Representation of higher-order statistical structures in natural scenes via spatial phase distributions

HaDi MaBouDi [a], Hideaki Shimazaki [b], Shun-ichi Amari [b], Hamid Soltanian-Zadeh [a,c,d,∗]

[a] School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
[b] RIKEN Brain Science Institute, Wako-Shi, Saitama, Japan
[c] Control and Intelligent Processing Center of Excellence (CIPCE), School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran
[d] Image Analysis Laboratory, Department of Radiology, Henry Ford Health System, Detroit, MI, United States

ABSTRACT

Natural scenes contain richer perceptual information in their spatial phase structure than their amplitudes. Modeling phase structure of natural scenes may explain higher-order structure inherent to the natural scenes, which is neglected in most classical models of redundancy reduction. Only recently, a few models have represented images using a complex form of receptive fields (RFs) and analyze their complex responses in terms of amplitude and phase. However, these complex representation models often tacitly assume a uniform phase distribution without empirical support. The structure of spatial phase distributions of natural scenes in the form of relative contributions of paired responses of RFs in quadrature has not been explored statistically until now. Here, we investigate the spatial phase structure of natural scenes using complex forms of various Gabor-like RFs. To analyze distributions of the spatial phase responses, we constructed a mixture model that accounts for multi-modal circular distributions, and the EM algorithm for estimation of the model parameters. Based on the likelihood, we report presence of both uniform and structured bimodal phase distributions in natural scenes. The latter bimodal distributions were symmetric with two peaks separated by about 180°. Thus, the redundancy in the natural scenes can be further removed by using the bimodal phase distributions obtained from these RFs in the complex representation models. These results predict that both phase invariant and phase sensitive complex cells are required to represent the regularities of natural scenes in visual systems.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Animals have to efficiently cope with natural environment for their survival. This imposes sensory systems to adapt to regularities of the natural environment. Barlow (Barlow, 1961) hypothesized that the goal of sensory systems is to reduce redundancy in the inputs, and represent them by statistically independent features to avoid ambiguity. More generally, Bayesian approaches for modeling perception seek for a suitable prior to represent environment in the brain. Thus, we gain insight into principles of neural systems by exploring the statistics of natural environment. This approach was most successfully applied to visual systems. For example, analysis of natural scene statistics has explained economy of neurons in the early visual systems (Field, 1987; Geisler,

2008; Hyvärinen, 2010; Simoncelli & Olshausen, 2001; van der Schaaf & van Hateren, 1996).

Statistical modeling studies of natural scenes also supported the efficient coding hypothesis. Celebrated examples are linear generative models that successfully replicated receptive fields (RFs) of simple cells in the primary visual cortex (Bell & Sejnowski, 1997; Hyvärinen, Hurri, & Hoyer, 2009; Olshausen, 1996). However, it is known that the early linear generative models do not fully explain the regularities in natural scenes. After training, residual dependency remains among the coefficients of the basis functions of these models (Simoncelli & Buccigrossi, 1997; Simoncelli & Olshausen, 2001; Zetzsche & Rhrbein, 2001). It has been reported that joint responses of a pair of neighboring oriented Gabor-like RFs with arbitrary angles exhibit spherical and symmetric dependency (Simoncelli & Olshausen, 2001; Zetzsche & Rhrbein, 2001). This indicates that the models did not completely remove higher-order statistics in the natural scenes despite the fact that perceptually salient features are often characterized by structured higher-order dependency (Hyvärinen, 2010; Karklin & Lewicki,

2003, 2006; Olshausen & Field, 1996; Thomson, 1999). Recently, several authors proposed nonlinear models that can account for the higher-order statistics of natural scenes and reproduced nonlinear properties of visual neurons (Cadieu & Olshausen, 2008, 2012; Hyvärinen & Hoyer, 2000; Hyvärinen, Hoyer, & Inki, 2001; Karklin & Lewicki, 2005; Lyu & Simoncelli, 2008; Sasaki, Gutmann, Shouno, & Hyvärinen, 2013; Sinz & Bethge, 2010; Wainwright, Schwartz, & Simoncelli, 2002).

What statistical structure underlies perceptually salient features in natural scenes? One important characteristics is illuminated by the fact that an image is not easily perceived when its phases are randomized, although our perception is rather intact after amplitude randomization (Oppenheim & Lim, 1981; Hansen & Hess, 2007; Wichmann, Braun, & Gegenfurtner, 2006). Thus, phases of an image contain significantly more perceptual information than its amplitudes. Indeed, our percept for presence or location of visually salient features are strongly correlated with their spatial phase information (Badcock, 1984; Burr, 1980; Morrone & Burr, 1988). More specifically, perceptually salient features such as edges, bars, and even more complex shapes in natural scenes are characterized by congruence between phases across different spatial frequencies (Kovesi, 1999; Morrone, Ross, Burr, & Owens, 1986; Wang & Simoncelli, 2003). A human discriminates an edge and bar based on the degree of the phase congruency (Badcock, 1984; Bennett & Banks, 1991; Burr, Morrone, & Spinelli, 1989; Field & Nachmias, 1984; Morrone et al., 1986) while such information is coded in the ventral visual streams (Henriksson, Hyvärinen, & Vanni, 2009; Mechler, Reich, & Victor, 2002; Perna, Tosetti, Montanaro, & Morrone, 2008). Neurophysiological studies demonstrated that both simple and complex cells in macaque V1 were sensitive to different degrees of the phase congruency (Mechler et al., 2002). Moreover, Portilla and Simoncelli developed a method for synthesizing textures from amplitude and phase statistics available from responses of complex, oriented wavelets to original textures (Portilla & Simoncelli, 2000). Applications of this method successfully demonstrated an involvement of monkey and human V2 area in encoding such textures (Freeman, Ziemba, Heeger, Simoncelli, & Movshon, 2013).

Despite the apparent importance of phase information, physiological mechanisms underlying detection of spatial phase (local phase) information are largely unknown. Image phases may be detected by combining responses of simple cells possessing two odd and even symmetric RFs. Indeed, nearby simple cells in the primary visual cortex exhibit similar Gabor-like RFs except that their phases are in quadrature (separated in phase by 90°) (Pollen & Ronner, 1981, see also DeAngelis, Ghose, Ohzawa, & Freeman, 1999 for an extended range of phase differences). Although simple and complex cells in the V1 area are typically classified by their phase invariance properties (Albrecht, De Valois, & Thorell, 1980; Movshon, Thompson, & Tolhurst, 1978; Skottun et al., 1991), recent studies reported evidence of some complex cells that are highly sensitive to spatial phase information (Crowder, Van Kleef, Dreher, & Ibbotson, 2007; Hietanen et al., 2013; Mechler & Ringach, 2002; Mechler et al., 2002). For example, spike-triggered covariance analysis of complex cells responding to natural scenes revealed two dominant subunits that were in quadrature phase (Touryan, Felsen, & Dan, 2005) and showed that they were more sensitive to image phases than amplitudes (Felsen, Touryan, Han, & Dan, 2005). These studies suggested that cells in an intermediate stage of visual systems might detect local phase structures by nonlinearly combining outputs of simple cells that have similar structure but are quadrature in a phase space.

Most previous nonlinear generative models of natural scenes accounted for information available only from their amplitudes, which resulted in reproducing response properties of *phase-invariant* complex cells (Hyvärinen & Hoyer, 2000; Hyvärinen &

Köster, 2006; Hyvärinen et al., 2001). However, recent studies suggested simultaneous yet separate representation of natural scenes with local amplitude and spatial phase using a complex form of RFs (Cadieu & Olshausen, 2008, 2012; Olshausen, Cadieu, & Warland, 2009). In this framework, joint responses of a pair of quadrature-phase RFs are represented in polar coordinates (amplitude and phase), which are conveniently described by complex values (Daugman, 1985). These models assume that the joint responses are factorial in the polar coordinates, and further assume that its phase distribution is uniform in order to parsimoniously explain the spherical dependency often observed in the joint responses of quadrature pairs of RFs. To our knowledge, Wegmann & Zetzsche first proposed that the polar representation of images is more suitable for the statistics of images than the Cartesian coordinates (Wegmann & Zetzsche, 1990). They introduced an efficient technique for encoding images by assuming that spatial amplitude and phase variables are independent, and that the phase variables are uniformly distributed. Recently, Laparra, Gutmann, Malo, and Hyvärinen (2011) considered a bimodal phase distribution as a prior for complex independent component analysis in modeling natural scenes (Laparra et al., 2011). The relation of the bimodal phase distribution to the higher-order dependency in natural scenes, however, must be carefully examined by eliminating artifacts caused by linear correlations in the responses, using a controlled set of RFs. While the statistical modeling studies of visual systems began to utilize the perceptually salient phase information to encode higher-order statistics of natural scenes, systematic investigation of the local phase structure of natural scenes based on the complex form of RFs has not been proposed.

In this study, we investigate spatial phase information of natural scenes using a variety of complex forms of RFs (complex RFs). Among the spectrum of Gabor-like complex RFs we tested, two thirds of the complex RFs exhibited a uniform phase distribution whereas one third was characterized by a symmetric, bimodal phase distribution with two peaks being separated by 180°. By a simulation approach, we further demonstrate that the uniform and bimodal phase distribution can arise from the interaction between the non-Gaussianity and higher-order dependency in the natural scenes. These results indicate that different complex RFs capture higher-order statistics of natural scenes via different types of phase distributions. Thus the uniform distribution alone is insufficient to remove the redundancy of natural scenes, but the redundancy can be further removed if we consider bimodal phase distributions.

2. Methods

2.1. Stimuli and preprocessing

We used whitened natural scenes provided by Olshausen and Field (1997) (The data is available online at http://redwood.berkeley.edu/bruno/sparsenet/). Their data sets were processed from the original Hans van Hateren's repository (van Hateren & van der Schaaf, 1998). To extract local phase information from natural scenes, we randomly selected 100,000 patches with size 32x32 from random locations in the preprocessed images. To ensure that the mean luminance of all patches were identical, we removed the DC component (the mean pixel value of the image patch) and rescaled the pixel values to make their variance equal to 1.

2.2. Complex representation of natural scenes

We model a two-dimensional image $I_m(\vec{x})$ using a set of complex basis functions and complex coefficients (Cadieu & Olshausen, 2008). Here, $\vec{x}$ represents a discrete spatial position,

or a pixel, in the image, and m (=1, 2,...) is an image index. Using d different complex basis functions, this model is described as

$$I_m(\vec{x}) = \sum_{i=1}^{d} \text{Re}\{s_{i,m}^* A_i(\vec{x})\}. \tag{1}$$

The operation Re{.} returns a real part of a complex value. In Eq. (1) each complex basis function is decomposed as $A_i(\vec{x}) = A_i^R(\vec{x}) + jA_i^I(\vec{x})$, where $A_i^R(\vec{x})$ and $A_i^I(\vec{x})$ are real and imaginary components and $j = \sqrt{-1}$. Similarly, the complex coefficient is denoted as $s_{i,m} = s_{i,m}^R + js_{i,m}^I$. The asterisk in Eq. (1) denotes the conjugate complex. In this paper, we term the complex basis function as a complex RF and the complex coefficient as a complex response of the complex RF. We define the real and imaginary components of the complex RF in the next subsection.

A series of complex responses, $s_{i,m}$ (m = 1, 2,...), of the i-th complex RF, $A_i(\vec{x})$, to the natural scene patches, $I_m(\vec{x})$, can be computed from Eq. (1) by taking an inner product of the complex RF with the natural scene patches (Fig. 1A). Fig. 1B illustrates simultaneous responses of the real and imaginary components of a Gabor-like complex RF to 100,000 natural scene patches randomly selected from whitened natural scenes. The responses of each component bear sparse distributions as reported previously for this type of RFs (Field, 1987; Field, 1994; Olshausen & Field, 2004) (see the marginal histograms next to the axes).

The responses of two Gabor filters that are close in space, orientation, or scale often exhibit circular dependency (Zetzsche & Rhrbein, 2001) as shown in an example in Fig. 1B. Thus, it was proposed that this dependency may be concisely described by using polar coordinates (Zetzsche, Krieger, & Wegmann, 1999; Zetzsche & Rhrbein, 2001). We extracted the spatial amplitude and phase from natural scenes as follows. In the polar coordinates, the complex response is written as $s_{i,m} = r_{i,m}e^{j\varphi_{i,m}}$, where $r_{i,m}$ and $\varphi_{i,m}$ are an amplitude and phase of the complex response, respectively. Eq. (1) is now written as

$$I_m(\vec{x}) = \sum_{i=1}^{d} r_{i,m}\left[\cos(\varphi_{i,m})A_i^R(\vec{x}) + \sin(\varphi_{i,m})A_i^I(\vec{x})\right]. \tag{2}$$

By taking the inner product of the complex RF with the natural scene patches represented in Eq. (2), the amplitude and phase responses are given respectively by

$$r_{i,m} = \sqrt{\left\langle A_i^R(\vec{x}), I_m(\vec{x})\right\rangle^2 + \left\langle A_i^I(\vec{x}), I_m(\vec{x})\right\rangle^2}, \tag{3}$$

and

$$\varphi_{i,m} = \arctan\frac{\left\langle A_i^I(\vec{x}), I_m(\vec{x})\right\rangle}{\left\langle A_i^R(\vec{x}), I_m(\vec{x})\right\rangle}. \tag{4}$$

The amplitude, $r_{i,m}$, in Eq. (3) represents the local energy of an image. Typical complex cells in the early visual cortex express phase-invariant nonlinearity. Such responses of the complex cells are well described by the energy model in the form of a squared-sum of simple cell responses as in Eq. (3) (Adelson, Bergen, Adelson, & Bergen, 1983; Emerson, Bergen, & Adelson, 1992; Hyvärinen & Köster, 2006). The phase, $\varphi_{i,m}$, in Eq. (4) represents a relative magnitude of the responses of the real and imaginary components of a complex RF, which are in a quadrature phase. Thus, the phase, $\varphi_{i,m}$, quantifies the local phase structure of natural scenes. The local amplitude and phase extraction in the form of Eqs. (3) and (4) are widely used in image processing (Knutsson & Granlund, 1983; Perona & Malik, 1990; Portilla & Simoncelli, 2000).

2.3. Construction of a collection of Gabor-like RFs

Simple cells in the early visual cortex respond to specific visual features such as particular orientation, spatial frequency and location in the visual field (Hubel & Wiesel, 1968; Movshon et al., 1978). They are also sensitive to spatial phase of a stimulus. It
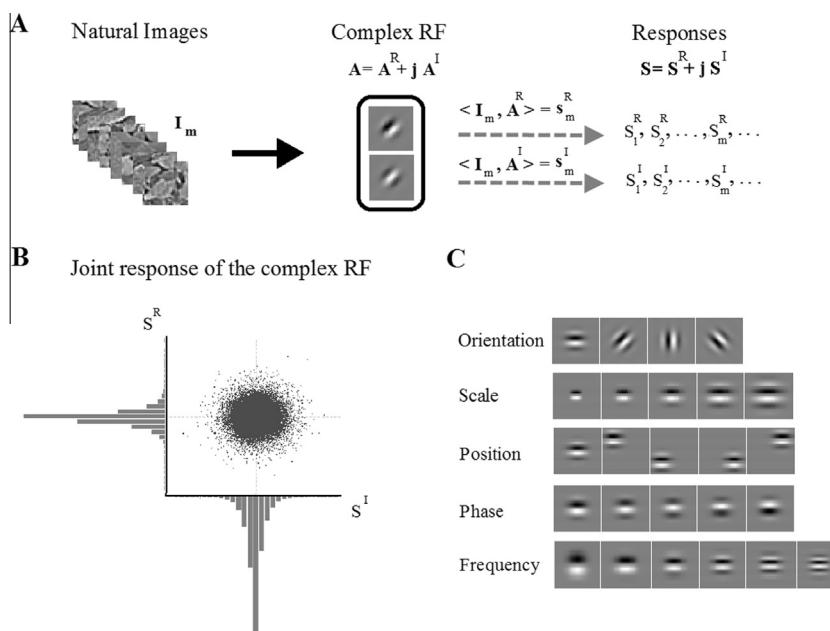
Fig. 1. Schematic illustration of phase extraction of natural scenes by complex RFs. (A) (Left) Samples from 100,000 whitened natural scene patches from Hans van Hateren's repository (van Hateren and van der Schaaf, 1998) (Middle) real and imaginary parts of an example complex RF as a pair of the same Gabor-like RFs that are separated in phase by 90°. (Right) A response sequence of the complex RF to the image patches. A complex response to an image patch is composed of inner products of the real and imaginary parts of the complex RF with the patch. (B) Joint responses of the real and imaginary parts of the complex RF. The abscissa and ordinate are the imaginary and real responses, respectively. The histograms next to the axes represent their sparse marginal distributions. (C) Examples of real components of complex RFs sorted using 5 different features of orientation, scale, position, phase, and frequency. Each row displays a spectrum of a single feature while the other 4 features are fixed. See Methods for the parameters of the RFs.

has been known that the Gabor-like filter is a suitable model of the RF of the simple cells (Daugman, 1985; Jones & Palmer, 1987). In this study, we model RFs of simple cells by using two-dimensional Gabor filters with five parameters that control location, frequency, phase and orientation of the filter. Thus the model covers a variety of RFs from non-oriented blob-like RFs to oriented ridge-like RFs with different scales, and can describe various selectivity of simple cells in the primary visual cortex. A subset of the similar RFs is obtained as a tight frame when representing natural images by sparse coding (Bell & Sejnowski, 1997; Hyvärinen & Hoyer, 2000; Olshausen, 1996). To extract local phase information from natural scene patches, we construct a complex RF (Olshausen et al., 2009; Pollen & Ronner, 1981) defined by a pair of Gabor filters that possess the same scales, orientations, and frequencies, but are separated in phase by $\pi/2$ (An example set of the complex RFs and their power spectra are shown in Figs. 1A and A1 respectively).

We constructed 3000 Gabor-like RFs with a patch size of 32×32 pixels. Each RF is modeled by a two-dimensional Gabor filter as follows:

$$g(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{\left(-\frac{\tilde{x}^2 + \rho^2 \tilde{y}^2}{2\sigma^2}\right)} \cos(2\pi\omega\,\tilde{x} + \psi), \tag{5}$$

where $\tilde{x} = (x - a)\cos(\alpha) - (y - b)\sin(\alpha)$ and $\tilde{y} = (x - a)\sin(\alpha) + (y - b)\cos(\alpha)$. Here the two-dimensional Gabor filter is defined as a sinusoidal grating component with the spatial frequency, ω, and absolute phase, ψ, modulated by an elliptical Gaussian window with the spatial aspect ratio, ρ, and the scaling parameter, σ. The Gabor filter with orientation α is obtained by rotating the coordinate system of the Gabor filter by an angle, $-\alpha$. The location of the peak of the Gaussian envelope is controlled by the position parameters, a and b. For the location of the Gabor filter, we tested five locations by placing the center of the Gaussian window, (a, b), at the center, top-left, top-right, bottom-left, and bottom-right within a patch. The parameter ρ is the spatial aspect ratio that determines the ellipticity of the receptive field. Jone & Palmer suggested the range given by $0.23 < \rho < 0.92$ for a simple cell (Jones & Palmer, 1987). Here we used $\rho = 0.60$ in the construction of the RFs.

The product $\sigma \cdot \omega$ controls the spatial frequency bandwidth of the RF. In neurophysiological literature, it has been reported that the half-response spatial frequency bandwidth, $\kappa = \log_2 \frac{\sigma\omega + \frac{1}{\pi}\sqrt{\frac{\log(2)}{2}}}{\sigma\omega - \frac{1}{\pi}\sqrt{\frac{\log(2)}{2}}}$ (in octaves), of simple cells is specified in the range of $0.4 < \kappa < 2.6$ (Daugman, 1985; De Valois, William Yund, & Hepler, 1982; Petkov & Kruizinga, 1997). Thus, we used a combination of different values of $\sigma = 2^{-n}$ ($n = 0, \ldots, 3$) and $\omega = 1, 1.5, \ldots, 4.5$ to cover this range of half-response spatial frequency bandwidth reported for the empirical data. We examined a half-magnitude orientation bandwidth of $45°$ for all RFs in our analysis. The absolute phase of the grating was divided into five values given by $\psi = -\pi/2, -\pi/4, 0, \pi/4$ and $\pi/2$. In order to eliminate the DC response of the Gabor filters, we subtracted the factor $e^{-\frac{\kappa^2}{2}}$ from the real component of complex RFs (Lee, 1996). Further, we normalized the individual RFs to a unit norm.

Neurophysiological experiments have shown that the half-magnitude orientation bandwidth of simple cells in macaques and rats is ranged from $10°$ to a non-sensitive orientation. However, different median ordination bandwidths about 30 to $60°$ were observed in simple cells (De Valois et al., 1982). Further, the absolute phases of many simple cells are not restricted to 0 or $\frac{\pi}{2}$, i.e., even or odd symmetric. Rather they are distributed uniformly (Burr et al., 1989; Daugman, 1985; Field & Nachmias, 1984; Movshon et al., 1978). We examined a wide range of RFs (in total 3000) chosen from combinations of these five parameters to represent empirically observed RFs of the simple cells.

Fig. 1C displays real components of complex RFs that differ in their orientation, scale, position, phase, or frequency. Each row in Fig. 1C displays a spectrum of a single feature while fixing the other 4 features. An imaginary component of each complex RF is identical to the real component shown in Fig. 1C except that their phase is different by $\pi/2$.

Although even and odd Gabor-like filter functions are orthogonal to each other, we found that many complex RFs represented on discretized pixels within a restricted patch were not orthogonal. Subsequently their real and imaginary responses were correlated (Fig. A4, Panel A). The RFs that could not provide a meaningful definition of the local amplitude and phase were mostly characterized by small scale or low frequency. Thus, we excluded complex RFs if inner products of their real and imaginary components were larger than 0.025. Accordingly, the real and imaginary responses of the selected complex RFs were uncorrelated or had very low correlations (linear correlation coefficients less than 0.043), as expected from uncorrelated inputs (Figs. A2 and A3). Among the 3000 complex RFs, 2010 (67%) were selected after controlling the orthogonality of the real and imaginary components.

2.4. A mixture model of a circular distribution

To model the observed phase distributions, we constructed a mixture model that can account for a wide spectrum of distributions for circular random variables. This model is composed of two von Mises distributions and a uniform circular distribution. The equation is given by

$$p(\varphi|\theta) = \sum_{m=1}^{2} P_m \cdot vM(\varphi|\kappa_m, \mu_m) + P_3 \cdot U(\varphi). \tag{6}$$

All model parameters are represented in a vector, $\theta = (\kappa_1, \kappa_2, \mu_1, \mu_2, P_1, P_2, P_3)$ Here, a von Mises distribution is given by $vM(\varphi|\kappa_m, \mu_m) = \frac{1}{2\pi I_0(\kappa_m)} e^{\kappa_m \cos(\varphi - \mu_m)}$ ($m = 1, 2$), where μ_m is a mean, κ_m ($\kappa_m > 0$) is a concentration parameter, and $I_0(\kappa_m)$ is the modified Bessel function with order zero. The uniform distribution is written as $U(\varphi) = \frac{1}{2\pi}$. The parameters $\{P_1, P_2, P_3\}$ are the mixing parameters for the component distributions. They must satisfy $\sum_{m=1}^{3} P_m = 1$ and $0 \leqslant P_m \leqslant 1$.

The parameters of the mixture model are estimated under the maximum likelihood principle, using the expectation–maximization (EM) algorithm developed in the Supplementary section. In this method, we used analytical solutions to update the parameters which were obtained numerically in Banerjee, Dhillon, Ghosh, Sra, & Ridgeway, 2005. In addition, our approach has advantages in that it can be easily extended to the analysis of multivariate circular random variables using a multivariate von Mises distribution (Mardia, Hughes, Taylor, & Singh, 2008), which is suitable for the analysis of multivariate phase distribution of natural scenes. Fig. 2 displays an example of the phase distribution fitted by this mixture model. In addition, we compared the Laparra's phase model (Laparra et al., 2011) with the mixture model, and found that the modified von Mises model (Laparra's phase model) is often inadequate to capture various bimodal distributions (Fig. A9).

3. Results

3.1. Higher-order statistics in natural scenes is encoded in phase distribution

In this section, we demonstrate circular dependency among responses of two Gabor filters that are close in space, orientation,
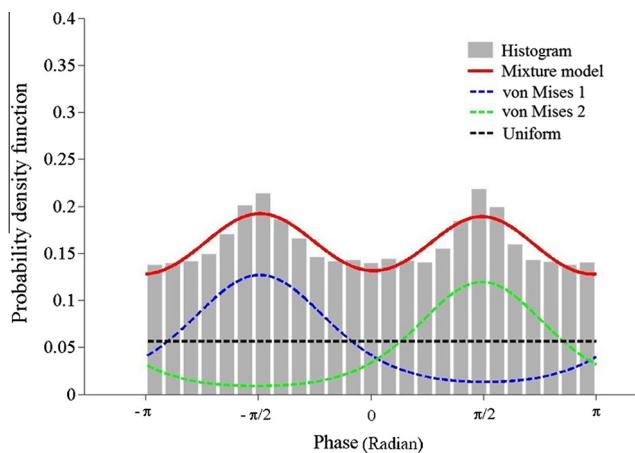
Fig. 2. Fitting a mixture model to circular bimodal data. The data (a gray histogram) is phase responses of the example complex RF. The red curve represents a mixture model fitted to the data according to the EM algorithm developed in Methods. This mixture model is composed of a uniform distribution (black dotted line) and two von Mises distributions (green and blue dotted lines) with different location and concentration parameters.

or scale (Fig. 1C), and show that the dependency can be parsimoniously described by using polar coordinates (Eqs. (3) and (4)).

Fig. 3A shows an example of complex responses from a representative complex RF. The left panel displays responses of a real and imaginary part of the complex RF (the structures of the complex RFs are shown next to the axes). Again, responses of each component are sparse. Furthermore, their joint responses reveal a spherical distribution, which is different from a *diamond-shape* joint distribution expected from the sparse marginal distributions under the assumption of their independence (Simoncelli & Olshausen, 2001; Zetzsche & Rhrbein, 2001). This spherical dependency manifests higher-order dependency contained in natural scenes (Hyvärinen, 2010; Hyvärinen & Hoyer, 2000; Simoncelli & Buccigrossi, 1997; Simoncelli & Olshausen, 2001; Zetzsche & Rhrbein, 2001). The spherical distribution may be concisely described by the polar coordinates (amplitude and phase) using Eqs. (3) and (4). The middle and right panels in Fig. 2A display amplitude and phase distributions of the complex responses. Notably, the phase distribution of this complex RF is uniform. Previous studies (Cadieu & Olshausen, 2008, 2012; Olshausen et al., 2009; Wegmann & Zetzsche, 1990) reported this uniform phase distribution and employed it to parsimoniously describe higher-order statistics in natural scenes.

However, here we report that responses of many complex RFs do not express the uniform phase distribution. Fig. 3B and C illustrate other representative examples of complex responses. The joint responses of the complex RFs are characterized by dependency that is visually similar to the one found in Fig. 3A. However, distinct difference between the two complex RFs can be found in their phase distributions. In these examples, the phase responses are characterized by its bimodal structure rather than the uniform structure. The energy model in Eq. (3) is sufficient to describe the responses of the complex RF with the uniform phase distribution shown in Fig. 3A. This is a classical model of a *phase invariant* complex cell that may receive inputs from simple cells possessing different phase sensitivities (Adelson et al., 1983; Emerson et al., 1992; Hyvärinen & Hoyer, 2000; Hyvärinen & Köster, 2006). To the contrary, the energy model is insufficient to explain the complex RFs with the bimodal phase distributions shown in Fig. 3B and C.

The real and imaginary responses of these complex RFs have equal variances and are uncorrelated. Therefore, these examples

demonstrate that different complex RFs encode higher-order statistics in natural scenes with different phase distributions. In the analysis of various complex RFs, we noticed phase distributions of different complex RFs varied while their amplitude distributions were similar. We thus investigate structures of the phase responses of the complex RFs in greater detail.

3.2. One thirds of complex RFs exhibit symmetric bimodal phase distributions

We analyzed phase responses of 2010 complex RFs whose real and imaginary components were orthogonal and their real and imaginary responses had often equal variances (Fig. A5), using the mixture model defined in Eq. (6). We fitted the mixture model to each complex RF, and obtained the likelihood for the mixture model (Eq. (A1) in Supplementary). We then tested the fitted mixture model against a null hypothesis of a uniform phase distribution by a likelihood ratio test:

$$D = -2 \log \left(\frac{\text{likelihood for a null model}}{\text{likelihood for the mixture model}} \right). \tag{7}$$

Here the likelihood for a null model is $100{,}000/2\pi$. If the data were sampled under the null hypothesis, the test statistic, D, follows a χ^2-distribution with 6 degree of freedom (the difference in the number of free parameters in the two models). To test the goodness-of-fit of the mixture model, we computed p-values of the test statistic, D, using $1 - F_{\chi^2}(D; 6)$, where $F_{\chi^2}(x; 6)$ is a cumulative distribution function of the χ^2-distribution with degree 6. The distribution of the p-values is displayed in Fig. 4A. After multiple comparison correction, the test was rejected in 37% of the complex RFs as shown in an inset of Fig. 4A (the Holm–Bonferroni multiple comparison correction with a family wise error rate (FWER) =0.05). Thus the phase distributions of nearly one third of the complex RFs that we examined significantly deviate from a uniform distribution.

The structured phase response of the complex RFs resulted from interactions between the complex RFs and natural scenes. In order to examine the contribution solely from the phase statistics in natural scenes, we applied the complex RF to surrogate data sets in which phases of reference natural scenes were shuffled while their amplitude components were kept intact (phase shuffled scenes). To construct phase shuffled scenes, the reference natural scenes were Fourier transformed, and then the phase component of each image was shuffled while we kept its amplitude. Finally, the phase shuffled scenes were obtained by computing the inverse Fourier transform of the combined shuffled phase and amplitude components. For the surrogate phase shuffled scenes, the number of complex RFs that exhibited significant bimodal phase structure was significantly reduced to 7% (Fig. 4B). For example, the phase responses of the complex RFs shown in Fig. 3 are all uniformly distributed when they are applied to the phase shuffled natural scenes (Fig. A6). Fig. 4C displays scatter plots of the log-likelihood ratios applied to the natural scenes and the phase shuffled scenes. Red dots in Fig. 4C are complex RFs that exhibit significant bimodal phase distributions when they are applied to the natural scenes whereas their phase responses are uniformly distributed when they are applied to the phase shuffled scenes. Overall, the log-likelihood of the phase shuffled scenes was significantly smaller than those of the natural scenes ($p < 1e-10$, signed-rank test). Similarly, only 5% of the complex RFs exhibited a significant bimodal phase structure when they are applied to white Gaussian signals (Fig. A8). Thus, we confirmed the contribution of the spatial phase information of the natural scenes to the observed bimodal phase structure. In particular, the portion of bimodal phase distributions increased to 54% when complex RFs restricted to zero absolute
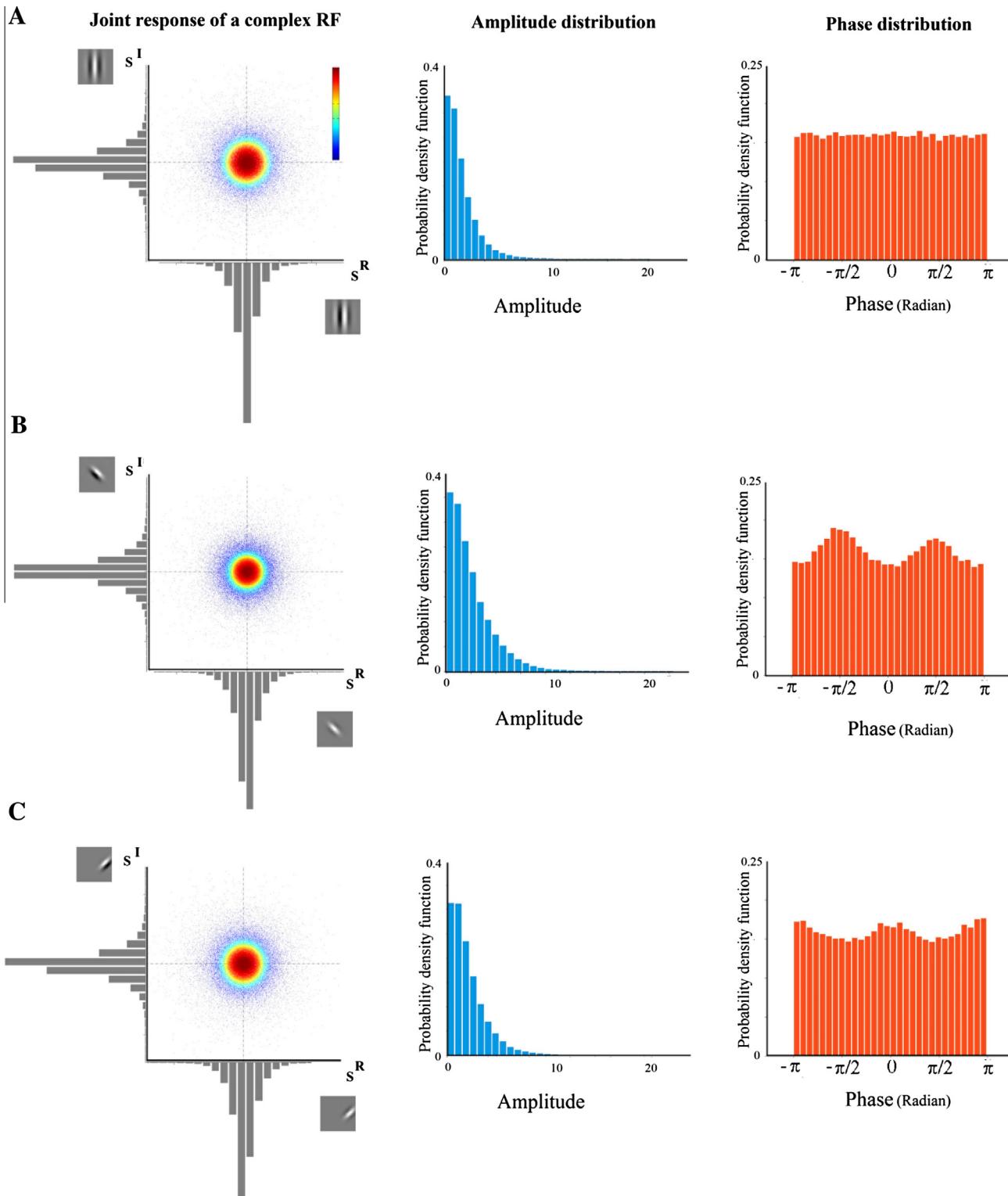
Fig. 3. Responses of three representative complex RFs. (A) A complex RF that exhibits spherical dependency and a uniform phase distribution. Parameters of the complex RFs are $\kappa = 0.75$ ocatave; $\psi = 0$ radian; $\alpha = \frac{\pi}{2}$ radian. The left panel displays a joint distribution of the responses of the complex RF to 100,000 patches randomly selected from natural scenes. The correlation coefficient of the real and imaginary responses is 0.0017. The middle and right panels respectively display distributions of the amplitude and phase variables (i.e., complex representations of the joint responses in the left panel). (B) A complex RF that exhibits a bimodal phase distribution. The presentation format is same as A. Parameters of the complex RFs are $\kappa = 1.65$ ocatave; $\psi = 0$ radian; $\alpha = \frac{3\pi}{4}$ radian. The correlation coefficient is 0.008. (C) Another example of a complex RF that exhibits a bimodal phase distribution. Parameters of the complex RFs are $\kappa = 1.25$ ocatave; $\psi = \frac{3\pi}{4}$ radian; $\alpha = \frac{\pi}{4}$ radian. The correlation coefficient is 0.0009.

phase ($\psi = 0$) and low spatial frequencies ($\omega < 2.5$)) were applied to natural scenes (Fig. A7). Thus, the bimodal phase structure is observed more frequently in the responses of the complex RFs with the low spatial frequencies than the high spatial frequency. Note that the real and imaginary components of the complex RFs with the low spatial frequencies are ridge and edge filters, respectively.
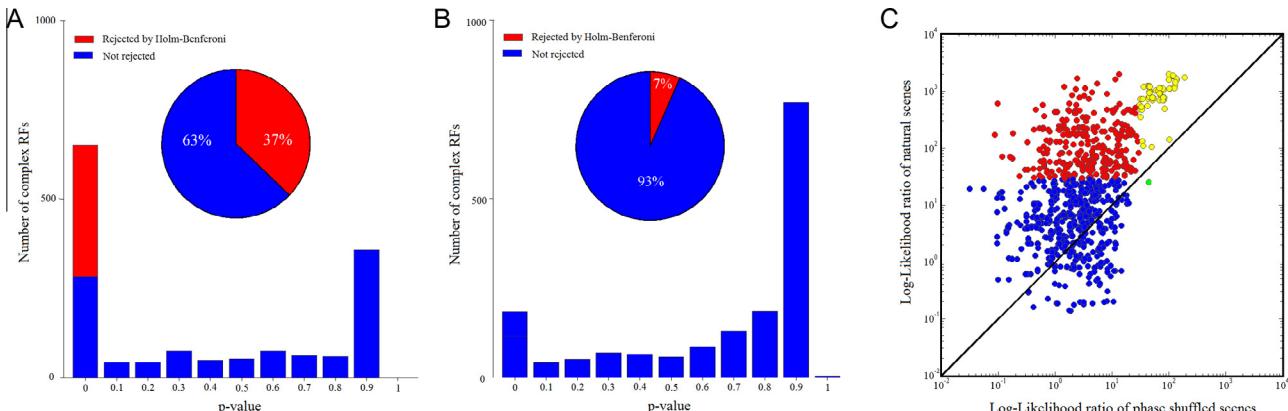
Fig. 4. Responses of one third of the complex RFs express bimodal phase distributions when they are applied to natural scenes. (A) Histogram of *p*-values for likelihood ratio tests when the mixture model is applied to natural scenes. The fraction of the complex RFs for which the null hypothesis of a uniform distribution is rejected by the test is marked in red (37%, *p* < 0.05 with Holm–Bonferroni multiple comparison correction). (B) Histogram of *p*-values for phase shuffled scenes (7% rejected, *p* < 0.05 with Holm–Bonferroni multiple comparison correction). (C) Comparison of the log-likelihood ratios of the mixture model applied to natural scenes (abscissa) with those obtained from phase shuffled scenes (ordinate). Red points represent complex RFs that exhibit significant bimodal phase structure when they are applied to natural scenes while they do not show a significant bimodal structure (the null hypothesis was not rejected) when they are applied to phase shuffled scenes. Yellow points represent complex RFs for which the tests were rejected for both natural and phase shuffled scenes. Blue points are complex RFs that did not show significant bimodal structure in response to both natural scenes and phase shuffled scenes. A green dot is a rare sample that exhibited significant bimodal structure in response to only the phase shuffled scenes.
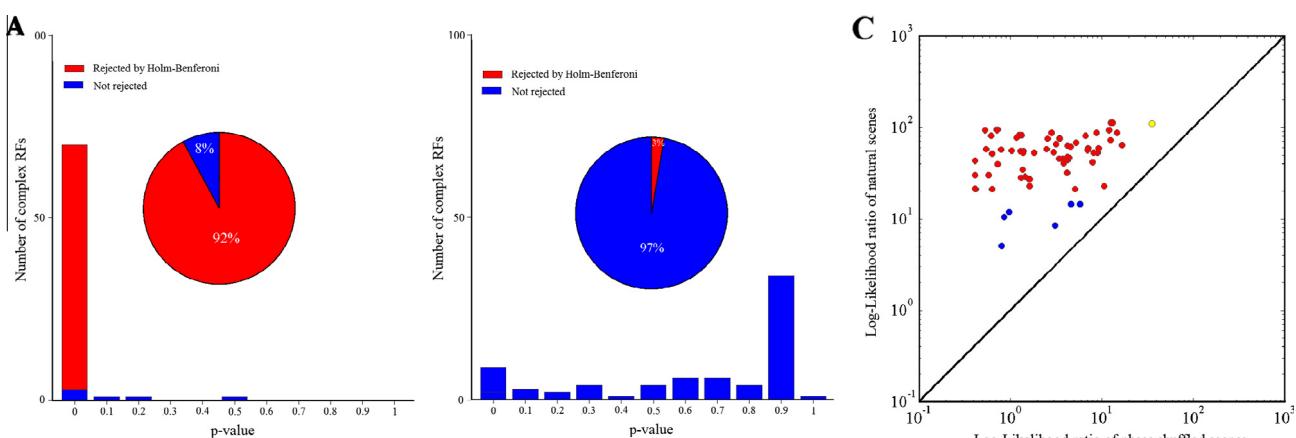
Therefore, the fact that we observe the significant bimodal phase distributions for these low spatial frequency complex RFs indicates unequal responses of the ridge and edge filters when they are applied on natural scene. Further, if we restrict the complex RFs within 1 octave of a frequency bandwidth, significant bimodal phase distributions are observed in 92% of the complex RFs when they are applied to natural scenes whereas this percentage decreases to 3% for phase shuffled scenes (see Fig. 5 and Table 1 for other ranges of frequency bandwidths).

In order to determine structures of the significant mixture models, we investigated the parameters of the fitted mixture models. First, the peak positions (μ_1 or μ_2) of the mixture models are widely distributed in the parameter space (Fig. 6A), but they are localized if we restrict complex RFs to those having the zero absolute phase (Fig. 6A insets, see below for further explanation). Second, the two peaks are separated by π (Fig. 6B). Further, the two estimated concentration parameters κ_1 and κ_2 of the significant mixture model are mostly on a diagonal line (Fig. 6C; correlation coefficient, 0.76). Similarly, the scatter plots of the estimated

mixing parameters, P_1 and P_2, express close values (Fig. 6D; correlation coefficient, −0.27). These results confirm that shapes (i.e., widths and heights) of the two von Mises distributions in the mixture model are similar. We thus conclude that the phase responses of one third of the complex RFs we examined are characterized by a symmetric, bimodal distribution. The inset histograms in Fig. 6A display the results when the complex RFs with the zero absolute phase ($\psi = 0$) are grouped by either low spatial frequencies ($\omega < 2.5$, left inset) or high spatial frequencies ($\omega \geqslant 2.5$, right inset). The peak locations for the low spatial frequencies (left panel) are distributed around $\pm\pi/2$. This result implies that the

Table 1
The percentage of the complex RFs characterized by the bimodal distribution.

Half-response spatial frequency bandwidth (octave)	0.4–0.6	0.6–0.9	0.9–1.2	1.2–1.5	1.5–2.5
Natural scenes	0%	15%	94%	98%	99%
Phase shuffled scenes	0%	0%	<1%	16%	53%

Fig. 5. Phase responses of the restricted complex RFs within 1 octave of a frequency bandwidth. (A) Histogram of *p*-values for likelihood ratio tests of the mixture model when complex RFs are applied to natural scenes. The significant bimodal phase distributions are observed in 92% (*p* < 0.05 with Holm–Bonferroni multiple comparison correction) of the restricted complex RFs. (B) Histogram of *p*-values for phase shuffled scenes (3% rejected, *p* < 0.05 with Holm–Bonferroni multiple comparison correction). (C) Comparison of the log-likelihood ratios of the mixture model for restricted complex RFs applied to natural scenes (abscissa) with those obtained from phase shuffled scenes (ordinate). Color codes are the same as in Fig. 4C.
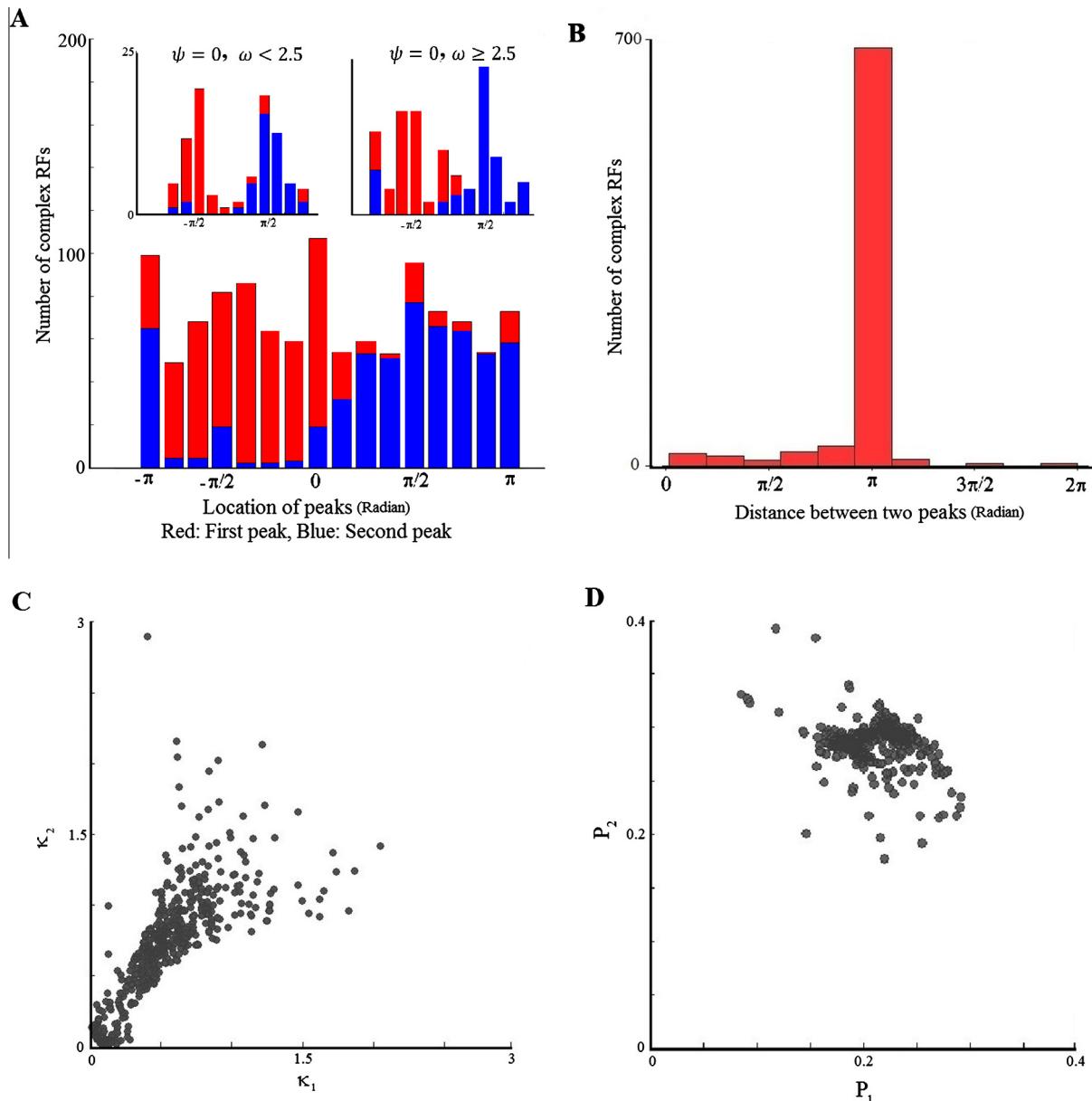
Fig. 6. Symmetric structure of the significant bimodal phase distributions. (A) Stacked histogram of peak positions of the mixture models fitted to phase distributions obtained from natural scenes. The red and blue graphs indicate distributions of the two peaks, μ_1 and μ_2, respectively. The histograms include only the parameters of the mixture models rejected by the likelihood ratio test. Left and right insets in the panel show similar stacked histograms for complex RFs with zero absolute phase ($\psi = 0$), grouped by either low spatial frequencies ($\omega < 2.5$, left inset) or high spatial frequencies ($\omega \geqslant 2.5$, right inset). (B) Distribution of distances between the two estimated peaks, $\mu_1 - \mu_2$. (C) Scatter plots of the estimated concentration parameters, κ_1 and κ_2. (D) Scatter plots of the estimated mixing parameters, P_1 and P_2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

edge structures are more prevalent than the ridge structures in the natural scenes.

Finally, we examine if the mixture model is sufficient to represent the observed phase structure. To this end, we assessed goodness-of-fit of the model using Kolmogorov–Smirnov (KS) statistics. Fig. 7A displays a distribution of the *p*-values of the KS statistics. Only a fraction of samples (15%) marked in black were rejected by the KS test whereas majority of the samples (gray, 85%) expressed phase distributions that are indistinguishable from the mixture model (the Holm–Bonferroni multiple comparison correction with a FWER = 0.05). The goodness-of-fit of the rejected samples may be improved by a more complex phase model or by improving a fitting algorithm. However, here we emphasize that most of the rejected samples exhibited the bimodal structure.

Fig. 7B illustrates the similarity of cumulative distribution functions (CDFs) of a sample data and a fitted mixture model that expressed a low *p*-value (0.041). In summary, most phase distributions of the complex RFs were effectively characterized by the mixture model.

3.3. Higher-order dependency of whitened sparse signals induces uniform and bimodal phase distributions

We confirmed the presence of higher-order dependency in natural scenes that are represented by the uniform or bimodal phase distributions. In this subsection, we demonstrate that simple mixtures of sparse source signals replicate the observed higher-order dependency.

Probabilistic models of neurons in early visual cortex such as independent component analysis (ICA) (Bell & Sejnowski, 1997; Hyvärinen, 2010) and sparse coding model (Olshausen, 1996; Olshausen & Field, 1997) assume that responses are independent and marginally follow super-Gaussian distributions. First, we demonstrate that the independent sparse responses indicate the presence of four peaks in the phase distribution when we consider a complex RF model. Assume that real, s_i^R, and imaginary, s_i^I, components of a complex response $s_i = s_i^R + js_i^I$ in Eq. (1) are independent and sampled from the Laplace distribution, $p(x) = \frac{\beta}{2}e^{-\beta|x|}$ where the scale parameter $\beta > 0$ controls the sparseness of the distribution. Then, the 2-dimensional distribution of s_i is given by

$$p(s_i^R, s_i^I) = \frac{\beta^2}{4}e^{-\beta(|s_i^R|+|s_i^I|)}. \tag{8}$$

Moreover, this probability density function can be rewritten in the polar coordinates as

$$p(r_i, \varphi_i) = \frac{\beta^2}{4}e^{-\beta r_i(|\sin(\varphi_i)|+|\cos(\varphi_i)|)}, \tag{9}$$

where r_i and φ_i are the amplitude and phase of s_i, respectively. By integrating over the amplitude, we obtain the marginal phase distribution,

$$p(\varphi_i) = \int_0^\infty p(r_i, \varphi_i)dr_i = \frac{\beta}{4(|\sin(\varphi_i)| + |\cos(\varphi_i)|)}. \tag{10}$$

This probability density confirms 4 peaks in the distribution of the phase variable. In Fig. 8A, we show that simulated joint responses are characterized by a diamond shape, and the corresponding phase distribution has four peaks. In general, it is expected to observe four peaks if the real and imaginary responses are independent and sparse.

Any departure from the four-peaks distribution implies the presence of higher-order dependency between the real and imaginary components. In order to demonstrate this, we generate dependent responses by superposing sparse source signals. Consider multiple independent complex source signals $V = (v_1, v_2,\ldots, v_N)^T$. Here, $v_i = v_i^R + jv_i^I$ ($i = 1, 2,\ldots, N$) where v_i^R and v_i^I are independent random variables sampled from the Laplace distribution. We construct dependent responses $S = (s_1, s_2,\ldots, s_N)^T$ by linearly mixing v_i by $S = PV$, where P is an $N \times N$ non-singular mixing matrix. We then whiten the joint responses to eliminate the second-order dependency of S and normalize variances of the real and imaginary components to one.

In Fig. 8B, we provide examples of the joint responses in which we mixed 4 source complex signals. In this example, the phase distribution is characterized by the bimodal distribution (Fig. 8B Left, Middle). Another example that exhibits a uniform phase distribution is shown in the right panels of Fig. 8B. Note that the marginals of the mixture signals in these examples are sparse (kurtosis are larger than 3) although the sparseness reduced from that of the original Laplace distribution because the Laplace distribution is not a stable distribution. In order to eliminate the possibility that the observed bimodal and uniform phase distribution is an artifact of the reduced sparseness due to the mixing, we constructed surrogate data sets by removing the higher-order dependency in the data by independently shuffling the real and imaginary components while keeping their marginal distributions. We confirmed that the phase distributions of the surrogate data exhibited the characteristic four peaks (Fig. 8B Bottom panels). Together with the whitening and variance normalization, these simulation results confirm that the uniform and bimodal phase distributions represent the higher-order dependency between the real and imaginary components, and further suggest that sparseness and higher-order
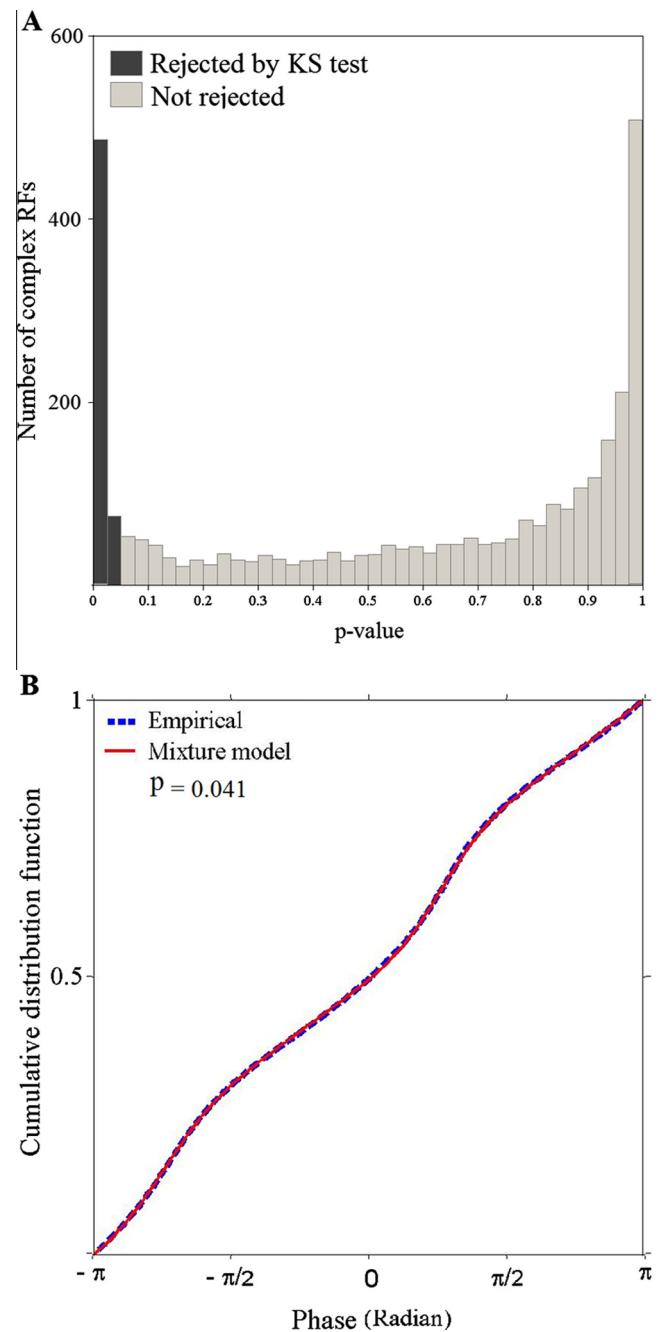

Fig. 7. The goodness-of-fit of the mixture models. (A) A distribution of the *p*-values for the KS statistics from 3000 complex RFs. A bar graph marked in black represents a fraction of samples rejected by the KS test (multiple comparison correction by the Holm–Bonferroni method with a family-wise error rate = 0.05). Phase distributions of these complex RFs were significantly different from the mixture model. (B) A cumulative distribution of an example data that was rejected by the KS test. The *p*-value of the KS statistic for this example is 0.041. Even the rejected sample exhibits the basic bimodal phase structure.

dependency of the natural scenes are represented by not only the uniform but also by the bimodal phase distribution.

4. Discussion

In this study, we examined spatial phase information of natural scenes in the framework of complex representation of images using Gabor-like complex RFs (Eq. (1) and Fig. 1). Typically phase distributions were characterized by a uniform or a spectrum of
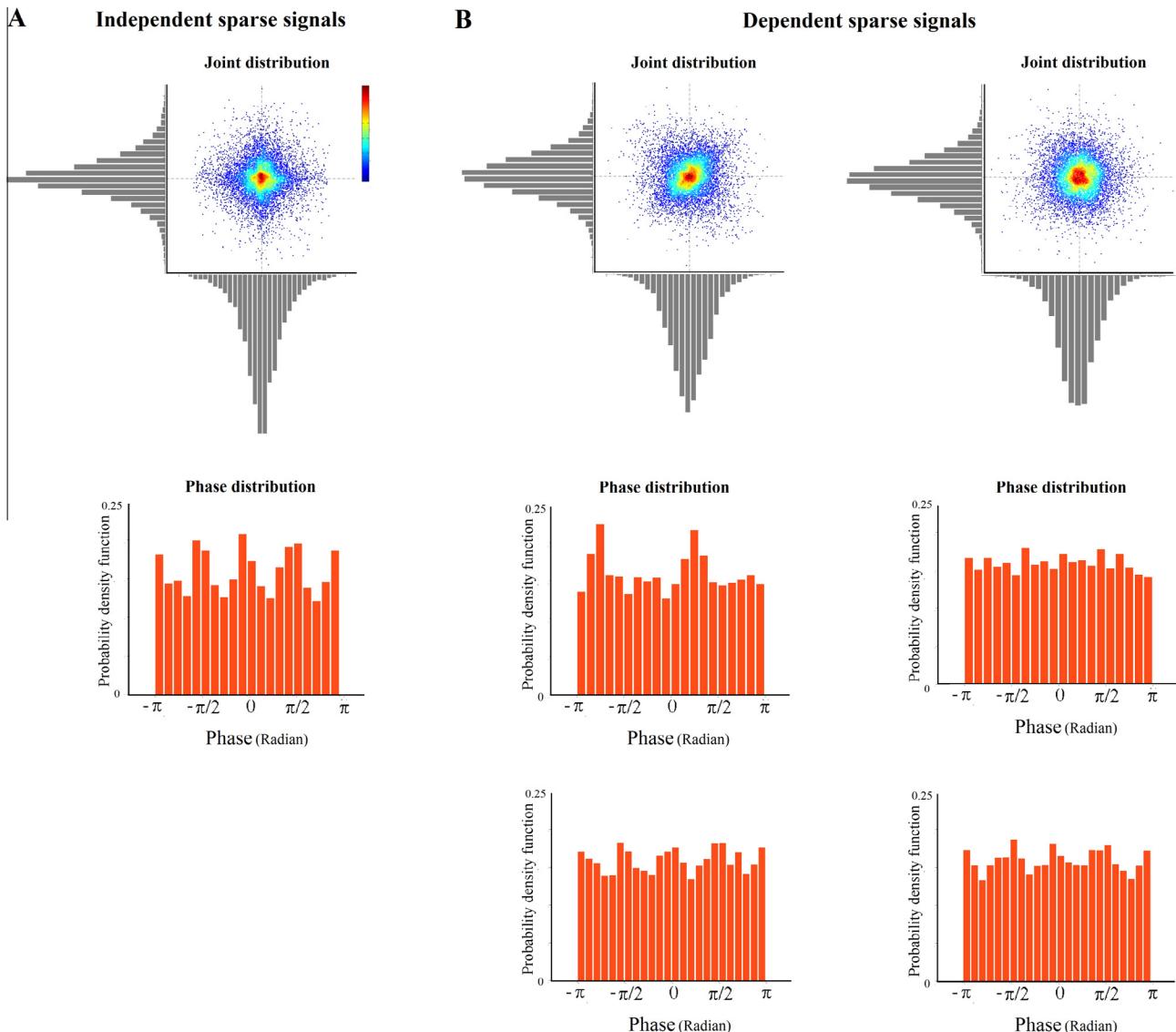
Fig. 8. Higher-order dependency in mixtures of sparse signals can be represented by the bimodal phase distribution. (A) Independent sparse signals. (Top) A joint distribution of the real and imaginary components. Each component is independently sampled from the Laplace distribution with $\beta = 2$. (Bottom) A phase distribution is characterized by four peaks. (B) Dependent sparse signals. (Left column) (Top) The whitened joint distribution of the real and imaginary components. Each component is a linear mixture of the independent sparse signals sampled from the Laplace distribution. These components are whitened to have a zero correlation and unit variance. (Middle) A phase distribution is characterized by the bimodal structure. (Bottom) A phase distribution exhibits four peaks after eliminating its higher-order dependency by randomizing pairs of the real and imaginary components in the data. (Right column) Another example of the mixture signals that exhibits a uniform phase distribution.

bimodal structure (Fig. 3). The structured phase distributions were analyzed by the mixture model of the von-Mises and uniform distributions (Eq. (6)). We estimated the model parameters using the EM algorithm (Fig. 2, see also Supplementary). Based on the likelihood for the model, the distributions of phase responses from one third of the examined complex RFs were determined to be bimodal and symmetrical whereas the phase distributions of the other two thirds of the complex RFs were uniform (Figs. 4, 5 and A8). The simulation study suggests that the uniform and bimodal phase distribution arises from the higher-order dependency of sparse response of complex RFs applied to natural scenes. Thus, the higher-order statistics of natural scene are described by the different types of phase distributions, depending on features of the complex RFs.

Conventional methods for efficient image representation assumed that the phase outputs of filters have uniform distributions (Cadieu & Olshausen, 2008, 2012; Olshausen et al., 2009;

Wegmann & Zetzsche, 1990). Thus our findings make a striking contrast to this previous assumption. For example, Wegmann & Zetzsche developed a technique for encoding images given that amplitude and phase responses are independent, and that phase responses are distributed uniformly (Wegmann & Zetzsche, 1990). They reported uniform phase distributions in the joint outputs of the even and odd filters, but did not report the bimodal phase distributions. An apparent difference might come from the fact that they investigated a set of images different from those used in this study that were more common data sets for natural image studies. In addition, they did not directly test if the phase responses are deviated from a uniform distribution. More generally, none of the previous analyses systematically explored the structure of phase distributions observed in a wide range of biologically plausible filters. By such systematic investigation here we report that a significant fraction of the complex RFs are characterized by the bimodal phase distributions.
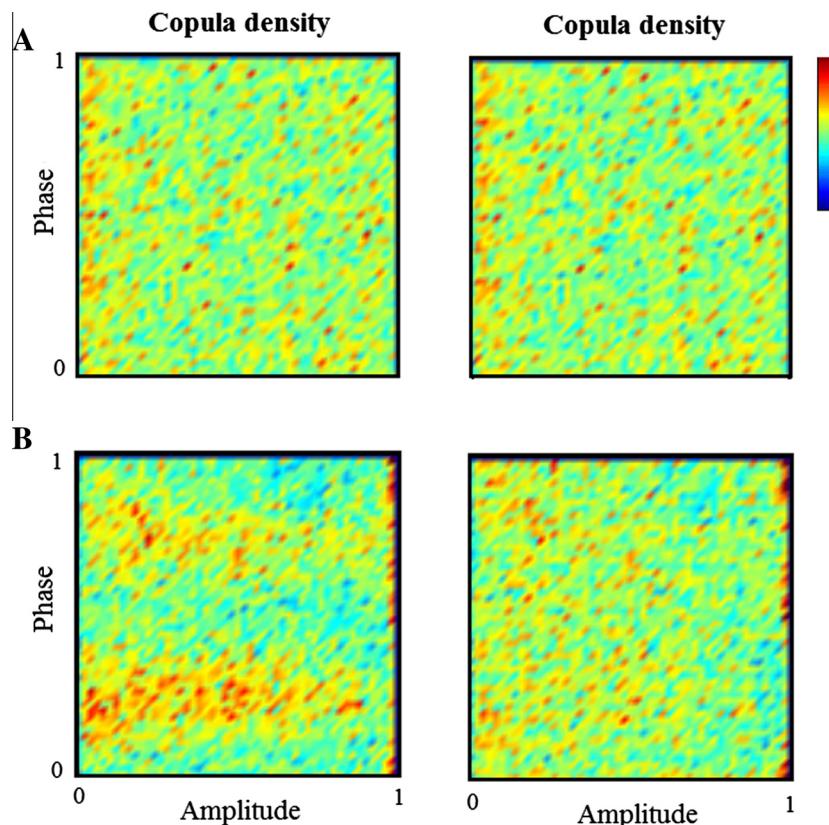
Fig. 9. Residual dependencies that were not explained by a factorial model of amplitude and phase. (A) (Left) A copula density function for the joint amplitude and phase responses of a complex RF shown in Fig. 2A. The density was computed from points whose x- and y-components are given by probability integral transforms of observed amplitudes and phases using a gamma distribution function fitted to the data and a uniform phase distribution function, respectively. (Right) A copula density function as in left panel, whereas the ordinate now represents probability integral transforms of observed phases using a mixture model (Eq. (6)) fitted to the data. (B) Example of copula density functions for a complex RF illustrated in Fig. 2B, which exhibits the bimodal phase distribution. The dependencies were partially removed by using the mixture model as seen in the less structured copula density in the right column than in the left column. However, the dependency was not removed completely.

In this analysis, we carefully controlled the set of complex RFs. Although even and odd Gabor-like filter functions are orthogonal to each other, the filter components represented on discredited pixels were sometimes not orthogonal. We thus excluded complex RFs if inner products of the real and imaginary components of the complex RFs are larger than 0.025. This threshold value was chosen rather arbitrarily. However, the observed bimodal phase distributions for the selected RFs are not the artifact of their small correlations. The panel A of Fig. A4 displays the relation between inner products of real and imaginary RFs (x-axis) and correlation coefficients of real and imaginary responses (y-axis), for natural and phase-shuffled scenes. Within ± 0.025 of the inner products, the same level of correlation coefficients (y-axis) were observed for responses to both natural and phases shuffled scenes. However the bimodal structures were not observed in the responses to the phase shuffled scenes (Fig. 4C). Therefore higher absolute values of the correlation coefficients of real and imaginary responses do not necessarily cause bimodal phase distributions. A more illustrative example may be found in Fig. 3. In these examples, inner products of the components of the exemplary complex RF that exhibited bimodal phase distribution (Fig. 3C) is smaller than the one that exhibited uniform phase distribution (Fig. 3A). Finally, the simulation study (Fig. 8) demonstrated that the bimodal phase distributions appear when real and imaginary responses are precisely uncorrelated.

Our study revealed the importance of the structured phase distributions in characterizing higher-order statistical structure of natural images, and encourages researchers to construct extended statistical models that exploit the phase information. However

existing models of such kind are scarce. In recent Bayesian modeling approaches, Olshausen and colleagues modeled nonlinear patterns in the sequence of natural scenes by introducing complex representations of the images (Cadieu & Olshausen, 2008, 2012; Olshausen et al., 2009). In this study, they used a prior for temporal variation of the image phase, but did not consider a prior for spatial phase (i.e., uniform phase prior). Laparra et al. proposed an extension of complex independent component analysis for modeling natural scenes using a bimodal phase distribution as a prior (Laparra et al., 2011). While their basic assumption was supported by the current study, their model of a phase distribution lacks the uniform component. Our results indicate that these Bayesian models can better describe regularities in natural scenes using a more appropriate phase prior, namely the proposed mixture model of a bimodal phase distribution.

These Bayesian models and our analysis assume that response dependency between the real and complex parts of RFs is factorial in the space of amplitude and phase. This assumption needs to be tested empirically. Dependency of two random variables is described by a copula function (Nelsen, 1999). A copula function and the corresponding copula density function are a joint CDF and density of uniform random variables to which any random variables can be transformed using its own marginal CDF. The copula functions allow us to visualize dependency structure regardless of their marginal distributions in a standardized format. Fig. 9 displays copula density functions of amplitude and phase responses from the complex RFs shown in Fig. 3. In each column, the abscissa represents probability integral transforms of amplitudes using a gamma distribution fitted to the data. The ordinate represents

probability integral transforms of phases using a uniform distribution (Left column) or a mixture model (Right column). The copula density functions should express no structure if the parametric marginal models faithfully represent the amplitude and phase distributions and if they are independent. Any structure in the copula density indicates incorrect marginal models and/or dependency in the two random variables. Visual inspection of the copula densities in the left and right columns reveals that the dependency is reduced by using the mixture model when the underlying phase distributions are non-uniform (Fig. 9B). However, structured copula density remains in the right column of Fig. 9B. Thus the factorial code in amplitude and phase with the proposed marginal models may be only approximation, and future hierarchical models should describe this dependency for explaining higher-order regularities in natural scenes.

Neurophysiological implications of the structured phase distributions are the presence of phase-sensitive cells beyond V1 simple cells. While a predominant view on the complex cells and cells in the higher visual area is that they are phase invariant (Albrecht et al., 1980;Movshon et al., 1978; Skottun et al., 1991), recent studies report evidence of phase sensitive cells (Crowder et al., 2007; Felsen et al., 2005; Hietanen et al., 2013; Mechler & Ringach, 2002; Mechler et al., 2002). Responses of phase invariant complex cells are described by an energy model (Eq. (3)) (Adelson et al., 1983; Hyvärinen & Hoyer, 2000; Hyvärinen & Köster, 2006; Pollen, Gaska, & Jacobson, 1988). Indeed, the two thirds of the complex RFs examined in this study exhibited uniform phase distributions (see Figs. 4 and A6). However, we also demonstrated the presence of bimodal phase distributions in one third of the complex RFs. According to the efficient coding, the observed complex RFs with the bimodal phase distribution may predict the presence of phase sensitive cells beyond V1 simple cells. Thus, the results suggest that we need an extended model beyond the energy model that utilizes the phase distribution to explain non-classical *phase sensitive* complex cells recently found in early visual cortices (Crowder et al., 2007; Felsen et al., 2005; Hietanen et al., 2013; Mechler & Ringach, 2002; Mechler et al., 2002).

In summary, we found that a spectrum of phase distributions is required to characterize spatial phase information of natural scenes. According to the efficient coding hypothesis, our result suggests the need of both phase invariant and phase sensitive complex cells in visual systems in order to describe the regularities of natural scenes.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.visres.2015.06.009.

References

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. *Sensory Communication*, 217–234.

Bell, A. J., & Sejnowski, T. J. (1997). The "independent components" of natural scenes are edge filters. *Vision Research, 37*(23), 3327–3338.

Bennett, P. J., & Banks, M. S. (1991). The effects of contrast, spatial scale, and orientation on foveal and peripheral phase discrimination. *Vision Research, 31*(10), 1759–1786.

Burr, D. C. (1980). Sensitivity to spatial phase. *Vision Research, 20*(5), 391–396.

Burr, D. C., Morrone, M. C., & Spinelli, D. (1989). Evidence for edge and bar detectors in human vision. *Vision Research, 29*(4), 419–431.

Cadieu, C. F., & Olshausen, B. A. (2008). Learning transformational invariants from natural movies. In: *NIPS* (pp. 209–216).

Cadieu, C. F., & Olshausen, B. A. (2012). Learning intermediate-level representations of form and motion from natural movies. *Neural Computation, 24*(4), 827–866.

Crowder, N. A., Van Kleef, J., Dreher, B., & Ibbotson, M. R. (2007). Complex cells increase their phase sensitivity at low contrasts and following adaptation. *Journal of Neurophysiology, 98*(3), 1155–1166.

Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A, 2*(7), 1160–1169.

De Valois, R. L., William Yund, E., & Hepler, N. (1982). The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research, 22*(5), 531–544.

DeAngelis, G. C., Ghose, G. M., Ohzawa, I., & Freeman, R. D. (1999). Functional micro-organization of primary visual cortex: Receptive field analysis of nearby neurons. *The Journal of Neuroscience, 19*(10), 4046–4064.

Emerson, R. C., Bergen, J. R., & Adelson, E. H. (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Research, 32*(2), 203–218.

Felsen, G., Touryan, J., Han, F., & Dan, Y. (2005). Cortical sensitivity to visual features in natural scenes. *PLoS Biology, 3*(10), e342.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, 4*(12), 2379–2394.

Field, D. (1994). What is the goal of sensory coding? *Neural Computation, 6*(4), 559–601.

Field, D. J., & Nachmias, J. (1984). Phase reversal discrimination. *Vision Research, 24*(4), 333–340.

Freeman, J., Ziemba, C. M., Heeger, D. J., Simoncelli, E. P., & Movshon, J. A. (2013). A functional and perceptual signature of the second visual area in primates *Nature Neuroscience, 16*(7), 974–981.

Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology, 59*, 167–192.

Hansen, B. C., & Hess, R. F. (2007). Structural sparseness and spatial phase alignment in natural scenes. *JOSA A, 24*(7), 1873–1885.

Henriksson, L., Hyvärinen, A., & Vanni, S. (2009). Representation of cross-frequency spatial phase relationships in human visual cortex. *The Journal of Neuroscience, 29*(45), 14342–14351.

Hietanen, M. A., Cloherty, S. L., van Kleef, J. P., Wang, C., Dreher, B., & Ibbotson, M. R. (2013). Phase sensitivity of complex cells in primary visual cortex. *Neuroscience, 237*, 19–28.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology, 195*(1), 215–243.

Hyvärinen, A. (2010). Statistical models of natural images and cortical visual representation. *Topics in Cognitive Science, 2*(2), 251–264.

Hyvärinen, A., & Köster, U. (2006). FastISA: A fast fixed-point algorithm for independent subspace analysis. In: *ESANN* (pp. 371–376).

Hyvärinen, A., & Hoyer, P. (2000). Emergence of phase-and shift-invariant features by decomposition of natural images into independent feature subspaces. *Neural Computation, 12*(7), 1705–1720.

Hyvärinen, A., Hoyer, P. O., & Inki, M. (2001). Topographic independent component analysis. *Neural Computation, 13*(7), 1527–1558.

Hyvärinen, A., Hurri, J., & Hoyer, P. O. (2009). *Natural image statistics: A probabilistic approach to early computational vision* (Vol. 39). Springer.

Jones, J. P., & Palmer, L. A. (1987). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *Methods, 5*, 24.

Karklin, Y., & Lewicki, M. S. (2003). Learning higher-order structures in natural images. *Network: Computation in Neural Systems, 14*(3), 483–499.

Karklin, Y., & Lewicki, M. S. (2005). A hierarchical Bayesian model for learning nonlinear statistical regularities in nonstationary natural signals. *Neural Computation, 17*(2), 397–423.

Karklin, Y., & Lewicki, M. S. (2006). Is early vision optimized for extracting higher-order dependencies? *Advances in Neural Information Processing Systems, 18*, 635.

Knutsson, H., & Granlund, G. H. (1983). Texture analysis using two-dimensional quadrature filters. In: *IEEE Computer Society Workshop on Computer Architecture for Pattern Analysis and Image Database Management-CAPAIDM: Pasadena*.

Kovesi, P. (1999). Image features from phase congruency *Videre: Journal of Computer Vision Research, 1*(3), 1–26.

Laparra, V., Gutmann, M. U., Malo, J., & Hyvärinen, A. (2011). Complex-valued independent component analysis of natural images. In: *Artificial Neural Networks and Machine Learning–ICANN 2011* (pp. 213–220).

Lee, T. S. (1996). Image representation using 2D Gabor wavelets. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 18*(10), 959–971.

Lyu, S., & Simoncelli, E. P. (2008). Nonlinear image representation using divisive normalization. In: *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on* (pp. 1–8). IEEE.

Adelson, E. H., Bergen, J. R., Adelson, E. H., & Bergen, J. R. (1983). Spatio-temporal energy models for the Perception of Motion. *Journal of the Optical Society of America*.

Albrecht, D. G., De Valois, R. L., & Thorell, L. G. (1980). Visual cortical neurons: Are bars or gratings the optimal stimuli? *Science, 207*(4426), 88–90.

Badcock, D. R. (1984). How do we discriminate relative spatial phase? *Vision Research, 24*(12), 1847–1857.

Banerjee, A., Dhillon, I. S., Ghosh, J., Sra, S., & Ridgeway, G. (2005). Clustering on the unit hypersphere using von Mises-Fisher distributions. *Journal of Machine Learning Research, 6*(9).

Mardia, K. V., Hughes, G., Taylor, C. C., & Singh, H. (2008). A multivariate von Mises distribution with applications to bioinformatics. *Canadian Journal of Statistics, 36* (1), 99–109.

Mechler, F., Reich, D. S., & Victor, J. D. (2002). Detection and discrimination of relative spatial phase by V1 neurons. *The Journal of Neuroscience, 22*(14), 6129–6157.

Mechler, F., & Ringach, D. L. (2002). On the classification of simple and complex cells. *Vision Research, 42*(8), 1017–1033.

Morrone, M. C., & Burr, D. C. (1988). Feature detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society of London. Series B, Biological Sciences,* 221–245.

Morrone, M. C., Ross, J., Burr, D. C., & Owens, R. (1986). Mach bands are phase dependent. *Nature, 324*(6094), 250–253.

Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978). Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex. *The Journal of Physiology, 283*(1), 101–120.

Nelsen, R. B. (1999). *An introduction to copulas*. Springer.

Olshausen, B. A. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature, 381*(6583), 607–609.

Olshausen, B. A., Cadieu, C. F., & Warland, D. K. (2009). Learning real and complex overcomplete representations from the statistics of natural images. In: *SPIE Optical Engineering + Applications* (p. 74460S–74460S). International Society for Optics and Photonics.

Olshausen, B. A., & Field, D. J. (1996). Natural image statistics and efficient coding*. *Network: Computation in Neural Systems, 7*(2), 333–339.

Olshausen, B. A., & Field, D. J. (1997). Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision Research, 37*(23), 3311–3325.

Olshausen, B. A., & Field, D. J. (2004). Sparse coding of sensory inputs *Current Opinion in Neurobiology, 14*(4), 481–487.

Oppenheim, A. V., & Lim, J. S. (1981). The importance of phase in signals. *Proceedings of the IEEE, 69*(5), 529–541.

Perna, A., Tosetti, M., Montanaro, D., & Morrone, M. C. (2008). BOLD response to spatial phase congruency in human brain. *Journal of Vision, 8*(10), 15.

Perona, P., & Malik, J. (1990). Detecting and localizing edges composed of steps, peaks and roofs. In: *Computer Vision, 1990. Proceedings, Third International Conference on* (pp. 52–57).

Petkov, N., & Kruizinga, P. (1997). Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: Bar and grating cells. *Biological Cybernetics, 76*(2), 83–96.

Pollen, D. A., Gaska, J. P., & Jacobson, L. D. (1988). Responses of simple and complex cells to compound sine-wave gratings. *Vision Research, 28*(1), 25–39.

Pollen, D. A., & Ronner, S. F. (1981). Phase relationships between adjacent simple cells in the visual cortex. *Science, 212*(4501), 1409–1411.

Portilla, J., & Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision, 40*(1), 49–70.

Sasaki, H., Gutmann, M. U., Shouno, H., & Hyvärinen, A. (2013). Correlated topographic analysis: Estimating an ordering of correlated components. *Machine Learning, 92*(2–3), 285–317.

Simoncelli, E. P., & Buccigrossi, R. W. (1997). Embedded wavelet image compression based on a joint probability model. In: *Image Processing, 1997. Proceedings, International Conference on* (Vol. 1, pp. 640–643). IEEE.

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience, 24*(1), 1193–1216.

Sinz, F. H., & Bethge, M. (2010). Lp-nested symmetric distributions. *Journal of Machine Learning Research, 11*, 3409–3451.

Skottun, B. C., De Valois, R. L., Grosof, D. H., Movshon, J. A., Albrecht, D. G., & Bonds, A. B. (1991). Classifying simple and complex cells on the basis of response modulation. *Vision Research, 31*(7), 1078–1086.

Thomson, M. G. (1999). Higher-order structure in natural scenes. *Journal of the Optical Society of America A, 16*(7), 1549–1553.

Touryan, J., Felsen, G., & Dan, Y. (2005). Spatial structure of complex cell receptive fields measured with natural images. *Neuron, 45*(5), 781–791.

Van der Schaaf, A. van, & van Hateren, J. van (1996). Modelling the power spectra of natural images: Statistics and information. *Vision Research, 36*(17), 2759–2770.

Van Hateren, J. H., & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 265*(1394), 359–366.

Wainwright, M. J., Schwartz, O., & Simoncelli, E. P. (2002). 10 Natural Image Statistics and Divisive Normalization. *Probabilistic Models of the Brain, 203*.

Wang, Z., & Simoncelli, E. P. (2003). Local phase coherence and the perception of blur. In: *Advances in neural information processing systems*.

Wegmann, B., & Zetzsche, C. (1990). Visual-system-based polar quantization of local amplitude and local phase of orientation filter outputs. In B. E. Rogowitz & J. P. Allebach (Eds.), (pp. 306–317).

Wichmann, F. A., Braun, D. I., & Gegenfurtner, K. R. (2006). Phase noise and classification of natural images. *Vision Research, 46*, 1520–1529.

Zetzsche, C., Krieger, G., & Wegmann, B. (1999). The atoms of vision: Cartesian or polar? *Journal of the Optical Society of America A, 16*(7), 1554–1565.

Zetzsche, C., & Rhrbein, F. (2001). Nonlinear and extra-classical receptive field properties and the statistics of natural scenes. *Network: Computation in Neural Systems, 12*(3), 331–350.