Hermite-Gaussian functions for image synthesis

Pablo E. Román*, Takeshi Asahi*, and Simon Casassus‡
*Center for Mathematical Modeling, Universidad de Chile, Av. Blanco Encalada 2120 Piso 7, Santiago, CHILE.
E-mail: proman@ing.uchile.cl Tel: +56-2-29784859
tasahi@dim.uchile.cl Tel: +56-2-2978 4602
‡Astronomy Department, Universidad de Chile, Camino El Observatorio 1515, Las Condes, Santiago, CHILE.
E-mail: scasassus@u.uchile.cl Tel: +56-2-29771137

***Abstract*—The advent of high fidelity observatory interferometers such ALMA, invites new opportunities in algorithmic research for image processing. Interferometer telescopes perform irregular sampling on the Fourier transform of the sky images called "visibilities." The inverse problem resolution is called "Image Synthesis". This inverse problem is "Ill-posed" and thus regularization technique should be applied. State-of-the-art image synthesis algorithm (e.g. CLEAN) do provide an inversion of data in Fourier space to produce sky images. Current algorithm are based on spatial regular grid for image processing requiring interpolation for estimating values on grid from values on arbitrary sampling points. Numeric implementation also considers a set of basis function for representing 2-d field. We propose to use a non-local representation for the inverse problem. Our method is based on Hermite-Gaussian (HG) functions. HG functions are a complete set of eigenvectors for the Fourier operator, which we use for representing the problem. The HG set also approximately diagonalize the image's expansion coefficient correlation matrix. We present a proof of concept expressed on images.**

I. Introduction

Modern radio interferometers such as ALMA are collecting large data set measurements for high-fidelity imaging. This novel instrumental capabilities challenges current interferometric imaging algorithms. Science requires an accurate image reconstruction and assessment.

A. Problem definition

The Image Synthesis [1] or Fourier Synthesis [2] problem corresponds to estimate the best image approximation from a sparse sampling of its Fourier transform. This is the statement of an inverse problem [3]. Equation (1) state the simple functional relationship from the image $I(x,y)$ and its associated interferometric signal [4] called "Visibility function" $V(u,v)$.

$$V(u,v) = \int_{\mathbb{R}^2} A(x,y)I(x,y)e^{-2\pi i(ux+vy)}dxdy \qquad (1)$$

The kernel $A(x,y)$ is called "the primary beam (PB)" and corresponds to the solid angle reception pattern of the antenna array also modelled as a Gaussian [5]. For simplicity in this paper, we consider the requested image $I(x,y) \equiv A(x,y)I(x,y)$ and restricted to decay as the PB does. In this case the visibility function is simply the Fourier transform of the image $V = \mathcal{F}\{I\}$ as defined in (Eq. 1). Current interferometer performs large sparse sampling $\{V_k^o\}_{k=1}^N$ in the visibility space at points $\{Z_k = (u_k, v_k)\}_{k=1}^N$ in order to fill frequency spacing and pursuing to enter into the high fidelity regime. The measurement's error deviations are non-negligible [6] and comparable in magnitude to expected values. Visibilities (Eq. 2) should be modelled as random field with expectation $E(V(Z)) = \overline{V}(Z)$ and variance $\sigma^2(Z)$. We consider $\epsilon(Z)$ as a 2-D delta-correlated random field (Eq. 3) restricted to have null expected value and independent for $\forall Z = (u,v) \in \mathbb{R}^2$. A delta-correlated random field is characterized by a correlation represented by a Dirac's delta function [7], and a common example of such field is the white Gaussian noise. It is common in the literature (e.g. [8]) to consider measurement as independent Gaussian variables. Nevertheless, Gaussian assumption for visibilities contradict the fact that the random field $I(X)$ must have a positive image intensity support $\forall X = (x,y) \in \mathbb{R}^2$. Furthermore, a Gaussian distribution for image always implies negatives values cases. Our simplifying assumption is that the visibility field model has delta-correlated noise [7] and independent value.

$$V(Z) = \overline{V}(Z) + \sigma(Z)\epsilon(Z),\ Z = (u,v)\ \sigma(Z) \gneqq 0 \qquad (2)$$

$$\begin{aligned} E(\epsilon(Z)) &= 0, \\ E(\epsilon(Z)\epsilon(Z')) &= \delta(Z - Z') \end{aligned} \qquad (3)$$

The problem is to have a good estimation of the expectation $E(I(Z))$ and variance of the random field according to the observed visibilities $\{V_k^O\}_{k=1}^N$ and deviation $\{\sigma_k^O\}_{k=1}^N$. Current image synthesis methods are only related to $E(I)$ and variance requirement is interpreted from residual files. The observation set is irregularly sampled so traditional inversion scheme cannot be applied. Furthermore, Fourier transform in (Eq. 1) induce a high correlation effect from independent visibilities. We highlight in this work the decorrelation effect in the coefficient expansion in the Hermite-Gaussian representation.

B. Current image synthesis methods

The image synthesis problem is ill-posed ([2], [9]) because possible solutions are not unique and do not depend continuously on input data. If the sampling is taken from a regular square grid then this inverse problem is well posed and the FFT algorithm reconstruct efficiently the image. However, if the sampling is irregular then high amplitude artifact appears when using the discrete Fourier sum. A regularization procedure should be used in order to exhaust the potential of high-fidelity imaging in large sparse sampling.

978-616-361-823-8

We distinguish two main approaches used for solving the image synthesis problem in radio astronomy. The first, is a class of greedy heuristics consisting in algorithms that accumulate function components for building the image. The algorithm iterates choosing the place in residual image with maximal intensity where a new image component is included. The second class of algorithms is based on formal optimization problem including regularization terms which are commonly solved using non-linear optimization methods.

The 1974's CLEAN algorithm [10] is an image synthesis heuristic that originate many variant (see [11], [12], [13], [14], [15], [16]). The CLEAN algorithm is currently the *de facto* standard used for image synthesis processing in radio astronomy. The CASA package [17] implement several version of the CLEAN method for processing ALMA data. The first version of this heuristics can be interpreted as a matching pursuit optimization technique [18] for minimizing the L_1 norm conforming a sparse problem formulation [19]. Therefore, the resulting model image is commonly spanned by a small number (hundreds) of 2-D Gaussian. Furthermore, CLEAN can be interpreted as an earlier example of compressed sensing [20] method as long as the final image representation is sparse.

Nevertheless, several criticisms have been issued to the CLEAN method. The first related to the non-existence of a formal statistical assessment for the resulting image, such as error bar on the image intensity and the spacial resolution of the image in order to assess identified feature on image. In practice, the way for circumventing this issue is by considering the elliptical Gaussian fit to the full-width-half-max ellipse of the discrete Fourier sum's PSF. This is also called the clean beam [21] and then interpreting the FWHM ellipsis as the resolution of the resulting images. However, as long as the model image is a discrete sum of functional components it do not represent natural signal as the background noise. The concept of "restored image" corresponds to add the image of the residuals to the convolved CLEAN image with a elliptical Gaussian (clean beam). In this sense the restored image will have a sense of interpretation with natural background noise and resolution. This pedestrian interpretation of the image is the main criticism of the CLEAN method. Another criticism over the method is also an advantage. CLEAN can be iterated with human assistance by indicating regions where to include more intensively the CLEAN functional components. If the morphology of the object is theoretically known in advance then observing the size of the residuals also helps to improve the image synthesis process. Nevertheless, this could introduce an obvious bias on final images. CLEAN iteration is commonly non-convergent since residual images can reach a minimal level but with high levels of oscillations. Those algorithms include the number of iterations to perform as a parameter. CLEAN is best suited for point-wise sources, but novel CLEAN version [15] are enhanced for managing extended sources. In the practice, CLEAN is the most used method on astrophysical papers possibly by its faster execution and memory management.

Another class of image synthesis algorithms are those based on optimization formulation of the problem. In this case the inverse problem is formulated as an optimization problem and additional constrains are included in order to transform the problem to a well posed convex problem. The regularized problem is solved by an optimization algorithm suited on the size and form of the problem. The maximum likelihood method and compress sensing enter in this category. Inverse problem theory [22] formalize computational solution of this problem in base of the optimization problem framework.

The maximum entropy method (MEM) ([23], [24], [8]) resolve a Bayesian maximum log-likelihood criteria. This method has benefits such as to have a statistical basis that could theoretically help the image's assessment. The size of the non-linear optimization problem is comparable to the number of pixels (10^6) of a common image. This method has a lower performance than the CLEAN-based heuristics. The MEM assumes visibilities measurement distributes as Gaussian then the log-likelihood cost function is approximately equal to a weighted least square plus an entropy penalization term $-\int I(x,y)Log(I(x,y)/M)dxdy)$. Images in intensity setting are restricted to have positive values $I(x,y) \geqslant 0$. The entropy term induce a reduction in total intensity of the image, and information about variance in results is missing ([25], [26]).

Recent research on image synthesis are focused on applying compressed sensing techniques ([19], [20]) to the radio-interferometric problem ([27], [28], [16], [29]). The assumption behind this regularization is the existence of a basis for image representation in which few functional component are sufficient for a fair visualization. Sparse representations have the advantage to be simpler than others in the context quantity of information and they have direct application to data mining as feature quantization [30]. Sparsity is considered an enabling characteristics of many intelligent biological system [31], [32]. Sparse representation also enable efficient indexing of complex object as images [33]. There are several attempts to formulate requirement conditions for evidencing this phenomena, the most popular is the incoherence property [19]. Examples of basis used by this method correspond to wavelets [27]. This method is empirically shown to have improved image definition. However, the best basis for sparse representation in image synthesis is still not defined. Statistical assessment of the resulting images is also a matter of discussion.

The only method proved to be capable of producing statistical information about image synthesis result are Monte Carlo based approaches. Gibbs sampling has been proposed for resolving the inverse problem [34] giving information about uncertainty of the image.

C. Some common problems in image synthesis

In practice, all image synthesis method require a two dimensional array to represent image data in computer memory and this representation is called "gridding" [35]. The Fourier transform in equation (1) is implemented as a Fast Fourier Transform (FFT) algorithm using Cartesian grid. However, the gridding procedure requires a function approximation from

an arbitrary sampling point to the regular grid nodes. Image synthesis algorithm require to iterate over all grid points in order to find the best approximation function. Thus, the computational complexity of the image synthesis process depends at least polynomially on the size of the grid. Original CLEAN method extrapolates the visibility function from sampling point to the regular grid by mean of a convolution operation. The gridding procedure includes an approximation error that is added to the sampling error.

All methods can be described by a set of functions for image's representation. The CLEAN method uses a set of Gaussian function with different coordinate centroid and variances (e.g. multiscale CLEAN [15]). The MEM method uses pixels as function representation (e.g. [8]). Compressed sensing methods for image synthesis are using Daubechies' wavelets [27]. Furthermore, the functional representation in image synthesis has strong influence on the stochastic properties of the resulting image. The visibility random field have the independence property per coordinate, but the Fourier transform induces strong correlations in images. The image representation coefficients will have a variance depending on the choice of the basis set. The correlation of a random field is defined by the covariance kernel (Eq. 4). Nevertheless, the discrete version of the covariance kernel is difficult to analyse with strong correlations. If we consider mega-pixel images (10^6 pix) then its covariance matrix requires of the order of 10^{12} components to compute.

$$K(X_1, X_2) = E([(I(X_1) - E(I(X_1))][I(X_2) - E(I(X_2))]) \quad X_1, X_2 \in \mathbb{R}^2 \tag{4}$$

The Karhunen-Loeve theorem (e.g. [36]) can be applied to the image random field by expanding (Eq. 6) on eigenvectors $\{\phi_{\vec{k}}\}$ of the covariance operator. 2-D basis are commonly indexed by two indexes, which we denote $\vec{k} = (k_1, k_2)$. The covariance operator is represented by the homogeneous Fredholm integral equation in (Eq. 5). The theorem states that coefficient $\{c_{\vec{k}}\}$ in equation (6) are uncorrelated random variables ($E(c_{\vec{k}} c'_{\vec{k}}) = 0,\ \forall \vec{k} \neq \vec{k}'$) with expectation $E(c_{\vec{k}}) = 0$. We interpret $E(I)$ as the model image resulting from the image synthesis algorithm and $\hat{I} = I - E(I)$ as the residual image.

$$\int_{\mathbb{R}^2} K(X, Y)\phi_{\vec{k}}(X)dX = \lambda_{\vec{k}}\phi_{\vec{k}}(Y) \tag{5}$$

$$I(X) - E(I(X)) = \sum_{\vec{k}} c_{\vec{k}}\phi_{\vec{k}}(X) \tag{6}$$

This representation has important simplifying consequences, since it decorrelates the image field. In this work we developed further such concepts and demonstrated that the Hermite-Gaussian basis ([37], [38]) is a set of approximated eigenvector for the covariance operator in equation (5).

We propose to avoid the cost of having a grid in image synthesis pipeline by calculating the Fourier transform in other way than traditional FFT. Our approach consist in considering harmonic properties of the interferometric measurements (Eq. 1). The Fourier transform has a convenient diagonal representation in an orthogonal eigenvector basis.

We considered representing visibilities using the Hermite-Gaussian [37][38] set of 1-D orthonormal basis of functions (Eq. 7) that are eigenfunction of the Fourier transform operator (Eq. 8) and simultaneously of the covariance operator (Eq. 5). This base functions have been explored on the signal processing field as applications to the fractional Fourier transform [37], chirplets analysis [39] , and optimal filtering [40].

$$\phi_k(x) = [2^k \pi^{1/2} k!]^{-1/2} H_k(x) e^{-x^2/2} \tag{7}$$

$$\mathcal{F}\{\phi_k\} = (-i)^k \phi_k \tag{8}$$

The set $\{\phi_k\}$ is a complete orthogonal basis of function in $L^2(\mathbb{R})$ [41]. Thus the visibility function (Eq. 1) has a unique representation in the 2-D version of this basis (Eq. 9). If coefficient $v_{\vec{k}} = v_{k_1,k_2}$ are known then the image is easily obtained by calculating the series with a phase factor $(i)^{\vec{k}} v_{\vec{k}} \equiv (i)^{k_1+k_2} v_{\vec{k}} = c_{\vec{k}}$ (Eq. 10).

The Hermite-Gaussian basis has been applied in astronomical imaging for galaxy's shape characterization analysis [42]. The previous work claims empirically that few series term of the series ($n \leq 20$) are sufficient to represent astronomical objects like elliptical galaxy. A large scale galaxy images survey [43] exhibits a similar order of magnitude for an optimal fit.

$$V(Z) = \sum_{\vec{k}} v_{\vec{k}}\phi_{\vec{k}}(Z),\ Z \in \mathbb{R}^2 \tag{9}$$

$$I(x) = \mathcal{F}^{-1}\{V\}(x) = \sum_{\vec{k}} (i)^{\vec{k}} v_{\vec{k}}\phi_{\vec{k}}(x) \tag{10}$$

The Hermite-Gaussian basis functions has been noticed on signal processing for machine vision and image compression under the study of the "Hermite Transform" [44]. It was shown that this basis of function can be adapted to image structure observing high level of energy compaction [45]. The observed compression was claimed to be comparable to wavelets compaction. Those results suggest that Hermite-Gaussian representation should be sparse for the image synthesis problem. The accuracy of the Hermite-Gaussian representation has been tested on image processing of rotated images [46]. A fast implementation comparable to FFT of the Hermite-Gaussian expansion has been reported in [47].

The novelty of our proposal is the non-existence of a grid for calculating the Fourier transform in this approach. Thus the complexity of further algorithms on top of this expansion relates with the number of terms considered in the series (Eq. 10) that has before be claimed to be much fewer than the number of pixels in image and by energy compaction argument. Additionally, we show that the Hermite-Gaussian basis approximately decorrelate representation covariance.

Section II describes the complex valued function for approximation and its harmonic properties. Section III present the way to find the approximation to the image in Hermite-

Gaussian functions. Section IV describes the proof of concept. Finally, section V states main conclusion and future works.

II. A SUITABLE BASIS SET

We describe the functional setting and argumentation for the convenience of using the proposed representation.

A. The complex Hermite-Gaussian function

We considers the $C^\infty(\mathbb{R}^2)$ set of complex functions (e.g. [48], [49]) called Hermite-Gaussian basis (Eq. 11). This basis is orthogonal and complete on $L_2(\mathbb{C})$ [48]. We use by convenience the complex number notation $z = u+iv$ and $\bar{z} = u-iv$.

$$\phi_{k_1,k_2}(z,\bar{z}) = H_{k_1,k_2}(z,\bar{z})e^{-\frac{1}{2}z\bar{z}} \tag{11}$$

$$\exp(-\frac{1}{2}z\bar{z} + sz + t\bar{z} - st) = \sum_{k_1,k_2=0}^{\infty} \phi_{k_1,k_2}(z,\bar{z})\frac{s^{k_1}}{k_1!}\frac{t^{k_2}}{k_2!} \tag{12}$$

$$\mathcal{F}\{\phi_{k_1,k_2}\} = (-i)^{k_1+k_2}\phi_{k_1,k_2} \tag{13}$$

Polynomials $H_{k_1,k_2}(z,\bar{z})$ [50] from equation (11) are complex valued Hermite polynomials, which are defined from a generating function (Eq. 12). Fourier eigenfunction property (Eq. 13) is derived by applying Fourier transform at both sides of equation (12). Numerical evaluation of such polynomials [50] are performed by recurrence relationship (Eq. (14)-(16)).

$$H_{k_1+1,k_2}(z,\bar{z}) = zH_{k_1,k_2}(z,\bar{z}) - k_2H_{k_1,k_2-1}(z,\bar{z}) \tag{14}$$

$$H_{k_1,k_2+1}(z,\bar{z}) = \bar{z}H_{k_1,k_2}(z,\bar{z}) - k_1H_{k_1-1,k_2}(z,\bar{z}) \tag{15}$$

$$H_{0,0}(z,\bar{z}) = 1,\ H_{k_1<0,k_2<0}(z,\bar{z}) = 0 \tag{16}$$

The approximation expansion to an order N for visibility data is given by equation (17).

$$V(u,v) = \sum_{k_1,k_2=0}^{k_1+k_2=N/2} v_{k_1,k_2}\phi_{k_1,k_2} \tag{17}$$

However, the image function is real requiring an Hermitean conjugate symmetry $\overline{V}(-u,-v) = V(u,v)$. Complex valued Hermite function have symmetries under conjugation and index permutation (see [50]) (Eq. 18 and 19). This implies a restriction over expansion coefficients v_{k_1,k_2}.

$$H_{k_1,k_2}(z,\bar{z}) = \overline{H_{k_2,k_1}(z,\bar{z})} \tag{18}$$

$$H_{k_1,k_2}(-z,-\bar{z}) = (-1)^{k_1+k_2}H_{k_1,k_2}(z,\bar{z}) \tag{19}$$

The coefficients must be hermitian or anti-hermitian depending on the parity of the integer number $(k_1 + k_2)$ (Eq. 20). This roughly reduce the number of representing coefficient to a half.

$$v_{k_1,k_2} = v_{\vec{k}} = (-1)^{k_1+k_2}\ \overline{v_{\vec{k}}} \tag{20}$$

This representation implicitly enforce a Fourier transform in the image and support multiple way to enable the visibility interpolation. The image is rendered by expanding in the Hermite-Gaussian basis and using the Fourier eigenvector property from equation (13).

B. The Karhunen-Loeve framework

We describe a well known representation theorem for random fields, which has the property of being decorrelated, separating coordinates from stochastic behavior, and then simplifying the numeric and statistical treatment. We adopt the notation $\phi_{k_1,k_2}(z,\bar{z}) \equiv \phi_{\vec{k}}(Z)$, $Z = (u,v)$.

Theorem II.1. *(Karhunen-Loeve) If $\hat{I}(X)$ is a real random field square integrable, defined over $\mathbb{R}^2$, with $E(\hat{I}(X)) = 0$ (or more generally $\hat{I} = I - E(I)$), and continuous correlation kernel $K_{\hat{I}}(X,X') = E(\hat{I}(X)\hat{I}(X'))$. Then the random field $\hat{I}(X)$ admit the following representation.*

$$\hat{I}(X) = \sum_{\vec{k}=0}^{\infty} c_{\vec{k}}\phi_{\vec{k}}(X) \tag{21}$$

Where $\{\phi_{\vec{k}}\}$ is a set of orthogonal basis functions that fulfil the following eigenvalue equation and resulting orthogonal properties.

$$\int_{\mathbb{R}^2} K_{\hat{I}}(X,X')\phi_{\vec{k}}(X')dX' = \lambda_{\vec{k}}\phi_{\vec{k}}(X) \tag{22}$$

$$c_{\vec{k}} = \int_{\mathbb{R}^2} \hat{I}(X')\phi_{\vec{k}}(X)dX' \tag{23}$$

Then the representation's coefficient c_k are complex random variables that have the following properties:

$$E(c_k) = 0, \qquad \forall k \in \mathbb{N} \tag{24}$$

$$E(c_k^* c_{k'}) = \lambda_k\delta_{k,k'}, \qquad \forall k,k' \in \mathbb{N} \tag{25}$$

Proof: See [36] for more technical details on the existence of the $\phi_{\vec{k}}(X)$ orthogonal basis and regularity of the representation. However, properties (24) and (25) are easily deduced from the required property $E(\hat{I}(X)) = 0$ and basis orthogonality.

$$\begin{aligned} E(c_{\vec{k}}) &= E(\int_{\mathbb{R}^2} \hat{I}(X')\phi_{\vec{k}}(X)dX') = \\ &= \int_{\mathbb{R}^2} E(\hat{I}(X'))\phi_{\vec{k}}(X)dX' = 0 \end{aligned} \tag{26}$$

$$\begin{aligned} E(c_{\vec{k}}^* c_{\vec{k}'}) &= \\ &= E(\int_{\mathbb{R}^2} \hat{I}(X)\phi_{\vec{k}}(X)^*dX \int_{\mathbb{R}^2} \hat{I}(X')\phi_{\vec{k}'}(X')dX') \\ &= \int_{\mathbb{R}^2} E(\hat{I}(X)\hat{I}(X'))\phi_{\vec{k}}(X)^*\phi_{\vec{k}'}(X')dXdX' \\ &= \int_{\mathbb{R}^2} K_{\hat{I}}(X,X')\phi_{\vec{k}}(X)^*\phi_{\vec{k}'}(X')dXdX' \\ &= \int_{\mathbb{R}^2} [\int_{\mathbb{R}^2} K_{\hat{I}}(X,X')\phi_{\vec{k}}(X)dX]^*\phi_{\vec{k}'}(X')dX' \\ &= \int_{\mathbb{R}^2} [\lambda_{\vec{k}}\phi_{\vec{k}}(X)]^*\phi_{\vec{k}'}(X')dX' \\ &= \lambda_k\delta_{\vec{k},\vec{k}'} \end{aligned} \tag{27}$$

■

The next lemma explains the strong diagonal dominance of the covariance operator in the Hermite-Gaussian basis

representation, which implies a strong approximation for decorrelation by a exponential decay factor. The correlation is near to diagonal by a negative exponential error.

Lemma II.2. *Given the model expressed on Equations (1) and (2) the correlation matrix $E(v_{\vec{k}}v_{\vec{k}'})$ of the coefficient of expansion v_k of $\hat{V}(Z) = V(Z) - E(V(Z))$ in Hermite-Gaussian functions is approximately diagonal in the sense of equation (28).*

$$E(v_{\vec{k}}v_{\vec{k}'}) = \begin{cases} \int_{\mathbb{R}^2} |\sigma(Z)\phi_{\vec{k}}(Z)|^2 dZ & k = k' \\ O(|\vec{k}'|^{-|\vec{k}|}) & |\vec{k}| > |\vec{k}'| > 1 \\ O(|\vec{k}|^{-|\vec{k}'|}) & |\vec{k}'| > |\vec{k}| > 1 \\ O((1+1/M)^{-|\vec{k}|}) & |\vec{k}| > |\vec{k}'| = 1 \text{ or } 0 \\ O((1+1/M)^{-|\vec{k}'|}) & |\vec{k}'| > |\vec{k}| = 1 \text{ or } 0 \end{cases} \tag{28}$$

Proof: The visibility field $V(Z)$ is defined as in in the model (Eq. 2), where $\epsilon(Z)$ is an independent field with expectation 0 and delta-correlated $E(\epsilon(Z)\epsilon(Z')) = \delta(Z - Z')$ as defined in the white noise process [7]. Projecting the field $\hat{V}(Z)$ on the HG basis results in equation (29).

$$E(v_{\vec{k}}v_{\vec{k}'}) = E(\int_{\mathbb{R}^2} \hat{V}^*(Z)\hat{V}(Z')\phi^*_{\vec{k}}(Z)\phi_{\vec{k}'}(Z')dZdZ') \tag{29}$$

$$= \int_{\mathbb{R}^2} \sigma^*(Z)\sigma(Z')\delta(Z - Z')\phi^*_{\vec{k}}(Z)\phi_{\vec{k}'}(Z')dZdZ'$$

$$= \int_{\mathbb{R}^2} |\sigma(Z)|^2\phi^*_{\vec{k}}(Z)\phi_{\vec{k}'}(Z)dZ \tag{30}$$

However, integral in equation (30) can be approximated using inherited properties of the noise from the primary beam attenuation function (Eq. 1). If noise component is attenuated as a Gaussian within a decay scale parameter M in image the noise component in visibility (after Fourier transform) decay also as a Gaussian with scale parameter of $1/M$. We are considering as that the typical decay scale of the observed object is $m = 1$ and $m << M$, this is for ensuring minimal attenuation of the observed object. Then the term $|\sigma(Z)|^2\phi^*_{\vec{k}}(Z)$ in the integral has Gaussian decay with characteristic scale of $s \sim 1/M + \sqrt{2|\vec{k}|}$ according to [51], [52] (see Theorem III.2). This implies that coefficient of the HG expansion (Eq. 31) of the function $|\sigma(Z)|^2\phi^*_{\vec{k}}(Z)$ decay as fast as $\beta_k \sim O(e^{-p_{k'}k})$ [53], [54] (see Theorem III.1) considering $|\vec{k}'| > |\vec{k}|$. From [53] and using integral tables [55] is possible to estimate the order of $p_k \preceq O(\log(s))$ by replacing the term $|\sigma(Z)|^2\phi^*_{\vec{k}}(Z)$ by a Gaussian of the same scale. Replacing p_k value we obtains the order estimation (Eq. 32) and its symmetric version (Eq. 33).

$$|\sigma(Z)|^2\phi^*_{\vec{k}}(Z) = \sum_{\vec{k}'} \beta_{\vec{k}'}\phi_{\vec{k}'}(Z) \tag{31}$$

$$(|\vec{k}| > |\vec{k}'| > 1)\ ; \beta_{\vec{k}'} \sim O(e^{-\log(|\vec{k}'|)|\vec{k}|}) \sim O(|\vec{k}'|^{-|\vec{k}|}) \tag{32}$$

$$(|\vec{k}'| > |\vec{k}| > 1)\ ; \beta_{\vec{k}} \sim O(e^{-\log(|\vec{k}|)|\vec{k}'|}) \sim O(|\vec{k}|^{-|\vec{k}'|}) \tag{33}$$

The case with $|\vec{k}'| < |\vec{k}|$ is symmetric. The cases with $|\vec{k}| = 0$ *or* 1 considers a constant scale $s \sim 1/M + 1$ greater than one, then allowing a decay of the correlation. When $|\vec{k}| = |\vec{k}|$ the integral (Eq. 30) reduce to the case in (Eq. 28). ■

The previous lemma states the level of approximation that HG basis diagonalize the covariance matrix of coefficient representation. Furthermore, continuous correlation kernel operator $K(X, X')$ is also approximately diagonalized by this basis by backtracking the Karhunen-Loeve theorem's proof and using this order estimation.

III. Image synthesis method

In our case, the purpose of the image synthesis method is to find the best linear combination of Hermite-Gaussian function for representing visibilities. The image is rendered by using the Fourier eigenvector property (Eq. 13) and evaluating on pixel coordinates. Several methods based on optimization exist in order to find a good approximation for coefficients.

$$V^m(Z) = \sum_k^N v_k\phi_k(Z) \tag{34}$$

$$\min_{\{v_k\}} \sum_l |(V_l^O - V^m(X_l))/\sigma_l|^2 \tag{35}$$

In practice, interpolation with polynomials in more than one dimension results in ill posed/conditioned linear systems. Regularization approaches should be used for solving this issue. Commonly those regularization are implemented as additional term in the functional to be minimized.

Our main assumption is the sparsity of the Hermite-Gaussian representation when the observed image fade radially as fast as an exponential. This assumption is supported by the antenna reception pattern called "primary beam" which is usually modelled as a Gaussian [5]. Therefore, if the image function decay as a simple Gaussian then its coefficient in series expansion decay faster than any polynomials in $\vec{k}$. Thus, sparsity is interpreted as long as first coefficients contain most of the image's information and the rest ignored by its fast decay.

Theorem III.1. *If the inverse problem of radio interferometry is affected by a Gaussian primary beam, then the coefficient of expansion in Hermite-Gaussian basis decay faster than any finite inverse power of the basis polynomial degree.*

Proof: See [54]. ■

With the previous theorem we can claim that our proposed expansion can be considered sparse within an error that decay exponentially with the degree for interferometric observations.

The Lasso [56] regularization is a common way to model an sparse problem. It consist to add a penalization term proportional to the L_1 norm. This method has the disadvantage of being parametrized by α as in equation (37).

$$V^m(X) = \sum_k^N v_k \phi_k(X) \tag{36}$$

$$\min_{\{v_k\}} \sum_l |(V_l^O - V^m(X_l))/\sigma_l|^2 + \alpha \sum_k |v_k| \tag{37}$$

We consider scaling consideration for conditioning the problem. Polynomial interpolation are known to be bad conditioned. A method for solving this issue was defining a scale for the problem in order to reduce the number coefficient to compute [51].

Theorem III.2. *If we consider an image with radial decay as a Gaussian then the approximation is numerically stable to order N if the scale of the problem ($I(X) = \sum_k^N c_k \phi(sX)$) is defined to be*

$$s \sim \sqrt{2N}/M \tag{38}$$

with M the radial limit in which the function $I(X)$ decay.

Proof: See [51] and considers that the primary beam decay at a radius M. Also considers that the last zero of the Hermite polynomial of order N is approximately $\sqrt{2N}$ [52]. ■

In this setting a one dimensional HG set of function are built normalized and working scaled to the interval $[-\sqrt{2N}, +\sqrt{2N}]$. Therefore, normalized HG functions are bounded by a single constant as the Cramer's inequality states (Eq. 39). This remarkable fact indicate that it is much numerically stable to compute $\phi_k(X)$ than calculating first the Hermite polynomial and then building the normalized HG functions.

$$\phi_k(X) \leq K\pi^{-\frac{1}{4}} \tag{39}$$

We implemented the computation of the HG basis by mean of an improved version of the recurrence relations (14-16) [57] for normalized ϕ_k instead of the polynomial. Thus this constitute a numerically stable implementation of the HG basis for higher order.

IV. Experimental results

As a proof-of-concept we present the Lasso method on a T1 weighted Magnetic Resonance Image (T1-MRI). We use we use this image (128x128 pix) (Fig. 1) as a model image and considering it decay to zero intensity to the border of the canvas.

A first experiment for inspecting the degree of interpolation was to interpolate the previous image by Hermite-Gaussian (HG) functions. The method used was to find directly the coefficient by orthogonality property of HG. We obtains very similar image (Fig. 2) than the original with some subtle ringing artifacts on borders. However as long as the image fills the canvas the interpolation lower its quality as expected from theorem III.2.

A second experiment is to see the effect of interpolating on Fourier space but on a regular grid. The Fourier transformation is made by FFT and interpolation method is again the simple projection method. As long as the scaling chosen from theorem III.2 contains the image the discretized version of the Hermite transform becomes approximately unitary and so its representing matrix. A degradation by effect of very small ringing artifact is noticed (Fig. 3), but image remains very near to original and almost equal than the interpolated image.

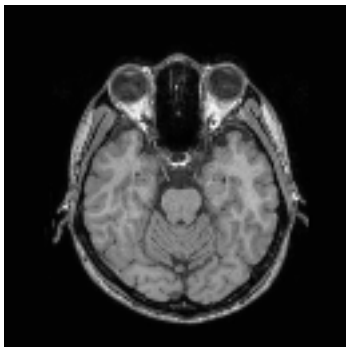

Fig. 1. MRI image model.

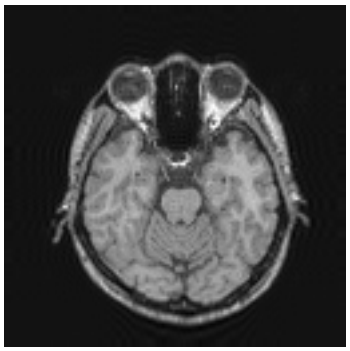

Fig. 2. Hermite-Gaussian interpolation (N=100).

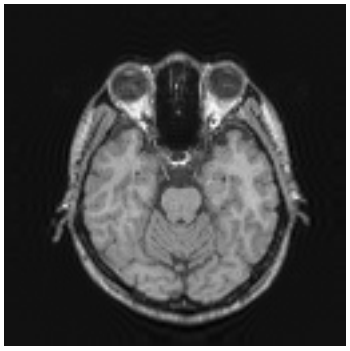

Fig. 3. The interpolation on regular grid in complex Fourier space (N=100).

It is interesting to show the real coefficient representation showing the sparsity. Most high valued coefficient are near the zero degree. If we only consider the 20% of the higher valued coefficient the resulting interpolated image remains the same. We notice in this representation the even and odd coefficient. If a coefficient has degree even then its associated Fourier eigenvalue is real and imaginary if not. Thus as log as the

image is real odd coefficient must be zero.

The third experiment is about to generate a random sampling in Fourier space. In this case (Fig. 4) we use the Lasso procedure for fitting the complex coefficient. We recover image with a sampling of 4000 points randomly placed (24% of this regular sampling), it clearly degrade the image but observed main features remains.

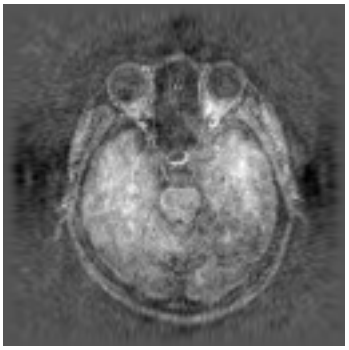

Fig. 4. Recovered image with the best alpha parameter.

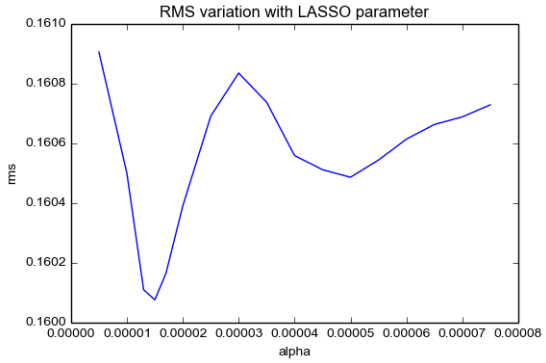

Fig. 5. Search of an optimal alpha for LASSO method. Best alpha found $\alpha = 1.5e-05$.

The alpha parameter of the Lasso model was estimated by minimizing the rms of the original figure in order to identify a best image. We show an image resulting from processing (Fig. 5) using an optimal alpha of $1.5e-05$ and having a rms of 0.16 ± 0.03 over a normalized peak signal of 1.0.

Coefficient for the recovered image has a 15% of low level amplitude (less than 1.0) considering a maximal value of 630 in real part and 15 in imaginary part. Low level amplitude spectra has been removed from the image and the results is shown in (Fig. 7). The resulting Hermite-Gaussian spectra (coefficient in lexicographic order by power) is shown (Fig. 6).

The last spectra shows many low amplitude could be discarded, then we test the case for an amplitude less than 1 (Fig. 7). That results in an almost equal image than the best image (Fig. 4).

The next experiment is to include noise in the sampled signal (Fig. 8). In this case some part of the image appears

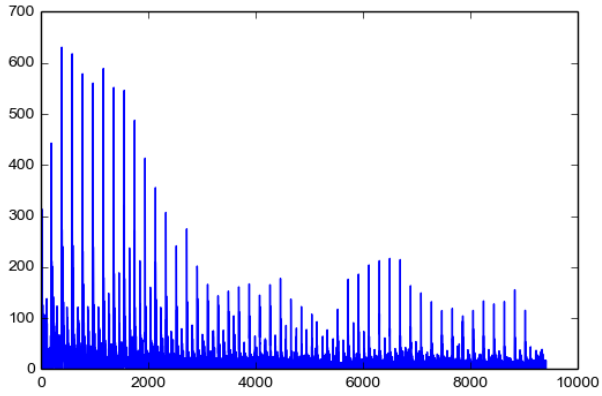

Fig. 6. Hermite-Gaussian amplitude spectra of LASSO method.

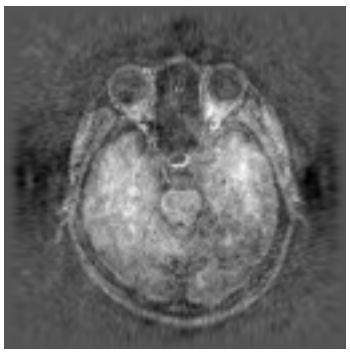

Fig. 7. LASSO recovered image with a cut-off of the 15% lower amplitude spectra.

blurred due to the injected noise. In spite of the artifacts observed on the resulting image, fine grain structure are still distinguishable as a difference with the clean method.

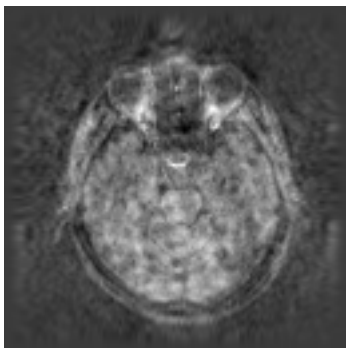

Fig. 8. The interpolation by Lasso method on a sparse sampling with 10% of noise amplitude (N=100).

The final experiment shows the effect of a real sampling from a real interferometer such as ALMA. The data set is from 12500 visibility of 80000 ALMA sampling from the protoplanetary disk HD142527 band9 continuum frequency range (ALMA data serial code: ADS/JAO.ALMA#2011.0.00465.S) [58]. The Lasso method (Fig. 9) reach similar image than

the famous "horse shoe" (Fig. 10) using the CASA software [17]. The CLEAN image reduction was configured using niter=1000, weighting="briggs", imsize=[128,128], and cell="0.063arcsec" in the automatic setting. The CLEAN image exhibit a signal-to-noise ratio of $SNR = 0.13$ and the LASSO interpolated image with only a 16% of the sampling reach $SNR = 0.07$. The clean image expose a dotted artifact structure on the "Horse Shoe" signal that is not observed at the interpolation side.

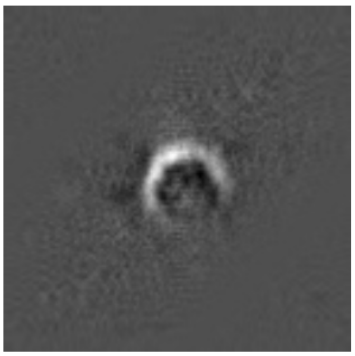

Fig. 9. The horse shoe: Interpolation results image from a 16% of real sampling of the HD12527 protoplanetary disk system of cycle-0 radio observatory ALMA.

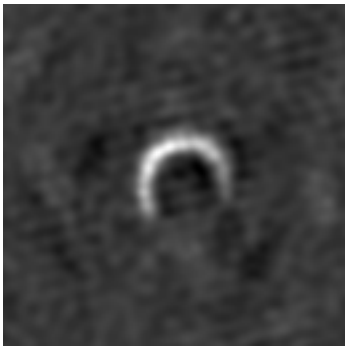

Fig. 10. The horse shoe: CLEAN image obtained using the CASA software.

V. Discussion and conclusions

A grid-less method for imaging is discussed into the framework of image synthesis. The method aims to overcome the errors induced by the gridding method and impose on the representation structure the Fourier transform. If the approximation method performs the computation on visibility domain then this representation ought to be superior in relation to traditional gridding. The representation has theoretical reason for being superior under noise in data. There is also proven theoretical reason for expecting sparse representation of the image in this basis.

Hermite-Gaussian atoms give representation from point structure to localized oscillations. We believe that further improvement of image representation on this basis is by considering redundancy by different scaling of atoms. This leads to a redundant dictionary representation expecting a sparsity. However, larger number of atoms drastically lowers computation performance. Therefore, we have the hope to find an efficient method for rendering the image based in the summation property of the complex Hermite polynomials [50]. This further work aims to calculate rendered regions based on nearby ones.

Artifact observed in recovered images with noise and sparse sampling seems cleaner than common reconstruction using traditional methods. The interpolation in the regular case is not as good as the FFT algorithm, but as long as the sparsity and noise is incorporated on test images the representation becomes superior. Image in radio astronomy are restricted to fade after certain radius. This is the reason to use the test image. However, as smaller is the image support with respect to the canvas then sparser are the coefficient. The Lasso method recover this property and impose sparsity on reconstruction as a regularization method. Lasso images recover the peak of intensity and noise remains lower.

We recover better SNR from Lasso interpolation than CLEAN images. Therefore, this proof-of-concept pave the way for further research in image synthesis algorithm based on the HG basis. Future work not only should test several convex optimization formulations and methods, but also statistical analysis based on component enabled by diagonalized covariance.

Acknowledgment

We thanks to the Chilean Postdoctoral Fondecyt project number 3140634. We acknowledge support from the Millenium Science Initiative (Chilean Ministry of Economy), through the grant Nucleus P10-022-F. We thanks to the Chilean Fondecyt project number 1111012. This research was partially supported by the supercomputing infrastructure of the NLHPC (ECM-02).

References

[1] A. Richard Thompson, James M. Moran, and Jr. George W. Swenson, *Interferometry and Synthesis in Radio Astronomy*, Wiley, 2008.

[2] P. Marechal and D. Wallach, "Fourier synthesis via partially finite convex programming.," *Mathematical and Computer Modelling*, vol. 49, no. 11-12, pp. 2206–2212, 2009.

[3] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics, 2005.

[4] B. G. Clark, "Coherence in Radio Astronomy," in *Synthesis Imaging in Radio Astronomy II*, G. B. Taylor, C. L. Carilli, and R. A. Perley, Eds., 1999, vol. 180 of *Astronomical Society of the Pacific Conference Series*, pp. 1–10.

[5] G. B. Taylor, C. L. Carilli, and R. A. Perley, *Synthesis Imaging in Radio Astronomy II*, Number 180 in Astronomical Society of the Pacific Conference Series. ASP, 1999.

[6] B. Mason and C. Brogan, "Relative integration times for the alma cycle 1 12-m, 7-m, and total power arrays," *NAASC Memo Series/ALMA Memo*, vol. 113, no. 598, 2013, The North American ALMA Science Center (NAASC).

[7] V.I. Klyatskin, *Stochastic Equations through the Eye of the Physicist: Basic Concepts, Exact Results and Asymptotic Approximations*, Elsevier Science, 2005.

[8] T. J. Cornwell and K. F. Evans, "A simple maximum entropy deconvolution algorithm," *Astronomy and Astrophysics*, vol. 143, pp. 77–83, 1985.

[9] W. Chen, "The ill-posedness of the sampling problem and regularized sampling algorithm," *Digital Signal Processing*, vol. 21, no. 2, pp. 375 – 390, 2011.

[10] J. A. Högbom, "Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines," *Astronomy and Astrophysics Supplement*, vol. 15, pp. 417, 1974.

[11] B. G. Clark, "An efficient implementation of the algorithm clean," *Astronomy and Astrophysics*, vol. 89, pp. 377–378, 1980.

[12] F. R. Schwab, "Relaxing the isoplanatism assumption in self-calibration; applications to low-frequency radio interferometry," *Astronomical Journal*, vol. 89, pp. 1076–1081, 1984.

[13] J. L. Starck, A. Bijaoui, B. Lopez, and C. Perrier, "Image reconstruction by the wavelet transform applied to aperture synthesis," *Astronomy and Astrophysics*, vol. 283, pp. 349–360, 1994.

[14] S. Bhatnagar and T. J. Cornwell, "Scale sensitive deconvolution of interferometric images. i. adaptive scale pixel (asp) decomposition," *Astronomy and Astrophysics*, vol. 426, pp. 747–754, 2004.

[15] U. Rau and T. J. Cornwell, "A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry," *Astronomy and Astrophysics*, vol. 532, 2011.

[16] C. Fannjiang, "Optimal arrays for compressed sensing in snapshot-mode radio interferometry," *Astronomy and Astrophysics*, vol. 559, pp. A73, Nov. 2013.

[17] J. P. McMullin, B. Waters, D. Schiebel, W. Young, and K. Golap, "CASA Architecture and Applications," in *Astronomical Data Analysis Software and Systems XVI*, R. A. Shaw, F. Hill, and D. J. Bell, Eds., Oct. 2007, vol. 376 of *Astronomical Society of the Pacific Conference Series*.

[18] A. Lannes, E. Anterrieu, and P. Marechal, "Clean and Wipe," *Astronomy and Astrophysics*, vol. 123, pp. 183–198, 1997.

[19] E. J. Candes and J. Romberg, "Sparsity and incoherence in compressive sampling. inverse problems," *Inverse Problems*, vol. 23, pp. 969–985, 2006.

[20] D. L. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, 2006.

[21] G. B. Taylor, C. L. Carilli, and R. A. Perley, Eds., *Synthesis Imaging in Radio Astronomy II*, vol. 180 of *Astronomical Society of the Pacific Conference Series*, Astronomical Society of the Pacific, San Francisco, 1999.

[22] C.R. Vogel, *Computational Methods for Inverse Problems*, Frontiers in Applied Mathematics. Society for Industrial and Applied Mathematics, 2002.

[23] J. G. Ables, "Maximum Entropy Spectral Analysis," *Astronomy and Astrophysics*, vol. 15, pp. 383, June 1974.

[24] S. F. Gull and G. J. Daniell, "Image reconstruction from incomplete and noisy data," *Nature*, vol. 272, pp. 686–690, Apr. 1978.

[25] J. L. Starck, F. Murtagh, P. Querre, and F. Bonnarel, "Entropy and astronomical data analysis: Perspectives from multiresolution analysis," *Astronomy and Astrophysics*, vol. 368, pp. 730–746, Mar. 2001.

[26] E. C. Sutton and B. D. Wandelt, "Optimal Image Reconstruction in Radio Interferometry," *The Astrophysical Journal Supplement Series*, vol. 162, pp. 401–416, Feb. 2006.

[27] R. E. Carrillo, J. D. McEwen, Dimitri Van De Ville, J-P Thiran, and Y. Wiaux, "Sparsity averaging for compressive imaging.," *IEEE Signal Process. Lett.*, vol. 20, no. 6, pp. 591–594, 2013.

[28] R. E. Carrillo, J. D. McEwen, D. Van De Ville, J. P. Thiran, and Y. Wiaux, "Sparsity Averaging for Compressive Imaging," *IEEE Signal Processing Letters*, vol. 20, pp. 591–594, June 2013.

[29] H. Garsden, J. N. Girard, and et al. Starck, "LOFAR Sparse Image Reconstruction," *ArXiv eprints*, June 2014.

[30] E. Oja, A. Hyvrinen, and P. Hoyer, "Image feature extraction and denoising by sparse coding," *Pattern Analysis & Applications*, vol. 2, no. 2, pp. 104–110, 1999.

[31] Peter Lennie, "The cost of cortical computation," *Current Biology*, vol. 13, no. 6, pp. 493 – 497, 2003.

[32] E. P. Simoncelli, "Statistical modeling of photographic images," in *Handbook of Image and Video Processing*, A. Bovik, Ed., chapter 4.7, pp. 431–441. Academic Press, May 2005, 2nd edition.

[33] Jianchao Yang, Kai Yu, Yihong Gong, and Thomas S. Huang, "Linear spatial pyramid matching using sparse coding for image classification.," in *CVPR*. 2009, pp. 1794–1801, IEEE.

[34] P.M. Sutter, B.D. Wandelt, J.D. McEwen, E.F. Bunn, A. Karakci, A. Korotkov, P. Timbie, G.S. Tucker, and L. Zhang, "Probabilistic image reconstruction for radio interferometers," *Monthly Notices of the Royal Astronomical Society*, vol. 00, pp. 0, 2014.

[35] D. S. Briggs, F. R. Schwab, and R. A. Sramek, "Imaging," in *Synthesis Imaging in Radio Astronomy II*, G. B. Taylor, C. L. Carilli, and R. A. Perley, Eds., 1999, vol. 180 of *Astronomical Society of the Pacific Conference Series*, pp. 127–149.

[36] R. Ghanem and P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, World Publishing Corporation, 1991.

[37] C. Candan, M. A. Kutay, and H. M. Ozaktas, "The discrete fractional fourier transform," *Signal Processing, IEEE Transactions on*, vol. 48, no. 5, pp. 1329–1337, 2000.

[38] B. Santhanam and T. S. Santhanam, "On discrete gausshermite functions and eigenvectors of the discrete fourier transform," *Signal Processing*, vol. 88, no. 11, pp. 2738 – 2746, 2008.

[39] H. M. Ozaktas, D. Mendlovic, L. Onural, and B. Barshan, "Convolution, filtering, and multiplexing in fractional fourier domains and their relation to chirp and wavelet transforms," *J. Opt. Soc. Am. A*, vol. 11, no. 2, pp. 547–559, Feb 1994.

[40] M.A. Kutay, H.M. Ozaktas, O. Ankan, and L. Onural, "Optimal filtering in fractional fourier domains," *Signal Processing, IEEE Transactions on*, vol. 45, no. 5, pp. 1129–1143, May 1997.

[41] G. B. Arfken and H. J. Weber, *Mathematical Methods For Physicists International Student Edition*, Elsevier Science, 2005.

[42] A. Refregier, "Shapelets i. a method for image analysis," *Monthly Notices of the Royal Astronomical Society*, vol. 338, no. 1, pp. 35–47, 2003.

[43] P. Melchior, M. Meneghetti, and M. Bartelmann, "Reliable shapelet image analysis," *Astronomy and Astrophysics*, vol. 463, no. 3, pp. 1215–1225, 2007.

[44] A. M. van Dijk and J. B. Martens, "Image representation and compression with steered Hermite transforms," *Signal Processing*, vol. 56, no. 1, pp. 1–16, 1997.

[45] B. Escalante-Ramirez, "The hermite transform as an efficient model for local image analysis: An application to medical image fusion," *Computers & Electrical Engineering*, vol. 34, no. 2, pp. 99 – 110, 2008.

[46] W. Park, G. Leibon, D. N. Rockmore, and G. S. Chirikjian, "Accurate image rotation using hermite expansions," *Trans. Img. Proc.*, vol. 18, no. 9, pp. 1988–2003, Sept. 2009.

[47] G. Leibon, Daniel N. Rockmore, W. Park, R. Taintor, and G. S. Chirikjian, "A fast hermite transform," *Theoretical Computer Science*, vol. 409, no. 2, pp. 211 – 228, 2008, Symbolic-Numerical Computations.

[48] A. Intissar and A. Intissar, "Spectral properties of the cauchy transform on l2," *Journal of Mathematical Analysis and Applications*, vol. 313, no. 2, pp. 400 – 418, 2006.

[49] A. Ghanmi, "A class of generalized complex hermite polynomials," *Journal of Mathematical Analysis and Application*, , no. 304, pp. 1395 – 1406, 2008.

[50] A. Ghanmi, "Operational formulae for the complex hermite polynomials hp, q(z, z)," *Integral Transforms and Special Functions*, vol. 24, no. 11, pp. 884–895, 2013.

[51] T. Tang, "The hermite spectral method for gaussian-type functions," *SIAM J. Sci. Comput.*, vol. 14, no. 3, pp. 594–606, May 1993.

[52] G. Szegö, *Orthogonal Polynomials*, Number v. 23 in American Mathematical Society colloquium publications. American Mathematical Society, 1959.

[53] J. P. Boyd, "The rate of convergence of hermite function series," *Mathematics of Computation*, , no. 35, pp. 1309–1316, 1980.

[54] J. P. Boyd, "Asymptotic coefficients of hermite function series," *Journal of Computational Physics*, , no. 54, pp. 382 – 410, 1984.

[55] I.S. Gradshtein and D. Zwillinger, *Table of Integrals, Series, and Products*, Academic Press. Elsevier Science & Technology Books, 2014.

[56] R. Tibshirani, "Regression shrinkage and selection via the lasso: a retrospective," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 73, pp. 273282, 2011.

[57] L. Aceto, H. R. Malonek, and G. Tomaz, "A unified matrix approach to the representation of Appell polynomials," *ArXiv e-prints*, jun 2014.

[58] S. Casassus, G. van der Plas, M, S. P., W. R. F. Dent, E. Fomalont, J. Hagelberg, A. Hales, A. Jordán, D. Mawet, F. Ménard, A. Wootten, D. Wilner, A. M. Hughes, M. R. Schreiber, J. H. Girard, B. Ercolano, H. Canovas, P. E. Román, and V. Salinas, "Flows of gas through a protoplanetary gap," *Nature*, vol. 493, pp. 191–194, Jan. 2013.