We present a theorem that specifies general assumptions under which source separation can be performed in the latent space. We then discuss these assumptions for the deglitching on Mars and show how our implementation relates to these assumptions. For sake of simplicity we take $a_1 = 1$.

Theorem 1. *Let $\mathbf{x} = \mathbf{s}_1 + \mathbf{n}$ with $\mathbf{s}_1$ and $\mathbf{n}$ two independent processes. Let us assume we have two processes $\widetilde{\mathbf{s}}_1$ and $\widetilde{\mathbf{n}}$ with $\mathbf{x} = \widetilde{\mathbf{s}}_1 + \widetilde{\mathbf{n}}$.*

Under the following assumptions:

(i) $\mathbf{n}$ has a maximum entropy distribution under moment constraints $\mathbb{E}\{\Phi(\mathbf{n})\}$

(ii) $\widetilde{\mathbf{n}}$ has a maximum entropy distribution under moment constraints $\mathbb{E}\{\Phi(\widetilde{\mathbf{n}})\}$

(iii) $\mathbb{E}\{\Phi(\widetilde{\mathbf{n}})\} = \mathbb{E}\{\Phi(\mathbf{n})\}$

(iv) $\widetilde{\mathbf{s}}_1$ and $\widetilde{\mathbf{n}}$ are independent

(v) The Fourier transform $\widehat{p}_{\mathbf{n}}$ of the distribution $p_{\mathbf{n}}$ of $\mathbf{n}$ is non-zero everywhere.

one has $\mathbf{n} \overset{d}{=} \widetilde{\mathbf{n}}$ and $\mathbf{s}_1 \overset{d}{=} \widetilde{\mathbf{s}}_1$ where the equality is on the distribution of the processes.

Discussion and implementation.

- Assumption (i) is the main assumption. It implies that the processes $\mathbf{n}$ is fully determined by the values $\mathbb{E}\{\Phi(\mathbf{n})\}$, since there is a unique distribution satisfying (i). A maximum entropy process $\mathbf{n}$ under correlation constraints $\mathbb{E}\{\mathbf{n}\mathbf{n}^\top\}$ is a Gaussian process. A wavelet Scattering Covariance captures non-linear correlations, assumption (i) tells us that process $\mathbf{n}$ is a non-Gaussian noise fully characterized by $\mathbb{E}\{\Phi(\mathbf{n})\}$. Now, the Scattering Covariance $\mathbb{E}\{\Phi(\mathbf{n})\}$ was shown to characterize a wide range of non-Gaussian noises (Morel et al. [2022]). In our case, the Mars seismic background noise $\mathbf{n}$ may not be fully characterized by its Scattering Covariance $\mathbb{E}\{\Phi(\mathbf{n})\}$, so that assumption (i) is only verified approximately, depending on the descriptive power of the representation $\mathbb{E}\{\Phi(\mathbf{n})\}$ for $\mathbf{n}$.
- Assumption (ii) is approximately verified, requiring the entropy of $\mathbf{x}$ to be close to the entropy of $\mathbf{n}$, which is typically the case of time-localized signals such as glitch, of comparable amplitude than $\mathbf{n}$. The gradient descent algorithm implements (ii), reconstructed $\widetilde{\mathbf{n}}$ is initialized to $\mathbf{x}$ and is updated until $\Phi(\mathbf{x})$ matches the $\Phi(\mathbf{n}_k)$.
- Assumption (iii) is imposed through the loss term $\mathcal{L}_{\text{prior}}$, up to estimation error of $\Phi(\mathbf{n})$ on a finite number of realizations.
- Assumption (iv) relates to the loss term $\mathcal{L}_{\text{cross}}$ that imposes statistical independence up to the cross-Scattering Covariance.
- Assumption (v) is a technical assumption satisfied for a Gaussian noise $\mathbf{n}$ for which the Fourier transform of $p_{\mathbf{n}}$ is a Gaussian. A non-Gaussian noise $\mathbf{n}$ satisfying (i) has a distribution of the form $p_{\mathbf{n}}(\cdot) = Z_{\boldsymbol{\theta}}^{-1} e^{-\boldsymbol{\theta}\Phi(\cdot)}$. Apart from the coefficients $\text{Ave}(S(\mathbf{n}))$, the scattering covariance Φ is quadratic in $\mathbf{n}$, thus we may assume (v) is still satisfied.

Proof. Part I. One can prove that there exists a unique process $\mathbf{n}$ that maximises entropy under moment constraint $\mathbb{E}\{\Phi(\mathbf{n})\}$, its distribution takes the form $p_{\mathbf{n}}(\cdot) = Z_{\boldsymbol{\theta}}^{-1} e^{-\boldsymbol{\theta}^T\Phi(\cdot)}$ for certain Lagrange multipliers $\boldsymbol{\theta} \in \mathbb{R}^M$ where M is the dimension of Φ. Assumptions (i), (ii), (iii) imply that $\mathbf{n}$ and $\widetilde{\mathbf{n}}$ are the same unique process, meaning $p_{\mathbf{n}} = p_{\widetilde{\mathbf{n}}}$.
Part II. Due to the independence of $\mathbf{s}_1, \mathbf{n}$ and $\widetilde{\mathbf{s}}_1, \widetilde{\mathbf{n}}$ (iv) we have $p_{\mathbf{x}} = p_{\mathbf{s}_1} \star p_n$ and $p_{\mathbf{x}} = p_{\widetilde{s}_1} \star p_{\widetilde{n}}$. Since $p_{\widetilde{\mathbf{n}}} = p_{\mathbf{n}}$ we get $p_{\mathbf{s}_1} \star p_{\mathbf{n}} = p_{\widetilde{\mathbf{s}}_1} \star p_{\widetilde{\mathbf{n}}}$. This is a measure deconvolution problem. Taking the Fourier transform on measures yields

$$(\widehat{p}_{\mathbf{s}_1} - \widehat{p}_{\widetilde{\mathbf{s}}_1})\,\widehat{p}_{\mathbf{n}} = 0.$$

Under assumption (v) we get $p_{\widetilde{\mathbf{s}}_1} = p_{\mathbf{s}_1}$. □

References

Rudy Morel, Gaspar Rochette, Roberto Leonarduzzi, Jean-Philippe Bouchaud, and Stéphane Mallat. Scale dependencies and self-similarity through wavelet scattering covariance. *arXiv preprint arXiv:2204.10177*, 2022.