Article

Geometric Shrinkage Priors for Kählerian Signal Filters [†]

Jaehyung Choi * and Andrew P. Mullhaupt

Department of Applied Mathematics and Statistics, State University of New York (SUNY), Stony Brook, NY 11794, USA

[†] This paper is an extended version of our paper published in MaxEnt 2014, Amboise, France, 21–26 September 2014.

* Author to whom correspondence should be addressed; E-Mail: jaehyung.choi@stonybrook.edu; Tel.: +1-631-632-8370.

Abstract: We construct geometric shrinkage priors for Kählerian signal filters. Based on the characteristics of Kähler manifolds, an efficient and robust algorithm for finding superharmonic priors which outperform the Jeffreys prior is introduced. Several ansätze for the Bayesian predictive priors are also suggested. In particular, the ansätze related to Kähler potential are geometrically intrinsic priors to the information manifold of which the geometry is derived from the potential. The implication of the algorithm to time series models is also provided.

Keywords: Kähler manifold; information geometry; Bayesian prediction; superharmonic prior

1. Introduction

In information geometry, signal processing is one of the most important applications. In particular, an information geometric approach to various linear time series models has been also well-known [1–7]. The geometric description of the linear systems is not confined to the pursuit of mathematical beauty. Komaki's work [8] is in the line of developing practical tools for Bayesian inference. Using the Kullback–Leibler divergence as a risk function for estimation, he found that superharmonic shrinkage priors outperform the Jeffreys prior in the viewpoint of information theory. Better prediction in the Bayesian framework is attainable by the Komaki priors.

However, a difficult part of Komaki's idea in practice is verifying whether or not a prior function is superharmonic. In particular, when high-dimensional statistical manifolds are considered, it is technically tricky to test the superharmonicity of prior functions because Laplace–Beltrami operators on the manifolds are non-trivial. Although some superharmonic priors for the autoregressive (AR) models were found not only in the two-dimensional cases [5,7] but also in arbitrary dimensions [6], there is no clue about the Bayesian shrinkage priors of more complicated models such as the autoregressive moving average (ARMA) models, the fractionally integrated ARMA (ARFIMA) models, and any arbitrary signal filters. Additionally, generic algorithms for systematically obtaining the information shrinkage priors are not known yet.

The connection between Kähler manifolds and information geometry has been reported [4,9–12] and the mathematical correspondence between a Kähler manifold and the information geometry of a linear system is recently revealed. It is found that the information geometry of a signal filter with a finite complex cepstrum norm is a Kähler manifold [7]. In particular, the Hermitian condition on the Kählerian information manifolds is clearly seen under conditions on the transfer function of the linear system. Moreover, many practical aspects of introducing Kähler manifolds to information geometry for signal processing were also reported in the same literature [7]. One of the benefits in the Kählerian information geometry is that the simpler form of the Laplace–Beltrami operator on the Kähler manifold is beneficial to finding the Komaki priors.

In this paper, we construct Komaki-style shrinkage priors for Kählerian signal filters. By introducing an algorithm which is based on the characteristics of Kähler manifolds, the Bayesian predictive priors outperforming the Jeffreys prior can be obtained in a more efficient and more robust way. Several prior ansätze are also suggested. Among the ansätze, the geometric shrinkage priors related to Kähler potential are intrinsic priors on the information manifold because the geometry is given by the Kähler potential. We also provide the geometric priors for the ARFIMA models where the Komaki priors have not been reported. The structure of this paper is as follows. In next section, theoretical backgrounds of Kählerian information geometry and superharmonic priors are introduced. In Section 3, an algorithm and ansätze for the geometric shrinkage priors are suggested. The implication of the algorithm to the ARFIMA models is given in Section 4. We conclude the paper in the last section.

2. Theoretical Backgrounds

2.1. Kählerian Filters

A linear filter with n-dimensional complex parameters $\boldsymbol{\xi}$ is characterized by a transfer function $h(w; \boldsymbol{\xi})$ in the frequency domain w with

$$y(w) = h(w; \boldsymbol{\xi})x(w)$$

where y and x are complex output and input signals, respectively. A spectral density function $S(w; \boldsymbol{\xi})$ is defined as the absolute square of the transfer function

$$S(w; \boldsymbol{\xi}) = |h(w; \boldsymbol{\xi})|^2$$

and it is a real-valued measurable quantity.

In information geometry, it is well-known by Amari and Nagaoka [1] that the geometry of a linear system is determined by the spectral density function $S(w; \boldsymbol{\xi})$ under the stability condition, minimum phase, and

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |\log S(w; \boldsymbol{\xi})|^2 dw < \infty.$$

The last condition is also known as the finite unweighted norm of the power cepstrum of a filter [13,14]. For a linear system with the spectral density function satisfying the above conditions, the metric tensor of the information geometry is given by

$$g_{\mu\nu}(\boldsymbol{\xi}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\partial_\mu \log S)(\partial_\nu \log S) dw$$

where the partial derivatives are taken with respect to the model parameters $\boldsymbol{\xi}$.

The metric tensor can be expressed in a complexified coordinate system and the Z-transformed transfer function. With the Z-transformation, the holomorphic transfer function can be written in the form of series expansion of z

$$h(z; \boldsymbol{\xi}) = \sum_{r=0}^{\infty} h_r(\boldsymbol{\xi}) z^{-r} \tag{1}$$

where h_r is an impulse response function. The Z-transformed power spectrum is also defined in the similar way. In this case, the conditions on the transfer function for constructing information geometry are identical to the spectral density function representation except for

$$\frac{1}{2\pi i} \oint_{|z|=1} |\log h(z; \boldsymbol{\xi})|^2 \frac{dz}{z} < \infty$$

and it is a necessary condition for the finite power cepstrum norm. The condition indicates that the Hardy norm of the logarithmic transfer function, also known as the unweighted complex cepstrum norm [14,15], is finite. The metric tensor of the geometry is given by the transfer function,

$$g_{ij}(\boldsymbol{\xi}) = \frac{1}{2\pi i} \oint_{|z|=1} \partial_i \log h(z; \boldsymbol{\xi}) \partial_j \log h(z; \boldsymbol{\xi}) \frac{dz}{z} \tag{2}$$

$$g_{i\bar{j}}(\boldsymbol{\xi}) = \frac{1}{2\pi i} \oint_{|z|=1} \partial_i \log h(z; \boldsymbol{\xi}) \partial_{\bar{j}} \log \bar{h}(\bar{z}; \bar{\boldsymbol{\xi}}) \frac{dz}{z} \tag{3}$$

where i, j run from 1 to n and $g_{\bar{i}\bar{j}}, g_{\bar{i}j}$ are the complex conjugates of g_{ij} and $g_{i\bar{j}}$, respectively.

After plugging the Z-transformed transfer function, Equation (1), into the metric tensor expressions, Equations (2) and (3), the metric tensor is expressed with the series expansion coefficients in z of the logarithmic transfer function by

$$g_{ij} = \partial_i \eta_0 \partial_j \eta_0$$

$$g_{i\bar{j}} = \partial_i \eta_0 \partial_{\bar{j}} \bar{\eta}_0 + \sum_{r=1}^{\infty} \partial_i \eta_r \partial_{\bar{j}} \bar{\eta}_r$$

where η_r is the coefficient of z^{-r} in the series expansion of the logarithmic transfer function, also known as a complex cepstrum coefficient [15]. It is obvious that $\eta_0 = \log h_0$.

Recently, it is found by Choi and Mullhaupt [7] that the information geometry of a linear system with a finite Hardy norm of a logarithmic transfer function (or the complex cepstrum norm) is the Kähler

manifold that is the Hermitian manifold with the closed Kähler two-form: $g_{ij} = g_{\bar{i}\bar{j}} = 0$ for the Hermitian manifold and $\partial_i g_{j\bar{k}} = \partial_j g_{i\bar{k}}$, $\partial_{\bar{i}} g_{k\bar{j}} = \partial_{\bar{j}} g_{k\bar{i}}$ for the closed Kähler two-form. Additionally, the Hermitian structure can be explicitly seen in the metric tensor if and only if the impulse response function with the highest degree in z, *i.e.*, h_0 in the unilateral transfer function case, is a constant in model parameters $\boldsymbol{\xi}$. In this paper, for simplicity, we only consider unilateral transfer functions with non-zero h_0 and the Kähler manifolds with the explicit Hermitian conditions on the metric tensors because complex manifolds are always Hermitian manifolds [16]. In this case, the necessary and sufficient condition for being a Kähler manifold is that $h_0(\boldsymbol{\xi})$ is a constant in $\boldsymbol{\xi}$ [7].

According to Choi and Mullhaupt [7], the benefits of the Kählerian description are the followings. First of all, geometric objects are straightforwardly computed on a Kähler manifold. The non-trivial metric tensor component is simply derived from the following formula

$$g_{i\bar{j}} = \partial_i \partial_{\bar{j}} \mathcal{K} \tag{4}$$

where $\mathcal{K}$ is the Kähler potential of the geometry. The Kähler potential in the information geometry of a linear filter is the square of the Hardy norm (or H^2-norm) of the logarithmic transfer function (or the square of the complex cepstrum norm) on the unit disk $\mathbb{D}$

$$\mathcal{K} = \frac{1}{2\pi i} \oint_{|z|=1} |\log h(z; \boldsymbol{\xi})|^2 \frac{dz}{z} = ||\log h(z; \boldsymbol{\xi})||_{H^2}^2 \tag{5}$$

and the details of the derivation are given in the literature [7]. The non-trivial components of the Levi–Civita connection are expressed as

$$\Gamma_{ij,\bar{k}} = \partial_i g_{j\bar{k}} = \partial_i \partial_j \partial_{\bar{k}} \mathcal{K} \tag{6}$$

and the other connection components are all vanishing. Notice that it is much simpler than the connection components on a non-Kähler manifold given by

$$\Gamma_{ij,k} = \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ik} - \partial_k g_{ij})$$

and it is obvious that the number of calculation steps is significantly reduced in the Kähler case. The Riemann curvature tensor of the linear system geometry is also represented in the simpler form which is given in Choi and Mullhaupt [7]. The Ricci tensor on the Kähler manifold is obtained as

$$R_{i\bar{j}} = -\partial_i \partial_{\bar{j}} \log \mathcal{G} \tag{7}$$

where $\mathcal{G}$ is the determinant of the metric tensor. It is evident that we can skip the calculation of the Riemann curvature tensor in order to compute the Ricci tensor on a Kähler manifold.

Additionally, the α-generalization of the geometric objects is linear in α on Kähler manifolds. Since the Riemann curvature tensor on a Kähler manifold is linear in the α-connection which is α-linear, the Riemann tensor also exhibits the α-linearity which leads to the α-linear Ricci tensor and scalar curvature.

In addition to these advantages, any submanifolds of a Kähler manifold are also Kähler manifolds. If the information geometry of a given statistical model is a Kähler manifold, its submodels also have Kähler manifolds as the information geometry and all the properties of the ambient manifold are also equipped with the submanifolds.

Lastly, the Kählerian information geometry is also useful to find superharmonic priors because of the simpler Laplace–Beltrami operators on the manifolds. We will cover the details of the superharmonic priors soon.

2.2. Superharmonic Priors

For further discussions, we need to introduce the superharmonic priors suggested by Komaki [8]. When we want to find the true probability distribution $p(y|\boldsymbol{\xi})$ based on given samples x of size N, one of the best approaches is using Bayesian predictive density $p_\pi(y|x^{(N)})$ with a prior $\pi(\boldsymbol{\xi})$:

$$p_\pi(y|x^{(N)}) = \frac{\int p(y|\boldsymbol{\xi})p(x^{(N)}|\boldsymbol{\xi})\pi(\boldsymbol{\xi})d\boldsymbol{\xi}}{\int p(x^{(N)}|\boldsymbol{\xi})\pi(\boldsymbol{\xi})d\boldsymbol{\xi}}.$$

The superharmonic priors π_I are derived from the difference between two risk functions with respect to the true probability density, one from the Jeffreys prior and another from the superharmonic prior:

$$\mathbb{E}[D_{KL}(p(y|\boldsymbol{\xi})||p_{\pi_J}(y|x^{(N)}))|\boldsymbol{\xi}] - \mathbb{E}[D_{KL}(p(y|\boldsymbol{\xi})||p_{\pi_I}(y|x^{(N)}))|\boldsymbol{\xi}]$$
$$= \frac{1}{2N^2}g^{ij}\partial_i \log\left(\frac{\pi_I}{\pi_J}\right)\partial_j \log\left(\frac{\pi_I}{\pi_J}\right) - \frac{1}{N^2}\frac{\pi_J}{\pi_I}\Delta\left(\frac{\pi_I}{\pi_J}\right) + o(N^{-2})$$

where D_{KL} is the Kullback–Leibler divergence and π_J is the Jeffreys prior which is the volume form of the statistical manifold. Each risk function indicates how far a given Bayesian predictive density is from the true distribution in the Kullback–Leibler divergence in average. Sine better priors are obtained from smaller risk functions, the priors outperforming the Jeffreys prior make the above expression greater than zero. Since the first term on the right-hand side is non-negative, the risk function of the Komaki prior is decreased with respect to the risk function of the Jeffreys prior if a prior function $\psi = \pi_I/\pi_J$ is superharmonic. If a superharmonic prior function ψ can be found, it is possible to do better Bayesian prediction in the viewpoint of information theory. In the same paper, Komaki also pointed out that shrinkage priors are information-theoretically more improved in prediction than the Jeffreys prior if and only if the square root of a prior function is superharmonic.

Since Komaki's paper [8], several superharmonic priors for the AR models have been found [5–7]. The Komaki prior for the AR(2) model in the pole coordinates [5] is given by

$$\psi = 1 - \xi^1\xi^2$$

where ξ^i is a pole of the transfer function. Tanaka [6] generalized the two-dimensional case to superharmonic priors for the AR model in an arbitrary dimension p. The shrinkage prior function for the AR(p) model is in the form of

$$\psi = \prod_{i<j}^{p}(1 - \xi^i\xi^j)$$

where ξ^i is a pole of the AR transfer function.

As mentioned before, one of the advantages in the Kählerian description is that finding the Komaki prior functions becomes more efficient than those in non-Kähler description because the

Laplace–Beltrami operators on Kähler manifolds are in the simpler forms. For a differentiable function ψ, the Laplace–Beltrami operator in the Kähler geometry is represented with

$$\Delta\psi = 2g^{i\bar{j}}\partial_i\partial_{\bar{j}}\psi.$$

Meanwhile, the Laplace–Beltrami operator on a non-Kähler manifold is expressed as

$$\Delta\psi = \frac{1}{\sqrt{\mathcal{G}}}\partial_i\left(\sqrt{\mathcal{G}}g^{ij}\partial_j\psi\right)$$
$$= g^{ij}\partial_i\partial_j\psi + \frac{1}{2}g^{ij}\partial_i\log\mathcal{G}\partial_j\psi + \partial_ig^{ij}\partial_j\psi$$

where $\mathcal{G}$ is the determinant of the metric tensor. It is obvious that additional calculations for the latter two terms in the right-hand side are indispensable in the non-Kähler cases.

With the computational benefits on the Kählerian information manifolds, the superharmonic prior function for the Kähler-AR(2) model [7] is found

$$\psi = (1 - |\xi^1|^2)(1 - \xi^1\bar{\xi}^2)(1 - \xi^2\bar{\xi}^1)(1 - |\xi^2|^2)$$

where ξ^i is the i-th pole of the transfer function and $\bar{\xi}^i$ is the complex conjugate of ξ^i. However, its generalization to any arbitrary dimensions has been unknown. Moreover, the Komaki priors for the ARMA models and the ARFIMA models are not reported yet.

3. Geometric Shrinkage Priors

As shown in the previous section, Kähler manifolds in information geometry are useful in order to obtain the superharmonic priors. In this section, we introduce an algorithm to find the geometric shrinkage priors by using the properties of Kähler geometry. Moreover, several ansätze for the priors are suggested.

For further discussions, let us set $\tau = u^* - \kappa(\boldsymbol{\xi}, \bar{\boldsymbol{\xi}})$ where u^* is a constant in $\boldsymbol{\xi} = (\xi^1, \xi^2, \cdots, \xi^n)$ and its complex conjugate $\bar{\boldsymbol{\xi}}$. The following lemma is worthwhile when the algorithm for the prior functions is constructed.

Lemma 1. *On a Kähler manifold, a function $\psi(\boldsymbol{\xi}, \bar{\boldsymbol{\xi}})$ is superharmonic if $\psi(\boldsymbol{\xi}, \bar{\boldsymbol{\xi}})$ is in the form of $\psi(\boldsymbol{\xi}, \bar{\boldsymbol{\xi}}) = \Psi(u^* - \kappa(\boldsymbol{\xi}, \bar{\boldsymbol{\xi}}))$ such that κ is subharmonic (or harmonic) and $\Psi'(\tau) > 0, \Psi''(\tau) \leq 0$ (or $\Psi'(\tau) > 0, \Psi''(\tau) < 0$).*

Proof. The Laplace–Beltrami operator on ψ is given by

$$\Delta\psi = 2g^{i\bar{j}}\partial_i\partial_{\bar{j}}\psi = 2g^{i\bar{j}}\partial_i\left(\left(-\partial_{\bar{j}}\kappa\right)\Psi'\right)$$
$$= 2\Psi''g^{i\bar{j}}\partial_i\kappa\partial_{\bar{j}}\kappa - 2\Psi'g^{i\bar{j}}\partial_i\partial_{\bar{j}}\kappa$$
$$= 2\Psi''||\partial\kappa||_g^2 - \Psi'\Delta\kappa$$

where the derivatives on Ψ are taken with respect to τ. It is obvious that if κ is subharmonic (or harmonic) and if $\Psi'(\tau) > 0, \Psi''(\tau) \leq 0$ (or $\Psi'(\tau) > 0, \Psi''(\tau) < 0$), then the right-hand side is negative, *i.e.*, ψ is a superharmonic function. $\square$

According to Lemma 1, superharmonic functions are easily obtained from subharmonic or harmonic functions by simply plugging the (sub-)harmonic functions as κ into Lemma 1.

By considering that a prior function should be positive, it is able to utilize Lemma 1 for obtaining the superharmonic prior functions. Let us confine the function ψ in Lemma 1 to be positive.

Theorem 1. *On a Kähler manifold, a positive function $\psi = \Psi(u^* - \kappa)$ is a superharmonic prior function if κ is subharmonic (or harmonic) and $\Psi'(\tau) > 0$, $\Psi''(\tau) \leq 0$ (or $\Psi'(\tau) > 0$, $\Psi''(\tau) < 0$).*

Proof. Since this is a special case of Lemma 1, the proof is obvious. $\square$

Although any (sub-)harmonic function κ can be used for constructing superharmonic priors, restriction on κ makes finding the ansätze of the geometric priors easier. From now on, upper-bounded functions are only our concerns. Additionally, we assume that κ and u^* are real. With these assumptions, it is possible to set u^* as a constant greater than the upper bound of κ in order for τ to be positive.

Ansätze for Ψ can be found in the following example.

Example 1. *Given subharmonic (or harmonic) κ and positive τ, i.e., upper-bounded κ, the following functions are candidates for Ψ*

$$\Psi_1(\tau) = \tau^a$$
$$\Psi_2(\tau) = \log(1 + \tau^a)$$

where $0 < a \leq 1$ (or $0 < a < 1$).

Proof. We only cover a subharmonic case for κ here and it is also straightforward for the harmonic case. First of all, Ψ_1 and Ψ_2 are all positive. For Ψ_1, it is easy to verify the followings:

$$\Psi_1'(\tau) = a\tau^{a-1} > 0$$
$$\Psi_1''(\tau) = a(a-1)\tau^{a-2} \leq 0$$

for $0 < a \leq 1$. The similar calculation is repeated for Ψ_2:

$$\Psi_2'(\tau) = \frac{a\tau^{a-1}}{(1 + \tau^a)} > 0$$
$$\Psi_2''(\tau) = \frac{a\tau^{a-2}(a - (1 + \tau^a))}{(1 + \tau^a)^2} \leq 0$$

for $0 < a \leq 1$.

Both functions Ψ_1 and Ψ_2 satisfy the conditions for Ψ in Lemma 1. $\square$

It is also possible to find ansätze for upper-bounded subharmonic κ. The following functions are candidates for upper-bounded and subharmonic κ.

Example 2. *For positive real numbers a_r and b_i, the following subharmonic functions are candidates for κ in the cases that those are upper-bounded:*

$$\kappa_1 = \mathcal{K}$$
$$\kappa_2 = \sum_{r=0}^{\infty} a_r |h_r(\boldsymbol{\xi})|^2$$
$$\kappa_3 = \sum_{i=1}^{n} b_i |\xi^i|^2.$$

Proof. Let us assume that the ansätze are upper-bounded in given domains. For κ_1, it is easy to show that the Kähler potential $\mathcal{K}$ is subharmonic:

$$\Delta \kappa_1 = \Delta \mathcal{K} = 2g^{i\bar{j}} \partial_i \partial_{\bar{j}} \mathcal{K}$$
$$= 2g^{i\bar{j}} g_{i\bar{j}} = 2n > 0.$$

The proof for subharmonicity of κ_2 is as follows:

$$\Delta \kappa_2 = \Delta \left(\sum_{r=0}^{\infty} a_r |h_r(\boldsymbol{\xi})|^2 \right) = 2g^{i\bar{j}} \partial_i \partial_{\bar{j}} \left(\sum_{r=0}^{\infty} a_r |h_r(\boldsymbol{\xi})|^2 \right)$$
$$= \sum_{r=0}^{\infty} 2a_r g^{i\bar{j}} \partial_i h_r \partial_{\bar{j}} \bar{h}_r = \sum_{r=0}^{\infty} 2a_r ||\partial h_r||_g^2 > 0.$$

The subharmonicity of κ_3 is tested by

$$\Delta \kappa_3 = \Delta \left(\sum_{i=1}^{n} b_i |\xi^i|^2 \right) = 2g^{i\bar{j}} \partial_i \partial_{\bar{j}} \left(\sum_{i=1}^{n} b_i |\xi^i|^2 \right) = \sum_{i=1}^{n} 2b_i g^{i\bar{i}} > 0.$$

If the upper-boundedness is satisfied, the above subharmonic functions are ansätze for κ. $\square$

Superharmonic prior functions on the Kähler manifolds are efficiently constructed from the following algorithm which exploits Theorem 1 and the ansätze for Ψ and κ. When we find positive and superharmonic functions, it is automatically the Komaki-style prior functions as usual. If positive, upper-bounded, and (sub-)harmonic functions are found, those functions are plugged into Theorem 1 in order to obtain superharmonic prior functions. Multiplying the Jeffreys prior by the superharmonic prior functions, we finally acquire the geometric shrinkage priors. Additionally, since the ansätze are already given, there is no extra cost to find the Komaki prior functions except for verifying whether or not the information geometry is a Kähler manifold. Comparing with the literature on the Komaki priors of the time series models [5–7], obtaining the geometric priors on the Kähler manifolds becomes more efficient and more robust.

4. Example: ARFIMA Models

The ARFIMA model is the generalization of the ARMA model with a fractional differencing parameter in order to model the long memory process. The transfer function of the ARFIMA(p, d, q) model with parameters $\boldsymbol{\xi} = (\xi^{-1}, \xi^0, \xi^1, \cdots, \xi^{p+q}) = (\sigma, d, \lambda_1, \cdots, \lambda_p, \mu_1, \cdots, \mu_q)$ is given by

$$h(z; \boldsymbol{\xi}) = \frac{\sigma^2}{2\pi} \frac{(1 - \mu_1 z^{-1})(1 - \mu_2 z^{-1}) \cdots (1 - \mu_q z^{-1})}{(1 - \lambda_1 z^{-1})(1 - \lambda_2 z^{-1}) \cdots (1 - \lambda_p z^{-1})} (1 - z^{-1})^d$$

where d is the differencing parameter and μ_i, λ_i, σ are a pole, a root, and a gain in the ARMA model, respectively. It is noteworthy that the transfer function of the ARFIMA model is decomposed into the ARMA model part and the fractionally integration part. Additionally, every poles and roots of the linear system are located inside the unit disk, *i.e.*, $|\lambda_i| < 1$ for $i = 1, \cdots, p$ and $|\mu_i| < 1$ for $i = 1, \cdots, q$.

Similar to the ARMA case [7], the full geometry of the ARFIMA model is a Kähler manifold and the submanifold of a constant gain σ is also Kähler geometry. This submanifold also exhibits the explicit

Hermitian condition on the metric tensor. It is easy to cross-check the Hermitian structure by fixing $h_0 = 1$ up to the gain of the signal filter. We will work on this submanifold.

Since the information geometry of the ARFIMA model is a Kähler manifold, the Kähler potential of the ARFIMA geometry is obtained from the square of the Hardy norm of the logarithmic transfer function (or the square of the complex cepstrum norm), Equation (5), represented with

$$\mathcal{K} = \sum_{r=1}^{\infty} \left| \frac{d + (\mu_1^r + \cdots + \mu_q^r) - (\lambda_1^r + \cdots + \lambda_p^r)}{r} \right|^2. \tag{8}$$

It is obvious that the Kähler potential for the ARFIMA model, Equation (8), is reducible to the Kähler potential of the ARMA geometry by setting $d = 0$. It is easy to verify that the Kähler potential of the ARFIMA geometry is upper-bounded by $(d + p + q)^2 \frac{\pi^2}{6}$.

By using Equation (4), the metric tensor of the Kähler geometry is simply derived from the Kähler potential. The metric tensor of the Kähler-ARFIMA geometry is given by

$$g_{i\bar{j}} = \begin{pmatrix} \frac{\pi^2}{6} & \frac{1}{\bar{\lambda}_j} \log\left(1 - \bar{\lambda}_j\right) & -\frac{1}{\bar{\mu}_j} \log\left(1 - \bar{\mu}_j\right) \\ \frac{1}{\lambda_i} \log\left(1 - \lambda_i\right) & \frac{1}{1 - \lambda_i \bar{\lambda}_j} & -\frac{1}{1 - \lambda_i \bar{\mu}_j} \\ -\frac{1}{\mu_i} \log\left(1 - \mu_i\right) & -\frac{1}{1 - \mu_i \bar{\lambda}_j} & \frac{1}{1 - \mu_i \bar{\mu}_j} \end{pmatrix}$$

and it is easy to show that the metric tensor contains the pure ARMA metric. The metric tensor is also in the similar form to the ARFIMA geometry in non-complexified coordinates [3]. The metric tensor indicates that the ARMA geometry is embedded in the ARFIMA geometry and corresponds to the submanifold of the ARFIMA manifold. The ARMA part of the metric tensor is the same metric with the Kähler-ARMA geometry in Choi and Mullhaupt [7]. In addition to that, we can cross-check the fact that the ARMA geometry is also a Kähler manifold based on a property of a Kähler manifold that a submanifold of the Kähler geometry is Kähler.

Other geometric objects can be derived from the metric tensor. For example, the non-trivial components of the 0-connection are given by Equation (6). It is noteworthy that any connection components with the d-coordinate in the first two indices of the connection are trivially zero and the others might not be vanishing. Similar to the 0-connection, the Ricci tensor components along the fractionally integrated direction are also zero because there is no dependence on d in the metric tensor. Considering the Schur complement, the non-vanishing Ricci tensor components are decomposed into the Ricci tensor from the pure ARMA part and the term from the mixing between the ARMA part and the fractionally integrated (FI) part:

$$R_{i\bar{j}} = R_{i\bar{j}}^{ARMA} + R_{i\bar{j}}^{ARMA-FI}$$

where i and j are not along the d-coordinate.

It is the time to be back to the geometric shrinkage priors. Since the Kähler potential of a given ARFIMA model is upper-bounded by a constant $u^* = (d + p + q)^2 \frac{\pi^2}{6}$, the intrinsic priors on the Kähler manifold can be found as it is proven in the previous section. By using the algorithm and the ansätze related to the Kähler potential, some geometric shrinkage prior functions for the ARFIMA model are constructed as

$$\psi_1 = (u^* - \mathcal{K})^a$$
$$\psi_2 = \log\left(1 + (u^* - \mathcal{K})^a\right)$$

where $0 < a \leq 1$. It is also noteworthy that when $d = 0$ in the Kähler potential, superharmonic priors of the ARMA (or AR/MA) models are obtained and finding the priors becomes much simpler than the literature on the Komaki priors of the AR models [5–7]. Similarly, κ_2 and κ_3 are also utilized for the superharmonic prior function ansätze in the ARFIMA models because the both functions are upper-bounded on the ARFIMA manifold. Moreover, if we set $d = 0$ for κ_2 or $b_0 = 0$ for κ_3, the ansätze for the ARFIMA models are reducible to the Komaki priors of the ARMA models.

5. Conclusion

In this paper, we build up an algorithm and ansätze for the geometric shrinkage priors of Kählerian signal filters. By using the properties of Kähler manifolds, an algorithm to find the Komaki priors is constructed and ansätze for the prior functions are suggested. Additionally, some ansätze associated with the Kähler potential are geometrically intrinsic to Kählerian information manifolds because the geometry is derived from the Kähler potential which is the square of the complex cepstrum norm of a linear system.

Comparing with the literature on the Komaki priors of the time series models, verification of the geometric priors is much easier on the Kähler manifold and it is also possible to acquire the geometric shrinkage priors for highly complicated models in the more efficient and robust way. For example, Bayesian predictive priors for the ARFIMA model are obtained from the algorithm and ansätze for the prior functions. The shrinkage priors of the ARMA cases are simply found from the geometric shrinkage priors of the ARFIMA models by using the property of submanifolds in the Kähler geometry.

Author Contributions

Both authors contributed equally to the main idea. The research was conducted out by both authors. Jaehyung Choi wrote the paper. Both authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Amari, S.; Nagaoka, H. *Methods of information geometry*; Oxford University Press: Oxford, UK, 2000.
2. Ravishanker, N.; Melnick, E. L.; Tsai, C. Differential geometry of ARMA models. *J. Time Ser. Anal.* **1990**, *11*, 259–274.
3. Ravishanker, N. Differential geometry of ARFIMA processes. *Commun. Stat. Theory Methods* **2001**, *30*, 1889–1902.
4. Barbaresco, F. Information intrinsic geometric flows. *AIP Conf. Proc.* **2006**, *872*, 211–218.

5. Tanaka, F.; Komaki, F. A superharmonic prior for the autoregressive process of the second order. *J. Time Ser. Anal.* **2008**, *29*, 444–452.

6. Tanaka, F. *Superharmonic priors for autoregressive models*; Mathematical Engineering Technical Reports; University of Tokyo: Tokyo, Japan, 2009.

7. Choi, J.; Mullhaupt, A. P. Kählerian information geometry for signal processing. arXiv:1404.2006.

8. Komaki, F. Shrinkage priors for Bayesian prediction. *Ann. Stat.* **2006**, *34*, 808–819.

9. Barndorff-Nielsen, O. E.; Jupp, P. E. Statistics, yokes and symplectic geometry. *Annales de la faculté des sciences de Toulouse 6 série* **1997**, *6*, 389–427.

10. Barbaresco, F. Information geometry of covariance matrix: Cartan-Siegel homogeneous bounded domains, Mostow/Berger fibration and Fréchet median. In *Matrix Information Geometry*; Bhatia, R., Nielsen, F., Eds.; Springer: Berlin and Heidelberg, Germany, 2012; pp. 199–256.

11. Zhang, J.; Li, F. Symplectic and Kähler structures on statistical manifolds induced from divergence functions. *Geom. Sci. Inf.* **2013**, *8085*, 595–603.

12. Barbaresco, F. Koszul information geometry and Souriau geometric temperature/capacity of Lie group thermodynamics. *Entropy* **2014**, *16*, 4521–4565.

13. Bogert, B.; Healy, M.; Tukey, J. The quefrency alanysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum and saphe cracking. In Proceedings of the Symposium on Time Series Analysis, Brown University, Providence, RI, USA, 11–14 June 1963; pp. 209–243.

14. Martin, R.J. A metric for ARMA processes. *IEEE Trans. Signal Process.* **2000**, *48*, 1164–1170.

15. Oppenheim, A. V. Superposition in a class of nonlinear systems. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 1965.

16. Nakahara, M. *Geometry, Topology and Physics*; Institute of Physics Publishing: Bristol, UK and Philadelphia, PA, USA, 2003.