The Astrophysical Journal, 638:488–498, 2006 February 10

HIGH-CONTRAST IMAGING FROM SPACE: SPECKLE NULLING IN A LOW-ABERRATION REGIME

Pascal J. Bordé[1]
Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138; pborde@cfa.harvard.edu

and

Wesley A. Traub[2]
Jet Propulsion Laboratory, MS 301-451, 4800 Oak Grove Drive, Pasadena, CA 91109; wtraub@jpl.nasa.gov
Received 2005 April 1; accepted 2005 October 9

ABSTRACT

High-contrast imaging from space must overcome two major noise sources to successfully detect a terrestrial planet angularly close to its parent star: photon noise from diffracted starlight and speckle noise from starlight scattered by instrumentally generated wave front perturbation. Coronagraphs tackle only the photon noise contribution by reducing diffracted starlight at the location of a planet. Speckle noise should be addressed with adaptive optics systems. Following the tracks of Malbet, Yu, and Shao, we develop in this paper two analytical methods for wave front sensing and control that aims at creating "dark holes," i.e., areas of the image plane cleared of speckles, assuming an ideal coronagraph and small aberrations. The first method, "speckle field nulling," is a fast FFT-based algorithm that requires the deformable-mirror influence functions to have identical shapes. The second method, "speckle energy minimization," is more general and provides the optimal deformable mirror shape via matrix inversion. With an $N \times N$ deformable mirror, the size of the matrix to be inverted is either $N^2 \times N^2$ in the general case or only $N \times N$ if the influence functions can be written as the tensor product of two one-dimensional functions. Moreover, speckle energy minimization makes it possible to trade off some of the dark hole area against an improved contrast. For both methods, complex wave front aberrations (amplitude and phase) are measured using just three images taken with the science camera (no dedicated wave front sensing channel is used); therefore, there are no noncommon path errors. We assess the theoretical performance of both methods with numerical simulations including realistic speckle noise and experimental influence functions. We find that these speckle-nulling techniques should be able to improve the contrast by several orders of magnitude.

Subject headings: instrumentation: adaptive optics — planetary systems — techniques: high angular resolution

1. INTRODUCTION

The field of extrasolar planet research has recently made a leap forward with the direct detection of extrasolar giant planets (EGPs). Using the *Spitzer Space Telescope*, Charbonneau et al. (2005) and Deming et al. (2005) have detected infrared photons from two transiting planets, TrES-1 and HD 209458b, respectively. Chauvin et al. (2004, 2005) have reported the infrared imaging of an EGP orbiting the nearby young brown dwarf 2M 1207 (2MASSW J1207334−393254) with NACO (association of NAOS [Nasmyth Adaptive Optics System] and CONICA [High Resolution Near Infrared Camera]) at VLT (Very Large Telescope), whereas Neuhäuser et al. (2005) have collected evidence for an EGP companion to the T Tauri star GQ Lup using VLT NACO as well.

Although there are claims that the direct detection of terrestrial planets could be performed from the ground with—yet to come—extremely large telescopes (Angel 2003; Chelli 2005), it is widely believed that success will be more likely in space. Direct detection is the key to spectroscopy of planetary atmospheres and discovery of biomarkers, namely, indirect evidence of life developed at the planetary scale (e.g., Des Marais et al. 2002).

Both NASA and ESA have space mission studies well underway to achieve this task. *Darwin*, the European mission to be launched in 2015, will be a thermal infrared nulling interferometer with three 3.5 m free-flying telescopes (Karlsson et al. 2004). *Terrestrial Planet Finder*, the American counterpart, will feature two missions: *Terrestrial Planet Finder–Coronagraph* (*TPF-C*), an 8 m × 3.5 m monolithic visible telescope equipped with a coronagraph to be launched in 2015, and *Terrestrial Planet Finder–Interferometer* (*TPF-I*), an analog to *Darwin* to be launched sometime in 2015–2019 (Coulter 2004).

The direct detection of the photons emitted by a terrestrial planet is made very challenging by the angular proximity of the parent star and by the very high contrast (i.e., luminosity ratio) between the planet and its star: about 10^6 in the thermal infrared and about 10^{10} in the visible. Both wavelength ranges have their scientific merits and technical difficulties, and both of them are thought to be necessary for an unambiguous detection of habitability and signs of life (e.g., Des Marais et al. 2002). In this paper, we deal with the visible range only.

In the visible, planet detection faces two fundamental noise sources: (1) quantum noise of the diffracted starlight and (2) speckle noise due to the scattering of the starlight by optical defects. Labeyrie (1995) proposed a technique based on "dark speckles" to overcome speckle noise: random fluctuations of the atmosphere cause the speckles to interfere destructively and disappear at certain locations in the image, thus creating localized dark spots suitable for planet detection. The statistical analysis of a large number of images then reveals the planet as a spot persistently brighter than the background.

Malbet et al. (1995) proposed using a deformable mirror (DM) instead of the atmosphere to make speckles interfere destructively

[1] Michelson Postdoctoral Fellow; current address: Michelson Science Center, California Institute of Technology, 770 South Wilson Avenue, Pasadena, CA 91125.

[2] Research Associate, Harvard-Smithsonian Center for Astrophysics.

in a targeted region of the image called a "search area" or "dark hole" (DH or $\mathcal{H}$). Following the tracks of these authors, this paper discusses methods for reducing the speckle noise below the planet level by using a DM and an ideal coronagraph. However, unlike Malbet et al. (1995), we propose noniterative algorithms, in order to limit the number of long exposures needed for terrestrial planet detection. We refer to these methods as "speckle-nulling" techniques, as Trauger et al. (2004) call them. Technical aspects of this work are inspired by the High Contrast Imaging Testbed (HCIT; Trauger et al. 2004), a speckle-nulling experiment hosted at the Jet Propulsion Laboratory, specifically designed to test *TPF-C* related technology.

After reviewing the process of speckle formation to establish our notations (§ 2), we derive two speckle-nulling methods in the case of small aberrations (§ 3). The speckle-nulling phase is preceded by the measurement of the electric field in the image plane (§ 3.4). The performance of both methods is then evaluated with one- and two-dimensional simulations (§ 4), first with white speckle noise (§ 4.1), then with nonwhite speckle noise (§ 4.2). Various effects and instrumental noises are considered in § 5. Finally, we conclude and discuss some future work (§ 6).

2. SPECKLE FORMATION

This paper is written in the framework of Fourier optics considering a single wavelength, knowing that a more sophisticated theory (scalar or vectorial) in polychromatic light will eventually be needed. Fourier transforms (FTs) are signaled by a hat.

Let us consider a simple telescope with an entrance pupil $\mathcal{P}$. In the pupil plane, we use the reduced coordinates $(u, v) = (x/\lambda, y/\lambda)$, where (x, y) are distances in meters and λ is the wavelength. We define the pupil function as

$$P(u,v) \equiv \begin{cases} 1 & \text{if } (u,v) \in \mathcal{P}, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Even in space, i.e., when not observing through a turbulent medium such as the atmosphere, the optical train of the telescope is affected by phase and amplitude aberrations. Phase aberrations are wave front corrugations that typically originate in mirror roughness caused by imperfect polishing, while amplitude aberrations are typically the result of a heterogeneous transmission or reflectivity. Moreover, Fresnel propagation turns phase aberrations into amplitude aberrations and the reverse (e.g., Guyon 2005). Regardless of where they originate physically, all phase and amplitude aberrations can be represented by a complex aberration function ϕ in a reimaged pupil plane, so that the aberrated pupil function is now $Pe^{i\phi}$.

The electric field associated with an incident plane wave of amplitude unity is then

$$E(u,v) = P(u,v)e^{i\phi(u,v)}. \tag{2}$$

Exoplanet detection requires that we work in a regime in which aberrations are reduced to a small fraction of the wavelength. Once in this regime, we can replace $e^{i\phi}$ by its first-order expansion $1 + i\phi$ (we discuss in § 5.2 the validity of this approximation). Since the electric field in the image plane is the FT of equation (2), we get

$$\hat{E}(\alpha,\beta) = \hat{P}(\alpha,\beta) + i\widehat{P\phi}(\alpha,\beta), \tag{3}$$

where (α, β) are angular coordinates in the image plane.

The physical picture is as follows. The first term ($\hat{P}$) is the direct image of the star. The second term ($\widehat{P\phi}$) is the field of speckles surrounding the central star image, where each speckle is generated by the equivalent of first-order scattering from one of the sinusoidal components of the complex aberration ϕ. Each speckle is essentially a ghost of the central point-spread function (PSF).

In the remainder of this paper, we focus on means of measuring and correcting the speckles in a coronagraphic image. Following Malbet et al. (1995), we leave out the unaberrated PSF term by assuming that it was canceled out by a coronagraph of some sort (see Quirrenbach 2005 for a review on coronagraphs). Thus, we clearly separate the gain in the contrast that can be obtained, on the one hand, by reducing the diffracted light with the coronagraph and, on the other hand, by fighting the scattered light with the speckle-nulling technique.

3. SPECKLE-NULLING THEORY

The purpose of speckle nulling is to reduce the speckle noise in a central region of the image plane. This region, the dark hole, then becomes dark enough to enable the detection of companions much fainter than the original speckles. Speckle nulling is achieved by way of a servo system that has a deformable mirror as an actuator. Because our sensing method requires DM actuation and is better understood with the knowledge of command control theory, we first model the deformable mirror (§ 3.1), then present two algorithms for the command control (§§ 3.2 and 3.3), and finally conclude with the sensing method (§ 3.4).

3.1. *Deformable Mirror*

The deformable mirror (DM) in Trauger et al. (2003) consists of a continuous face sheet supported by $N \times N$ actuators arranged in a square pattern of constant spacing. This DM format is well adapted to either square or circular pupils, the only pupil shapes that we consider in this paper.[3] We assume that the DM is physically located in a plane that is conjugate to the entrance pupil. However, what we call "DM" in the following is the projection of this real DM in the entrance pupil plane. The projected spacing between actuators is denoted by d. We assume that the optical magnification is such that the DM projected size is matched to the entrance pupil, i.e., $Nd = D$, where D is either the pupil side length or its diameter. The DM surface deformation in response to the actuation of actuator $(k, l) \in \{0, \ldots, N-1\}^2$ is described by an "influence function," denoted by f_{kl}. The total phase change introduced by the DM (DM phase function) is

$$\psi(u,v) \equiv \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} a_{kl} f_{kl}(u,v), \tag{4}$$

where a_{kl} are actuator strokes (measured in radians). Note that contrary to the complex aberration function ϕ, the DM phase function is purely real.

With an ideal coronagraph and a DM, the image-plane electric field formerly given by equation (3) becomes

$$\hat{E}'(\alpha,\beta) = i\widehat{P\phi}(\alpha,\beta) + i\widehat{P\psi}(\alpha,\beta). \tag{5}$$

In the next two sections, we explore two approaches for speckle nulling. In § 3.2, we begin naively by trying to cancel $\hat{E}'$. Because there is a maximum spatial frequency that the DM can correct for, the DH necessarily has a limited extension. Any energy at higher spatial frequencies will be aliased in the DH and

[3] Two square DMs can be assembled to accommodate an elliptical pupil such as that envisioned for *TPF-C*.

limit its depth. Therefore, the DM cannot be driven to cancel $\hat{E}'$ unless $\widehat{P\phi} = 0$ outside the DH (i.e., unless there are already no speckles outside the DH). With this in mind, we start over in § 3.3 with the idea that speckle nulling is better approached by minimizing the field energy.

3.2. *Speckle Field Nulling*

The speckle field nulling approach consists in trying to null out $\hat{E}'$ in the DH region ($\mathcal{H}$), meaning we seek a solution to the equation

$$\forall(\alpha,\beta)\in\mathcal{H},\quad \widehat{P\phi}(\alpha,\beta)+\widehat{P\psi}(\alpha,\beta)=0, \tag{6}$$

although, as we show, this equation has no exact solution unless $\widehat{P\phi}$ happens to be a band-limited function within the controllable band of the DM.

By replacing ψ with its expression (4), we obtain

$$\forall(\alpha,\beta)\in\mathcal{H},\quad \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} a_{kl}\widehat{Pf_{kl}}(\alpha,\beta)=-\widehat{P\phi}(\alpha,\beta). \tag{7}$$

We recognize in system (7) a linear system in a_{kl} that could be solved using various techniques such as singular value decomposition (SVD; Press et al. 2002, § 2.6). Although it is general, this solution does not provide much insight into the problem of speckle nulling. For this reason, let us now examine a different solution, less general but with more explanatory power. We comment on the use of SVD at the end of this section.

We consider a square pupil. In this case, all DM actuators receive light and the pupil function has no limiting effect on the DM phase function, i.e., $P\psi = \psi$. Moreover, we assume that all influence functions are identical in shape and write $f_{kl}(u,v) = f(u-k(d/\lambda), v-l(d/\lambda))$. Under these hypotheses,

$$P\psi(u,v)=f(u,v)*\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} a_{kl}\delta\left(u-k\frac{d}{\lambda}, v-l\frac{d}{\lambda}\right), \tag{8}$$

where δ is Dirac's bidimensional distribution and the asterisk denotes the convolution.

Substituting $\widehat{P\psi}$ with the FT of equation (8) in equation (6) yields

$$\forall(\alpha,\beta)\in\mathcal{H},\ \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} a_{kl}e^{-i(2\pi d/\lambda)(k\alpha+l\beta)}=-\frac{\widehat{P\phi}(\alpha,\beta)}{\hat{f}(\alpha,\beta)}. \tag{9}$$

We recognize in the left-hand side of equation (9) a truncated Fourier series. If we choose the a_{kl} to be the first N^2 Fourier coefficients of $-\widehat{P\phi}/\hat{f}$, i.e.,

$$a_{kl}=\frac{2d^2}{\lambda^2}\int\int_{[-\lambda/2d,\,\lambda/2d]^2} -\frac{\widehat{P\phi}(\alpha,\beta)}{\hat{f}(\alpha,\beta)}e^{i(2\pi d/\lambda)(k\alpha+l\beta)}\,d\alpha\,d\beta, \tag{10}$$

then according to Fourier theory, we minimize the mean-square error between both sides of the equation (see, e.g., Hsu 1967, § 1.5). This error cannot be reduced to zero unless the Fourier coefficients of $-\widehat{P\phi}/\hat{f}$ happen to vanish for $k,l<0$ and $k,l>N$. At this point, we have reached the important conclusion that *perfect speckle cancellation cannot be achieved with a finite-size DM unless the wave front aberrations are band-limited*. Moreover, we can assert that the maximum DH extension is the square domain $\mathcal{H}\equiv[-\lambda/2d,\lambda/2d]^2=[-(N/2)(\lambda/D),(N/2)(\lambda/D)]^2$.

Solution (10) is physically acceptable only if the Fourier coefficients are real numbers, which means mathematically that $\widehat{P\phi}/\hat{f}$ should be Hermitian.[4] If there are phase aberrations only, then $P\phi$ is real, $\widehat{P\phi}/\hat{f}$ is Hermitian, and the a_{kl} are real. This is no longer true if there are amplitude aberrations as well, reflecting the fact that the DM alone cannot correct both phase and amplitude aberrations in $\mathcal{H}$. However, by considering the Hermitian function that is equal to $\widehat{P\phi}/\hat{f}$ in one-half of the DH, say, $\mathcal{H}^+\equiv[0,\lambda/2d]\times[-\lambda/2d,\lambda/2d]$, we obtain the real coefficients,

$$a_{kl}=4d^2\int\int_{\mathcal{H}^+} -\frac{\widehat{P\phi}(\alpha,\beta)}{\hat{f}(\alpha,\beta)}\cos\left[\frac{2\pi d}{\lambda}(k\alpha+l\beta)\right]d\alpha\,d\beta, \tag{11}$$

that correct both amplitude and phase aberrations in $\mathcal{H}^+$. As we have $\lambda/2d=(N/2)(\lambda/d)$, the DH has a size of $N\times N$ resolution elements (resels) with phase aberrations only and $(N/2)\times N$ resels with phase and amplitude aberrations. Therefore, *a DM can correct both amplitude and phase aberrations in the image plane, albeit in a region that is either the left, right, upper, or lower half of the phase-corrected region.*

As Malbet et al. (1995) pointed out, let us remind the reader that $\lambda/2d$ is equal to the Nyquist frequency for a sampling interval d/λ. Therefore, we find that the maximum extension for the DH corresponds to the range in which the sampling theorem applies to the wave front at the DM actuator scale. Indeed, taking the inverse FT of equation (9) leads to the wave front reconstruction formula

$$P\phi(u,v)=-\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} a_{kl}f\left(u-k\frac{d}{\lambda}, v-l\frac{d}{\lambda}\right). \tag{12}$$

Again, this reconstruction cannot be perfect unless the spectrum of $P\phi$ is contained in $\mathcal{H}$. Note that because $\hat{f}$ is generally not a flat function (as would be the case if influence functions were, for instance, two-dimensional sinc functions), the actuator strokes are not equal to the negative of the wave front values sampled at the actuator locations.

Our Fourier solution was derived by assuming (1) that all influence functions are identical in shape and (2) that the pupil has a square shape. Hypothesis 1 appears to be reasonable at least for the DM in use on the HCIT (J. Green 2005, private communication), but this remains to be precisely measured. If hypothesis 2 is relaxed, then (*a*) some actuators do not receive any light and play no role, so there are effectively fewer terms in the summation in equation (9), and (*b*) the fact that influence functions on the pupil boundary are only partly illuminated is ignored.

Now that we have two methods to solve equation (6), Fourier expansion and SVD, let us compare their solutions. We deal here with functions belonging to the Hilbert space of square integrable functions $f:\mathcal{H}\to\mathbb{C}$. This space has $\langle f,g\rangle\equiv\int\int_{\mathcal{H}} fg^*$ for the dot product and $\|f\|\equiv(\int\int_{\mathcal{H}}|f|^2)^{1/2}$ for the norm. As mentioned above, Fourier expansion minimizes the mean-square error between both sides of equation (9), i.e., $||(\widehat{P\phi}+\widehat{P\psi})/\hat{f}||^2$. By contrast, SVD has the built-in property of minimizing the norm of the residuals of equation (7), i.e., $||\widehat{P\phi}+\widehat{P\psi}||$. In other words, SVD minimizes $||\hat{E}'||^2$, the speckle field energy, which

[4] A function f is said to be Hermitian if $\forall(x,y),\ f(x,y)=f^*(-x,-y)$. The FT of a real function is Hermitian and vice versa.

seems more satisfactory from a physical point of view. To find out which is best, we have performed one-dimensional numerical simulations. It turns out that SVD yields dark holes that are 50% deeper (median value) than does the Fourier expansion. In addition, SVD does not require all influence functions to have the same shape.

However, considering four detector pixels per resel in two dimensions (critical sampling), SVD would require us to manipulate matrices as large as $N^2 \times 4N^2$ (or even $N^2 \times 8N^2$ when real and imaginary parts are separated). Such matrices would occupy 537 MB of memory space for 64×64 actuators and single-precision floating-point numbers. By contrast, Fourier expansion would be straightforwardly obtained with fast FTs (FFTs) of $2N \times 2N$ arrays at critical sampling, but again at the cost of a 50% shallower dark hole and a strong hypothesis on the influence functions. In the next section, we seek to find a computationally less intensive solution that still minimizes the speckle energy in the dark hole but does not require any hypothesis on the influence functions.

3.3. *Speckle Energy Minimization*

Let us start with the idea that the best solution is defined as that *minimizing the total energy of the speckle field in the DH*. For the sake of simplicity, we assume once again a square pupil but not necessarily a common shape for the influence functions. The total energy in the speckle field reads

$$\mathcal{E} \equiv \iint_{\mathcal{H}} \left| \widehat{P\phi}(\alpha, \beta) + \hat{\psi}(\alpha, \beta) \right|^2 d\alpha \, d\beta = \left\langle \widehat{P\phi} + \hat{\psi}, \widehat{P\phi} + \hat{\psi} \right\rangle, \tag{13}$$

using the same notation as in § 3.2.

Given that $\partial \hat{\psi} / \partial a_{kl} = \hat{f}_{kl}$, the energy is minimized when

$$\forall (k, l) \in \{0, \ldots, N-1\}^2, \quad \frac{\partial \mathcal{E}}{\partial a_{kl}} = 0 \Longleftrightarrow \mathrm{Re}\left(\left\langle \widehat{P\phi} + \hat{\psi}, \hat{f}_{kl} \right\rangle \right) = 0, \tag{14}$$

where "Re" stands for the real part. Note that this is less demanding than equation (6), as equation (6) implies equation (14) but the reverse is not true.

Using definition (4) for ψ and realizing that $\langle \hat{f}_{nm}, \hat{f}_{kl} \rangle$ is a real number,[5] we finally get

$$\forall (k, l) \in \{0, \ldots, N-1\}^2, \quad \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} a_{nm} \langle \hat{f}_{nm}, \hat{f}_{kl} \rangle = -\mathrm{Re}\left(\left\langle \widehat{P\phi}, \hat{f}_{kl} \right\rangle \right). \tag{15}$$

As in equation (7), we find a system that is linear in the actuator strokes. By replacing double indices with single ones, e.g., (k, l) becomes $s = kN + l$, equation (15) can be solved in matrix format by inverting an $N^2 \times N^2$ real matrix. This is already an improvement with respect to the $N^2 \times 4N^2$ complex matrix required by SVD in § 3.2.

It appears that the same solution can be obtained with a much less demanding $N \times N$ matrix inversion, provided that two-dimensional influence functions can be written as the tensor product of two one-dimensional functions (separation of variables), i.e., $f_{kl}(u, v) = g_k(u) g_l(v)$. This would be the case for box functions or bidimensional Gaussians and is good at the 5% level for the DM in use on the HCIT. This property also holds in the image plane, since the FT of the previous equation yields $\hat{f}_{kl}(\alpha, \beta) = \hat{g}_k(\alpha) \hat{g}_l(\beta)$.

By separating variables, equation (15) becomes

$$\forall (k, l) \in \{0, \ldots, N-1\}^2, \quad \sum_{n=0}^{N-1} \langle \hat{g}_n, \hat{g}_k \rangle \sum_{m=0}^{N-1} a_{nm} \langle \hat{g}_m, \hat{g}_l \rangle = -\mathrm{Re}\left(\left\langle \widehat{P\phi}, \hat{f}_{kl} \right\rangle \right). \tag{16}$$

As the left-hand side happens to be the product of three $N \times N$ matrices, equation (16) can be rewritten as an equality between square matrices:

$$GAG = \Phi, \quad \text{where} \begin{cases} G_{kl} = \langle \hat{g}_k, \hat{g}_l \rangle, \\ A_{kl} = a_{kl}, \\ \Phi_{kl} = -\mathrm{Re}\left(\left\langle \widehat{P\phi}, \hat{f}_{kl} \right\rangle \right). \end{cases} \tag{17}$$

For square-box and actual HCIT influence functions, numerical calculations show that G is diagonally dominant[6] and therefore invertible by regular Gaussian elimination. The solution to equation (17) is then

$$A = G^{-1} \Phi G^{-1}. \tag{18}$$

Note that G^{-1} can be precomputed and stored, so that computing the strokes effectively requires only two matrix multiplications. As shown in the Appendix, an equivalent result can be obtained by working with pupil-plane quantities.

As for the field-nulling approach, correcting amplitude errors also implies restricting the dark hole either to $\mathcal{H}^+ = [0, (N/2)(\lambda/D)] \times [-(N/2)(\lambda/D), (N/2)(\lambda/D)]$ or else to $\mathcal{H}^- = [-(N/2)(\lambda/D), 0] \times [-(N/2)(\lambda/D), (N/2)(\lambda/D)]$. To account for amplitude errors and keep the formalism we have presented so far, it is sufficient to replace $\widehat{P\phi}$ with a function equal to $\widehat{P\phi}(\alpha, \beta)$ in either $\mathcal{H}^+$ or $\mathcal{H}^-$ (depending on the half where one wishes to create the dark hole) and equal to $\widehat{P\phi}^*(-\alpha, -\beta)$ in the other half (Hermitian symmetry). Because its FT is Hermitian, the new aberration function in the pupil plane is real, and thus the algorithm processes amplitude and phase errors at the same time as if there were phase errors only.

Let us derive the residual total energy in the DH after the correction has been applied. Starting from definition (13) and rewriting condition (14) as $\mathrm{Re}(\langle \widehat{P\phi} + \hat{\psi}, \hat{\psi} \rangle) = 0$, we find

$$\mathcal{E}_{\min} = \left\langle \widehat{P\phi}, \widehat{P\phi} \right\rangle - \langle \hat{\psi}, \hat{\psi} \rangle. \tag{19}$$

The former term is the initial speckle energy in the DH, while the latter is the speckle energy decrease gained with the DM. Mathematically, $\mathcal{E}_{\min}^{1/2}$ measures the distance (according to the norm we have defined) between the speckle field and its approximation with the DM inside the DH. Because there is no exact solution to equation (6), the residual energy cannot be made equal to zero in $\mathcal{H}^+$ or $\mathcal{H}^-$. However, the energy approach offers an additional degree of freedom: by concentrically reducing the domain over which the energy is minimized, the speckle energy can be further decreased (see § 4).

[5] This property stems from the Hermitian character of $\hat{f}_{kl}$ together with the symmetry of $\mathcal{H}$.

[6] A matrix $A = [a_{ij}]$ is said to be diagonally dominant if $\forall i, |a_{ii}| > \sum_{j \neq i} |a_{ij}|$.

3.4. *Speckle Field Measurement*

So far, our speckle-nulling theory has presupposed knowledge of the speckle field $\widehat{P\phi}$ or, equivalently, of the phase and amplitude aberrations across the pupil, embodied in the complex phase function $P\phi$. In this section, we show how the speckle field can be measured directly in the image plane. As the detector measures an intensity, a single image yields only the modulus of the speckle field. The phase of the speckle field can be retrieved by perturbing the phase function $P\phi$ in a controlled way and by recording the corresponding images, a process analogous to "phase diversity" (e.g., Löfdahl & Scharmer 1994). In our system, the DM provides the natural means for creating this controlled perturbation.

As we show, exactly three images obtained with well-chosen DM settings provide enough information to measure $\widehat{P\phi}$. Let us call image 0 the original image recorded with setting ψ_0, whereas images 1 and 2 are recorded with settings $\psi_0 + \delta\psi_1$ and $\psi_0 + \delta\psi_2$.

To be general, we consider in the field of view the presence of an exoplanet and an exozodiacal cloud (hereafter exozodi), in addition to the star itself. The electric fields of these objects are incoherent with that of the star, so their intensities should be added to the star's intensity. Because they are much fainter than the star, the speckles they produce are negligible with respect to the star speckles, and their intensities can be considered as independent of ϕ and ψ. The total intensity of every image pixel (α, β) then takes the successive values

$$\begin{cases} I_0 = \left|\widehat{P\phi} + \hat{\psi}_0\right|^2 + I_p + I_z, \\ I_1 = \left|\widehat{P\phi} + \hat{\psi}_0 + \widehat{\delta\psi_1}\right|^2 + I_p + I_z, \\ I_2 = \left|\widehat{P\phi} + \hat{\psi}_0 + \widehat{\delta\psi_2}\right|^2 + I_p + I_z, \end{cases} \tag{20}$$

where I_p and I_z are the exoplanet and exozodi intensities, respectively.

System (20) can be reduced to the linear system

$$\begin{cases} \left(\widehat{\delta\psi_1}\right)^* \left(\widehat{P\phi} + \hat{\psi}_0\right) + \widehat{\delta\psi_1}\left(\widehat{P\phi} + \hat{\psi}_0\right)^* = I_1 - I_0 - \left|\widehat{\delta\psi_1}\right|^2, \\ \left(\widehat{\delta\psi_2}\right)^* \left(\widehat{P\phi} + \hat{\psi}_0\right) + \widehat{\delta\psi_2}\left(\widehat{P\phi} + \hat{\psi}_0\right)^* = I_2 - I_0 - \left|\widehat{\delta\psi_2}\right|^2, \end{cases} \tag{21}$$

where the asterisk in the exponent denotes the complex conjugate. Note how the exoplanet and exozodi intensities have disappeared from the equations, demonstrating that faint objects do not affect the measurement process of stellar speckles. However, note that because of quantum noise, the planet detection can still be problematic if the exozodi is much brighter than the planet.

Now, system (21) admits a unique solution if its determinant,

$$\Delta \equiv \left(\widehat{\delta\psi_1}\right)^* \widehat{\delta\psi_2} - \widehat{\delta\psi_1}\left(\widehat{\delta\psi_2}\right)^*, \tag{22}$$

is not zero, that is, if

$$\left|\widehat{\delta\psi_1}\right|\left|\widehat{\delta\psi_2}\right| \sin\left[\arg\left(\widehat{\delta\psi_2}\right) - \arg\left(\widehat{\delta\psi_1}\right)\right] \neq 0. \tag{23}$$

Condition (23) tells us that the DM setting changes, $\delta\psi_1$ and $\delta\psi_2$, should modify the speckles differently in any given pixel; otherwise, not enough information is secured to measure $\widehat{P\phi}$ unambiguously in this pixel. It should be expected for this method to work practically that the magnitude of the speckle modification be greater than the photon noise level.

We have not yet found a rigorous derivation of the optimum values for the amplitude $|\widehat{\delta\psi_1}|$ and $|\widehat{\delta\psi_2}|$, but a heuristic argument suggests to us that the optimum perturbations may be that $I_1 \approx I_0$ and $I_2 \approx I_0$. That is to say, the DM-induced speckle intensity pattern, taken by itself, should be approximately the same as the original speckle intensity pattern. Thus, at each pixel we choose $|\widehat{\delta\psi_1}| \approx |\widehat{\delta\psi_2}| \approx I_0^{1/2}$, with the caveat that neither should be zero to keep equation (23) valid.

The phase of $\widehat{\delta\psi_1}$ does not matter, but the phase difference between $\widehat{\delta\psi_1}$ and $\widehat{\delta\psi_2}$ should be made as close to $\pi/2$ as possible to keep Δ from zero. Practically, this can be realized as follows:

1. Compute $\delta\psi_1$ stroke changes from equation (11) or equation (18) by replacing $\widehat{P\phi}$ with $I_0^{1/2} e^{i\theta}$, where θ is a random phase.
2. Compute $\delta\psi_2$ stroke changes from equation (11) or equation (18) by replacing $\widehat{P\phi}$ with $\widehat{\delta\psi_1} e^{i\pi/2}$.

Now that we have made sure that $\Delta \neq 0$, we finally derive

$$\widehat{P\phi} = \frac{\widehat{\delta\psi_2}\left(I_1 - I_0 - \left|\widehat{\delta\psi_1}\right|^2\right) - \widehat{\delta\psi_1}\left(I_2 - I_0 - \left|\widehat{\delta\psi_2}\right|^2\right)}{\Delta} - \hat{\psi}_0. \tag{24}$$

Equation (24) shows that the initially unknown speckle field $(\widehat{P\phi})$ can be experimentally measured in just three exposures taken under identical circumstances but with different shapes imposed on the DM.

4. SPECKLE-NULLING SIMULATIONS

4.1. *White Speckle Noise*

In this section, we perform one- and two-dimensional simulations for the theoretical case of white speckle noise caused by phase aberrations only. The DM has 64 actuators and top-hat influence functions. Smoother influence functions have been tested and do not lead to qualitatively different results. A simulation with actual HCIT influence functions is presented in § 4.2. The simulated portion of the pupil plane is made twice as large as the pupil by zero padding, so that every element of resolution in the image plane would be sampled by two detector pixels. This corresponds to the realistic case of a photon-starved exoplanet detection for which readout noise must be minimized.

4.1.1. *One-dimensional Simulations*

Figure 1 shows a complete one-dimensional simulation including speckle field measurement (§ 3.4) and speckle nulling with field nulling (§ 3.2) and energy minimization (§ 3.3). The standard deviation of the phase aberrations is set to $\lambda/1000$. Intensities are scaled with respect to the maximum of the star PSF in the absence of a coronagraph. Ideal conditions are assumed: no photon noise, a noiseless detector, and perfect precision in the control of the DM actuators. Under these conditions, the speckle field is perfectly estimated, and the mean intensity in the DH is 5.8×10^{-11} with field nulling and 1.4×10^{-11} with energy minimization, i.e., about 1500 and 6500 times lower than the mean intensity outside the DH, respectively. Repeated simulations with different noise sequences show that energy minimization always

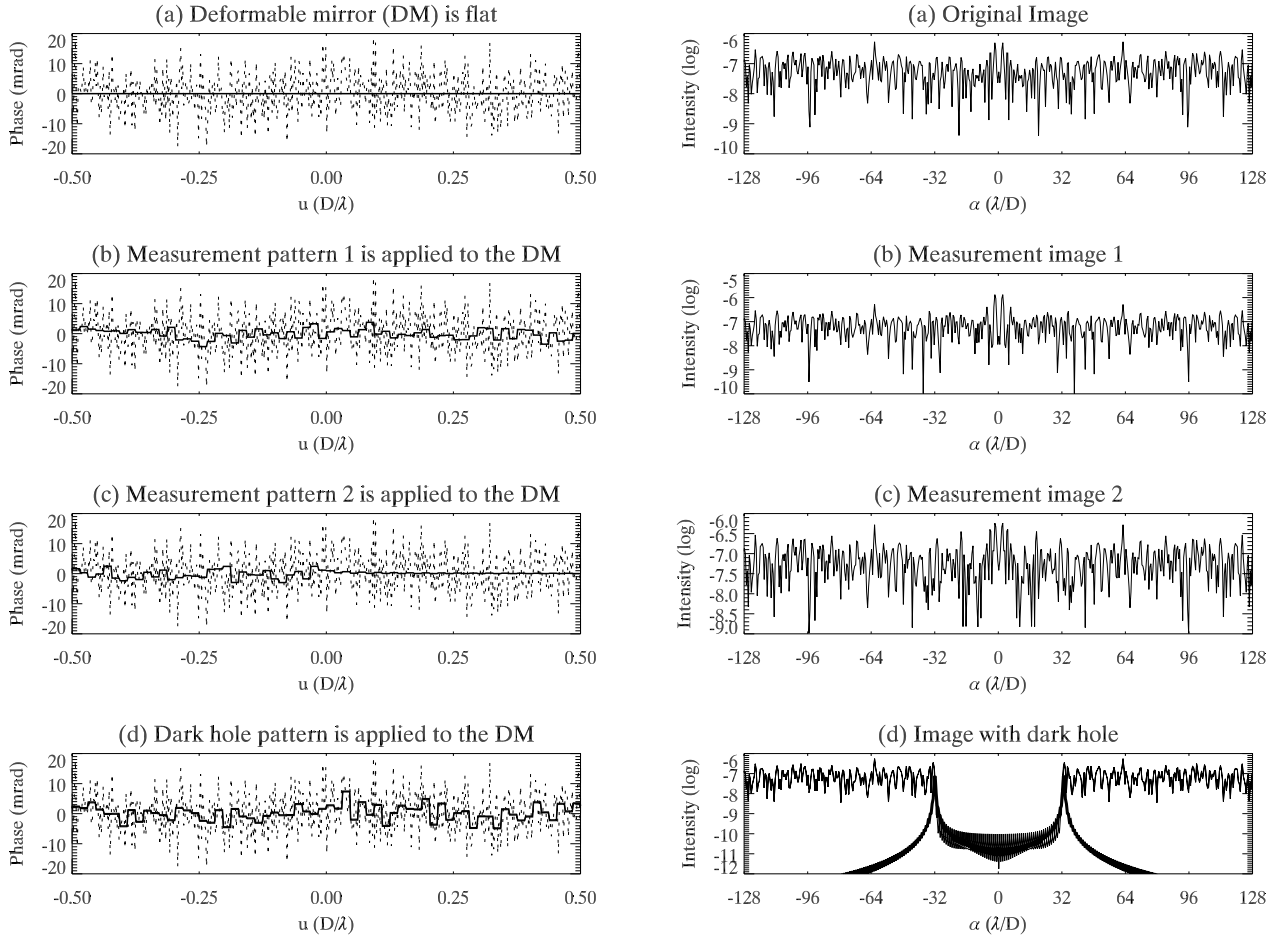

FIG. 1.—Full one-dimensional speckle-nulling simulation for a one-dimensional pupil with 64 actuators. *Left:* Phase aberrations (*dotted lines*) and their low spatial frequency approximations at the actuator scale (*solid lines*; negative of the patterns applied to the DM). *Right:* Corresponding images. The full algorithm is a four-step process: (*a*) Original speckles are measured with the current DM shape (taken here to be flat). (*b*, *c*) Speckles are modified by driving the DM. At this point, enough information has been gathered to deduce the phase and amplitude aberrations of the wave front. (*d*) Low spatial frequency aberrations are corrected with the DM, canceling out speckles in the center of the image. Dark holes produced by the field-nulling and energy-minimization approaches (§§ 3.2 and 3.3) do not differ noticeably in the figure (*thin solid line*). A rough estimate based on the mean intensity prior to the dark hole creation (§ 4.1) is superimposed as a thick solid line.

performs better than field nulling by a factor of a few. Field nulling solved with SVD yields the same numerical solution as that of energy minimization (they differ by the last digit only), in agreement with the idea that they both minimize speckle energy.

4.1.2. *Dark Hole Depth Estimate in One Dimension*

In the one-dimensional case, it is easy to roughly predict the shape and the depth of the DH. The function $\widehat{P\phi} + \hat{\psi}$ is band-limited, since the pupil has a finite size. As the pupil linear dimension is D/λ, the maximum spatial frequency of $\widehat{P\phi} + \hat{\psi}$ is $D/2\lambda$. Let us apply the sampling theorem at the Nyquist sampling frequency D/λ and write

$$\left(\widehat{P\phi} + \hat{\psi}\right)(\alpha) = \sum_{n=-\infty}^{+\infty} \left(\widehat{P\phi}_n + \hat{\psi}_n\right) \mathrm{sinc}\left(\frac{\alpha D}{\lambda} - n\right), \quad (25)$$

where the subscript n denotes the function value for $\alpha = n(\lambda/D)$.

Substituting α with $n(\lambda/D)$ and d with D/N leads to

$$\widehat{P\phi}_n + \hat{\psi}_n = \widehat{P\phi}_n + \hat{f}_n \sum_{k=0}^{N-1} a_k e^{-i(2\pi kn/N)}. \quad (26)$$

The field-nulling equation (6) here takes the discrete form

$$\forall n \in \{0, \ldots, N-1\},$$
$$\widehat{P\phi}_n + \hat{\psi}_n = 0 \Longleftrightarrow a_k = \sum_{n=0}^{N-1} \left(-\frac{\widehat{P\phi}_n}{\hat{f}_n}\right) e^{i(2\pi kn/N)}; \quad (27)$$

i.e., the actuator strokes are computed thanks to an inverse FFT.

Let us now turn to the residual speckle field

$$\left(\widehat{P\phi} + \hat{\psi}\right)(\alpha) = \sum_{n=-\infty}^{-1} \widehat{P\phi}_n \,\mathrm{sinc}\left(\frac{\alpha D}{\lambda} - n\right) + \sum_{n=N}^{+\infty} \widehat{P\phi}_n \,\mathrm{sinc}\left(\frac{\alpha D}{\lambda} - n\right). \quad (28)$$

Because the sinc function decreases rapidly with α, the terms flanking the DH ($n = -1$ and N) should by themselves give the order of magnitude of the residual speckle field in the DH. In the case of phase aberrations only and white noise, we have $|\widehat{P\phi}_{-1}|^2 \approx |\widehat{P\phi}_N|^2 \approx \overline{I}_0$, where $\overline{I}_0$ is the mean intensity in the

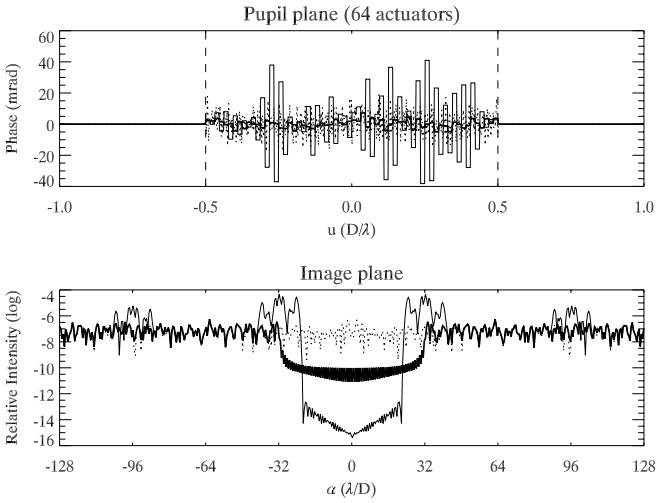

FIG. 2.—Speckle-nulling simulation for a one-dimensional pupil with 64 actuators. The energy-minimization algorithm makes it possible to push down the dark hole floor at the cost of some search area: the thick solid line shows the original dark hole, while the thin solid line shows the deeper and narrower dark hole. In this particular example with phase aberrations only, the average floor is decreased from 1.4×10^{-11} to 2.7×10^{-15} (a factor of about 5200 in contrast) by reducing the dark hole size from $64\lambda/D$ to $44\lambda/D$ (31% reduction). As explained in the text, this trade-off is obtained by increasing the amplitude of the highest spatial frequencies on the DM, making the dark hole's rim brighter at the same time.

image plane prior to the DH creation. Therefore, a crude estimate of the intensity profile in the DH should be

$$I_{\rm DH}(\alpha) \approx \overline{I}_0 \left[\mathrm{sinc}\left(\frac{\alpha D}{\lambda} + 1 \right) + \mathrm{sinc}\left(\frac{\alpha D}{\lambda} - N \right) \right]^2 . \quad (29)$$

We have superimposed this approximation as a thick line in Figure 1. In this case the match is remarkable, but more simulations show that it is generally good to within only a factor of 10. Nevertheless, it demonstrates that the DH depth depends critically on the residual speckle field at its edges and hence on the decreasing rate of the complex aberration spectrum with spatial frequency. In that respect, a white spectrum is certainly the worst case. Equation (29) further indicates that the DH depth depends on the number of actuators: as N is increased, the DH widens and gets deeper. With 8, 16, 32, and 64 actuators, equation (29) predicts $\overline{I}_0/\overline{I}_{\rm DH}$ to reach about 100, 300, 1000, and 4500, respectively.

4.1.3. *Dark Hole Depth versus Search Area*

As mentioned in § 3.3, speckle nulling by energy minimization can be performed in a region narrower than the maximum DH. Figure 2 illustrates this point: by reducing the search area from 64 to 44 resels (31% reduction), the DH floor was decreased from 1.4×10^{-11} to 2.7×10^{-15}, i.e., a gain of about 5200 in contrast (further reducing the search area does not bring any significant gain). By giving up search space, one frees the degrees of freedom corresponding to the highest spatial frequency components on the DM pattern. These can be used to improve the DH depth at lower spatial frequency because of the PSF angular extension (this is essentially the same reason why high spatial frequency speckles limit the DH depth). As the search space is reduced, the leverage of these highest spatial frequency components decreases (PSF wings falling off). The energy-minimization algorithm compensates by putting more energy at high frequency (see Fig. 2, *top*) and increasingly brighter spots in the image (around $\pm 32\lambda/D$ and $\pm 96\lambda/D$ in the bottom panel of Fig. 2).

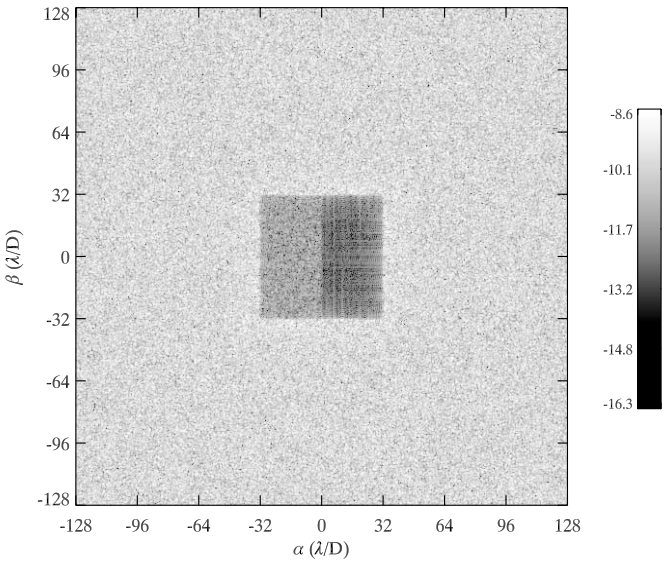

FIG. 3.—Speckle nulling with the energy-minimization algorithm (§ 3.3) for a square pupil in two dimensions (64×64 actuators). The gray scale shows the logarithm of the intensity. Speckles are cleared from the central part of the image, creating a dark hole (or search area) suitable for the detection of faint companions. Phase aberrations are corrected in the full dark hole, while amplitude aberrations are corrected in the right part only and made worse by a factor of 4 in intensity in the left part. Thus, the difference in intensity between the two sides of the dark hole gives a measure of the wave front amplitude errors. In this simulation, the standard deviations of the phase and amplitude aberrations are $\lambda/10^3$ and $\lambda/10^4$, respectively. The dark hole shape is a result of the actuator grid geometry on the DM, not of the pupil shape.

Thus, the trade-off range might be limited in practice by the maximum actuator stroke (currently 0.6 μm on the HCIT) and/or by the detector's dynamic range.

In two dimensions, the trade-off limits are well illustrated by the following example: considering a 64×64 DM and a random wave front, we find that the DH floor can be decreased from 2.4×10^{-12} to 1.4×10^{-13} (a factor of 17) if the search area is reduced from 64×64 to 60×60 resels (12% reduction in area). This implies a maximum actuator stroke of 10 nm and a detector dynamic range of 10^6. A further reduction to 58×58 resels does not feature a lower DH floor (2.1×10^{-13}) and would imply a maximum actuator stroke of 10 μm and a detector dynamic range of 10^{10}. In this case, the leverage of the additionally freed high spatial frequency components is so weak that the algorithm starts diverging.

4.1.4. *Two-dimensional Simulations with Phase and Amplitude Aberrations*

In Figures 3–4, we show an example of two-dimensional speckle nulling with phase and amplitude aberrations for a square pupil. To reflect the fact that phase aberrations dominate amplitude aberrations in real experiments (see Trauger et al. 2004), the rms amplitude of the amplitude aberrations is made 10 times smaller than that of the phase aberrations (the choice of a factor of 10 is arbitrary). The DH is split into two regions: in the right one ($\mathcal{H}^+$), the amplitude and phase aberrations are corrected, whereas in the left one ($\mathcal{H}^-$), the phase aberrations are corrected and the amplitude aberrations are made worse by a factor of 4 in intensity.

4.2. *Realistic Speckle Noise*

4.2.1. *Power Spectral Density of Phase Aberrations*

With the 3.5 m $\times$ 8 m *TPF-C* primary mirror in mind, we have studied the phase aberration map of an actual 8 m mirror: the

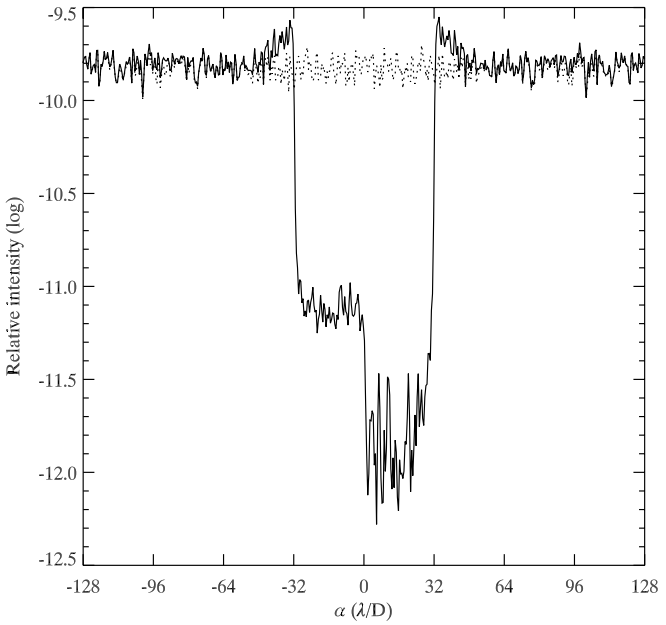

FIG. 4.—Average cut through the dark hole of the speckle nulling with the energy-minimization algorithm (§ 3.3) in two dimensions. The solid curve is an average of the intensity in Fig. 3 over β. The dotted line represents the state prior to correction with the DM. Note that the dark hole has a rim that is brighter than the background.

primary mirror of Antu, the first 8.2 m unit telescope of ESO's Very Large Telescope (VLT). This phase map[7] was obtained with the active optics system on and is characteristic of zonal errors (aberrations that cannot be fitted by low-order Zernike-type polynomials). It can be seen in Figure 5 that the azimuthally

[7] The phase map of Antu's primary mirror can be found courtesy of ESO at http://www.eso.org/projects/vlt/unit-tel/m1unit.html.

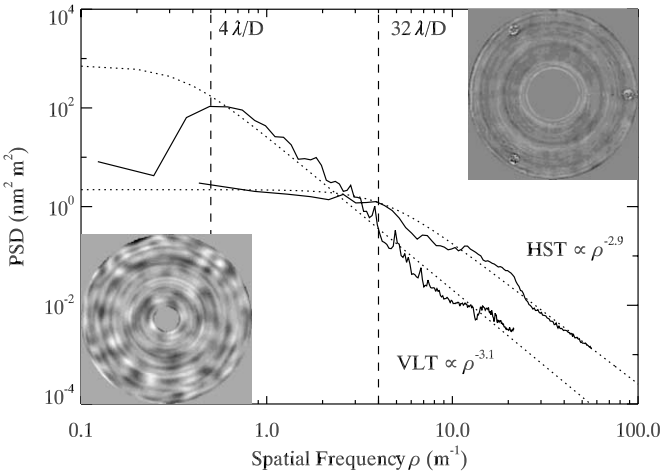

FIG. 5.—Azimuthally averaged power spectral density (PSD) of phase aberrations for the 8.2 m primary mirror of a VLT unit telescope (Antu) and the combination of the primary and secondary mirrors of *HST*. Both data sets appear as solid lines and are fitted with model (30) drawn with dotted lines. The model parameters are listed in Table 1. The dashed lines indicate the boundaries for spatial frequencies leading to speckles in the $(4\text{–}32)\lambda/D$ region of the image plane for an 8 m mirror. The bottom left and top right insets are the VLT phase map (courtesy of ESO) and *HST* phase map (Krist & Burrows 1995), respectively.

TABLE 1
PARAMETER VALUES FOR THE AZIMUTHALLY AVERAGED PSD MODEL OF *HST* AND VLT PHASE MAPS

Telescope	PSD_0 (nm^2 m^2)	ρ_c (m^{-1})	x
HST	2.2	4.3	2.9
VLT	720	0.35	3.1

averaged power spectral density (PSD) of such a map is well represented by

$$\mathrm{PSD}(\rho) = \frac{\mathrm{PSD}_0}{1 + (\rho/\rho_c)^x}, \tag{30}$$

where $\rho = (\alpha^2 + \beta^2)^{1/2}$. Values for PSD_0, ρ_c, and x are listed in Table 1. For comparison, the same treatment has been applied to the *Hubble Space Telescope* (*HST*) zonal error map from Krist & Burrows (1995).

We conclude from this study that *a realistic phase aberration PSD for an* 8 m *mirror decreases as the third power of the spatial frequency.* The standard deviation of the VLT phase map is 20.9 nm (18.5 nm for *HST*). The square root of the power of the phase aberrations in the 0.5–4 m^{-1} spatial frequency range [$(4\text{–}32)\lambda/D$ for an 8 m mirror] is 19.4 nm, i.e., about $\lambda/25$ at 500 nm, clearly not in the validity domain of our linear approximation.

4.2.2. *One-dimensional Simulation*

Figure 6 shows a simulation performed under the same conditions as for Figure 1 but with a VLT-like PSD. The PSD is scaled so that the standard deviation of the phase aberrations is equal to $\lambda/1000$. The average DH floor is now 5.3×10^{-12}, 6 orders of magnitude below the intensity peak in the original image! In agreement with § 4.1, we find that *the DH's depth depends critically on the magnitude of the speckle field at the edge of the DH and hence on the decrease of the phase aberration PSD with spatial frequency.*

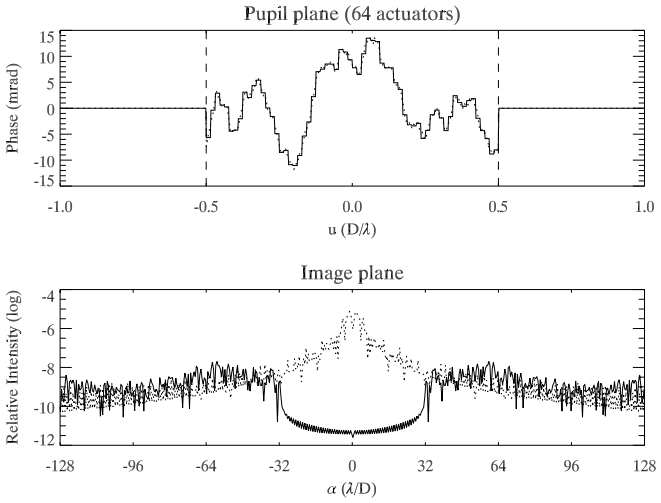

FIG. 6.—One-dimensional speckle nulling as computed by energy minimization (§ 3.3) with a phase aberration PSD given by equation (30). The values for x and ρ_c are the same as for the VLT; PSD_0 is such that the standard deviation of the phase aberrations is the same as in Fig. 1 ($\lambda/1000$). These realistic phase aberrations lead to deeper dark holes than in the hypothetical white-noise case of Fig. 1*d*.

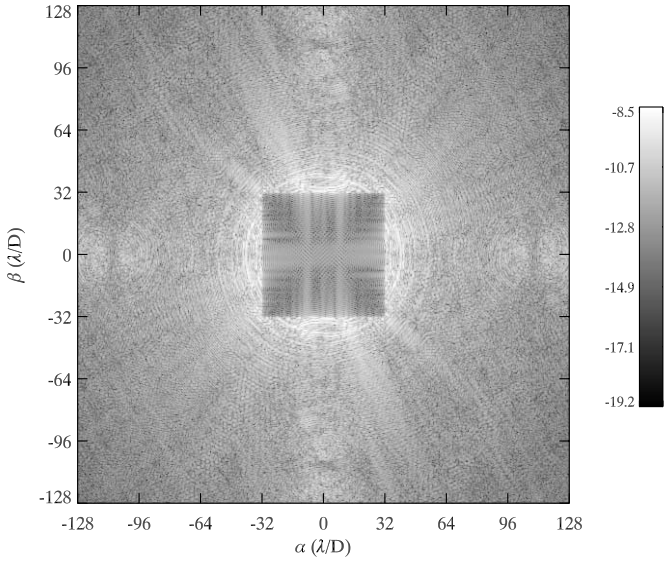

FIG. 7.—Speckle nulling with the field-nulling algorithm (§ 3.2) with a VLT-like PSD instead of white noise (64×64 actuators). The standard deviation of the phase aberrations is fixed to $\lambda/1000$. Actual HCIT influence functions are used. The average DH floor is 5.9×10^{-12}.

4.2.3. *Two-dimensional Simulation*

For the two-dimensional simulation in Figures 7–8, we have kept the original VLT phase map and circular pupil but scaled the standard deviation of the phase aberrations to $\lambda/1000$. In addition, we have used the actual HCIT influence functions from Trauger et al. (2003). The average DH floor is then 5.9×10^{-12} with field nulling (case shown) and 7.1×10^{-11} with energy minimization. The worse performance of the second method reflects the cost of the variable separation hypothesis, only accurate to within 5% for the HCIT. Note that the DH retains its square shape with a circular pupil, as the DH shape is fixed by the actuator grid geometry on the DM (a square grid of constant spacing in our case).

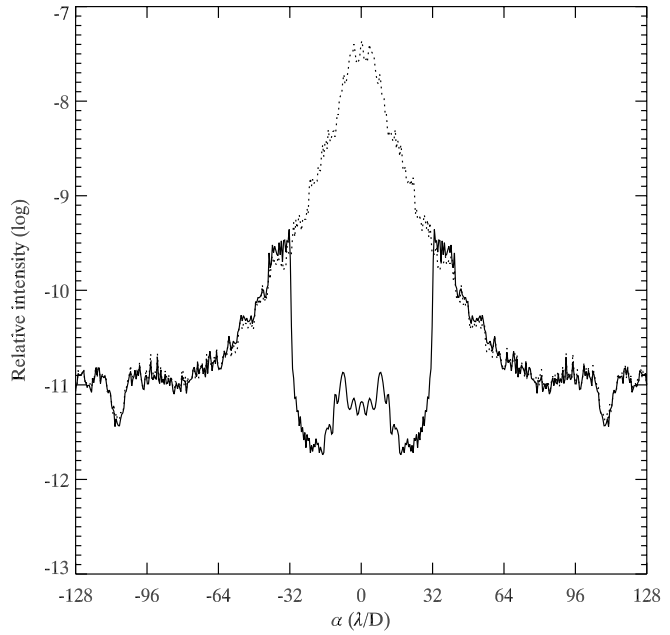

FIG. 8.—Average cut through the dark hole of speckle nulling with the field-nulling algorithm (§ 3.2) with a VLT-like PSD instead of white noise. The solid curve is an average of the intensity in Fig. 7 over β. The dotted line represents the state prior to correction with the DM.

5. DISCUSSION

5.1. *Quantum and Readout Noise*

In § 4, we presented noise-free simulations. To give an idea of the effect of quantum and readout noises, let us consider a sunlike star at 10 pc observed by a 3.5×8 m space telescope with a 5% overall efficiency. In a 100 nm bandwidth centered at 600 nm, the telescope collects about 2×10^{12} photoelectrons in 1 hr exposures. Considering the quantum noise, a 1 e^- read noise, and ignoring chromatic effects, simulations of sequences of four 1 hr exposures show that the average DH floor in Figure 1 would jump from 1.4×10^{-11} to 2.7×10^{-10}, whereas the average DH floor in Figure 6 would jump from 5.2×10^{-12} to 3.2×10^{-11}.

5.2. *Validity of the Linear Approximation*

In practice, our speckle-nulling process will work as stated provided equation (3) holds, that is, if $|P\phi + \psi| \gg \frac{1}{2}|P\phi^2|$. If c is the improvement in contrast with respect to the speckle floor and σ_ϕ the standard deviation of the wave front aberrations in radians, this condition translates into $\sigma_\phi/\sqrt{c} \gg \sigma_\phi^2/\sqrt{2}$, or $\sigma_\phi \ll (2/c)^{1/2}$. In terms of optical path difference, the standard deviation should then be much less than $\lambda/[\pi(2c)^{1/2}] = \lambda/140$ for $c = 10^3$. This is why we considered $\lambda/1000$ rms wave fronts in our simulations. As the wave front will probably not be of this quality at the start, the speckle-nulling method presented here is intended to be used in the course of observations, after a first phase for which the bulk of the aberrations have been taken out.

When the linear approximation breaks down, three images with different DM settings still provide enough information about the aberrations that a DH could be created thanks to a global nonlinear analysis of these images (P. Bordé, W. Traub, & J. Trauger 2004, private communication). Malbet et al. (1995) also explored nonlinear solutions, but with many more iterations ($\approx$20).

5.3. *Real Coronagraphs*

Dwelling on the validity of equation (3), real coronagraphs would not only remove the direct image of the star ($\hat{P}$), they would also modify the speckle field ($\widehat{P\phi}$) and the DM phase function ($\widehat{P\psi}$). This can be easily incorporated in the theory. A more delicate point is that real coronagraphs are not translation-invariant systems. As a consequence, effective influence functions as seen from behind the coronagraph will vary over the pupil. For image-plane coronagraphs with band-limited sinc masks (Kuchner & Traub 2002, § 4), we estimate this variation to be of the order of 10%, assuming $\epsilon = 0.1$ and 64 actuators. Only energy minimization, not field nulling (unless solved with SVD), can accommodate this effect.

5.4. *Actuator Stroke Precision*

What about the precision at which actuators should be controlled? As a consequence of the linearity of equation (15), the DH depth depends quadratically on the precision of the actuator strokes. We deduce—and this is confirmed by simulations—that a 4 orders of magnitude deep DH can only be obtained if the strokes are controlled at a 1% precision, i.e., 6 pm rms with $\lambda/1000$ aberrations at 600 nm. This precision corresponds to the current resolution of the actuator drivers on the HCIT.

5.5. *Instrumental Stability*

Regarding instrumental stability, we assumed that the instrument would remain perfectly stable during the four-step process.

However, despite the foreseen thermal and mechanical controls of the spacecraft, very slow drifts during the few hours of single exposures should be expected. Therefore, we intend to study in a subsequent paper how to incorporate a model of the drifts in our method. The exact parameters of this model would be derived from a learning phase after the launch of the spacecraft.

5.6. *Chromaticity*

We have not considered the effect of chromaticity. Let us point out that phase aberrations due to mirror surface errors scale with wavelength, so the correction derived from one wavelength would apply to all wavelengths. This is unfortunately not the case for amplitude aberrations. Although these are weaker than phase aberrations, a degradation of the correction should be expected in polychromatic light. Moreover, polychromatic wave front sensing will require a revised theory, as speckles will move out radially in proportion to the wavelength.

6. CONCLUSION AND FUTURE WORK

In this paper, we presented two techniques to optimally null out speckles in the central field of an image behind an ideal coronagraph in space. The measurement phase necessitates only three images, the fourth image being fully corrected. Depending on the number of actuators and the desired search area, the gain in contrast can reach several orders of magnitude.

These techniques are intended to work in a low-aberration regime, such as in the course of observations after an initial correction phase. They are primarily meant to be used in space but could be implemented in a second-stage adaptive optics (AO) system on ground-based telescopes. Of these two methods, the speckle energy minimization approach seems to be more powerful and flexible: (1) it offers the possibility to trade off some search area against an improved contrast, and (2) it can accommodate influence function variations over the pupil (necessary with real coronagraphs). If influence functions feature the required symmetry (variable separation), it is computationally very effective, but it is otherwise still better than SVD.

Since the principles underlying these speckle-nulling techniques are general, it should be possible to use them in conjunction with most coronagraph designs, including those with band-limited masks (Kuchner & Traub 2002), pupil mapping (Guyon et al. 2005; Vanderbei & Traub 2005), and shaped pupils (Kasdin et al. 2003). It is our intent to complete our work by integrating models of these coronagraphs in our simulations and to carry out experiments with the HCIT. In addition, we will seek to incorporate in the measurement theory a linear model for the evolution of aberrations, and we will work toward a theory accommodating the spectral bandwidth needed for the detection and spectroscopy of terrestrial planets.

We wish to thank the anonymous referee for insightful comments that helped to greatly improve the content of this paper. We acknowledge many helpful discussions with Chris Burrows, John Trauger, Joe Green, and Stuart Shaklan. This work was performed in part under contract 1256791 with the Jet Propulsion Laboratory (JPL), funded by NASA through the Michelson Fellowship Program, and in part under contract 1260535 from JPL. JPL is managed for NASA by the California Institute of Technology. This research has made use of NASA's Astrophysics Data System.

APPENDIX

FORMULATION OF ENERGY MINIMIZATION IN THE PUPIL PLANE

In this appendix, we show that energy minimization can be formulated in the pupil plane as well. Note that no measurement takes place in the pupil plane: the aberration function $P\phi$ is obtained by the inverse FT of $\widehat{P\phi}$, which is still measured as described in § 3.4. Although we here present a solution for the phase aberrations, the amplitude aberrations can be corrected in half of the domain without changing the formalism, exactly as explained in § 3.3.

By replacing $\hat{E}$ with its expression as a FT, the dark hole energy reads

$$\mathcal{E} = \int\int_{\mathcal{H}} \left[\int\int_{\mathcal{P}} E(u,v) e^{-i2\pi(u\alpha+v\beta)}\, du\, dv \int\int_{\mathcal{P}} E^*(u',v')\, e^{i2\pi(u'\alpha+v'\beta)}\, du'\, dv' \right] d\alpha\, d\beta. \tag{A1}$$

Now we invert the integration order and integrate over $\mathcal{H}$ to get

$$\mathcal{E} = \frac{1}{4d^2} \int\int_{\mathcal{P}} E(u,v) \int\int_{\mathcal{P}} E^*(u',v') h(u'-u) h(v'-v)\, du\, dv\, du'\, dv', \tag{A2}$$

where $h(x-y) \equiv \mathrm{sinc}[\lambda(x-y)/2d]$. The energy is minimized when

$$\forall (k,l), \qquad \int\int_{\mathcal{P}} f_{kl}(u,v) \int\int_{\mathcal{P}} [\phi(u',v') + \psi(u',v')] h(u'-u) h(v'-v)\, du\, dv\, du'\, dv' = 0. \tag{A3}$$

In the next steps, we first replace ψ with equation (4) then $f_{kl}(u,v)$ with its tensor product $g_k(u) g_l(v)$ in order to separately integrate the variables (u, u') and (v, v'). The final result reads

$$\forall (k,l) \in \{0, \ldots, N-1\}^2, \qquad \sum_{n=0}^{N-1} G_{kn} \sum_{m=0}^{N-1} a_{nm} G_{ml} = \Phi_{kl},$$

$$\text{where} \qquad G_{ij} = \int\int_{\mathcal{P}} g_i(x) g_j(y) h(x-y)\, dx\, dy, \qquad \Phi_{kl} = \int\int_{\mathcal{P}} g_k(u) g_l(v) \int\int_{\mathcal{P}} \phi(u',v') h(u'-u) h(v'-v)\, du\, dv\, du'\, dv'. \tag{A4}$$

System (A4) has a form identical to system (17) and can be solved with the same technique.

REFERENCES

Angel, R. 2003, in Proc. Conf. Towards Other Earths: *DARWIN*/*TPF* and the Search for Extrasolar Terrestrial Planets, ed. M. Fridlund & T. Henning (ESA-SP 539; Noordwijk: ESA), 221
Charbonneau, D., et al. 2005, ApJ, 626, 523
Chauvin, G., Lagrange, A.-M., Dumas, C., Zuckerman, B., Mouillet, D., Song, I., Beuzit, J.-L., & Lowrance, P. 2004, A&A, 425, L29
———. 2005, A&A, 438, L25
Chelli, A. 2005, A&A, 441, 1205
Coulter, D. 2004, Proc. SPIE, 5487, 1207
Deming, D., Seager, S., Richardson, L., & Harrington, J. 2005, Nature, 434, 740
Des Marais, D., et al. 2002, Astrobiology, 2, 153
Guyon, O. 2005, ApJ, 629, 592
Guyon, O., Pluzhnik, E., Galicher, R., Martinache, F., Ridgway, S., & Woodruff, R. 2005, ApJ, 622, 744
Hsu, H. 1967, Outline of Fourier Analysis (New York: Unitech Division)
Karlsson, A., Wallner, O., Perdigues Armengol, J., & Absil, O. 2004, Proc. SPIE, 5491, 831
Kasdin, N., Vanderbei, R., Spergel, D., & Littman, M. 2003, ApJ, 582, 1147
Krist, J., & Burrows, C. 1995, Appl. Opt., 34, 4951
Kuchner, M., & Traub, W. 2002, ApJ, 570, 900
Labeyrie, A. 1995, A&A, 298, 544
Löfdahl, M., & Scharmer, G. 1994, A&AS, 107, 243
Malbet, F., Yu, J., & Shao, M. 1995, PASP, 107, 386
Neuhäuser, R., Guenther, E., Wuchterl, G., Mugrauer, M., Bedalov, A., & Hauschildt, P. 2005, A&A, 435, L13
Press, W., Teukolsky, S., Vetterling, W., & Flannery, B. 2002, Numerical Recipes in C (Cambridge: Cambridge Univ. Press)
Quirrenbach, A. 2005, preprint (astro-ph/0502254)
Trauger, J., Moody, D., Gordon, B., Gürsel, Y., Ealey, M., & Bagwell, R. 2003, Proc. SPIE, 4854, 1
Trauger, J., et al. 2004, Proc. SPIE, 5487, 1330
Vanderbei, R., & Traub, W. 2005, ApJ, 626, 1079