Computationally efficient coherent-mode representations

Brynmor J. Davis* and Robert W. Schoonover

The Beckman Institute for Advanced Science and Technology and the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA
**Corresponding author: bryn@illinois.edu*

Received December 9, 2008; accepted January 16, 2009;
posted February 19, 2009 (Doc. ID 104952); published March 18, 2009

The numerical calculation of traditional coherent-mode representations (CMRs) involves an eigenvalue decomposition of the cross-spectral density matrix. An efficient alternative modal representation of a partially coherent field can be realized using an LDL† decomposition. Storage requirements are reduced by an amount on the order of the ratio between the coherence length and the source width. The efficiency of calculations requiring a CMR (e.g., numerical evaluation of partially coherent propagation effects) may thus be significantly improved, particularly when low-coherence fields are considered. © 2009 Optical Society of America

OCIS codes: 030.0030, 030.4070.

Light is an electromagnetic phenomenon and is thus governed by Maxwell's equations. However, it is often the case that in order to make useful predictions, the sources and fields must be assumed to be stochastic. Statistically stationary and ergodic models usually provide a good description of the real world [1–3], with stationary statistics described to second order by the cross-spectral density tensor (CSD),

$$W_{\alpha\beta}(\mathbf{r}_1,\mathbf{r}_2,\omega) = \int \langle \mathcal{E}^*_\alpha(\mathbf{r}_1,t)\mathcal{E}_\beta(\mathbf{r}_2,t+\tau)\rangle e^{i\omega\tau}\mathrm{d}\tau, \quad (1)$$

where the angle brackets denote an ensemble expectation, $\mathbf{r}_j$ is a spatial position, $\mathcal{E}_\alpha$ is the field in the direction given by α (e.g., the x, y, or z direction), t is time, τ is a temporal delay, and ω can be identified as a temporal oscillation frequency of the field. When considering primary electromagnetic fields and sources, the spatial variables $(\mathbf{r}_1,\mathbf{r}_2)$ and field directions (α,β) are each three dimensional. Each may be reduced to two dimensions for secondary planar sources with a low propagation angle. When the electromagnetic field is well approximated by a scalar field, the CSD tensor is replaced by a scalar CSD.

The laws governing changes in the statistics of the radiated fields upon propagation are well understood—the CSD for a propagating field obeys a double Helmholtz equation (Section 6.6.3 of [3]). The CSD is a function of variables in either four or six dimensions, while the field is a function of variables in either two or three dimensions. The numerical complexity of the propagation calculation is exponential in the number of dimensions, and so direct calculation of the propagation of the CSD may be impracticable. This problem can be ameliorated by use of the coherent-mode representation (CMR) [4,5].

For electromagnetic fields, the CMR of the CSD is expressed as a sum of weighted modes, i.e., $\vec{\vec{W}}(\mathbf{r}_1,\mathbf{r}_2,\omega)=\Sigma_n\lambda_n(\omega)\boldsymbol{\psi}_n(\mathbf{r}_2,\omega)\boldsymbol{\psi}_n^\dagger(\mathbf{r}_1,\omega)$, where $\boldsymbol{\psi}_n$ is a coherent mode of the field, λ_n is the associated spectral weighting function, and $\dagger$ is the Hermitian conjugate operator [6,7]. Using this expansion, the spectral representation of the random field can be written as $\mathbf{E}(\mathbf{r},\omega)=\Sigma_n\chi_n(\omega)\boldsymbol{\psi}_n(\mathbf{r},\omega)$, where, for fixed value of ω, χ_n is a random variable with $\langle\chi_m^*(\omega)\chi_n(\omega)\rangle=\lambda_n(\omega)\delta_{nm}$. Thus, the mode coefficients are uncorrelated and have a variance determined by the corresponding spectral weight. For scalar fields, the same representation holds, except W and ψ_n are scalar functions [4,5]. In both cases, stationarity ensures that the field is uncorrelated across frequencies ω. Thus the CSD can be treated one frequency at a time.

Operations on the field (e.g., propagation and focusing) can be applied to each mode of the CMR individually and the results resummed with weights according to the CMR. Deterministic numerical propagation techniques, such as finite-element methods [8], may therefore be applied to each of the modes in the CMR. This gives the postoperation CSD from calculations on a series of fields—a less numerically expensive procedure than direct propagation of the CSD.

The CMR is typically found via a Mercer (eigenvalue) expansion of the Hermitian and nonnegative-defined CSD. In certain cases the CMR may be calculated analytically through this method, as may the effects of certain simple operations such as free-space propagation [9–11]. However, the CMR of more complicated fields may be difficult to find analytically. In such cases the CMR may be calculated numerically. This work shows that the eigenvector-based CMR is not optimally efficient in terms of the number of operations or the storage resources required. An LDL† factorization (Theorem 4.1.2 of [12]) is proposed as the basis of an alternative approach. This method reduces calculation and storage costs at the expense of mode orthogonality. It should be noted, however, that mode orthogonality is also often lost upon propagation of the traditional CMR anyway. An exception to this is propagation within the limits of the Fresnel or Fraunhofer approximations; beyond these simple cases, light propagation through absorptive, anisotropic, optically active, and/or inhomogeneous media generally results in the loss of mode orthogonality.

0146-9592/09/070923-3/$15.00

In both the scalar and vector cases, the CSD can be represented digitally by first projecting the field on to a set of N basis vectors $\{\mathbf{v}^{(n)}(\mathbf{r},\omega)\}_{n=1}^{N}$. For example, the basis vectors may be chosen as a partition of pixels/voxels with elemental area/volume dictated by the Nyquist criterion. In the electromagnetic case each spatial position will be associated with multiple pixels/voxels, one for each field direction axis considered. The field is then represented discretely by $\mathbf{E}(\mathbf{r},\omega)\approx\Sigma_n e_n(\omega)\mathbf{v}^{(n)}(\mathbf{r},\omega)$, where e_n is the random coefficient associated with $\mathbf{v}^{(n)}$. The discrete CSD is a matrix,

$$\bar{W}=\langle \mathbf{e}\mathbf{e}^{\dagger}\rangle, \tag{2}$$

where the element $\bar{W}_{nm}$ gives the correlation between the coefficients e_n and e_m. For the pixelated basis set described above, $\bar{W}$ may be found by simply sampling W and appropriately rearranging the results.

A general modal expansion of the field can be found by decomposing the CSD matrix as

$$\bar{W}=\bar{\Psi}\bar{\Lambda}\bar{\Psi}^{\dagger}, \tag{3}$$

where $\bar{\Psi}$ and $\bar{\Lambda}$ are $N\times N$ matrices. Each column of $\bar{\Psi}$ is a mode, $\bar{\boldsymbol{\psi}}_n$, of the discrete field, and Λ_{mn} represents the covariance of the mode coefficients for $\bar{\boldsymbol{\psi}}_n$ and $\bar{\boldsymbol{\psi}}_m$.

In general there are an infinite number of factorizations fitting the form of Eq. (3), as this expression represents a simple change of basis. The basis may be chosen to facilitate computation, e.g., the basis may be the eigenmodes of a fiber optic cable or a resonator cavity, or the basis may be found through methods designed to quickly calculate the modal structure. For example, an alternative electromagnetic CMR [13] has been proposed that uses uncoupled decompositions on each scalar component of the vector field and then calculates component-to-component (e.g., x field to y field) correlations with respect to these modes. For this representation, the matrix $\bar{\Lambda}$ is not necessarily diagonal, and not all of the modes are statistically uncorrelated.

The propagation of the modes may be carried out generally through the expression $\bar{\Phi}=\bar{G}(\bar{\Psi})$, where $\bar{G}$ is the propagator and acts on each column of $\bar{\Psi}$ independently and $\bar{\Phi}$ is a matrix containing the propagated modes. The CSD upon propagation is then given by the expression $\bar{W}'=\bar{\Phi}\bar{\Lambda}\bar{\Phi}^{\dagger}$. The utility of the traditional CMR arises from the choice of a factorization that diagonalizes (or partially diagonalizes) the matrix $\bar{W}$. When $\bar{\Lambda}$ has a diagonal structure, the partially coherent field can be regarded as an incoherent sum of fully coherent modes.

Determination of the most efficient decomposition of the CSD depends on the speed with which the decomposition may be performed, how efficient the resultant modes may be propagated, and the speed at which the CSD may be recomposed. A diagonal $\bar{\Lambda}$ offers an advantage factor of $\mathcal{O}(N)$ in the recomposition step. The remainder of this Letter addresses an optimized diagonalization scheme giving diagonal $\bar{\Lambda}$.

The traditional CMR is based on an eigenvector decomposition, but there are several alternative methods for diagonalizing the matrix $\bar{W}$ [12]. Here the LDL† diagonalization method is chosen, as it is computationally efficient, well studied, and can exploit any sparsity present in the CSD matrix. There are two distinct computational advantages achieved by using the LDL† factorization: the number of operations needed to calculate the factorization is reduced, and there is a reduction in the amount of memory needed to store the calculated modes.

The LDL† factorization decomposes a Hermitian matrix in the form of Eq. (3) with the properties that $\bar{\Psi}$ is lower triangular (i.e., $\bar{\Psi}_{nm}=0\ \forall\ m>n$), $\bar{\Lambda}$ is nonzero only on the main diagonal (so that the modes are uncorrelated), and the elements of $\bar{\Lambda}$ are real and nonnegative (guaranteed by a nonnegative definite CSD matrix).

In many applications where a set of random variables (with a given covariance) must be simulated, a Cholesky decomposition is used [14,15] to give a factorization similar to Eq. (3). Here, the very closely related LDL† decomposition is chosen, as the result is of the exact form of Eq. (3) and because the numerical stability of LDL† algorithms has been thoroughly addressed. Discretization of the CSD can lead to representation errors and an indefinite matrix (especially for fine spatial sampling). Methods have been developed to address this problem [16].

For an $N\times N$ matrix, the LDL† decomposition can be computed with $\mathcal{O}(N^3)$ operations. The eigenvector decomposition can also be computed with $\mathcal{O}(N^3)$, for example, by using approaches based on the symmetric QR algorithm (Section 8.3 of [12]). However the prefactor for these calculation costs is higher for the eigenvector decomposition (by an amount dependent on the specific implementation chosen), with the result that the LDL† factorization is more computationally efficient. Additionally, reduced storage resources are required. Since $\bar{\Psi}$ is lower triangular for the LDL† decomposition, there will be a savings factor of approximately two when compared to the eigenvector-based CMR, where no $\bar{\Psi}_{nm}$ is guaranteed to be zero. More importantly, the LDL† algorithm can be optimized to exploit sparsity in the CSD. Consider a two-dimensional discretized field that has a width of M pixels (so the CSD matrix is of dimension $N=M^2$) and zero correlation between points separated by more than M/γ pixels in either direction. A simple one-dimensional pixel ordering results in a CSD matrix $\bar{W}$ that is zero beyond the (N/γ)th diagonal. This type of matrix is known as "banded" and gives significant computational simplifications. While each mode of the Mercer expansion will have a full N nonzero entries, the modes in the LDL† system will have only N/γ nonzero entries (Sec. 4.3 of [12]). The storage requirements are thus reduced by a factor of γ, i.e., by the ratio between the beam width and the coherence

length. The banded structure may also be exploited to give computational cost improvements in both the Mercer and LDL† expansions. The improvement will depend on the algorithms used.

Some models of partially coherent fields (e.g., the Gaussian Schell-model beam [10]) have CSD functions that decay with $|\mathbf{r}_1-\mathbf{r}_2|$ but do not go to zero. In this case it is possible to sparsely approximate the CSD as

$$\tilde{W}(\mathbf{r}_1,\mathbf{r}_2,\omega)=A(\mathbf{r}_1-\mathbf{r}_2)W(\mathbf{r}_1,\mathbf{r}_2,\omega), \quad (4)$$

where $A(\mathbf{r}_1-\mathbf{r}_2)$ is a window function that is zero beyond some threshold of $|\mathbf{r}_1-\mathbf{r}_2|$. It is important to retain a nonnegative definite $\tilde{W}$, which can be done by choosing a window function A that has a nonnegative Fourier spectrum. To create such a function, one can take the autocorrelation of any finite-width window function. Using Eq. (4) it is possible to set many values in the CSD matrix to zero with little deviation from the original CSD. This produces a banded matrix and, with the LDL† factorization, sparse modes. A matrix that cannot be described as banded (e.g., a matrix resulting from a field where all significant points are correlated) will not result in sparse modes; however, a memory savings of approximately a factor of two will still be obtained via the LDL† implementation.

As an example, a scalar Gaussian Schell-model source will be defined by beam width σ_s and coherence length σ_g and discretely represented in a CSD matrix. While analytic expressions exist for the CMR of a Gaussian Schell-model source, the CMR will be evaluated numerically here. The autocorrelation of a Blackman window was used to generate A ($A=h*h[|\mathbf{r}_2-\mathbf{r}_1|/2\sigma_g]$, where $*$ denotes convolution and h is the Blackman window). All calculations were carried out using the MATLAB (The Mathworks Inc.) computation platform in double precision on a Mac Mini with a 2 GHz Intel Core 2 Duo processor and 2 GB of RAM. In particular, the EIG and LDL algorithms were used, as was the SPARSE data storage organization. The storage requirements listed are those reported by MATLAB and the calculation times were measured using the TIC and TOC commands.

In Fig. 1 the storage requirements can be seen to depend approximately linearly on γ, as expected. For the algorithms used in this experiment, the improvement in computational cost also appears to be linear in γ. It should be noted that the times plotted in Fig. 1 include normalizing the modes from the LDL† algorithm. It should also be noted that the memory savings reported in Fig. 1 depend only on the sparsity of the CSD matrix, so that while the calculations were carried out for a Gaussian Schell-model example, the results will be generally applicable to any CSD that can be represented in a matrix with bandedness given by γ.

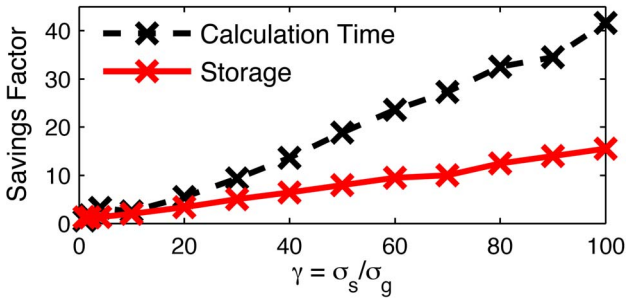

Fig. 1. (Color online) A comparison of the LDL† and Mercer decompositions for the CSD of a Gaussian Schell-model source. Mercer-to-LDL† quotients of both calculation time and storage memory are plotted as a function of the ratio between the source width and the source coherence.

In this Letter, it has been shown that the eigenvalue decomposition traditionally used for calculating coherent-mode decompositions of the CSD is not always optimal from a computational standpoint. The LDL† method for decomposing matrices may be used to efficiently calculate a CMR for a CSD matrix. This method is faster and requires less memory than the eigenvalue decomposition (by a factor on the order of the source width-to-coherence ratio).

This work was supported in part by the Grainger Foundation for Emerging Technologies. The authors would also like to thank P. Scott Carney for encouraging this research and for useful discussions.

References

1. C. Brosseau, *Fundamentals of Polarized Light: a Statistical Optics Approach* (Wiley, 1998).
2. J. W. Goodman, *Statistical Optics* (Wiley-Interscience, 1985).
3. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge U. Press, 1995).
4. E. Wolf, Opt. Commun. **38**, 3 (1981).
5. E. Wolf, J. Opt. Soc. Am. **72**, 343 (1982).
6. F. Gori, M. Santarsiero, R. Simon, G. Piquero, R. Borghi, and G. Guattari, J. Opt. Soc. Am. A **20**, 78 (2003).
7. J. Tervo, T. Setälä, and A. T. Friberg, J. Opt. Soc. Am. A **21**, 2205 (2004).
8. J. Jin, *The Finite Element Method in Electromagnetics*, 2nd ed. (Wiley-IEEE, 2002).
9. F. Gori, Opt. Commun. **34**, 301 (1980).
10. A. Starikov and E. Wolf, J. Opt. Soc. Am. **72**, 923 (1982).
11. K. Sundar, N. Mukunda, and R. Simon, J. Opt. Soc. Am. A **12**, 560 (1995).
12. G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hopkins U. Press, 1996).
13. K. Kim and E. Wolf, Opt. Commun. **261**, 19 (2006).
14. J. H. Michels, P. K. Varshney, and D. D. Weiner, IEEE Trans. Signal Process. **42**, 367 (1994).
15. B. J. Davis, Opt. Express **15**, 2837 (2007).
16. C. Ashcraft, R. G. Grimes, and J. G. Lewis, SIAM J. Matrix Anal. Appl. **20**, 513 (1998).