}\Delta\phi_{mn}\right\rangle\Big], \tag{3}$$

where we have interchanged the partial derivative and the time average operators, and used

$$\left(\left\langle \cos(\Delta\phi_{mn})\right\rangle^2 + \mathrm{i}\left\langle \sin(\Delta\phi_{mn})\right\rangle^2\right)^{1/2} = \left|\left\langle e^{\mathrm{i}\Delta\phi_{mn}}\right\rangle\right|.$$

Since ϕ_m is the phase of the m-th measurement, its derivative with respect to any $\mathbf{w}_j$ is zero unless $m = j$ or $n = j$. In the former case, a reasoning similar to Appendix A shows that

$$\nabla_{\mathbf{w}_j}\Delta\phi_{jk} \equiv \nabla_{\mathbf{w}_j}\phi_j - \nabla_{\mathbf{w}_j}\phi_k = \nabla_{\mathbf{w}_j}\phi_j =$$
$$= \frac{[\mathbf{z_h}\cdot\mathbf{z} - \mathbf{z}\cdot\mathbf{z_h}]\cdot\mathbf{w}_j}{Y_j^2} = \frac{\boldsymbol{\Gamma}_z\cdot\mathbf{w}_j}{Y_j^2}, \tag{4}$$

where $\boldsymbol{\Gamma}_z(t) = \mathbf{z_h}(t)\mathbf{z}^\mathsf{T}(t) - \mathbf{z}(t)\mathbf{z_h}^\mathsf{T}(t)$. It is easy to see that $\nabla_{\mathbf{w}_j}\Delta\phi_{jk} = -\nabla_{\mathbf{w}_j}\Delta\phi_{kj}$. Furthermore, $p_{mm} = 1$ by definition, hence $\nabla_{\mathbf{w}_j} p_{mm} = 0$ for all m and j. From these considerations, the only nonzero terms in the derivative of P are of the form

$$\nabla_{\mathbf{w}_j} p_{jk} = \nabla_{\mathbf{w}_j} p_{kj} = 2|\varrho_{jk}| \left|\left\langle e^{\mathrm{i}(\phi_j-\phi_k)}\right\rangle\right|^{-1} \times$$
$$\times \left[-\left\langle \cos(\Delta\phi_{jk})\right\rangle \left\langle \sin(\Delta\phi_{jk})\frac{\boldsymbol{\Gamma}_z\cdot\mathbf{w}_j}{Y_j^2}\right\rangle + \right.$$
$$\left. + \left\langle \sin(\Delta\phi_{jk})\right\rangle \left\langle \cos(\Delta\phi_{jk})\frac{\boldsymbol{\Gamma}_z\cdot\mathbf{w}_j}{Y_j^2}\right\rangle\right]. \tag{5}$$

We now define $\Psi_{jk} \equiv \langle\phi_j - \phi_k\rangle = \langle\Delta\phi_{jk}\rangle$. Plugging in this definition into Eq. (5) we obtain

$$\nabla_{\mathbf{w}_j} p_{jk} = 2|\varrho_{jk}| \times$$
$$\times \left[-\cos(\Psi_{jk}) \left\langle \sin(\Delta\phi_{jk})\frac{\boldsymbol{\Gamma}_z\cdot\mathbf{w}_j}{Y_j^2}\right\rangle + \right.$$
$$\left. + \sin(\Psi_{jk}) \left\langle \cos(\Delta\phi_{jk})\frac{\boldsymbol{\Gamma}_z\cdot\mathbf{w}_j}{Y_j^2}\right\rangle\right]$$
$$= 2|\varrho_{jk}| \left[\left\langle -\cos\Psi_{jk}\sin\Delta\phi_{jk}\frac{\boldsymbol{\Gamma}_z\cdot\mathbf{w}_j}{Y_j^2}\right\rangle + \right.$$
$$\left. + \left\langle \sin\Psi_{jk}\cos\Delta\phi_{jk}\frac{\boldsymbol{\Gamma}_z\cdot\mathbf{w}_j}{Y_j^2}\right\rangle\right]$$
$$= 2|\varrho_{jk}| \left\langle \sin\left(\Psi_{jk} - \Delta\phi_{jk}\right)\frac{\boldsymbol{\Gamma}_z}{Y_j^2}\right\rangle\cdot\mathbf{w}_j,$$

where we again used $\sin(a-b) = \sin a\cos b - \cos a\sin b$ in the last step. Finally,

$$\nabla_{\mathbf{w}_j} P = \frac{1-\lambda}{N^2}\sum_{m,n}\nabla_{\mathbf{w}_j} p_{mn} = 2\frac{1-\lambda}{N^2}\sum_{m<n}\nabla_{\mathbf{w}_j} p_{mn} =$$
$$= 4\frac{1-\lambda}{N^2}\sum_{k=1}^{N}|\varrho_{jk}|\left\langle \sin\left[\Psi_{jk} - \Delta\phi_{jk}(t)\right]\frac{\boldsymbol{\Gamma}_z(t)}{Y_j(t)^2}\right\rangle\cdot\mathbf{w}_j.$$

which is Eq. 8 of [1].

Appendix C
Gradient of J in pSCA

In this section we derive Eq. 10 of [1] for the gradient of J. Recall that J is given by

$$J \equiv \sum_{j=1}^{P}\left|\sum_{i=1}^{N} u_{ij}\right| = \sum_{j=1}^{P}\left|\sum_{i=1}^{N}\sum_{k=1}^{P} v_{ik}w_{kj}\right|$$

where the w_{kj} are real coefficients that we want to optimize and the v_{ik} are fixed complex numbers. Also recall that Re$(.)$ and Im$(.)$ denote the real and imaginary parts.

We begin by expanding the complex absolute value:

$$\sum_j\left|\sum_{i,k} v_{ik}w_{kj}\right| = \sum_j\left\{\left[\mathrm{Re}\left(\sum_{i,k} v_{ik}w_{kj}\right)\right]^2 + \left[\mathrm{Im}\left(\sum_{i,k} v_{ik}w_{kj}\right)\right]^2\right\}^{1/2}.$$

When computing the derivative in order to w_{kj}, only one term in the leftmost sum matters. Thus,

$$\frac{\partial J}{\partial w_{kj}} =$$
$$= 2\left(\left[\mathrm{Re}\left(\sum_{i,k} v_{ik}w_{kj}\right)\right]^2 + \left[\mathrm{Im}\left(\sum_{i,k} v_{ik}w_{kj}\right)\right]^2\right)^{-1/2} \times$$
$$\times \left[\frac{\partial\left[\mathrm{Re}\left(\sum_{i,k} v_{ik}w_{kj}\right)\right]^2}{\partial w_{kj}} + \frac{\partial\left[\mathrm{Im}\left(\sum_{i,k} v_{ik}w_{kj}\right)\right]^2}{\partial w_{kj}}\right]$$
$$= \frac{1}{\left|\sum_i u_{ij}\right|}\left[\mathrm{Re}\left(\sum_{i,k} v_{ik}w_{kj}\right)\frac{\partial\ \mathrm{Re}\left(\sum_{i,k} v_{ik}w_{kj}\right)}{\partial w_{kj}} + \right.$$
$$\left. + \mathrm{Im}\left(\sum_{i,k} v_{ik}w_{kj}\right)\frac{\partial\ \mathrm{Im}\left(\sum_{i,k} v_{ik}w_{kj}\right)}{\partial w_{kj}}\right]. \tag{6}$$

In the sums inside the derivatives, the sum on k can be dropped as only one of those terms will be nonzero. Therefore,

$$\frac{\partial\ \mathrm{Re}\left(\sum_{i,k} v_{ik}w_{kj}\right)}{\partial w_{kj}} = \frac{\partial\ \mathrm{Re}\left(\sum_{i} v_{ik}w_{kj}\right)}{\partial w_{kj}}$$
$$= \frac{\partial\ \mathrm{Re}\left(\sum_i v_{ik}\right)w_{kj}}{\partial w_{kj}} = \mathrm{Re}\left(\sum_i v_{ik}\right) = \mathrm{Re}\left(\bar{v}_k\right),$$

where we used $\bar{v}_i \equiv \sum_k v_{ki}$ to denote the sum of the i-th column of $\mathbf{V}$. Similarly,

$$\frac{\partial \ \mathrm{Im}\Big(\sum_{i,k} v_{ik} w_{kj}\Big)}{\partial w_{kj}} = \mathrm{Im}\left(\bar{v}_k\right).$$

These results, with the notation $\bar{u}_j \equiv \sum_k v_{jk}$ as the sum of the j-th column of $\mathbf{U}$, can be plugged into Eq. (6) to yield

$$\mathbf{G}_{kj} = \frac{\partial J}{\partial w_{kj}} = \frac{1}{|\bar{u}_j|}\Big[\mathrm{Re}(\bar{v_k}) \times \mathrm{Re}(\bar{u}_j) + \mathrm{Im}(\bar{v_k}) \times \mathrm{Im}(\bar{u}_j)\Big].$$

Appendix D
Mean field

In this section we derive Eq. 9 of [1] for the interaction of an oscillator with the cluster it is part of. We will assume that there are N_j oscillators in this cluster, coupled all-to-all with the same coupling coefficient κ, and that all inter-cluster interactions are weak enough to be disregarded. We begin with Kuramoto's model (Eq. 1 of [1]) omitting the time dependency:

$$\begin{aligned}
\dot{\phi}_i &= \omega_i + \sum_{k\in c_j} \kappa_{ik} \sin(\phi_k - \phi_i) + \sum_{k\notin c_j} \kappa_{ik} \sin(\phi_k - \phi_i) \\
\dot{\phi}_i &= \omega_i + \sum_{k\in c_j} \kappa_{ik} \sin(\phi_k - \phi_i) \\
&= \omega_i + \sum_{k\in c_j} \kappa_{ik} \frac{e^{\mathrm{i}(\phi_j - \phi_i)} - e^{-\mathrm{i}(\phi_j - \phi_i)}}{2\mathrm{i}} \\
&= \omega_i + \frac{e^{-\mathrm{i}\phi_i}}{2\mathrm{i}} \sum_{k\in c_j} \kappa_{ik} e^{\mathrm{i}\phi_k} - \frac{e^{\mathrm{i}\phi_i}}{2\mathrm{i}} \sum_{k\in c_j} \kappa_{ik} e^{-\mathrm{i}\phi_k}.
\end{aligned}$$

We now plug in the definition of mean field $\varrho_{c_j} e^{\mathrm{i}\Phi_{c_j}} = \frac{1}{N_j}\sum_{k\in c_j} e^{\mathrm{i}\phi_k}$ to obtain

$$\begin{aligned}
\dot{\phi}_i &= \omega_i + N_j \frac{e^{-\mathrm{i}\phi_i}}{2\mathrm{i}} \kappa \varrho_{c_j} e^{\mathrm{i}\Phi_{c_j}} - N_j \frac{e^{\mathrm{i}\phi_i}}{2\mathrm{i}} \kappa \varrho_{c_j} e^{-\mathrm{i}\Phi_{c_j}} \\
&= \omega_i + N_j \kappa \varrho_{c_j} \left[\sin(\Phi_{c_j} - \phi_i) - \sin(\phi_i - \Phi_{c_j})\right] \\
&= \omega_i + 2N_j \kappa \varrho_{c_j} \sin(\Phi_{c_j} - \phi_i).
\end{aligned}$$

References

[1] M. Almeida, J.-H. Schleimer, J. Bioucas-Dias, and R. Vigário, "Source separation and clustering of phase-locked subspaces," *IEEE Transactions on Neural Networks* (accepted), 2011.