THE ASTROPHYSICAL JOURNAL, 380:L5–L8, 1991 October 10

CONSTRAINED REALIZATIONS OF GAUSSIAN FIELDS: A SIMPLE ALGORITHM

YEHUDA HOFFMAN[1] AND EREZ RIBAK[2,3]
Received 1991 February 5; accepted 1991 July 19

ABSTRACT

We present a straightforward method for the construction of constrained realizations of Gaussian fields. Consider a Gaussian random field and its ensemble mean field given a set of constraints. The residual of the field from its mean is statistically independent of the actual numerical value of the constraints. The algorithm is based on a simple construction of this residual field, which is then added to the analytically calculated mean field. This algorithm is exact and involves no iterations. The computational effort involves only setting one random realization in the inverse k-space, its FFT to real space, and the calculation of the mean field (given the set of constraints) over all grid points. Furthermore, it can be applied to a large number of constraints of arbitrary amplitudes and positions. The algorithm facilitates the generation of the initial conditions for N-body simulations which obey a few hundreds of constraints imposed by the observable universe.

Subject headings: cosmology — galaxies: clustering

1. INTRODUCTION

The canonical cosmological model assumes that structure in the universe has emerged from a primordial density perturbation field. It further assumes that this is a random Gaussian field (cf. Kolb & Turner 1990), which is defined by its power spectrum and a normalization constant. The power spectrum is determined by general theoretical considerations concerning the cosmological model, the physical properties of the early universe, and the nature of the dark matter. For example, in the popular "standard" cold dark matter model (Blumenthal et al. 1984) the power spectrum obeys the Harrison (1970)–Zel'dovich (1972) spectrum on the very large scales and goes asymptotically to k^{-3} on small scale (where k is the wavenumber). The theory specifies only the statistical properties of the amplitudes and phases in the plane waves representation. Thus, for Gaussian fields the real and imaginary parts are independently normally distributed around zero with a variance given by half the power spectrum, which implies that the phases are uniformly distributed. This is being used in setting up the initial conditions for large N-body numerical simulations, which have proved to be a major tool in studying the large-scale structure of the universe (cf. Efstathiou et al. 1985). Yet, in many interesting cosmological problems one is interested in generating special-purpose initial conditions, which are designed to obey some given constraints. One then needs an efficient algorithm for generating such realizations.

The aim here is to present the optimal algorithm for constructing constrained realizations of Gaussian fields. In the cosmological context its main advantage is to enable generating initial conditions for numerical simulations. These simulations are designed to include a priori many of the observed features of the observed universe. Such features are the dipole velocity relative to the microwave background (Lubin & Villela 1986), the local density field as derived by the POTENT algorithm (Bertschinger & Dekel 1989), the bulk velocity (Lynden-Bell et al. 1988), and the projected (two-dimensional) *IRAS* galaxy distribution (Scharf et al. 1991).

[1] Racah Institute of Physics, The Hebrew University, Jerusalem, Israel.
[2] Department of Physics, The Technion, Haifa 32000, Israel.
[3] Electro-Optics Research Division, Technion R&D Foundation, Haifa, Israel.

Bertschinger (1987) was the first to pose the problem of the constrained realizations and to develop an algorithm for their construction. It is based on an iterative simulated annealing technique which requires $O[(M^2 + 1)N]$ operations to generate one independent realization, where N is the number of degrees of freedom and M is the number of constraints. The number of degrees of freedom is essentially the number of grid elements required for the field in the computer memory. This method can be naively described as relaxing a system with constraints on the field and on its power spectrum into a plausible position in phase space. As the grid density grows and as the number of constraints increases to more than a few, the system converges so slowly that this algorithm becomes prohibitively expansive and impractical. A simple variant of this method was suggested by Binney & Quinn (1990), in which the random field is expanded in spherical harmonics rather than plane wave basis. For a localized set of constraints, such as a single local maximum which can be described by a few spherical harmonics, the construction simplifies dramatically and can be solved exactly with no iterations at all. The drawback of the latter algorithm is that it can be applied only in the case where the constraints are quite localized and define and obvious center of symmetry.

A different approach has been used by us recently (Hoffman & Ribak 1992, hereafter Paper I). In this approach we devised an approximate iterative algorithm in which only the Fourier *phases* are subjected to the imposed constraints. The rationale behind this approach is that random phases are essentially all that is required to make a random field Gaussian (in the random field the power spectrum amplitudes are χ^2 distributed with 2 degrees of freedom). It follows from the central limit theorem that in the case where the number of constraints is much smaller than the number of degrees of freedom, the subensemble of constrained-phase realizations properly samples the whole ensemble of constrained realizations. This algorithm converges rather rapidly and is very efficient and easy to construct. It is somewhat limited by the requirement of two Fourier transforms for each iteration (to impose the power spectrum and the field constraints alternatively). Also, as in the Bertschinger (1987) method it is formulated in terms of constraints imposed upon the field itself, rather than its linear functionals such as derivatives or integrals. Here we introduce

a much improved method which is noniterative and involves only the calculation of the correlation matrix and its inverse. This method can be used for any kind of constraints where the field and its linear functionals are imposed to have a specific numerical value. The method is very efficient and can be used to set a very large number of constraints.

2. CONSTRAINED GAUSSIAN FIELDS

We start with a short review of the properties of constrained Gaussian fields (for a more rigorous review see Adler 1981; Bardeen et al. 1986; see also Paper I for other aspects). Consider a random homogeneous and isotropic Gaussian field $f(\mathbf{r})$ with zero mean which is defined by its power spectrum, $P(k)$. This field is subjected to a set of M constraints, $\Gamma = \{C_i(\mathbf{r})|_{\mathbf{r}_i} = c_i;\ i = 1, \ldots, M\}$. The ith constraint C_i can be imposed on the field itself, that is, $f(\mathbf{r}_i) = c_i$, or any linear functional of it such as its derivative, $(\partial/\partial x)f(\mathbf{r})|_{\mathbf{r}_i} = c_i$, or a convolution over $f(\mathbf{r})$ with some function $g(\mathbf{r})$, $\int d^3r g(\mathbf{r} - \mathbf{r}_i)f(\mathbf{r}) = c_i$. The conditional probability distribution function of the field $f(\mathbf{r})$ is

$$\mathscr{P}[f(\mathbf{r})\,|\,\Gamma] = \frac{\mathscr{P}[f(\mathbf{r}), \Gamma]}{\mathscr{P}[\Gamma]} . \tag{1}$$

Here, $\mathscr{P}[\ldots]$ is the multivariate Gaussian of the appropriate variables. The conditional probability distribution function can be described as a shifted Gaussian around the ensemble mean field, $\bar{f}(\mathbf{r})$, defined as

$$\bar{f}(\mathbf{r}) = \langle f(\mathbf{r})\,|\,\Gamma\rangle = \xi_i(\mathbf{r})\xi_{ij}^{-1}c_j . \tag{2}$$

Summation over repeated indices is assumed throughout this paper, and $\langle\ldots\rangle$ denotes an ensemble average. Here $\xi_i(\mathbf{r}) = \langle f(\mathbf{r})C_i\rangle$ is the cross-correlation between the field and the ith constraint, $\xi_{ij} = \langle C_i C_j\rangle$ is the constraints' correlation matrix, and C_i is evaluated at $\mathbf{r}_i$. In the case where the constraints involve only the field itself, the correlation matrix is written in terms of the two-point autocorrelation function $\xi(r)$, that is, $\xi_{ij} = \xi(|\mathbf{r}_i - \mathbf{r}_j|)$ and $\xi_i(\mathbf{r}) = \xi(|\mathbf{r}_i - \mathbf{r}|)$.

The residual field has some unique properties which we will utilize. It is defined as the difference between the Gaussian field and the mean field (under constraints): $F(\mathbf{r}) \equiv f(\mathbf{r}) - \bar{f}(\mathbf{r})$. The variance of the residual field (Paper I) is

$$\langle F^2(\mathbf{r})\,|\,\Gamma\rangle = \sigma^2 - \xi_i(\mathbf{r})\xi_{ij}^{-1}\xi_j(\mathbf{r}) , \tag{3}$$

where $\sigma^2 = \xi(0)$. Note that the residual $F(\mathbf{r})$ is a random Gaussian field which is not homogenous nor isotropic whose variance is independent of the numerical value of the imposed constraints. In terms of the residual field $F(\mathbf{r})$ the constraint points are expressed as $\{F(\mathbf{r}_i) = 0;\ i = 1, \ldots, M\}$ (in the case where only the field itself is constrained). Any particular constrained realization can be written as the sum of the analytically calculated mean field given Γ, $\bar{f}(\mathbf{r}) = \langle f(\mathbf{r})\,|\,\Gamma\rangle$, and the random residual field, $F(\mathbf{r})$. The key point in the present algorithm is that the statistical properties of the residual field are all *independent* of the numerical values of the constraints c_i, and for any particular choice of the constraints a realization of the residual can be easily constructed.

We proceed now to the construction of a constrained realization of the field $f(\mathbf{r})$. This is done in five stages: (1) create a random realization; (2) find the values of the realization corresponding to the constraints; (3) calculate the mean of the realization; (4) evaluate the residual of the realization; (5) combine the residual with the required, constrained mean.

In step 1 we start with the production of an unconstrained realization of the field, $\tilde{f}(\mathbf{r})$, determined by the power spectrum only. For the particular realization we calculate in step 2 the actual values of the variables which are to be constrained. These variables can be looked upon as defining another set of constraints, $\tilde{\Gamma} = \{\tilde{c}_i\}$. This a posteriori set of constraints is evaluated at the positions of the original constraints and has the values of this specific realization. For this "random" selected $\tilde{\Gamma}$, we calculate in step 3 the corresponding mean field expected as if the set was chosen initially, $\tilde{f}(\mathbf{r}) = \langle\tilde{f}(\mathbf{r})\,|\,\tilde{\Gamma}\rangle$. From the given particular realization and the calculated mean field given $\tilde{\Gamma}$, the residual is easily written as $F(\mathbf{r}) = \tilde{f}(\mathbf{r}) - \tilde{f}(\mathbf{r})$ (step 4). The residual field thus generated is a particular realization subject to the desired constraints, Γ. In step 5 we evaluate the mean field (given Γ), according to equation 2, and add it to the residual $F(\mathbf{r})$ to yield

$$f(\mathbf{r}) = \tilde{f}(\mathbf{r}) + \xi_i(\mathbf{r})\xi_{ij}^{-1}(c_j - \tilde{c}_j) . \tag{4}$$

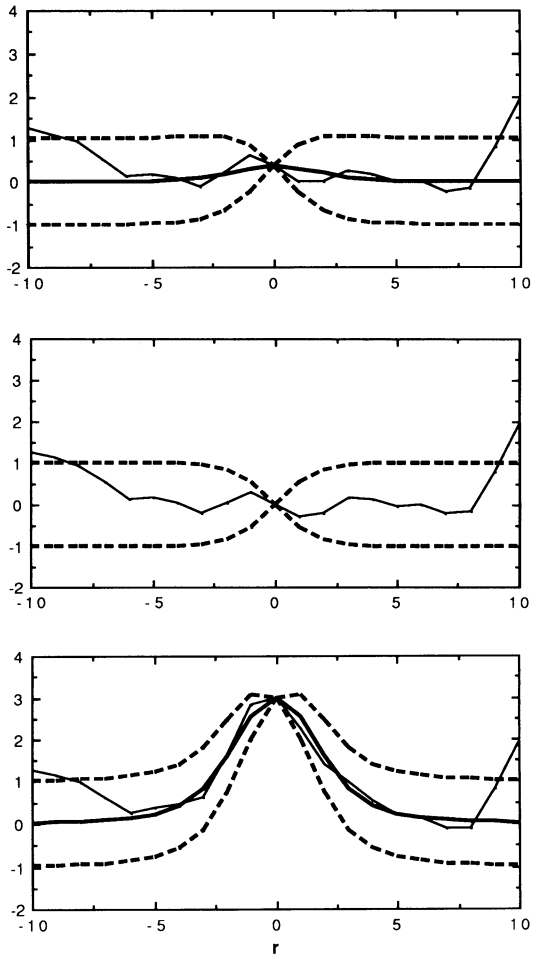

FIG. 1.—A single constraint, $f(0) = 3\ \sigma$, is imposed on a Gaussian field which is defined by its power spectrum, $P(k) \propto k^{-1}\exp(-k^2)$, and is evaluated on a 32^3 grid. *Upper panel:* A cut in the field through the constraint point (*thin line*) is compared with the theoretically calculated mean (*heavy line*), given the actual value at the origin, and the mean plus and minus one standard deviation (*dashed lines*). *Central Panel:* Residual of field from the mean (*thin line*) is compared with the (plus and minus) theoretical one standard deviation (*dashed lines*). *Lower panel:* Residual is added to the true mean (*heavy line*) to create the desired field (*thin line*). The bracket of one standard deviation is shown again.

In the particular case where only $f(\mathbf{r})$ itself is constrained equation 4 reads

$$f(\mathbf{r}) = \tilde{f}(\mathbf{r}) + \xi_i(\mathbf{r})\xi_{ij}^{-1}[c_j - \tilde{f}(\mathbf{r}_j)] \,. \quad (5)$$

The method is illustrated in Figure 1 where a single constraint is imposed on a three-dimensional field, out of which only a one-dimensional cut through one of the main axes is shown. The field is constrained at $r_1 = 0$ by $C_1 = 3\ \sigma$, and the power spectrum is the one defined in § 3. In the top panel the unconstrained realization is presented together with the mean field, given its actual value at r_1, which is given by $\bar{f}(\mathbf{r}) = \tilde{f}(0)\xi(\mathbf{r})/\sigma^2$. This mean profile is bracketed by the mean plus and minus one standard deviation. The central panel shows the residual constructed by subtracting the theoretical mean value from the actual one. This residual is then added to the mean field, given the actual constraint, $\bar{f}(\mathbf{r}) = C_1\xi(\mathbf{r})/\sigma^2$, as shown in the lower panel.

The constructed field $f(\mathbf{r})$ obeys the imposed constraints, and we substitute it for the unconstrained $\tilde{f}(\mathbf{r})$. Note that there is a one-to-one correspondence between the trial field $\tilde{f}(\mathbf{r})$ and the constructed one. Furthermore, the ensemble of realizations produced by the algorithm presented here properly samples the subensemble of all realizations constrained by Γ. The algorithm is an optimal one because it is exact and involves only one realization of an unconstrained random field and the calculation of the mean field under the given constraints. The desired field is then evaluated by performing $2M^2 + 1$ operations at each grid point (eq. 4), far less than what is required by all other methods mentioned above.

3. APPLICATION

To illustrate the present method we present here the following example. A Gaussian field whose power spectrum is $P(k) \propto k^{-1} \exp[-(kR_s)^2]$ is constrained to have a 3 σ peak (local maximum) at the origin. For simplicity the peak constraints are formulated by (spatially) expanding the mean field around the origin to second-order terms, and by constraining the six nearest neighbors. The peak parameters are taken from Bardeen et al. (1986). For this particular example a peak with the mean curvature (for 3 σ) and of equal second derivatives along the axes has been chosen. The realization is performed on a 32^3 grid with a smoothing length of $R_s = 1$ (grid units). This is shown in Figure 2, where the field $f(\mathbf{r})$ along the three axes is compared with the mean field, and a bracket of the mean field plus and minus one standard deviation, that is, $\bar{f}(\mathbf{r}) \pm \langle F^2(\mathbf{r})|\Gamma\rangle^{1/2}$. One sees that the deviation of $f(\mathbf{r})$ from the mean $\bar{f}(\mathbf{r})$ is of the order of the standard deviation, $\langle F^2(\mathbf{r})\rangle^{1/2}$ as expected. An ensemble of $\mathcal{N} = 100$ such realizations has been constructed, and its mean and standard deviation are compared with their theoretically calculated values. The agreement between the simulations is within the $\mathcal{N}^{-1/2}$ expected scatter. Figure 3 compares the theoretical standard deviation with the numerical one (along the three axes). The

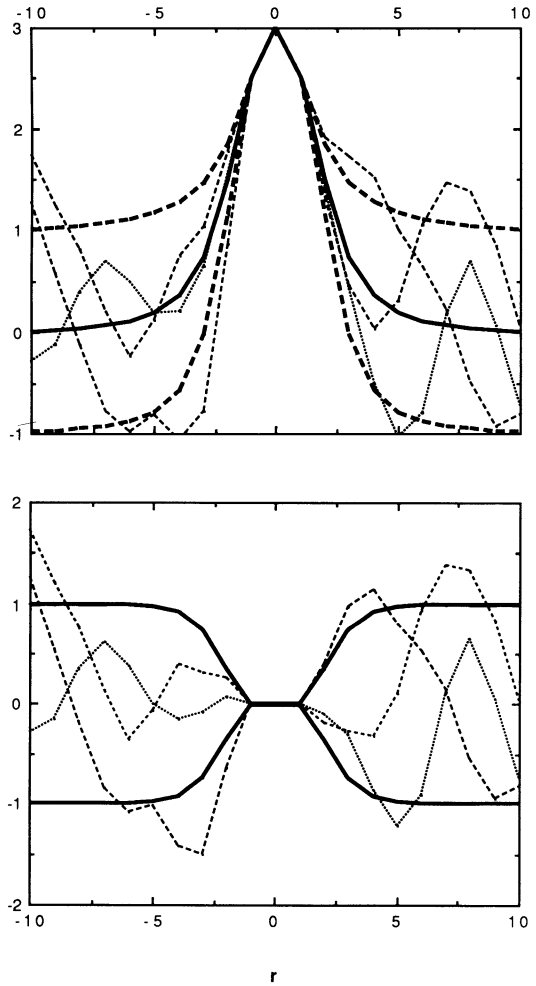

FIG. 2.—*Upper panel:* Constrained field profiles along the three major axes through the peak (*thin broken lines*) are compared with the theoretically calculated mean (*heavy line*) and the mean plus and minus one standard deviation (*upper and lower heavy dashed lines*) profiles. *Lower panel:* Residual of $f(\mathbf{r})$ from the mean $\bar{f}(\mathbf{r})$ (*thin broken lines*) are compared with the theoretical plus and minus one standard deviation (*heavy lines*).

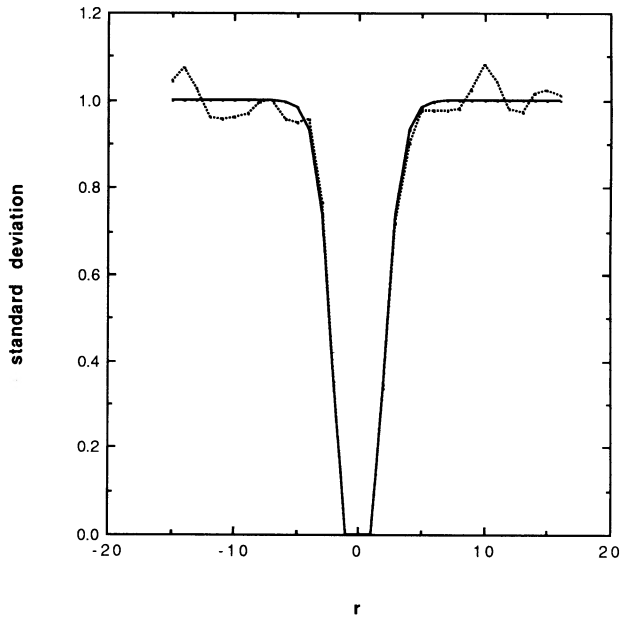

FIG. 3.—Standard deviation of an ensemble of 100 numerical constrained realizations is compared with its theoretical expected value. Simulated field is subjected to a 3 σ local maximum imposed at the origin. The numerical standard deviation is evaluated along the axes (*broken lines*) and is compared with the value calculated based on the analytical relation of eq. 3 (*continuous line*).

agreement is very good, and no systematic deviations have been found for different realizations, power spectra, or cutoffs.

4. SUMMARY

The algorithm suggested here provides an optimal solution to the problem of the construction of constrained realizations of Gaussian fields. The computational effort consists only of the creation of one random unconstrained realization and the calculation of the mean field (given the set of constraints) over all grid points. The number of constraints can be very large, and they can be imposed on the field itself or on any linear functional of it. The proposed method can be used to construct initial conditions for N-body simulations which are constrained to include some of the observed features of the locally observed universe.

The hospitality of the Institute of Astronomy (Cambridge University) and the Space Research Institute (Technion) where part of this work was done is gratefully acknowledged (Y. H.).

REFERENCES

Adler, R. J. 1981, The Geometry of Random Fields (Chichester: Wiley)
Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
Bertschinger, E. 1987, ApJ, 323, L103
Bertschinger, E., & Dekel, A. 1989, ApJ, 336, L5
Binney, J., & Quinn, T. 1990, preprint
Efstathiou, G., Davis, M., Frenk, C. S., & White, S. D. M. 1985, ApJS, 57, 241
Harrison, E. R. 1970, Phys. Rev. D, 1, 2726
Hoffman, Y., & Ribak, E. 1992, ApJ, in press (Paper I)
Kolb, E. W., & Turner, M. 1990, The Early Universe (New York: Addison-Wesley)
Lubin, P., & Villela, T. 1986, in Galaxy Distances and Deviations from Universal Expansion, ed. B. F. Madore & R. B. Tully (Dordrecht: Reidel), 169
Lynden-Bell, D., Faber, S. M., Burstein, D., Davies, R. L., Dressler, A., Terlevich, R. J., & Wegner, G. 1988, ApJ, 326, 19
Scharf, C., Hoffman, Y., Lahav, O., & Lynden-Bell, D. 1991, preprint
Zel'dovich, Ya. B. 1972, MNRAS, 160, 1P