

# CLEAN and WIPE

A. Lannes, E. Anterrieu, and P. Maréchal

Laboratoire d'Astrophysique (and CERFACS), Observatoire Midi-Pyrénées, 14, Avenue Edouard Belin, F-31400 Toulouse, France

Received March 14; accepted July 22, 1996

**Abstract.** CLEAN is a matching pursuit process which has two weak points: in situations of astrophysical interest, the “clean map” is not stable in the corresponding object representation space, and the “atoms” of this representation (translated versions of the “clean beam”) are not well suited for reconstructing the boundaries of the structuring entities of the object at the related resolution level. As a result, CLEAN must be interrupted before the best possible fit is reached. How does WIPE wipe CLEAN clean? First, by introducing a global regularization principle based on the concept of resolution; and second, by conducting the matching pursuit process at the level of the high-resolution basis functions of the object space.

**Key words:** methods: data analysis; numerical — techniques: image processing; interferometric

## 1. Introduction

WIPE is a Fourier synthesis method recently developed in radio imaging and optical interferometry (Lannes et al. 1994, 1996). The name of WIPE is associated with that of CLEAN, a deconvolution method intensively used in astronomy (Högbom 1974; Schwarz 1978). (WIPE can also be used as a deconvolution technique.) As a matter of fact, WIPE was devised, quite independently, on the grounds of well known properties in harmonic analysis and band-limited interpolation (Lannes et al. 1987a). There also exists a probabilistic interpretation of this method (Maréchal & Lannes 1997). Nevertheless, to some extent, WIPE can equally well be regarded as an updated version of CLEAN. This paper is aimed precisely at clarifying the relationship between these two methods.

In Sect. 2, we first show that CLEAN lies in the family of “matching pursuit” techniques (Mallat & Zhang 1993). In other words, the principle of the “iterative harmonic analysis” of CLEAN, as it is exhibited in Schwarz (1978), can be regarded as a special case of a more general approach. We then present the reconstruction methods that

can be implemented for solving the related optimization problems, first, without any control of the propagation of errors (Sect. 3), and afterwards, with a complete control of the stability parameters (Sect. 4). The first weak point of CLEAN can thus be exhibited: in situations of astrophysical interest (for example, when observing extended complex objects with dilute apertures), the matching process of CLEAN is ill-conditioned. As shown in Sect. 5, the regularization principle of WIPE, which can be applied to CLEAN as it is, remedies this lack of robustness. Unfortunately, CLEAN has another weak point: the basis functions used in the matching pursuit process, the ‘clean beams,’ are not well suited for reconstructing the boundaries of the structuring entities of the images at the selected resolution level. As a result, CLEAN must be interrupted before the best possible fit is reached. As indicated in Sect. 5, the basic version of WIPE overcomes this difficulty in a simple and efficient manner. The multiresolution extension of WIPE, underlying the main aspects of this paper, reinforces the arguments justifying this methodological choice.

The guiding idea of our analysis is based on the theoretical framework presented in Sect. 1.1. (The reader who is not familiar with the related basic notions is invited to consult Appendices 1 to 4 of Lannes et al. 1987a). As CLEAN was first used as a Fourier synthesis technique, it was natural to compare the principles of CLEAN and WIPE in this context. Such a comparative study can only be started from a common statement of the problem. The corresponding formulation, which makes the theoretical framework more attractive, is presented in Sect. 1.2. We then also specify the general conditions of the numerical simulations in support of our analysis.

### 1.1. Theoretical framework

In many inverse problems, the “reconstructed object” is described in an “object representation space”  $E$  generated by  $m$  vectors  $g_k$  selected among a family  $\mathcal{E}_o$  of  $M$  vectors  $g_1, g_2, \dots, g_M$ . The latter may be regarded as the “atoms” of the object representation under consideration. The linear space generated by all the vectors of  $\mathcal{E}_o$ ,  $E_o$ , is a subspace of some Euclidean space: the “object space”  $H_o$ .

The “data vector”  $\psi_d$  lies in another Euclidean space: the “data space”  $K_d$ . (When the data are complex quantities, it is always possible to work in the real linear space underlying the complex linear space directly involved in the analysis.) To a first approximation,  $\psi_d$  is related to the object (or image) to be reconstructed via a linear operator  $A$  from  $H_o$  into  $K_d$ . The basic problem is to choose the vectors  $g_k$  in  $\mathcal{E}_o$ , and thereby to construct the object representation space  $E$ . The solutions are then obtained by minimizing on  $E$  the quadratic functional:

$$q : H_o \rightarrow \mathbb{R}, \quad q(\phi) \triangleq \|\psi_d - A\phi\|_d^2. \quad (1)$$

Clearly  $\|\cdot\|_d = (\cdot | \cdot)_d^{1/2}$  is the norm on  $K_d$ ; the scalar products on  $K_o$  and  $K_d$ , as well as the norms, will be distinguished by the subscripts  $o$  and  $d$ , respectively.

Let  $F$  be the “image” of  $E$  by  $A$  (the space of the  $A\phi$ 's,  $\phi$  spanning  $E$ ),  $A_E$  be the operator from  $E$  onto  $F$  induced by  $A$ , and  $\psi_F$  be the projection of  $\psi_d$  onto  $F$  (see Fig. 1). The vectors  $\phi$  minimizing  $q$  on  $E$ , the solutions of the problem, are such that  $A_E\phi = \psi_F$ . They are identical up to a vector lying in the kernel of  $A_E$ . Note that  $\ker A_E = E \cap \ker A$ . (By definition, the kernel of  $A$ ,  $\ker A$ , also referred to as the null space of  $A$ , is the set of vectors  $\phi$  such that  $A\phi = 0$ ). The condition  $\dim E \leq \dim K_d$  is a necessary condition for  $\ker A_E$  to be reduced to  $\{0\}$ .

As  $\psi_d - \psi_F$  is orthogonal to  $F$ , the solutions  $\phi$  of the problem are characterized by the property:

$$\forall \varphi \in E, \quad (A\varphi | \psi_d - A\phi)_d = 0 \quad (\phi \in E).$$

On denoting by  $A^*$  the adjoint of  $A$ , this property can also be written in the form

$$\forall \varphi \in E, \quad (\varphi | r)_o = 0 \quad r \triangleq A^*(\psi_d - A\phi) \quad (\phi \in E),$$

where  $r$  is regarded as a residue. The solutions  $\phi$  are therefore characterized by the fact that the corresponding residue is orthogonal to  $E$ , i.e.,

$$\forall g_k \in \mathcal{E}, \quad (g_k | r)_o = 0. \quad (2)$$

This condition is of course equivalent to  $P_E r = 0$ , where  $P_E$  is the projection (operator) of  $H_o$  onto  $E$ . The solutions of the problem are therefore the solutions of the equation on  $E$ :

$$P_E A^* A_E \phi = P_E A^* \psi_d. \quad (3)$$

For any  $\varphi \in E$  and any  $\psi \in K_d$ , we have

$$(A\varphi | \psi)_d = (\varphi | A^*\psi)_o = (\varphi | P_E A^*\psi)_o,$$

hence  $A_E^* = P_E A^*$ . This explicitly shows that Eq. (3) is the “normal equation”:

$$A_E^* A_E \phi = A_E^* \psi_d. \quad (4)$$

When the problem is well posed,  $A_E$  is a one-to-one map ( $\ker A_E = \{0\}$ ); the solution is then unique: there

exists only one vector  $\phi \in E$  such that  $A_E\phi = \psi_F$ . This vector,  $\phi_E$ , is said to be the least-squares solution of the equation  $A_E\phi = \psi_F$ .

In this case, let  $\delta\psi_F$  be a variation of  $\psi_F$  in  $F$ , and  $\delta\phi_E$  be the corresponding variation of  $\phi_E$  (see Fig. 1). As shown in Appendix 1, the robustness of the reconstruction process is then governed by the inequality

$$\frac{\|\delta\phi_E\|_o}{\|\phi_E\|_o} \leq \kappa_E \frac{\|\delta\psi_F\|_d}{\|\psi_F\|_d} \quad (5)$$

where  $\kappa_E$  is the condition number of  $A_E$ :

$$\kappa_E \triangleq \frac{\sqrt{\mu'}}{\sqrt{\mu}}; \quad (6)$$

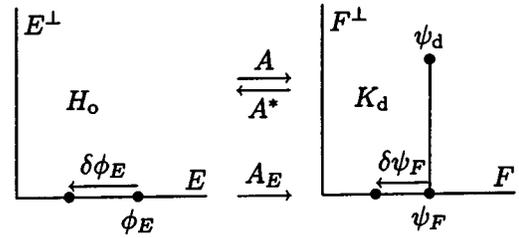
$\mu$  and  $\mu'$  respectively denote the greatest lower bound and the least upper bound of  $\|A_E\phi\|_d^2$  for the  $\phi$ 's with norm unity in  $E$ :

$$\mu \triangleq \inf_{\|\phi\|_o=1} \|A_E\phi\|_d^2 \quad \mu' \triangleq \sup_{\|\phi\|_o=1} \|A_E\phi\|_d^2. \quad (7)$$

The closer to 1 is the condition number, the easier and the more robust is the reconstruction process. As

$$\|A_E\phi\|_d^2 = (\phi | A_E^* A_E \phi)_o,$$

$\mu$  and  $\mu'$  are the smallest and largest eigenvalues of  $A_E^* A_E$ , respectively.



**Fig. 1.** Uniqueness of the solution and robustness of the reconstruction process. Operator  $A$  is an operator from the object space  $H_o$  into the data space  $K_d$ . The object representation space  $E$  is a particular subspace of  $H_o$  (see text). The image of  $E$  by  $A$ , the range of  $A_E$ , is denoted by  $F$ . In this Euclidean representation,  $\psi_F$  is the projection of the data vector  $\psi_d$  onto  $F$ . The inverse problem must be stated so that  $A_E$  is a one-to-one map from  $E$  onto  $F$ , the condition number  $\kappa_E$  having a reasonable value (see Eq. (5))

In the analysis of the different methods that can be devised and implemented for solving the problem, three main aspects must be examined:

- 1) the precise definition of  $\psi_d$ ,  $A$  and  $E$ ;
- 2) the selected minimization technique;
- 3) the robustness of the reconstruction process.

The first point is essential for the interpretation of the results. It is related to the statement of the problem and, if need be, to its “regularization;”  $\psi_d$  is not necessarily the vector of the experimental data. In such a case, the definition of  $A$  must of course take into account the corresponding manipulations. As regards  $E$ , which may also be involved in the regularization of the problem, it is important to note that this space may be constructed, in a global manner or step by step, interactively or automatically. As suggested by the statement of point 2, many different techniques can be used for minimizing  $q$  on  $E$ ; some of these are certainly more efficient than others, but in this case, the choice is not crucial. The last point concerns the study of the propagation of errors. The part played by inequality (5) in the development of the corresponding error analysis shows that a good reconstruction procedure must also provide, in particular, the condition number  $\kappa_E$ .

1.2. Formulation of the problems of Fourier synthesis in astronomy

In the problems of Fourier synthesis encountered in astronomy, the “object function” of interest,  $\phi_o$ , is a real-valued function of an angular position variable  $\xi = (\xi_1, \xi_2)$ . The “object model variable”  $\phi$  lies in some object space  $H_o$  whose vectors, the functions  $\phi$ , are defined at a very high level of resolution. In this sense, one may say that  $H_o$  emulates the Hilbert space of square integrable real-valued functions  $L^2_{\mathbb{R}}(\mathbb{R}^2)$ . More precisely,  $H_o$  is characterized by two key parameters: the extension  $\Delta\xi$  of its field, and its resolution scale  $\delta\xi$ ;  $\delta\xi$  is therefore much smaller than the resolution limit that can be reasonably expected at the end of the reconstruction process:  $\delta\xi \triangleq \Delta\xi/N$ , where  $N$  is some power of 2. (The larger is  $N$ , the more oversampled is the object field.) To define the object space more explicitly, we introduce the finite grid (see Fig. 2):

$$\mathbb{G} \triangleq \mathbb{L} \times \mathbb{L}, \quad \mathbb{L} \triangleq \left\{ p \in \mathbb{Z} : -\frac{N}{2} \leq p \leq \frac{N}{2} - 1 \right\}.$$

Let  $\mathbf{p}$  be an integer vector lying in  $\mathbb{Z}^2$ :  $\mathbf{p} = (p_1, p_2)$ . Clearly, the functions

$$e_{\mathbf{p}}(\xi) \triangleq e_o(\xi - \mathbf{p} \delta\xi) \tag{8}$$

where

$$e_o(\xi) \triangleq \text{sinc}\left(\frac{\xi_1}{\delta\xi}\right) \text{sinc}\left(\frac{\xi_2}{\delta\xi}\right) \tag{9}$$

form an orthogonal set with

$$\| e_{\mathbf{p}} \|_o^2 = (\delta\xi)^2 \quad (\forall \mathbf{p} \in \mathbb{Z}^2). \tag{10}$$

In our comparative analysis of CLEAN and WIPE,  $H_o$  is the Euclidean space generated by the basis vectors  $e_{\mathbf{p}}$ ,

$\mathbf{p}$  spanning  $\mathbb{G}$  (see Fig. 2). The functions  $\phi$  lying in this space can therefore be expanded in the form:

$$\phi = \sum_{\mathbf{p} \in \mathbb{G}} x_{\mathbf{p}} e_{\mathbf{p}} \quad (x_{\mathbf{p}} \in \mathbb{R}). \tag{11}$$

In the general context of this paper, the functions  $e_{\mathbf{p}}$ , which play the role of interpolation or scaling functions, can be regarded as the “elementary particles” of the object representation. Evidently, other orthogonal sets of scaling functions can be used (cf. Sect. 4.3 of Lannes et al. 1994). Let us finally note that  $H_o$  is a subspace of  $L^2_{\mathbb{R}}(\mathbb{R}^2)$ ; its inner product can be explicitly expressed in the form (cf. Eqs. (8)-(11)):

$$(\phi_1 | \phi_2)_o = (\delta\xi)^2 \sum_{\mathbf{p} \in \mathbb{G}} x_{1,\mathbf{p}} x_{2,\mathbf{p}}.$$

Thus,

$$\| \phi \|_o^2 = (\delta\xi)^2 \sum_{\mathbf{p} \in \mathbb{G}} x_{\mathbf{p}}^2.$$

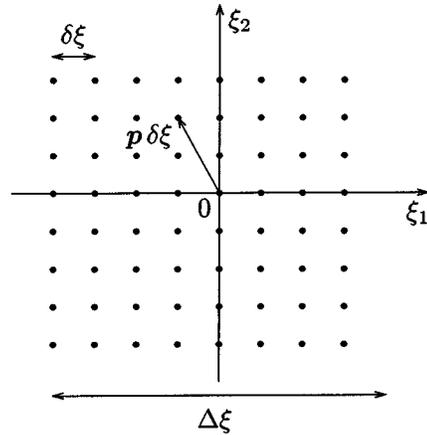


Fig. 2. Object grid  $\mathbb{G} \delta\xi$  for  $N = 8$

The Fourier transform of  $\phi$  is defined by the relationship

$$\widehat{\phi}(\mathbf{u}) \triangleq \int \phi(\xi) \exp(-2i\pi \mathbf{u} \cdot \xi) d\xi,$$

where  $\mathbf{u}$  is a two-dimensional angular spatial frequency:  $\mathbf{u} = (u_1, u_2)$ . According to Eqs. (11), (8) and (9), we therefore have:

$$\widehat{\phi} = \sum_{\mathbf{p} \in \mathbb{G}} x_{\mathbf{p}} \widehat{e}_{\mathbf{p}}$$

where

$$\widehat{e}_{\mathbf{p}}(\mathbf{u}) = \widehat{e}_o(\mathbf{u}) \exp\left(-2i\pi \mathbf{p} \cdot \frac{\mathbf{u}}{\Delta\xi}\right)$$

with  $\Delta u \triangleq 1/\delta\xi$  and

$$\widehat{e}_o(\mathbf{u}) = \frac{1}{(\Delta u)^2} \text{rect}\left(\frac{u_1}{\Delta u}\right) \text{rect}\left(\frac{u_2}{\Delta u}\right).$$

The “experimental data”  $\psi_e(\mathbf{u})$  are blurred values of  $\widehat{\phi}_o(\mathbf{u})$  on a finite list of frequencies in the Fourier domain:  $\mathcal{L}_e \triangleq \{\mathbf{u}_e(1), \mathbf{u}_e(2), \dots, \mathbf{u}_e(N_e)\}$ . As  $\phi_o$  is a real-valued function, it is natural to define  $\psi_e(-\mathbf{u})$  as the complex conjugate of  $\psi_e(\mathbf{u})$ . The ‘experimental frequency list’  $\mathcal{L}_e$  is defined consequently: if  $\mathbf{u} \in \mathcal{L}_e$ , then  $-\mathbf{u} \in \mathcal{L}_e$  (except for  $\mathbf{u} = \mathbf{o}$ : in the convention adopted in the sequel, either the null frequency does not lie in  $\mathcal{L}_e$ , or there exists only one occurrence of this point). The experimental frequency coverage generated by  $\mathcal{L}_e$  is therefore centrosymmetric (see Fig. 3a).

The experimental data vector  $\psi_e$  lies in the “experimental data space”  $K_e$ . By definition,  $K_e$  is the real Euclidean space underlying the space of complex-valued functions on  $\mathcal{L}_e$  such that  $\psi(-\mathbf{u}) = \overline{\psi(\mathbf{u})}$ . This space, whose dimension is equal to  $N_e$ , is endowed with the scalar product

$$(\psi_1 | \psi_2)_e \triangleq \sum_{\mathbf{u} \in \mathcal{L}_e} \overline{\psi_1(\mathbf{u})} \psi_2(\mathbf{u}) W(\mathbf{u}) (\delta u)^2, \quad (12)$$

where  $\delta u \triangleq \Delta u/N$ , and  $W$  is a given weighting function (see for instance Eqs. (17) and (18) further on). Note that this scalar product is real:  $(\psi_1 | \psi_2)_e = (\psi_2 | \psi_1)_e$ . In CLEAN, the data space  $K_d$  reduces to  $K_e$ , whereas in WIPE,  $K_d$  is a natural extension of  $K_e$  (cf. Sect. 5.1).

Let  $\mathcal{H}$  be the Fourier domain:  $\mathcal{H} \triangleq (-\Delta u/2, \Delta u/2)^2$ . In Fourier synthesis, the frequency coverage to be synthesized is a centrosymmetric region  $\mathcal{H}_s \subset \mathcal{H}$ . In Fig. 3a, this region is the disc centred on the origin. CLEAN and WIPE share a common objective, that of the image to be reconstructed. This image,  $\phi_s$ , is defined so that its Fourier transform is quadratically negligible outside  $\mathcal{H}_s$ . More explicitly,  $\phi_s$  is defined by a convolution relation of the form:

$$\phi_s \triangleq s \star \phi_o. \quad (13)$$

The “synthetic beam”  $s$  is a function resulting from the choice of  $\mathcal{H}_s$ : the “clean beam” in CLEAN, the “neat beam” in WIPE (see Fig. 3b). The Fourier data corresponding to  $\phi_s$  are then defined by the relationship:

$$\psi_s(\mathbf{u}) \triangleq \widehat{s}(\mathbf{u}) \psi_e(\mathbf{u}) \quad (\text{on } \mathcal{L}_e). \quad (14)$$

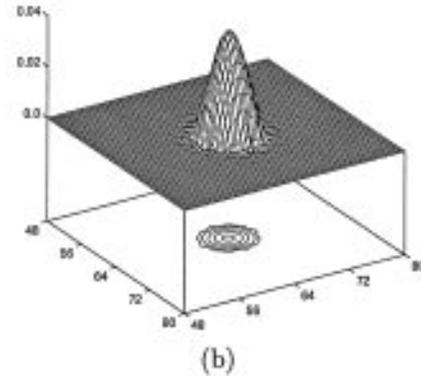
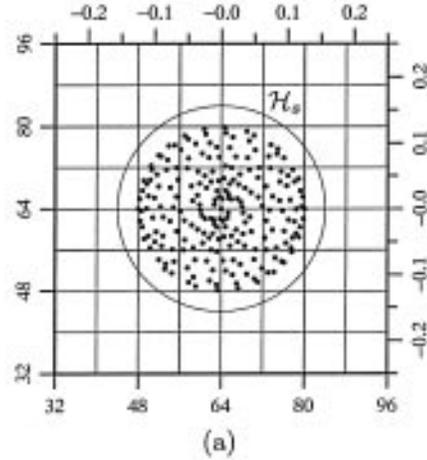
Clearly,  $\psi_s(\mathbf{u})$  lies in  $K_e \subseteq K_d$ .

The neat beam can be regarded as a sort of optimal clean beam: the optimal apodized point spread function that can be designed within the limits of the Heisenberg principle. (Apodization theory is concerned with the control of the distribution of light over the exit pupil of an optical system in order to achieve a suppression of side lobes of the diffraction pattern — Slepian 1965). More precisely, the neat beam  $s$  is a centrosymmetric function lying in the object space  $H_o$ , and satisfying the following properties:

- (i) The “energy” of  $\widehat{s}$  is concentrated in  $\mathcal{H}_s$ . In other words,  $\widehat{s}$  is small outside  $\mathcal{H}_s$  in the mean-square sense. Given  $\chi^2$  less than 1, but close to 1 (say  $\chi^2 = 0.97$ ),  $s$  satisfies a condition of the form:

$$\frac{1}{\|s\|^2} \int_{\mathcal{H}_s} |\widehat{s}(\mathbf{u})|^2 d\mathbf{u} = \chi^2 \quad (15)$$

where  $\|s\|^2 = \|\widehat{s}\|^2 = \int_{\mathbb{R}^2} |\widehat{s}(\mathbf{u})|^2 d\mathbf{u}$ .



**Fig. 3.** General conditions of the simulation; **a)** experimental frequency coverage and frequency coverage to be synthesized; **b)** representation of the corresponding neat beam (for  $\chi^2 = 0.97$ ). The experimental frequency list corresponding to the experimental frequency coverage shown in **a)** includes 211 frequency points ( $\dim K_e = 211$ ). The Fourier grid is of the form  $\mathbb{G} \delta u$  where  $\delta u = \Delta u/N$  with  $N = 128$ . The neat beam  $s$  represented in **b)** corresponds to the frequency coverage to be synthesized  $\mathcal{H}_s$ . It is centred in the object grid  $\mathbb{G} \delta \xi$  where  $\delta \xi = 1/\Delta u$ . The resolution limit of the reconstruction process (the full width of  $s$  at half maximum) is of the order of  $1.5/\Delta u_s$  where  $\Delta u_s$  is the diameter of  $\mathcal{H}_s$

- (ii) The effective support  $\mathcal{D}_s$  of  $s$  is as small as possible with respect to the choices of  $\mathcal{H}_s$  and  $\chi$ . The idea is of course to have the best possible resolution.

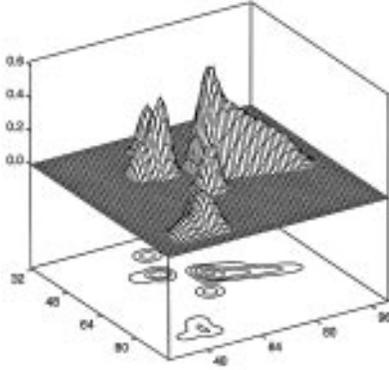
– (iii) The normalization condition  $\widehat{s}(\mathbf{0}) = \mathbf{1}$ , so that:

$$\int_{\mathbb{R}^2} \phi_s(\boldsymbol{\xi}) d\boldsymbol{\xi} = \int_{\mathbb{R}^2} \phi_o(\boldsymbol{\xi}) d\boldsymbol{\xi}. \quad (16)$$

In the terminology adopted in this paper, an atom such as  $s$  is a finite linear combination of elementary particles  $e_p$ . The integers  $p$  involved in this linear combination lie in a subset  $\mathbb{D}_s$  of  $\mathbb{G}$  (see Figs. 2 and 3b). As explicitly shown in Sect. 2 of Lannes et al. (1996), the computation of the corresponding coefficients does not raise any particular difficulty.

In a first approach to Fourier synthesis, Eq. (13) suggests that the basis functions of the object representation space  $E$  should be translated versions of  $s$ : a finite number of atoms centred on the nodes of grid  $\mathbb{G}$ . This very natural idea, which is exploited as it is, in the matching pursuit principle of CLEAN (cf. Sect. 2.2), may be completely relaxed in WIPE. More precisely, the matching pursuit process may be performed at the level of the elementary particles. The resolution limit of the reconstruction process is then kept under control thanks to an appropriate regularization principle (cf. Sect. 5.1).

The simulations presented in this paper correspond to the general conditions defined in Fig. 3. The “object function”  $\phi_o$  was defined as a set of “ $\delta$  functions” centred on the nodes of grid  $\mathbb{G}$   $\delta\xi$  with  $N = 128$ . The image to be reconstructed ( $\phi_s = s \star \phi_o$ ) then lies in  $H_o$ . Figure 4 gives an idea of the complexity of this image.



**Fig. 4.** Image to be reconstructed,  $\phi_s$ , at the resolution level defined in Fig. 3. Note that the portion of the field represented here is twice as large as that defined in Fig. 3b

The Fourier data were blurred by adding a Gaussian noise. More precisely, for all the  $\mathbf{u} \in \mathcal{L}_e$  (see Fig. 3a), the standard deviation of  $[\psi_e(\mathbf{u}) - \widehat{\phi}_o(\mathbf{u})]$  was set equal to 5% of the total flux of the object:  $\sigma_e(\mathbf{u}) = \widehat{\phi}_o(\mathbf{o})/20$ . The weighting function  $W$  introduced in Eq. (12) was explicitly defined by the formula

$$W(\mathbf{u}) \triangleq \frac{w(\mathbf{u})}{\sum_{\mathbf{u}' \in \mathcal{L}_e} \text{sinc}^2\left(\frac{u_1 - u'_1}{\delta u}\right) \text{sinc}^2\left(\frac{u_2 - u'_2}{\delta u}\right) w(\mathbf{u}')} \quad (17)$$

where

$$w(\mathbf{u}) \triangleq 1/\sigma_s^2(\mathbf{u}), \quad \sigma_s(\mathbf{u}) \triangleq \widehat{s}(\mathbf{u})\sigma_e(\mathbf{u}). \quad (18)$$

## 2. Reconstruction via matching pursuit methods

Among the various iterative methods that can be implemented for finding an approximation to the image (or the object) to be reconstructed, there exists a very slow algorithm which is based on a matching pursuit strategy. As will be clarified in this section, this algorithm is nothing but an aborted version of a particular algorithm minimizing  $q$  on  $E_o$  (see the introduction of Sect. 1.1). The corresponding iterative process must never be used in practice for solving the problem. Its slow convergence may however be of interest for initializing the choice of the object representation space  $E$ . It is therefore important to analyse its principle (Sect. 2.1), and in particular, to show that CLEAN is an algorithm of this type (Sect. 2.2).

### 2.1. Reconstruction principle

Let  $h_j \triangleq Ag_j$  be the virtual data vector corresponding to the object atom  $g_j$  (cf. Sect. 1.1), and  $Q_j$  be the projection (operator) of  $K_d$  onto the space generated by  $h_j$ :

$$Q_j \psi = \eta_j^2 (h_j | \psi)_d h_j \quad (\psi \in K_d) \quad (19)$$

where

$$\eta_j \triangleq \|h_j\|_d^{-1} = (g_j | A^*Ag_j)_o^{-1/2}. \quad (20)$$

The guiding idea is to determine the projection of  $\psi_d$  onto  $F_d \triangleq AE_o$  via the elementary projections  $Q_j$ .

Let us consider the iteration in  $F_d$ :

$$\psi_0 = 0, \quad \psi_{n+1} = \psi_n + \omega Q_k(\psi_d - \psi_n), \quad (\omega > 0); \quad (21)$$

$\omega$  is a relaxation parameter to be defined. At each iteration,  $k \equiv k_n$  is chosen so that

$$\|Q_k(\psi_d - \psi_n)\|_d = \max_{1 \leq j \leq M} \|Q_j(\psi_d - \psi_n)\|_d. \quad (22)$$

If  $Q_k(\psi_d - \psi_n) = 0$ , then  $\psi_n = \psi_{F_d}$  (the projection of  $\psi_d$  onto  $F_d$ ) and the problem is solved.

Let us set  $y_n \triangleq \psi_d - \psi_n$  and  $z_n \triangleq \psi_{F_d} - \psi_n$ . As  $Q_k y_n = Q_k z_n$ , we have from Eq. (21):

$$y_{n+1} = y_n - \omega Q_k y_n \quad z_{n+1} = z_n - \omega Q_k z_n.$$

It follows that

$$\begin{aligned} \|y_{n+1}\|_d^2 &= \|y_n\|_d^2 - 2\omega(y_n | Q_k y_n)_d + \omega^2 \|Q_k y_n\|_d^2 \\ &= \|y_n\|_d^2 - 2\omega \|Q_k y_n\|_d^2 + \omega^2 \|Q_k y_n\|_d^2, \end{aligned}$$

hence

$$\|y_{n+1}\|_d^2 = \|y_n\|_d^2 - \omega(2 - \omega) \|Q_k y_n\|_d^2. \quad (23)$$

Likewise,  $\|z_{n+1}\|_{\mathcal{d}}^2 = \|z_n\|_{\mathcal{d}}^2 - \omega(2-\omega) \|Q_k z_n\|_{\mathcal{d}}^2$ . Provided that  $\omega$  lies in the open interval  $(0, 2)$ ,  $\omega(2-\omega)$  is strictly positive. Then,  $\|z_{n+1}\|_{\mathcal{d}} \leq \|z_n\|_{\mathcal{d}}$ . The sequence  $\{\beta_n\}_{n=0}^{\infty}$ , where  $\beta_n \triangleq \|z_n\|_{\mathcal{d}}^2$ , therefore converges towards some nonnegative number  $\beta$ . As shown in Appendix 2,  $\beta$  proves to be equal to 0. As a result,  $z_n \rightarrow 0$  as  $n \rightarrow \infty$ . The iterates (21) then converge towards  $\psi_{F_d}$ .

The maximal value of  $\omega(2-\omega)$  is attained for  $\omega = 1$ . To increase the convergence speed of the projection operation,  $\omega$  may be set equal to this optimal value. The corresponding algorithm,  $\psi_{n+1} = \psi_n + Q_k(\psi_d - \psi_n)$ , is nothing but a traditional matching pursuit process (see Mallat & Zhang 1993).

As  $y_n = \psi_d - \psi_n$ , we have from Eq. (19),

$$Q_k y_n = \eta_k^2 (Ag_k | \psi_d - \psi_n)_d Ag_k,$$

hence

$$Q_k y_n = \eta_k^2 (g_k | r_n)_o Ag_k \quad (24)$$

where  $r_n \triangleq A^*(\psi_d - \psi_n)$ . The relaxed matching pursuit iteration (21) can therefore be written in the form

$$\psi_0 = 0, \quad \psi_{n+1} = \psi_n + \omega \eta_k^2 (g_k | r_n)_o Ag_k.$$

Clearly, this sequence is the image by  $A$  of the object sequence (in  $E_o$ ):

$$\phi_0 = 0, \quad \phi_{n+1} = \phi_n + \omega \eta_k^2 \rho_{n,k} g_k, \quad (25)$$

where

$$\rho_{n,k} \triangleq (g_k | r_n)_o \quad r_n = A^* \psi_d - A^* A \phi_n. \quad (26)$$

According to its definition, the residue  $r_n$  is obtained via the iteration:

$$r_0 = A^* \psi_d, \quad r_{n+1} = r_n - \omega \eta_k^2 \rho_{n,k} A^* Ag_k. \quad (27)$$

As, from Eqs. (24) and (26),  $Q_k y_n = \eta_k^2 \rho_{n,k} Ag_k$ , we have (cf. Eq. (20)):

$$\|Q_k y_n\|_{\mathcal{d}}^2 = \eta_k^4 \rho_{n,k}^2 (g_k | A^* Ag_k)_o = \eta_k^2 \rho_{n,k}^2.$$

On setting (cf. Eq. (1))

$$q_n \triangleq q(\phi_n) = \|\psi_d - A \phi_n\|_{\mathcal{d}}^2 = \|\psi_d - \psi_n\|_{\mathcal{d}}^2 = \|y_n\|_{\mathcal{d}}^2,$$

it follows from Eq. (23) that  $q_n$  is obtained through the iteration:

$$q_0 = \|\psi_d\|_{\mathcal{d}}^2, \quad q_{n+1} = q_n - \omega(2-\omega) \eta_k^2 \rho_{n,k}^2. \quad (28)$$

Provided that  $\omega$  lies in the open interval  $(0, 2)$ , the iterates  $q_n$  converge towards the minimal value of  $q$  on  $E_o$ . Sequence (25) then converges towards a solution  $\phi$  of the problem;  $\phi$  is the unique solution  $\phi_{E_o}$ , if and only if  $A_{E_o}$  is a one-to-one map.

## 2.2. Presentation of CLEAN as a matching pursuit algorithm

In our formulation of CLEAN, which essentially follows that of Högbom (1974), the object space is the Euclidean space  $H_o$  introduced in Sect. 1.2. The vectors  $g_j$  are then translated versions of the clean beam  $\text{CB} \equiv s$  (see Fig. 3b). More precisely, the elements of  $\mathcal{E}_o$  are the clean beams  $\text{CB}_{\mathbf{p}}$  centred on the nodes of the “clean box”  $\mathbb{G}_c \delta\xi$ :

$$\text{CB}_{\mathbf{p}} \equiv s_{\mathbf{p}}, \quad s_{\mathbf{p}}(\boldsymbol{\xi}) \triangleq s(\boldsymbol{\xi} - \mathbf{p} \delta\xi) \quad (\mathbf{p} \in \mathbb{G}_c \subset \mathbb{G}).$$

The data space  $K_d$  coincides with the experimental data space  $K_e$ , and  $A$  with the experimental Fourier sampling operator:  $(A_e \phi)(\mathbf{u}) \triangleq \widehat{\phi}(\mathbf{u})$  on  $\mathcal{L}_e$ . As the image to be reconstructed is defined as the convolution of the original object by the clean beam (Eq. (13)), the data vector  $\psi_d$  must be defined as the experimental data vector  $\psi_e$  damped by the Fourier transform of the clean beam:  $\psi_d = \psi_s$  (cf. Eq. (14)). We then have  $q \equiv q_e$  with (cf. Eqs. (1), (12) and (17)):

$$\begin{aligned} q_e(\phi) &\triangleq \|\psi_s - A_e \phi\|_e^2 \\ &= \sum_{\mathbf{u} \in \mathcal{L}_e} |\psi_s(\mathbf{u}) - \widehat{\phi}(\mathbf{u})|^2 W(\mathbf{u}) (\delta u)^2. \end{aligned} \quad (29)$$

As explicitly shown in Appendix 3, the “dirty map” is the map of the scalar components of  $A_e^* \psi_e$  in the basis of the elementary particles  $e_{\mathbf{p}}$ . In this context,  $A_e^* \psi_s$  may be referred to as the “dusty map”. For clarity, we set  $\text{DM}_e \equiv A_e^* \psi_e$  and  $\text{DM}_s \equiv A_e^* \psi_s$ . Likewise, the action of  $A_e^* A_e$  corresponds to a “discrete convolution” by the “dirty beam”  $\text{DB}$ :  $A_e^* A_e \phi = \text{DB} \star \phi$  (the precise definition of this operation is given in Appendix 3). Thus, from Eq. (20), the parameters  $\eta_{\mathbf{p}} \equiv \eta_j$  are all equal to:

$$\eta = (\text{CB} | \text{DB} \star \text{CB})_o^{-1/2}. \quad (30)$$

The relaxed matching pursuit iteration (25) can then be written in the form

$$\phi_0 = 0, \quad \phi_{n+1} = \phi_n + \omega \eta^2 \rho_{n,\mathbf{p}} \text{CB}_{\mathbf{p}}, \quad (31)$$

where (from Eq. (26))

$$\rho_{n,\mathbf{p}} = (\text{CB}_{\mathbf{p}} | r_n)_o \quad r_n = \text{DM}_s - \text{DB} \star \phi_n. \quad (32)$$

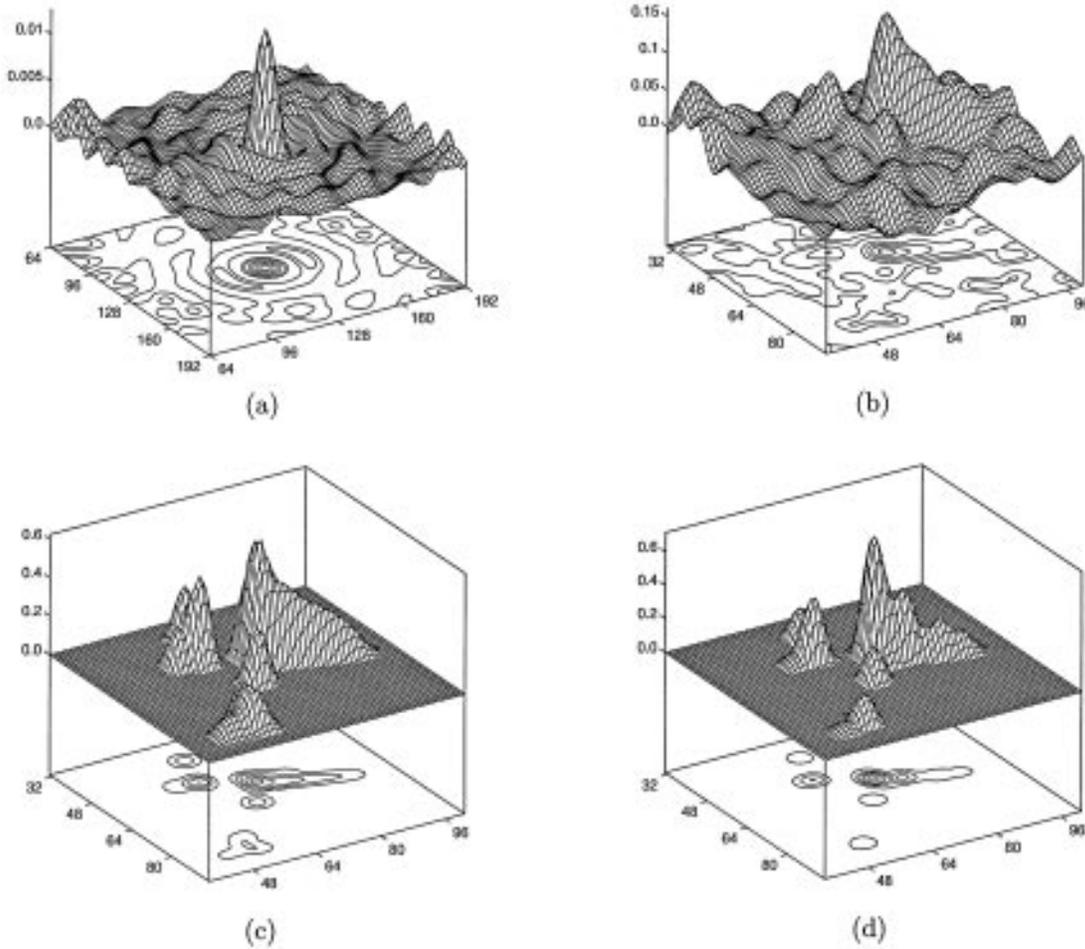
Clearly,  $\rho_n$  (the map of the  $\rho_{n,\mathbf{p}}$ ) is nothing but the “discrete intercorrelation” of  $r_n$  with  $\text{CB}$ .

The residue  $r_n$  and the quadratic errors  $q_n$  are respectively obtained via the iterations (27) and (28):

$$r_0 = \text{DM}_s, \quad r_{n+1} = r_n - \omega \eta^2 \rho_{n,\mathbf{p}} (\text{DB} \star \text{CB}_{\mathbf{p}}) \quad (33)$$

and

$$\begin{aligned} q_{e,0} &= \|\psi_s\|_e^2 = \sum_{\mathbf{u} \in \mathcal{L}_e} |\psi_s(\mathbf{u})|^2 W(\mathbf{u}) (\delta u)^2, \\ q_{e,n+1} &= q_{e,n} - \omega(2-\omega) \eta^2 \rho_{n,\mathbf{p}}^2. \end{aligned} \quad (34)$$



**Fig. 5.** Image reconstruction via CLEAN with  $\omega = 0.2$ ; **a)** dirty beam; **b)** dusty map; **c)** image to be reconstructed (Fig. 4); **d)** clean map for  $C_e = 1.99$  (the definition of the fit criterion  $C_e$  is given in Eq. (35)). In the conditions of this simulation (see Fig. 3), the optimal fit threshold of CLEAN is of the order of 1.75. For a lower threshold, the support of the clean map is no longer contained in that of the image to be reconstructed. In the framework of the analysis presented in this paper, the residual maps  $r_n$  or  $\rho_n$  must not be added to the clean map

Note that  $\text{DB} \star \text{CB}_{\mathbf{p}} = \text{DB}_{\mathbf{p}} \star \text{CB} = (\text{DB} \star \text{CB})_{\mathbf{p}}$ .

In the classical presentation of CLEAN, the convolution by the clean beam is performed a posteriori, whence some small differences in these iterations (cf. Appendix 4). In particular, in the version of CLEAN presented here,  $\mathbf{p}$  is chosen (at each iteration) so that  $|\rho_{n,\mathbf{p}}| = \max_{\mathbf{p}' \in \mathbb{G}_e} |\rho_{n,\mathbf{p}'}|$ .

The process is interrupted as soon as  $q_{e,n}$  is less than a threshold value related to the level of the noise in the Fourier domain. In our implementation of CLEAN, we introduce the “fit criterion” (cf. Eqs. (18) and (29)):

$$C_e(\phi) \triangleq \frac{\sqrt{q_e(\phi)}}{\|\sigma_s\|_e} \quad \|\sigma_s\|_e^2 = \sum_{\mathbf{u} \in \mathcal{L}_e} \sigma_s^2(\mathbf{u}) W(\mathbf{u}) (\delta u)^2. \quad (35)$$

As soon as  $C_e \triangleq C_e(\phi_n)$  is less than 2 (for example), the matching pursuit process is interrupted;  $\phi_n$  is the corresponding “clean map”.

In the simulation presented in Fig. 5, we show the clean map corresponding to the fit threshold 2. The relaxation parameter  $\omega$  was set equal to 0.2, and the clean box was defined as the support of  $\phi_s$  at a lower level of resolution (twice as low). In the conditions of this simulation, the optimal fit threshold of CLEAN is of the order of 1.75. For a lower threshold, the support of the clean map is no longer contained in that of the image to be reconstructed.

Let  $E$  be the object representation space generated by the  $\text{CB}_{\mathbf{p}}$  selected by CLEAN. Clearly, the clean map  $\phi_n$  does not minimize  $q_e$  on  $E$ . The same matching pursuit algorithm (with  $\omega = 1$ ) can be confined to  $E$  for performing the complete minimization on this space. This corresponds to the principle of what is referred to as “Window CLEAN” (Schwarz 1978). The algorithms presented in Sects. 3 and 4 are much more efficient for this purpose, but as specified in Sect. 5, they only reveal that (in situations of astrophysical interest) the solution

thus obtained is without any interest: the problem is ill-conditioned.

### 3. Optimization without control of robustness

Let  $\mathcal{E}$  be any subset of  $\mathcal{E}_o$ , say that generated by an aborted matching pursuit process;  $\mathcal{E}$  has  $m$  elements. Let us now consider the problem of minimizing  $q(\phi)$  on the space  $E$  generated by the  $g_k$ ,  $k$  spanning

$$\mathbb{J} \triangleq \{k : g_k \in \mathcal{E}\}.$$

By definition,  $E$  is the range of the operator:

$$S : \mathbb{R}^m \rightarrow H_o, \quad S\alpha \triangleq \sum_{k \in \mathbb{J}} \alpha_k g_k. \quad (36)$$

In the case where  $\mathbb{R}^m$  is equipped with its standard scalar product, the adjoint of  $S$  is explicitly defined by the relationship:

$$(S^* \phi)_k = (g_k | \phi)_o \quad (\text{for all } k \in \mathbb{J}).$$

Indeed, for any  $\phi \in H_o$ , we have from Eq. (36):

$$(S\alpha | \phi)_o = \sum_{k \in \mathbb{J}} \alpha_k (g_k | \phi)_o \equiv \sum_{k \in \mathbb{J}} \alpha_k (S^* \phi)_k.$$

In what follows,  $S$  is not necessarily a one-to-one map from  $\mathbb{R}^m$  onto  $E$ : the vectors  $g_k$  lying in  $\mathcal{E}$  are not necessarily linearly independent.

Let  $\alpha$  now be a vector minimizing on  $\mathbb{R}^m$  the quantity  $\|\psi_d - AS\alpha\|_d^2$ . Then, the vector  $\phi = S\alpha$  minimizes  $q$  on  $E$ . From Eq. (2), the vectors  $\alpha$  in question are such that

$$\forall k \in \mathbb{J}, \quad (g_k | A^*(\psi_d - AS\alpha))_o = 0. \quad (37)$$

These vectors are therefore the solutions of the normal equation

$$S^*A^*(\psi_d - AS\alpha) = 0 \quad (\alpha \in \mathbb{R}^m) \quad (38)$$

(the least-squares solutions of the equation  $AS\alpha = \psi_d$ ).

In most cases encountered in image reconstruction, the conjugate-gradients method is the best suited technique for solving Eq. (38). The version of this method presented below provides  $\phi = S\alpha$ .

ALGORITHM 1:

Step 0: Set  $\epsilon = 10^{-6}$  (for example) and  $n = 0$ ;  
 choose a natural starting point  $\phi_0$  in  $E$ ;  
 compute  $r_0 = A^*\psi_d - A^*A\phi_0$ ,  
 $q_0 = \|\psi_d - A\phi_0\|_d^2$ ;  
 compute  $\rho_{0,k} = (g_k | r_0)_o$  (for all  $k \in \mathbb{J}$ );  
 set  $\delta_{0,k} = \rho_{0,k}$  (for all  $k \in \mathbb{J}$ ).

Step 1: Compute

$$d_n = \sum_{k \in \mathbb{J}} \delta_{n,k} g_k,$$

$$z_n = A^*Ad_n,$$

$$\zeta_{n,k} = (g_k | z_n)_o \quad (\text{for all } k \in \mathbb{J}),$$

$$\omega_n = \sum_{k \in \mathbb{J}} |\rho_{n,k}|^2 / \sum_{k \in \mathbb{J}} \delta_{n,k} \zeta_{n,k},$$

$$q_{n+1} = q_n - (2r_n - \omega_n z_n | \omega_n d_n)_o,$$

$$\phi_{n+1} = \phi_n + \omega_n d_n,$$

$$r_{n+1} = r_n - \omega_n z_n,$$

$$\rho_{n+1,k} = \rho_{n,k} - \omega_n \zeta_{n,k} \quad (\text{for all } k \in \mathbb{J});$$

if  $\max_{k \in \mathbb{J}} \{|\rho_{n+1,k}| / \|g_k\|_o\} < \epsilon \|r_{n+1}\|_o$ ,  
 termination.

Compute

$$\gamma_n = \sum_{k \in \mathbb{J}} |\rho_{n+1,k}|^2 / \sum_{k \in \mathbb{J}} |\rho_{n,k}|^2,$$

$$\delta_{n+1,k} = \rho_{n+1,k} + \gamma_n \delta_{n,k} \quad (\text{for all } k \in \mathbb{J});$$

increment  $n$  and loop to step 1.  $\square$

Throughout this algorithm,  $r_n$  is the residue of the equation  $A^*\psi_d - A^*A\phi = 0$  for  $\phi = \phi_n$ . Likewise,  $q_n$  is the value of  $q(\phi)$  at the same iterate:

$$r_n = A^*\psi_d - A^*A\phi_n, \quad q_n = \|\psi_d - A\phi_n\|_d^2.$$

The iteration in  $q_n$  results from the identity:

$$\begin{aligned} q(\phi + \delta\phi) &= q(\phi) - 2(\psi_d - A\phi | A\delta\phi)_o + \|A\delta\phi\|_d^2 \\ &= q(\phi) - 2(A^*\psi_d - A^*A\phi | \delta\phi)_o \\ &\quad + (A^*A\delta\phi | \delta\phi)_o. \end{aligned}$$

The sequence of vectors  $\phi_n$  converges towards a solution of the problem with all the remarkable properties of the conjugate-gradients method (see Lannes et al. 1987b). In practice,  $E$  is chosen so that  $A_E$  is a one-to-one map. The uniqueness of the solution can easily be verified by modifying the starting point of the algorithm. The stopping criterion is based on the fact that the final residue must be practically orthogonal to all the  $g_k$ 's (Eq. (37)); the cosine of the angle between the vectors  $r_n$  and  $g_k$  is equal to  $\rho_{n,k} / (\|g_k\|_o \|r_n\|_o)$ . Here, as  $\mathbb{R}^m$  is endowed with its standard scalar product, this algorithm cannot provide the condition number of  $A_E$ . (The transposition of what is presented in Sect. 4.2 would give the "generalized condition number" of  $AS$ ). We therefore recommend to use algorithm 1 only when  $\kappa_E$  is approximately known.

REMARK 3. Let us consider the special case where  $A$  is the identity operator on  $K_d$  (which then coincides with  $H_o$ ). The problem is then to find  $P_E\phi_d$ , the projection of  $\phi_d = \psi_d$  onto  $E$ . Note that  $\kappa_E$  is then equal to unity. In this case, algorithm 1 reduces to

ALGORITHM 2:

Step 0: Set  $\epsilon = 10^{-6}$  (for example) and  $n = 0$ ;  
 set  $\phi_0 = 0$  and  $r_0 = \phi_d$ ;  
 compute  
 $q_0 = \|\phi_d\|_d^2$ ,

$$\begin{aligned} \rho_{0,k} &= (g_k | r_0)_o \quad (\text{for all } k \in \mathbb{J}); \\ \text{set } \delta_{0,k} &= \rho_{0,k} \quad (\text{for all } k \in \mathbb{J}). \end{aligned}$$

Step 1: Compute

$$z_n = \sum_{k \in \mathbb{J}} \delta_{n,k} g_k,$$

$$\zeta_{n,k} = (g_k | z_n)_o \quad (\text{for all } k \in \mathbb{J}),$$

$$\omega_n = \sum_{k \in \mathbb{J}} |\rho_{n,k}|^2 / \sum_{k \in \mathbb{J}} \delta_{n,k} \zeta_{n,k},$$

$$q_{n+1} = q_n - (2r_n - \omega_n z_n | \omega_n z_n)_o,$$

$$\phi_{n+1} = \phi_n + \omega_n z_n,$$

$$r_{n+1} = r_n - \omega_n z_n,$$

$$\rho_{n+1,k} = \rho_{n,k} - \omega_n \zeta_{n,k} \quad (\text{for all } k \in \mathbb{J});$$

if  $\max_{k \in \mathbb{J}} \{|\rho_{n+1,k}| / \|g_k\|_o\} < \epsilon \|r_{n+1}\|_o$ ,  
termination.

Compute

$$\gamma_n = \sum_{k \in \mathbb{J}} |\rho_{n+1,k}|^2 / \sum_{k \in \mathbb{J}} |\rho_{n,k}|^2,$$

$$\delta_{n+1,k} = \rho_{n+1,k} + \gamma_n \delta_{n,k} \quad (\text{for all } k \in \mathbb{J});$$

increment  $n$  and loop to step 1.  $\square$

This algorithm converges towards the projection of  $\phi_d$  onto  $E$  with all the properties of the conjugate-gradients method.

In principle, the projection operation can also be performed by using the matching pursuit iteration (25). In this case, on setting  $\omega$  equal to its optimal value, this iteration reduces to

$$\phi_0 = 0, \quad \phi_{n+1} = \phi_n + \eta_k^2 \rho_{n,k} g_k,$$

where

$$\eta_k = \|g_k\|_o^{-1} \quad \rho_{n,k} = (g_k | r_n)_o \quad r_n = \phi_d - \phi_n.$$

The residues  $r_n$  are then obtained via iteration (27):

$$r_0 = \phi_d, \quad r_{n+1} = r_n - \eta_k^2 \rho_{n,k} g_k;$$

and likewise for  $q_n$  (cf. Eq. (28)):

$$q_0 = \|\phi_d\|_o^2, \quad q_{n+1} = q_n - \eta_k^2 \rho_{n,k}^2.$$

At each iteration, it is then natural to choose  $k$  so that  $\eta_k |\rho_{n,k}| = \max_{j \in \mathbb{J}} \eta_j |\rho_{n,j}|$ . In the general case where the  $g_k$ 's ( $k \in \mathbb{J}$ ) do not form an orthogonal set, the conjugate-gradients algorithm is of course preferable.

#### 4. Optimization with control of robustness

For clarity, let us now assume that  $A_E$  is a one-to-one map. The method presented in Sect. 3 then yields a solution  $\alpha$  of Eq. (38), and thereby the solution of the problem:  $\phi_E = S\alpha$ . Unfortunately, as already mentioned, this method does not provide any information on the robustness of the reconstruction process. The most natural way

of obtaining this information is to find  $\phi_E$ , directly, as the solution of the normal Eq. (4):

$$B\phi = \phi_d, \quad (39)$$

where

$$B \triangleq A_E^* A_E = P_E A^* A_E \quad \phi_d \triangleq A_E^* \psi_d = P_E A^* \psi_d. \quad (40)$$

In this section, we present the corresponding developments. To conduct our analysis, the eigenvalues of  $B$  are ordered so as to form a nondecreasing sequence (cf. Eq. (7)):

$$\mu \triangleq \mu_1 \leq \mu_2 \leq \dots \leq \mu' \triangleq \mu_m. \quad (41)$$

As  $A_E$  is assumed to be a one-to-one map,  $\mu$  is strictly positive. In the general case where the  $g_k$  generating  $E$  do not form an orthogonal set, the reader must keep in mind the fact that the action of  $P_E$  can be performed with the aid of algorithm 2.

##### 4.1. Reconstruction algorithm

The problem is solved by using the conjugate-gradients method (cf. Sect. 2.3 of Lannes et al. 1987b). Starting from any  $\phi_0$  in  $E$ , the iterates  $\phi_n$  converge to  $\phi_E$  in at most  $m$  iterations,  $\phi_n$  getting closer to  $\phi_E$  at each iteration. In this algorithm,  $d_n$  is the ‘‘direction of research’’ in iteration  $n+1$ , whereas  $\omega_n$  is the corresponding ‘‘parameter of exact line search’’;  $r_n$  is the residue of the normal Eq. (39) for  $\phi = \phi_n$ :

$$r_n \triangleq \phi_d - B\phi_n.$$

As  $B\phi_E = \phi_d$ , we have  $r_n = B(\phi_E - \phi_n)$ , hence:

$$\|\phi_E - \phi_n\|_o \leq \frac{1}{\mu} \|r_n\|_o.$$

Denoting by  $\mu_e$  an estimate of  $\mu$ , we therefore have:

$$\frac{\|\phi_E - \phi_n\|_o}{\|\phi_n\|_o} \lesssim \epsilon_n, \quad \epsilon_n \triangleq \frac{\|r_n\|_o}{\mu_e \|\phi_n\|_o}.$$

Let us introduce an acceptable error threshold  $\epsilon_\phi$  for  $\|\phi_E - \phi_n\|_o / \|\phi_n\|_o$ . Clearly, the iterative process can be interrupted as soon as  $\epsilon_n$  is less than  $\epsilon_\phi$ ;  $\epsilon_n$  therefore plays the role of a convergence estimator. The estimate of  $\mu$  is refined throughout the iterative process as indicated in Sect. 4.2. The corresponding algorithm can then be summarized as follows.

ALGORITHM 3:

Step 0: Set  $\epsilon_\phi = 10^{-4}$  (for example) and  $n = 0$ ;  
choose a natural starting point  $\phi_0$  in  $E$ ;  
compute  $r_0 = \phi_d - B\phi_0$ ;  
set  $d_0 = r_0$ .

Step 1: Compute

$$\begin{aligned}
z_n &= B d_n, \\
\omega_n &= \| r_n \|_o^2 / (d_n | z_n)_o, \\
\phi_{n+1} &= \phi_n + \omega_n d_n, \\
r_{n+1} &= r_n - \omega_n z_n, \\
\epsilon_{n+1} &= \| r_{n+1} \|_o / (\mu_e \| \phi_{n+1} \|_o); \\
\text{if } \epsilon_{n+1} &\leq \epsilon_\phi, \text{ termination.}
\end{aligned}$$

Compute

$$\begin{aligned}
\gamma_n &= \| r_{n+1} \|_o^2 / \| r_n \|_o^2, \\
d_{n+1} &= r_{n+1} + \gamma_n d_n.
\end{aligned}$$

Increment  $n$  and loop to step 1.  $\square$

#### 4.2. Effective object representation space

In the conjugate-gradients method, the  $n$ -dimensional subspace of  $E$  generated by the conjugate directions  $d_0, \dots, d_{n-1}$ ,

$$E_n \triangleq \text{span}\{d_\ell\}_{\ell=0}^{n-1},$$

is referred to as the Krylov space of order  $n$ . According to a well known property (cf. properties 2 and 3.1 of Lannes et al. 1987b),  $\phi_n$  minimizes  $q(\phi)$  on  $E_n$ .

Provided that  $n$  is sufficiently large, the least-squares solutions in  $E$  and  $E_n$  are very close to one another. At the end of the reconstruction process,  $E_n$  is therefore the effective object representation space. The dimension of this space, as well as the robustness of the reconstruction process, depends upon the localization of the eigenvalues of  $B$ , and more precisely, on the relative weight of the projections of  $r_0$  onto the corresponding eigenspaces. We are thus led to consider the operator

$$B_n: E_n \rightarrow E_n, \quad B_n \phi \triangleq P_n A^* A \phi,$$

where  $P_n$  is the projection (operator) onto  $E_n$ .

The residues  $r_0, \dots, r_{n-1}$  form an orthogonal basis for  $E_n$  (see Appendix 4 of Lannes et al. 1987b). As established in Appendix 2 of Lannes et al. (1996), the matrix of  $B_n$  expressed in this basis is tridiagonal (this matrix is of course symmetric). Its diagonal and upper-diagonal elements are respectively given by the relationships

$$b_{n;\ell,\ell} = \begin{cases} \frac{1}{\omega_\ell} & (\ell = 0) \\ \frac{1}{\omega_\ell} + \frac{\gamma_{\ell-1}}{\omega_{\ell-1}} & (1 \leq \ell \leq n-1) \end{cases}$$

and

$$b_{n;\ell-1,\ell} = -\frac{\sqrt{\gamma_{\ell-1}}}{\omega_{\ell-1}} \quad (1 \leq \ell \leq n-1).$$

The eigenvalues of  $B_n$  can therefore be calculated very easily with the aid of the QR algorithm (cf. Sect. 11.3 of

Press et al. 1992). Let us order these eigenvalues as those of  $B$  (see Eq. (41)):

$$\mu_{n,1} \leq \mu_{n,2} \leq \dots \leq \mu_{n,n}.$$

By referring to the eigenvalue analysis based on the notion of “minmax numbers” (cf. Appendix 5 of Lannes et al. 1987a), it is easy to show that

$$\mu \leq \mu_{n+1,1} \leq \mu_{n,1}, \quad \mu_{n,n} \leq \mu_{n+1,n+1} \leq \mu'.$$

Provided that the projections of  $\phi_d$  onto the eigenspaces corresponding to  $\mu$  and  $\mu'$  are different from zero, a condition which is always numerically satisfied in practice,  $\mu_{n,1}$  and  $\mu_{n,n}$  respectively tend to  $\mu$  and  $\mu'$  as  $n$  tends to  $m$  (see Fig. 3 of Lannes et al. 1996).

In our reconstruction processes, the eigenvalues of  $B_n$  are computed at each iteration. (The cost for this is negligible compared to that of the action of  $B$ .) As soon as  $(\mu_{n,1} - \mu_{n+1,1})/\mu_{n,1}$  is less than say  $10^{-3}$ ,

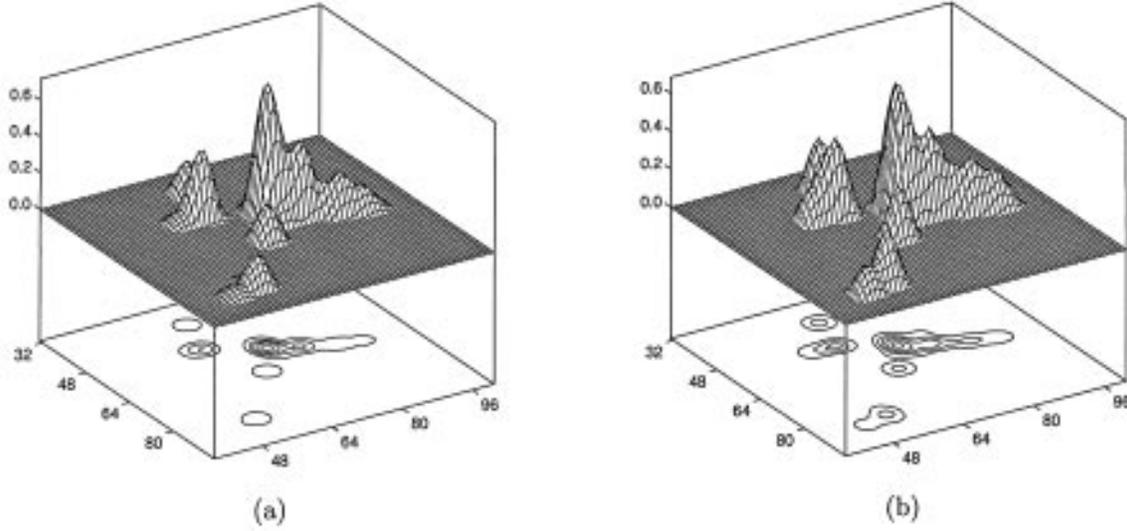
$$\mu_e = \mu_{n,1}, \quad \mu'_e = \mu_{n,n},$$

are very good approximations to  $\mu$  and  $\mu'$ , respectively. In most cases, the termination test of the basic algorithm is then satisfied (see Fig. 3 of Lannes et al. 1996).

## 5. How WIPE wipes CLEAN clean

We now have all the tools for analysing the weak points of CLEAN as well as the tricks of WIPE allowing the corresponding difficulties to be overcome.

In situations of astrophysical interest, CLEAN is implemented with a value of the relaxation parameter  $\omega$  much less than 1 (say 0.2). The basis vectors  $s_p$  selected in the matching pursuit process then define an acceptable object representation space  $E$ . Unfortunately, the problem is often ill conditioned;  $A_E$  is a one-to-one map, but its condition number is very large. For example, in the simulation presented in Fig. 5d,  $\kappa_E$  is equal to 45.08. As a result,  $\phi_E$  is a very perturbed version of the clean map. This is unsatisfactory. Indeed, in this situation, the clean map can only be regarded as an image for which the fit criterion  $C_e$  (introduced in Eq. (35)) is of the order of the threshold value (say 2). In other words, the clean map must be accepted as it is, without any reference to a stable optimization process. The interpretation of the results may then be doubtful. As indicated in Sect. 5.1, the regularization principle of WIPE remedies this lack of robustness, but the regularized version of CLEAN thus obtained is still different from WIPE. Indeed, CLEAN has another weak point: the boundaries of the structuring entities of the image may not be correctly represented in the clean map. In such situations, the matching pursuit strategy of CLEAN is not well suited for solving the problem. For example, in the particular case of the Fourier data of our simulation, the best possible fit corresponds



**Fig. 6.** Image reconstruction via the regularized version of CLEAN; **a)** traditional clean map for  $C_e = 1.99$ ; **b)** improved clean map  $\phi_E$  provided by the regularized version of CLEAN ( $C_e = 1.15$ ). These images have to be compared with the image to be reconstructed (Fig. 5c). As shown in Fig. 7b, the matching pursuit process of WIPE can still refine image (b)

to a value of  $C_e$  of about 0.9 (see Fig. 7 further on). Even with a good support constraint (a reasonable “clean box”), CLEAN cannot reach this fit threshold with a satisfactory representation of the image support. In the same situation, as shown in Sect. 5.2, WIPE reaches this threshold without any difficulty.

### 5.1. How WIPE regularizes CLEAN

In the basic version of CLEAN presented in Sect. 2.2, the functions  $\phi \in E$  are linear combinations of atoms  $s_p$ . In the sense defined in Sect. 1.2, the “energy” of the Fourier transform of each atom  $s_p$  is concentrated in  $\mathcal{H}_s$  (cf. Eq. (15)). The intrinsic instability of CLEAN is related to the fact that this property does not necessarily hold for any linear combination of such atoms. This difficulty arises any time the distances between these atoms are much smaller than the corresponding resolution limit. This is precisely the case when dealing with extended objects and a small relaxation parameter  $\omega$ .

To overcome this difficulty, WIPE suggests that CLEAN should define the reconstructed image as the function minimizing on  $E$  the functional

$$q(\phi) \triangleq q_e(\phi) + q_r(\phi) \quad (42)$$

where (see Eq. (29))

$$q_e(\phi) \triangleq \sum_{\mathbf{u} \in \mathcal{L}_e} |\psi_s(\mathbf{u}) - \widehat{\phi}(\mathbf{u})|^2 W(\mathbf{u}) (\delta u)^2$$

and

$$q_r(\phi) \triangleq \sum_{\mathbf{u} \in \mathcal{L}_r} |\widehat{\phi}(\mathbf{u})|^2 (\delta u)^2. \quad (43)$$

The experimental criterion  $q_e$  constrains  $\phi$  (the model) to be consistent with the damped Fourier data, while the regularization criterion  $q_r$  penalizes the high-frequency components of  $\phi$ . The elements of the regularization frequency list  $\mathcal{L}_r$  are located outside  $\mathcal{H}_s$ , at the nodes of grid  $\mathbb{H}_r \delta u$  where

$$\mathbb{H}_r \triangleq \{\mathbf{r} \in \mathbb{Z}^2 : \mathbf{r} \delta u \in \mathcal{H}, \mathbf{r} \delta u \notin \mathcal{H}_s\}.$$

In the traditional version of CLEAN,  $q$  reduces to  $q_e$ . The minimization of  $q_e$  on  $E$  then reveals the ill-conditioned character of the problem.

The regularized version of CLEAN corresponding to criterion (42) can be formulated in the theoretical framework presented in Sect. 1.1. To clarify this point, let us introduce the data vector:

$$\psi_d(\mathbf{u}) \triangleq \begin{cases} \psi_s(\mathbf{u}) & \text{on } \mathcal{L}_e; \\ 0 & \text{on } \mathcal{L}_r. \end{cases}$$

This vector lies in a data space  $K_d$  endowed with the scalar product:

$$\begin{aligned} (\psi_1 | \psi_2)_d &\triangleq \sum_{\mathbf{u} \in \mathcal{L}_e} \overline{\psi_1(\mathbf{u})} \psi_2(\mathbf{u}) W(\mathbf{u}) (\delta u)^2 \\ &+ \sum_{\mathbf{u} \in \mathcal{L}_r} \overline{\psi_1(\mathbf{u})} \psi_2(\mathbf{u}) (\delta u)^2. \end{aligned}$$

The Fourier sampling operator  $A$  is then the operator:

$$A: H_o \rightarrow K_d, \quad (A\phi)(\mathbf{u}) \triangleq \begin{cases} \widehat{\phi}(\mathbf{u}) & \text{on } \mathcal{L}_e; \\ \widehat{\phi}(\mathbf{u}) & \text{on } \mathcal{L}_r. \end{cases}$$

According to Eqs. (42), (29) and (43),  $q(\phi)$  can then be effectively written in the form  $\|\psi_d - A\phi\|_d^2$  (see Eq. (1)).

The problem is then stated in terms of Fourier interpolation (Lannes et al. 1987a and 1994). This is why  $q$  is of the form  $q_e + \alpha q_r$  with  $\alpha = 1$ . In this context, it is important to note that in the definition of  $q_e$  (Eq. (29)), where  $\psi_s$  is defined in Eq. (14), the weighting function  $W(\mathbf{u})$  takes into account the local redundancy of  $\mathbf{u}$  up to  $\delta u$  (Eq. (17)).

The algorithms described in Sects. 3 and 4 can be used for minimizing  $q$  on  $E$ . The action of  $A^*A$  is that of a convolutor. The corresponding point spread function, the “dusty beam”, has two components: the dirty beam and the “regularization beam”. The latter is induced by the regularization list  $\mathcal{L}_r$ . With regard to the dusty map, note that  $A^*\psi_d \equiv A_e^*\psi_s$  (cf. Sect. 2.2).

In the simulation presented in Fig. 6, we compare the clean map of Fig. 5d (for which  $C_e = 1.99$ ) with the image provided by this  $\text{opt}(C_e = 1.15)$ . The condition number is now reasonable:  $\kappa_E = 2.62$ . As shown in the next subsection, the construction of the object representation space  $E$  can, however, be refined.

### 5.2. How WIPE relaxes the matching pursuit process of CLEAN

In the regularized version of CLEAN described in Sect. 5.1, the calculation of the condition number  $\kappa_E$  requires the action of the projection  $P_E$  at each iteration of algorithm 3. This projection is performed at the cost of the conjugate-gradients iterations of algorithm 2. This first remark suggests that  $E$  should be redefined as the linear space generated by the elementary particles  $e_{\mathbf{p}'}$  of all the atoms  $s_{\mathbf{p}}$  selected in the matching pursuit process of CLEAN. The projection onto  $E$  is then trivial since these elementary particles form an orthogonal set. As the resolution limit of the reconstruction process is then controlled by the regularizer  $q_r$  (cf. Eq. (43)), the choice of such an object representation space proves to be very natural. Moreover, the definition of  $E$  can then be refined by continuing (or even by conducting) the matching pursuit process at the level of the elementary particles. As specified below, this is what is precisely done in WIPE.

Let  $\mathbb{D}$  then be the subset of  $\mathbb{G}$  corresponding to the choice of  $\mathcal{E}$ . (The elementary particles generating  $E$  are centred on the nodes of  $\mathbb{D}\delta\xi$ .) We say that  $\mathbb{D}$  is the “discrete field (or support)” associated with the definition of  $E$ . Depending on the particular problems to be solved, this discrete field may be fixed from the outset (for example, in an interactive manner), or constructed step by step in a matching pursuit strategy.

In this last case, which corresponds to the basic version of WIPE, let us denote by  $\mathbb{D}^{(i)}$  the discrete field obtained at the end of the  $i^{\text{th}}$  step of the construction of the object representation space. Let  $\phi^{(i)}$  then be the solution of the problem in the corresponding object representation space  $E^{(i)}$ . In the basis of the elementary particles  $e_{\mathbf{p}}$  (the

interpolation basis of  $H_o$ ), the scalar components of the residue  $r^{(i)}$  are the quantities:

$$\rho_{\mathbf{p}}^{(i)} \triangleq (e_{\mathbf{p}} | r^{(i)})_o \quad r^{(i)} \triangleq A^*\psi_s - A^*A\phi^{(i)}.$$

According to the definition of  $\phi^{(i)}$ , these coefficients vanish on  $\mathbb{D}^{(i)}$  (see Eq. (2) with  $g_k \equiv e_{\mathbf{p}}$ ,  $\mathbf{p} \in \mathbb{D}^{(i)}$ ). One then has to decide whether the current field has to be extended. The current values of  $C_e$  and  $\kappa_E$  play an essential role in this decision. When the reconstruction procedure is not interrupted at this stage, WIPE uses algorithm 3 for computing the solution of the problem in the object representation space relative to the union of  $\mathbb{D}^{(i)}$  with some set  $\mathbb{D}' \subset \mathbb{G}$ :

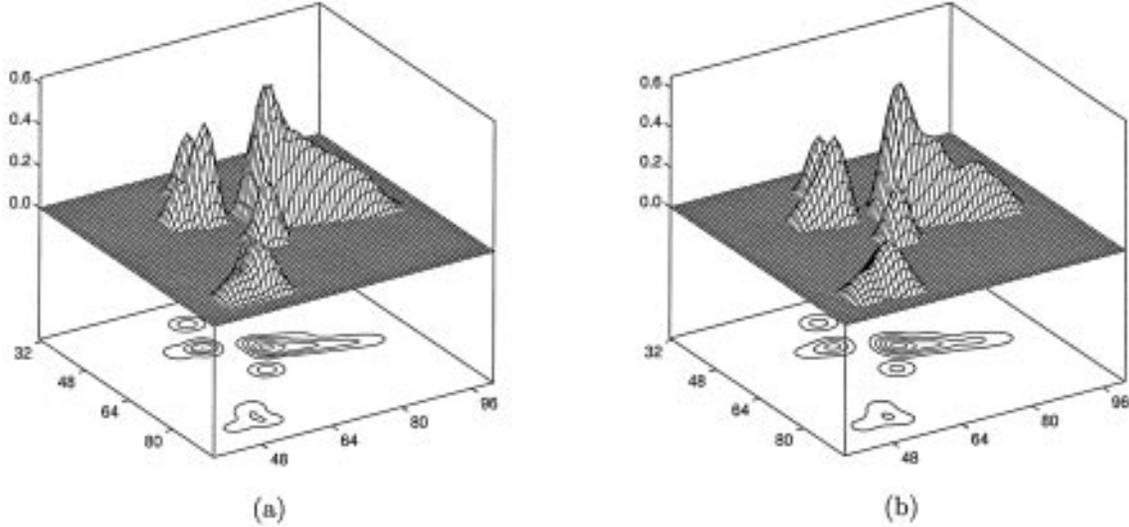
$$\mathbb{D}^{(i+1)} = \mathbb{D}^{(i)} \cup \mathbb{D}'.$$

There exist many ways of selecting  $\mathbb{D}'$ . All are based on the examination of the distribution of the coefficients  $\rho_{\mathbf{p}}^{(i)}$  outside  $\mathbb{D}^{(i)}$ . For example, one may try to define  $\mathbb{D}'$  as a connected region containing the “pixel”  $\mathbf{p}_{\text{max}}$  for which the maximum of these coefficients is attained. The simplest choice is then to define  $\mathbb{D}'$  as the discrete field of the atom  $s$  centred on this pixel. With regard to the construction of the object representation space, the corresponding version of WIPE is then very similar to that of CLEAN.

In the matching pursuit steps where the field of the reconstructed image must be refined, it is natural to choose the nodes of  $\mathbb{D}'$  along the boundaries of the structuring entities of the image. Let  $N_s$  be the number of particles involved in the linear combination defining the neat beam  $s$  (the number of nodes in  $\mathbb{D}_s$ ). In the basic version of WIPE, the size of  $\mathbb{D}'$ , expressed in number of nodes, is defined as a fraction of  $N_s$  (say  $N_s/2$ ), and the selected nodes are those for which the coefficients  $\rho_{\mathbf{p}}^{(i)}$  are the largest above some given threshold (half of the maximal value, for example). The field of the image (or object) to be reconstructed can thus be obtained in a natural manner.

The construction of the object representation space is interrupted as soon as the fit criterion  $C_e(\phi^{(i)})$  is sufficiently small, for instance, less than or of the order of 0.85. The current field is then refined by a morphological smoothing of its connected entities. In this classical operation of mathematical morphology, the discrete support of the neat beam,  $\mathbb{D}_s$ , is of course used as structuring element. The boundaries of the effective field of the “neat map” (the reconstructed image) are thus defined at the appropriate resolution. In particular, the connected entities of size smaller than that of  $\mathbb{D}_s$  are eliminated. As illustrated in Fig. 7, it is thus possible to reach the optimal value of  $C_e$  (0.88 in the simulation under consideration) with a satisfactory representation of the image field.

Let  $E$  be the object representation space at the end of the action of WIPE, and  $\mathbb{D}$  be the corresponding discrete field. There exists a variant of WIPE, in which the object representation space is a particular subspace of  $E$ , that



**Fig. 7.** Image reconstruction through WIPE; **a)** image to be reconstructed; **b)** neat map (reconstructed image). The latter, for which  $C_e = 0.88$  and  $\kappa_E = 3.83$  (and which was obtained without any clean box), is to be compared to image **a)** and to the maps presented in Fig. 6. The boundaries of the structuring entities of the image are now correctly restored, hence a better intensity distribution. The unreliable character of the oscillating perturbation along the main structuring entity of the reconstructed images is revealed by the image-eigenmode analysis provided by WIPE (see Fig. 6 of Lannes et al. 1996)

generated by all the atoms  $s_p$  whose discrete field is contained in  $\mathbb{D}$ . In the conditions of the simulation presented in Fig. 7, the corresponding solution is very close to that provided by WIPE. As expected, the condition number is then slightly smaller (here, 3.36 instead of 3.83).

From the outset, the discrete field  $\mathbb{D}$  may be taken equal to that of the clean box. One then uses the global version of WIPE in which the nonnegativity constraint is imposed (cf. Sect. 4.3 of Lannes et al. 1996). At the end of the corresponding reconstruction process, the fit criterion  $C_e$  is often smaller than its optimal value. As a result, the support of the image (or object) to be reconstructed is not well restored. A similar remark can be made for the Fourier synthesis methods in which the regularization principle is based on the concept of entropy. Moreover, the relative weights of the experimental and regularization criteria must then be carefully chosen (Cornwell 1983; Maréchal & Lannes 1997). The strategy adopted in the basic version of WIPE is therefore preferable; its implementation is simpler and more efficient.

The condition “ $C_e$  of the order of 1 with  $\kappa_E$  less than say 5, with a sufficiently small value of  $\|\sigma_s\|_e / \|\psi_s\|_e$ ” often suffices to ensure a good solution to the problem, but strictly speaking, this is not a sufficient condition. The complete control must be based on a multiresolution strategy. The corresponding developments will be presented in a forthcoming paper.

### Appendix 1. Notion of condition number

For any  $\phi \neq 0$  in  $E$ , we have from the definitions of  $\mu$  and  $\mu'$  given in Eq. (7):

$$\mu \leq \frac{\|A_E \phi\|_d^2}{\|\phi\|_o^2} \leq \mu'.$$

For  $\phi = \delta\phi_E$ , the first inequality gives

$$\mu \leq \frac{\|\delta\psi_F\|_d^2}{\|\delta\phi_E\|_o^2},$$

whereas for  $\phi = \phi_E$  the second yields

$$\frac{\|\psi_F\|_d^2}{\|\phi_E\|_o^2} \leq \mu'.$$

By combining these inequalities, it follows that

$$\mu \frac{\|\psi_F\|_d^2}{\|\phi_E\|_o^2} \leq \mu' \frac{\|\delta\psi_F\|_d^2}{\|\delta\phi_E\|_o^2},$$

hence:

$$\frac{\|\delta\phi_E\|_o}{\|\phi_E\|_o} \leq \frac{\sqrt{\mu'}}{\sqrt{\mu}} \frac{\|\delta\psi_F\|_d}{\|\psi_F\|_d}.$$

The square root of  $\mu'/\mu$  is referred to as the condition number of  $A_E$ .

## Appendix 2. Convergence property

Using the notation introduced in Sect. 2.1, we have

$$\|z_{n+1}\|_{\mathbb{d}}^2 = \|z_n\|_{\mathbb{d}}^2 - \omega(2 - \omega) \|Q_k z_n\|_{\mathbb{d}}^2,$$

i.e., from Eq. (19):

$$\beta_{n+1} = \beta_n - \omega(2 - \omega) \eta_k^2 (h_k | z_n)_{\mathbb{d}}^2.$$

Thus (cf. Eq. (22)),

$$\beta_{n+1} \leq \beta_n - C \max_{1 \leq j \leq M} (h_j | z_n)_{\mathbb{d}}^2,$$

where  $C \triangleq \omega(2 - \omega) \min_{1 \leq j \leq M} \eta_j^2$ . As the relaxation parameter  $\omega$  is supposed to lie in the open interval  $(0, 2)$ ,  $C$  is strictly positive. As shown in remark A2, there exists a positive constant  $C'$  such that for all  $z$  in  $F_{\mathbb{d}}$ , we have:

$$\max_{1 \leq j \leq M} (h_j | z)_{\mathbb{d}}^2 \geq C' \|z\|_{\mathbb{d}}^2.$$

As a result,

$$\beta_{n+1} \leq \beta_n - C'' \beta_n$$

with  $C'' \triangleq CC' > 0$ .

Let us now assume that  $\beta$  is different from 0. There then exists  $n$  such that

$$0 < \beta_n - \beta < C'' \beta,$$

hence

$$\beta_{n+1} \leq \beta_n - C'' \beta < \beta.$$

This is impossible, since  $\beta_{n+1}$  must be greater than  $\beta$ . Consequently,  $\beta = 0$ .

REMARK A2. The property in question can be established as follows. Consider the operator on  $F_{\mathbb{d}}$ :

$$Rz \triangleq \sum_{j=1}^M (h_j | z)_{\mathbb{d}} h_j.$$

For any  $z$  and  $z'$  in  $F_{\mathbb{d}}$ , we have:

$$\begin{aligned} (z | Rz')_{\mathbb{d}} &= \sum_{j=1}^M (h_j | z')_{\mathbb{d}} (z | h_j)_{\mathbb{d}} \\ &= \sum_{j=1}^M (h_j | z)_{\mathbb{d}} (z' | h_j)_{\mathbb{d}} = (z' | Rz)_{\mathbb{d}}. \end{aligned}$$

This identity shows that  $R$  is self-adjoint. Moreover, as

$$(z | Rz)_{\mathbb{d}} = \sum_{j=1}^M (h_j | z)_{\mathbb{d}}^2,$$

the condition  $(z | Rz)_{\mathbb{d}} = 0$  implies  $z = 0$ ;  $R$  is therefore positive definite. The fact that  $F_{\mathbb{d}}$  is of finite dimension

then implies that the smallest eigenvalue of  $R$  is strictly positive. Consequently, for any  $z \in F_{\mathbb{d}}$ ,

$$\sum_{j=1}^M (h_j | z)_{\mathbb{d}}^2 \geq \lambda \|z\|_{\mathbb{d}}^2 \quad (\lambda > 0).$$

It follows immediately that

$$\max_{1 \leq j \leq M} (h_j | z)_{\mathbb{d}}^2 \geq C' \|z\|_{\mathbb{d}}^2$$

with  $C' \triangleq \lambda/M$ .

## Appendix 3. Dirty map and dirty beam

We first show that the dirty map is the map of the scalar components of  $A_e^* \psi_e$  in the basis of the elementary particles  $e_{\mathbf{p}}$ . According to Eqs. (11) and (10),  $A_e^* \psi_e$  can be expanded in the form  $A_e^* \psi_e = \sum_{\mathbf{p} \in \mathbb{G}} x_{e,\mathbf{p}} e_{\mathbf{p}}$ , where

$$\begin{aligned} x_{e,\mathbf{p}} &= \frac{1}{(\delta\xi)^2} (e_{\mathbf{p}} | A_e^* \psi_e)_{\circ} \\ &= \frac{1}{(\delta\xi)^2} (A_e e_{\mathbf{p}} | \psi_e)_{\mathbb{d}}. \end{aligned}$$

As

$$\widehat{e}_{\mathbf{p}}(\mathbf{u}) = \widehat{e}_{\circ}(\mathbf{u}) \exp\left(-2i\pi\mathbf{p} \cdot \frac{\mathbf{u}}{\Delta u}\right)$$

with

$$\widehat{e}_{\circ}(\mathbf{u}) = \frac{1}{(\Delta u)^2} \text{rect}\left(\frac{u_1}{\Delta u}\right) \text{rect}\left(\frac{u_2}{\Delta u}\right),$$

it then follows from Eq. (12) that

$$x_{e,\mathbf{p}} = \frac{1}{(\delta\xi \Delta u)^2} \sum_{\mathbf{u} \in \mathcal{L}_e} W(\mathbf{u}) \psi_e(\mathbf{u}) \exp\left(2i\pi\mathbf{p} \cdot \frac{\mathbf{u}}{\Delta u}\right) (\delta u)^2,$$

hence, since  $\delta\xi \Delta u = 1$ :

$$x_{e,\mathbf{p}} = \sum_{\mathbf{u} \in \mathcal{L}_e} W(\mathbf{u}) \psi_e(\mathbf{u}) \exp\left(2i\pi\mathbf{p} \cdot \frac{\mathbf{u}}{\Delta u}\right) (\delta u)^2.$$

This explicitly shows that  $A_e^* \psi_e$  can be identified with the dirty map (see for example Fig. 5b).

The action of  $A_e^*$  corresponds to a “back Fourier sampling operation”. The dirty map looks like the inverse Fourier transform of  $W\psi_e$ , but from a mathematical point of view, it isn’t. Indeed,  $W\psi_e$  is a vector in the experimental data space  $K_e$  and not the distribution  $\sum_{\mathbf{u} \in \mathcal{L}_e} W(\mathbf{u}) \psi_e(\mathbf{u}) \delta_{\mathbf{u}}$ . When considering the basic versions of CLEAN and WIPE, this distinction may seem to be a “mathematical stylishness”, but this is not the case, for example, in multifrequency Fourier synthesis (see the context of Eq. (68) in Lannes et al. 1996).

Let us now consider the action of  $A_e^* A_e$  on any  $\phi \in H_o$ . Setting  $\Phi \triangleq A_e^* A_e \phi$ , and expanding  $\phi$  and  $\Phi$  in the forms

$$\phi = \sum_{\mathbf{p} \in \mathbb{G}} x_{\mathbf{p}} e_{\mathbf{p}}, \quad \Phi = \sum_{\mathbf{p} \in \mathbb{G}} X_{\mathbf{p}} e_{\mathbf{p}},$$

we have:

$$\begin{aligned} X_{\mathbf{p}} &= \frac{1}{(\delta\xi)^2} (e_{\mathbf{p}} | A_e^* A_e \phi)_o = \frac{1}{(\delta\xi)^2} \sum_{\mathbf{p}' \in \mathbb{G}} x_{\mathbf{p}'} (e_{\mathbf{p}} | A_e^* A_e e_{\mathbf{p}'})_o \\ &= \frac{1}{(\delta\xi)^2} \sum_{\mathbf{p}' \in \mathbb{G}} (A_e e_{\mathbf{p}} | A_e e_{\mathbf{p}'})_d x_{\mathbf{p}'}. \end{aligned}$$

By using the same arguments as above, it then follows that  $X_{\mathbf{p}}$  can be written in the form

$$X_{\mathbf{p}} = \sum_{\mathbf{p}' \in \mathbb{G}} h_{e, \mathbf{p}-\mathbf{p}'} x_{\mathbf{p}'}$$

where

$$h_{e, \mathbf{p}} \triangleq \frac{1}{(\Delta u)^2} \sum_{\mathbf{u} \in \mathcal{L}_e} W(\mathbf{u}) \exp\left(2i\pi \mathbf{p} \cdot \frac{\mathbf{u}}{\Delta u}\right) (\delta u)^2.$$

Note that  $h_{e, -\mathbf{p}} = h_{e, \mathbf{p}}$  and  $(\delta u / \Delta u)^2 = 1/N^2$ . Let  $\mathbb{G}'$  be the grid twice as large as  $\mathbb{G}$ :

$$\mathbb{G}' \triangleq \mathbb{L}' \times \mathbb{L}', \quad \mathbb{L}' \triangleq \{\mathbf{p} \in \mathbb{Z} : -N \leq p \leq N-1\}.$$

The map of the coefficients  $h_{e, \mathbf{p}}$  on  $\mathbb{G}'$  defines what is referred to as the dirty beam DB (see for example Fig. 5a). An expression such as

$$\text{DB} \star \phi$$

then denotes the vector (lying in  $H_o$ ) whose scalar components are given by the discrete convolution:

$$\sum_{\mathbf{p}' \in \mathbb{G}} h_{e, \mathbf{p}-\mathbf{p}'} x_{\mathbf{p}'} \quad (\mathbf{p} \in \mathbb{G}).$$

As a result, in the general case where the nonzero components of  $\phi$  are distributed all over grid  $\mathbb{G}$ , the operation  $\text{DB} \star \phi$  is performed by implementing the FFT algorithm on grid  $\mathbb{G}'$ .

When  $N$  is large and the experimental frequency list very long, the direct calculation of the dirty map and the dirty beam may be very time-consuming. To save computer time, it is then preferable to use appropriate Fast Fourier Sampling techniques. The complete description of these FFS algorithms is given in Sect. 3 of Lannes et al. (1996).

#### Appendix 4. On the traditional version of CLEAN

In the classical presentation of CLEAN (Högbom 1974), the Fourier data are not damped by  $\hat{s}$ , and the convolution by the clean beam is performed a posteriori. More

precisely, the successive clean maps of the traditional version of CLEAN are given by the convolutions

$$\phi_n = \text{CB} \star \varphi_n,$$

where the iteration in  $\varphi$  is defined by the formula:

$$\varphi_0 = 0, \quad \varphi_{n+1} = \varphi_n + \omega \frac{r_{e, n, \mathbf{p}}}{h_{e, \mathbf{o}}} e_{\mathbf{p}};$$

$r_{e, n, \mathbf{p}}$  denotes the scalar component of the ‘‘experimental residue’’

$$r_{e, n} \triangleq \text{DM}_e - \text{DB} \star \varphi_n$$

at pixel  $\mathbf{p}$ , and  $h_{e, \mathbf{o}}$  that of the dirty beam at the origin:

$$r_{e, n, \mathbf{p}} = \frac{1}{(\delta\xi)^2} (e_{\mathbf{p}} | r_{e, n})_o \quad h_{e, \mathbf{o}} = \frac{1}{(\delta\xi)^2} (e_{\mathbf{o}} | A_e^* A_e e_{\mathbf{o}})_o.$$

The relaxation parameter  $\omega$  is referred to as the ‘‘loop gain’’. At each iteration,  $\mathbf{p}$  is chosen so that

$$|r_{e, n, \mathbf{p}}| = \max_{\mathbf{p}' \in \mathbb{G}_c} |r_{e, n, \mathbf{p}'}|.$$

The experimental residue  $r_{e, n}$  is obtained iteratively according to the formula:

$$r_{e, 0} = \text{DM}_e, \quad r_{e, n+1} = r_{e, n} - \omega \frac{r_{e, n, \mathbf{p}}}{h_{e, \mathbf{o}}} \text{DB}_{\mathbf{p}}.$$

Note that  $\text{DB}_{\mathbf{p}} = \text{DB} \star e_{\mathbf{p}}$ .

The objective of the iteration in  $\varphi$  is to minimize the quadratic functional:

$$\|\psi_e - A_e \varphi\|_e^2 = \sum_{\mathbf{u} \in \mathcal{L}_e} |\psi_e(\mathbf{u}) - \hat{\varphi}(\mathbf{u})|^2 W(\mathbf{u}) (\delta u)^2.$$

This iteration is nothing but the matching pursuit iteration (25) in which the vectors  $g_k$  are basis vectors  $e_{\mathbf{p}}$ ,  $\mathbf{p}$  lying in the clean box (cf. Eqs. (20) and (26); see also Appendix 3). Let us note in passing that in the traditional version of CLEAN, ‘‘these basis vectors are more or less dealt with as  $\delta$  functions’’ centred on the nodes of grid  $\mathbb{G}$ .

As far as the connection with our formulation of CLEAN is concerned, the related iterations in  $\phi$  and  $r$  are of the form

$$\phi_0 = 0, \quad \phi_{n+1} = \phi_n + \omega \frac{r_{e, n, \mathbf{p}}}{h_{e, \mathbf{o}}} \text{CB}_{\mathbf{p}}$$

and

$$r_0 = \text{DM}_s, \quad r_{n+1} = r_n - \omega \frac{r_{e, n, \mathbf{p}}}{h_{e, \mathbf{o}}} (\text{DB} \star \text{CB}_{\mathbf{p}}).$$

These iterations slightly differ from those introduced in our presentation of CLEAN (Eqs. (31) and (33)). The main difference is that the selected successive pixels  $\mathbf{p}$  are not necessarily the same: in our version of CLEAN,  $\mathbf{p}$  is chosen, at each iteration, so that  $|\rho_{n, \mathbf{p}}| = \max_{\mathbf{p}' \in \mathbb{G}_c} |\rho_{n, \mathbf{p}'}|$  (see Eq. (32)). According to the analysis presented in Sect. 5, the weak points of CLEAN related to its intrinsic instability unfortunately remain the same. As the objective of CLEAN is to find an approximation to the object function convolved by the clean beam, it is therefore more natural to consider that, basically, CLEAN is the matching pursuit process described in Sect. 2.2.

**References**

- Cornwell T.J., 1983, A&A 121, 281  
Högbom J.A., 1974, A&AS 15, 417  
Lannes A., Roques S., Casanove M.-J., 1987a, J. Mod. Opt. 34, 161  
Lannes A., Casanove M.-J., Roques S., 1987b, J. Mod. Opt. 34, 321  
Lannes A., Anterrieu E., Bouyoucef K., 1994, J. Mod. Opt. 41, 1537  
Lannes A., Anterrieu E., Bouyoucef K., 1996, J. Mod. Opt. 43, 105  
Mallat S., Zhang Z., 1993, in: Progress on Wavelet Analysis and Applications, Meyer Y. and Roques S. (eds.). Éditions Frontières, Gif-sur-Yvette (France) p. 155  
Maréchal P., Lannes A., 1997, Inv. Prob. 13, 135  
Press W.H., Teukolsky S.A., Vetterling W.T., Slannery B.P., 1992, Numerical recipes in C. The art of scientific computing, second edition, Cambridge, New York, New Rochelle, Melbourne and Sydney: Cambridge University Press  
Schwarz U.J., 1978, A&A 65, 345  
Slepian D., 1965, J. Opt. Soc. Am. 55, 1110