

# Iterative LMMSE Image Restoration From Partially-Known Blur

by

Vladimir Z. Mesarović<sup>†</sup>, Nikolas P. Galatsanos<sup>‡\*</sup>, and Miles N. Wernick<sup>‡</sup>

<sup>†</sup>Crystal Semiconductor Products Division Cirrus Logic Corp.  
4210 S. Industrial Drive  
Austin, TX 78744.

<sup>‡</sup>Illinois Institute of Technology  
Department of Electrical and Computer Engineering  
3301 S. Dearborn St., Chicago, Illinois 60616  
E-mail: npg@ece.iit.edu  
Tel: (312) 567-5259

## Abstract

We address the problem of space-invariant image restoration when the blurring operator is not known exactly, a situation that arises regularly in practice. To account for this uncertainty, we model the point spread function as the sum of a known deterministic component and an unknown random one. Such an approach has been studied before, but the problem of estimating the parameters of the restoration filter has not been addressed systematically. We propose an approach based on a Gaussian statistical assumption and derive an iterative, expectation-maximization (EM) algorithm that simultaneously restores the image and estimates the required filter parameters. We obtain two versions of the algorithm based on two different models for the statistics of the image. The computations are performed in the discrete Fourier transform (DFT) domain, thus they are computationally efficient even for large images. We examine the convergence properties of the resulting estimators and evaluate their performance experimentally.

*Permission to publish this abstract separately is granted*

---

\* corresponding author † corresponding author.

# I Introduction

Conventional image restoration algorithms assume exact knowledge of the blurring operator whereas blind deconvolution methods assume that little or nothing is known about the blur (see Ref. 1 and 2 for a recent review). In most practical applications, the point-spread function (PSF) is neither totally unknown nor perfectly known. Instead, some PSF model is usually available, but this model is often inaccurate. To represent this situation, we will assume that the PSF can be modeled as the sum of a known deterministic component and an unknown random one. We refer to this as a partially-known blur model.

The partially-known blur model is appropriate for many realistic applications and has been studied previously (see Refs. 3, 4, 5, 6, 7, 8 and 9 ). For instance, in medical imaging techniques such as positron emission tomography (PET) and single-photon emission computed tomography (SPECT), the PSF is difficult to specify completely<sup>10</sup>, in part because it is object-dependent, owing to scattering and photon attenuation<sup>11</sup>. In astronomy, atmospheric turbulence yields a random time-varying PSF which is not known exactly. Blind deconvolution, which assumes that the PSF is completely unknown, does not accurately model situations such as these in which lack of knowledge of the PSF is only partial.

To our knowledge the partially-known blur model was first considered in Ref. 3. In that work, the linear minimum mean square-error (LMMSE) solution was derived for the continuous case only, the statistics of the PSF and noise were assumed to be known, and no experimental results were provided. In Refs. 4, 6 and 7 the LMMSE filter was derived for the discrete case and numerical experiments were reported.

LMMSE restoration filters require knowledge of the signal covariance, which is not usually available in practice. In Ref. 6 the signal covariance is assumed known, but in the method proposed in Refs. 4 and 7 it is estimated iteratively by using the current LMMSE signal estimate to update the signal covariance estimate. However, convergence of the proposed iterative algorithm in Refs.

4 and 7 was not shown. Furthermore, the filter and the estimation algorithm is not derived for the circulant case; thus, it cannot be efficiently calculated for large images. In this paper we correct the above mentioned shortcomings of previous work on this problem. In other words:

1. We propose an iterative algorithm to estimate the parameters of the derived restoration filter and we show the convergence of the proposed algorithm.
2. We derive both the LMMSE filter and the iterative algorithm in the discrete Fourier domain (DFT) using the circulant assumption. Thus, it can be efficiently applied to large images.

More specifically, based on two different image models we propose two approaches for simultaneous iterative identification of the parameters and restoration of the image. These approaches are based on the expectation-maximization (EM) algorithm in Ref. 14 and a Gaussian statistical model. Our first approach is based on a stationary image model with a circulant covariance matrix; the second uses a simultaneously autoregressive (SAR) image model (see Refs. 15, and 16.)

The EM algorithm iterates between the so-called expectation (E) and the maximization (M) steps. The M-step maximum-likelihood (ML) estimates of the image and noise model parameters are found. In the E-step the conditional mean of the image, given the data, is computed. Since, for Gaussian data, the conditional mean is the LMMSE estimate (see Ref. 17 pp. 389). In the E step of the algorithm we obtain the LMMSE estimate as an intermediate result. We derive both proposed approaches in the DFT domain, thus they are computationally efficient even for large images. We show using numerical experiments that the LMMSE filter we obtain performs better than the LMMSE filter derived in Ref. 4.

For linear models, there is a close relationship between LMMSE solutions, and statistical solutions based on Gaussian statistics. In this paper we exploit that relationship by using a ML approach and the EM algorithm to obtain a systematic approach to obtaining LMMSE estimates using ML techniques.

The rest of the paper is organized as follows. In Section II the observation and image models

are discussed. In Section III the EM algorithm is applied to the restoration problem from partially-known blurs for two image models. The convergence properties of the resulting iterative sequences are derived in Section IV. The proposed estimators are tested experimentally in Section V. In Section VI conclusions and suggestions for future research are given.

## II Observation and Image Models

### II.A Observation Model

In Ref. 4 and 6 the space-invariant PSF was represented as the sum of a deterministic component and a stochastic component of zero-mean, i.e.,

$$\mathbf{h} = \bar{\mathbf{h}} + \Delta\mathbf{h}, \quad (1)$$

where  $\bar{\mathbf{h}} \in \mathcal{R}^N$  and  $\Delta\mathbf{h} \in \mathcal{R}^N$  are the deterministic (known) and the random (unknown error) components of the PSF, respectively. This is a very general model that attempts to incorporate the random (unknown error) component of the PSF in the restoration algorithm. The unknown component of the PSF is modeled as stationary zero-mean white noise with  $N \times N$  covariance matrix  $\mathbf{R}_{\Delta h} = \beta\mathbf{I}$ , where  $\beta$  denotes the variance of the PSF noise and  $\mathbf{I}$  is the identity matrix. The observation vector  $\mathbf{g}$  is also contaminated by zero-mean additive white Gaussian noise with  $N \times N$  covariance matrix  $\mathbf{R}_{\Delta g} = \gamma\mathbf{I}$ , where  $\gamma$  denotes the variance of the observation noise. Furthermore, the noises in the observed data and the PSF are assumed independent of each other and independent from the source image  $\mathbf{f}$ . In this case, the image-degradation can be described by the model in Refs. 4, 5, 6, 9, 12 and 13

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \Delta\mathbf{g}, \quad (2)$$

in which

$$\mathbf{H} = \bar{\mathbf{H}} + \Delta\mathbf{H}, \quad (3)$$

and  $\mathbf{g}, \mathbf{f}, \Delta\mathbf{g} \in \mathcal{R}^N$  are lexicographically ordered representations of the observed degraded image, the source image, and the additive noise in the observed image, respectively. The matrix  $\bar{\mathbf{H}}$  is the

known (assumed, estimated or measured) component of the  $N \times N$  PSF matrix  $\mathbf{H}$ ;  $\Delta\mathbf{H}$  is the unknown component of the PSF matrix, generated by  $\Delta\mathbf{h}$  defined in (1). Throughout the rest of this paper a circulant approximation of Toeplitz matrices<sup>18</sup> will be used to allow calculations to be performed using the discrete Fourier transform (DFT); thus, the matrices  $\mathbf{R}_{\Delta h}$ ,  $\mathbf{R}_{\Delta g}$ ,  $\bar{\mathbf{H}}$  and  $\Delta\mathbf{H}$  are  $N \times N$  circulant matrices (see Ref. 18 pp. 224). We will furthermore assume a stationary image model. Therefore, the image autocorrelation

$$E\{\mathbf{f}\mathbf{f}^t\} = \mathbf{R}_f, \quad (4)$$

is also a  $N \times N$  circulant matrix, thus it is defined by  $N$  parameters. These parameters are either the vector containing the elements of a row or column of  $\mathbf{R}_f$  or the eigenvalues of this matrix which can be found by the DFT (see Ref. 18 pp. 224). We shall refer to this model for the rest of this paper as the *full  $\mathbf{R}_f$  model*.

In many practical situations  $\mathbf{R}_f$  in (4) may not be available and must be estimated from the blurred and noisy data. However, the  $N$  parameters required for the definition of  $\mathbf{R}_f$  may not be identifiable from a single degraded image. To avoid this problem, the simultaneously autoregressive (SAR) image models is used. The SAR model for image has been used previously in the context of Bayesian estimation (see Refs. 15, 16 and 19). However, the SAR model has been also used implicitly in the past in the context of regularized least-squares estimation (see Refs. 18 pp. 149 , and 20). For this model the covariance is assumed to be

$$[\mathbf{R}_f] = [\alpha\mathbf{Q}^t\mathbf{Q}]^{-1}, \quad (5)$$

where  $\mathbf{Q}$  represents the circulant convolution with the operator given by

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -(4 + \epsilon) & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (6)$$

where  $\epsilon$  a small positive number. In the limit as  $\epsilon$  goes to zero  $\mathbf{Q}$  becomes the Laplacian operator.

Since  $\mathbf{R}_f$  and  $\mathbf{Q}$  are circulant matrices their eigenvalues are given by the DFT (see Ref. 18 pp. 224). The SAR model is Eq. (5) implies that  $\mathbf{Q}\mathbf{f}$  the convolution of the image with the  $\mathbf{Q}$  operator

is an independent identically distributed (iid) signal. The parameter  $\alpha$  is positive, unknown and can be viewed as the variance of the iid signal. For natural images, that display a high-degree of spatial correlation, this is a good model.

In this paper for simplicity a zero mean image model is used. This is a usual assumption in image restoration applications. A zero mean image is estimated and then the mean of the observed data is added. This is justified by the fact that in most cases the noise is zero mean and the convolution by the PSF does not affect the mean.

### III Parameter Estimation and Restoration Using the EM Algorithm

In this section we apply the Expectation-Maximization (EM) in Ref. 14 algorithm to the problem of identifying, and restoring, from partially-known (random) blur. The EM algorithm is an iterative approach for computing maximum-likelihood (ML) estimates of unknown parameters. In the present application the EM algorithm is useful for its capability to identify the unknown parameters while simultaneously restoring the degraded image. We derive two algorithms based on the EM approach: one for each image model in (4) and (5). In the first algorithm, for the full  $\mathbf{R}_f$  model, we assume that the parameters that define  $\mathbf{R}_f$ ,  $\beta$ , and  $\gamma$ , are the unknown parameters that we want to estimate. In general, the covariance matrix  $\mathbf{R}_f$  has  $N$  unknown parameters: therefore, in this case, there are  $N + 2$  parameters to estimate. In our second algorithm, for the SAR model, the image covariance matrix contained in (5) is parameterized with a single parameter  $\alpha$  (matrix  $\mathbf{Q}$  is known), and that  $\beta$  and  $\gamma$  are also unknown; thus, there are a total of three parameters to estimate.

#### III.A EM Algorithm for the full $\mathbf{R}_f$ model

We begin by calculating the likelihood functional for our problem. Let  $\Psi' = \{\mathbf{R}_f, \beta, \gamma\}$  denote the set of unknown parameters we wish to estimate. We assume that  $\mathbf{f}$ ,  $\Delta\mathbf{H}$ , and  $\Delta\mathbf{g}$ , in (2) are

independent, and that the observed image  $\mathbf{g}$  is Gaussian distributed with PDF parameterized by  $\Psi'$ , i.e.,

$$P(\mathbf{g}|\Psi') = [\det(2\pi[\bar{\mathbf{H}}\mathbf{R}_f\bar{\mathbf{H}}^t + \mathbf{R}_T(\mathbf{f})])]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\mathbf{g}^t[\bar{\mathbf{H}}\mathbf{R}_f\bar{\mathbf{H}}^t + \mathbf{R}_T(\mathbf{f})]^{-1}\mathbf{g}\right\}, \quad (7)$$

where

$$\mathbf{R}_T(\mathbf{f}) = E\left\{(\Delta\mathbf{H}\mathbf{f} + \Delta\mathbf{g})(\Delta\mathbf{H}\mathbf{f} + \Delta\mathbf{g})^t\right\}. \quad (8)$$

ML estimation of the parameter set  $\Psi'$  involves the determination of the  $\Psi'_{ML}$  that maximizes the likelihood function  $P(\mathbf{g}|\Psi')$ . The likelihood function in (7) is nonlinear and non-convex with respect to  $\Psi'$ ; thus, it is very difficult to optimize. Therefore, we propose to use the numerically efficient iterative method of the EM algorithm in Ref. 14 to maximize the functional in (7).

In applying the EM algorithm, the observation  $\mathbf{g}$  represents the incomplete data and a set of complete data  $\mathbf{z}$  must be defined. The sample realization  $\mathbf{z}$  from the corresponding sample space is observed through the linear transformation  $\mathbf{T}$ , which relates the complete and incomplete data according to

$$\mathbf{g} = \mathbf{T}\mathbf{z}. \quad (9)$$

The particular form of  $\mathbf{T}$  depends on the choice of definition of the complete data (see (17)). Since  $\mathbf{g}$  is zero-mean Gaussian,  $\mathbf{z}$  is also zero-mean Gaussian with PDF parameterized by the set  $\Psi$  and equal to

$$P(\mathbf{z}|\Psi) = [\det(2\pi\mathbf{R}_z)]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\mathbf{z}^t\mathbf{R}_z^{-1}\mathbf{z}\right\}, \quad (10)$$

where  $\mathbf{R}_z$  is the covariance matrix of  $\mathbf{z}$ , and  $\Psi = \{\mathbf{R}_z\}$ . Taking the logarithm of both sides of (10) and discarding the additive and multiplicative constants, we obtain

$$\log P(\mathbf{z}|\Psi) \propto -\left\{\log[\det(\mathbf{R}_z)] + \mathbf{z}^t\mathbf{R}_z^{-1}\mathbf{z}\right\}. \quad (11)$$

The complete-data vector  $\mathbf{z}$  is not available and is only observable through the non-invertible (many-to-one) mapping (9).

The EM algorithm consists of two steps: the expectation step (E-step) and the maximization step (M-step). In the E-step of the EM algorithm, the conditional expectation of  $\log P(\mathbf{z}|\Psi)$  is

computed, conditioned upon the observed data  $\mathbf{g}$  and the current estimate of the parameter set  $\Psi$ . In the M-step, the result of the E-step is maximized with respect to unknown parameters.

In the present problem the conditional PDF of  $\mathbf{z}$  given  $\mathbf{g}$  is Gaussian<sup>17</sup> with conditional mean equal to

$$\mathbf{m}_{z|g} = E\{\mathbf{z}|\mathbf{g}\} = \mathbf{R}_{zg}\mathbf{R}_g^{-1}\mathbf{g} \quad (12)$$

and conditional covariance equal to

$$\mathbf{R}_{z|g} = E\{\mathbf{z}\mathbf{z}^t|\mathbf{g}\} = \mathbf{R}_z - \mathbf{R}_{zg}\mathbf{R}_g^{-1}\mathbf{R}_{gz}, \quad (13)$$

where  $\mathbf{R}_{zg} = \mathbf{R}_z\mathbf{T}^t$ ,  $\mathbf{R}_g = \mathbf{T}\mathbf{R}_z\mathbf{T}^t$  and  $\mathbf{R}_{gz} = \mathbf{R}_{zg}^t$ . Therefore, the E-step starts with an estimate of the parameter set  $\Psi^{(n)}$  (superscript  $(n)$  denotes the iteration index) and finds the conditional expectation of the log-likelihood (11) of the complete data given the incomplete data  $\mathbf{g}$  and the current parameter estimate  $\Psi^{(n)}$ :

$$E\{\log P(\mathbf{z}|\Psi) | \mathbf{g}, \Psi^{(n)}\} \propto -\left\{\log[\det(\mathbf{R}_z)] + \text{tr}[\mathbf{R}_z^{-1}(\mathbf{R}_{z|g}^{(n)} + \mathbf{m}_{z|g}^{(n)}\mathbf{m}_{z|g}^{(n)t})]\right\}, \quad (14)$$

where  $\text{tr}(\cdot)$  denotes the trace of a matrix. Clearly, maximizing (14) is equivalent to minimizing

$$F(\Psi; \Psi^{(n)}) = \log[\det(\mathbf{R}_z)] + \text{tr}[\mathbf{R}_z^{-1}\mathbf{R}_{z|g}^{(n)}] + \mathbf{m}_{z|g}^{(n)t}\mathbf{R}_z^{-1}\mathbf{m}_{z|g}^{(n)}. \quad (15)$$

Therefore, the E-step of the EM algorithm amounts to computing  $F(\Psi; \Psi^{(n)})$  in (15) with the use of (12) and (13), and the M-step reduces to minimizing the  $F(\Psi; \Psi^{(n)})$  with respect to  $\Psi$ .

The art of using the EM algorithm lies in choosing an appropriate complete-data specification  $\mathbf{z}$  (for details see for example Ref. 17). The difficulty of direct maximization of  $P(\mathbf{g}|\Psi)$  is overcome by embedding the sample space  $\mathbf{g}$  in a richer or larger sample space  $\mathbf{z}$  where optimization problems are easier to solve. We choose the complete data  $\mathbf{z}$  as a concatenation of the lexicographically ordered images  $\mathbf{f}$  and  $\mathbf{g}$  as in Ref. 21; i.e.,

$$\mathbf{z} = \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix}. \quad (16)$$

This particular choice of  $\mathbf{z}$  allows for parameter identification and easy computation of (15). The incomplete and complete data are now related via

$$\mathbf{g} = [\mathbf{0} \quad \mathbf{I}] \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix} = \mathbf{T}\mathbf{z}, \quad (17)$$

where  $\mathbf{0}$  and  $\mathbf{I}$  are the  $N \times N$  zero and identity matrices, respectively.

According to (16), the covariance matrix of  $\mathbf{z}$  is given by

$$\mathbf{R}_z = E\{\mathbf{z}\mathbf{z}^t\} = \begin{bmatrix} \mathbf{R}_f & \mathbf{R}_f \bar{\mathbf{H}}^t \\ \bar{\mathbf{H}}\mathbf{R}_f & \bar{\mathbf{H}}\mathbf{R}_f \bar{\mathbf{H}}^t + \mathbf{R}_T(\mathbf{f}) \end{bmatrix}, \quad (18)$$

where  $\mathbf{R}_T(\mathbf{f})$  is given in (8).

The inverse of  $\mathbf{R}_z$  is given by Ref. 17

$$\mathbf{R}_z^{-1} = \begin{bmatrix} \mathbf{R}_f^{-1} + \bar{\mathbf{H}}^t \mathbf{R}_T^{-1}(\mathbf{f}) \bar{\mathbf{H}} & -\bar{\mathbf{H}}^t \mathbf{R}_T^{-1}(\mathbf{f}) \\ -\mathbf{R}_T^{-1}(\mathbf{f}) \bar{\mathbf{H}} & \mathbf{R}_T^{-1}(\mathbf{f}) \end{bmatrix}. \quad (19)$$

Substituting (18) and (19) into (15) we obtain

$$\begin{aligned} F(\Psi; \Psi^{(n)}) &= \log[\det(\mathbf{R}_f)] + \text{tr}[\mathbf{R}_f^{-1} \mathbf{R}_{f|g}^{(n)}] + \mathbf{m}_{f|g}^{(n)t} \mathbf{R}_f^{-1} \mathbf{m}_{f|g}^{(n)} \\ &+ \log[\det(\mathbf{R}_T(\mathbf{f}))] + \text{tr}[\bar{\mathbf{H}}^t \mathbf{R}_T^{-1}(\mathbf{f}) \bar{\mathbf{H}} \mathbf{R}_{f|g}^{(n)}] + (\bar{\mathbf{H}} \mathbf{m}_{f|g}^{(n)} - \mathbf{g})^t \mathbf{R}_T^{-1}(\mathbf{f}) (\bar{\mathbf{H}} \mathbf{m}_{f|g}^{(n)} - \mathbf{g}), \end{aligned} \quad (20)$$

in which

$$\mathbf{m}_{f|g}^{(n)} = \mathbf{R}_f^{(n)} \bar{\mathbf{H}}^t (\bar{\mathbf{H}} \mathbf{R}_f^{(n)} \bar{\mathbf{H}}^t + \mathbf{R}_T^{(n)}(\mathbf{f}))^{-1} \mathbf{g} \quad (21)$$

and

$$\mathbf{R}_{f/g}^{(n)} = \mathbf{R}_f^{(n)} - \mathbf{R}_f^{(n)} \bar{\mathbf{H}}^t (\bar{\mathbf{H}} \mathbf{R}_f^{(n)} \bar{\mathbf{H}}^t + \mathbf{R}_T^{(n)}(\mathbf{f}))^{-1} \bar{\mathbf{H}} \mathbf{R}_f^{(n)}. \quad (22)$$

Using the DFT diagonalization properties for circulant matrices, the eigenvalues of  $\mathbf{R}_T(\mathbf{f})$  given in Appendix A and applying similar calculations as in Appendices A and B of Ref. 9, in the DFT domain the representation of (20), (21), and (22) is given by:

$$\begin{aligned} F(\Psi; \Psi^{(n)}) &= \sum_{i=0}^{N-1} \log(S_f(i)) + \frac{1}{S_f(i)} (S_{f|g}^{(n)}(i) + \frac{1}{N} |M_{f|g}^{(n)}(i)|^2) \\ &+ \log(N\beta S_f(i) + \gamma) + \frac{1}{N\beta S_f(i) + \gamma} \left\{ |\bar{H}(i)|^2 S_{f|g}^{(n)}(i) + \frac{1}{N} |\bar{H}(i) M_{f|g}^{(n)}(i) - G(i)|^2 \right\}, \end{aligned} \quad (23)$$

$$M_{f|g}^{(n)}(i) = \frac{\bar{H}^*(i)S_f^{(n)}(i)}{|\bar{H}(i)|^2 S_f^{(n)}(i) + N\beta^{(n)}S_f^{(n)}(i) + \gamma^{(n)}}G(i), \quad (24)$$

and

$$S_{f|g}^{(n)}(i) = \frac{S_f^{(n)}(i) [N\beta^{(n)}S_f^{(n)}(i) + \gamma^{(n)}]}{|\bar{H}(i)|^2 S_f^{(n)}(i) + N\beta^{(n)}S_f^{(n)}(i) + \gamma^{(n)}}, \quad (25)$$

where  $i$  is the frequency index,  $\bar{H}(i)$  are the eigenvalues of the known part of the PSF,  $G(i)$  are the DFT coefficients of the observed data, and  $S_f(i)$  are the power spectrum coefficients of the image (i.e., the eigenvalues of  $\mathbf{R}_f$ ).  $M_{f|g}^{(n)}(i)$  and  $S_{f|g}^{(n)}(i)$  represent the DFT coefficients of the conditional mean vector and the eigenvalues of the conditional covariance matrix, respectively. According to (24) the restored image-DFT coefficients  $M_{f|g}^{(n)}(i)$  are obtained as the output of a LMMSE filter based on the current estimate of  $\Psi$ , with the observed image as input. Equations (23), (24), and (25) are the DFT-domain expressions for the E-step of the EM algorithm for the full  $\mathbf{R}_f$  model.

In the M-step of the EM algorithm the likelihood functional in (23) is minimized with respect to the unknown parameters. In Appendix B we show that the M-step reduces to simple linear update equations in  $S_f(i)$ ,  $\beta$  and  $\gamma$ , i.e.,

$$S_f^{(n+1)}(i) = S_{f|g}^{(n)}(i) + \frac{1}{N}|M_{f|g}^{(n)}(i)|^2 \quad (26)$$

$$\beta^{(n+1)} = \beta^{(n)} \frac{v(\beta^{(n)})}{u(\beta^{(n)})} \quad (27)$$

$$\gamma^{(n+1)} = \gamma^{(n)} \frac{v(\gamma^{(n)})}{u(\gamma^{(n)})} \quad (28)$$

where  $u(\beta^{(n)})$ ,  $v(\beta^{(n)})$ ,  $u(\gamma^{(n)})$ , and  $v(\gamma^{(n)})$  are given in Appendix B in (B-4), (B-5), (B-10), and (B-11), respectively. Thus, the M-step of the EM algorithm for full  $\mathbf{R}_f$  model consists of calculating (26), (27), and (28) at each iteration, in which  $M_{f|g}^{(n)}(i)$  and  $S_{f|g}^{(n)}(i)$  are computed by (24) and (25).

The discrete form of the LMMSE filter for the random blur restoration problem was first derived in Ref. 4 and expressed in term of autocorrelation matrices. This filter can be represented in the DFT domain in terms of power spectra as

$$F_{lmmse}^{(n)}(i) = \frac{\bar{H}^*(i)S_f^{(n)}(i)}{|\bar{H}(i)|^2 S_f^{(n)}(i) + \beta S_f^{(n)}(i) + \gamma}G(i), \quad (29)$$

where  $\beta$  and  $\gamma$  were assumed known. Clearly, although (29) is similar in form to our solution (24) it differs by a factor of  $N$  in the denominator. This observation is based on the assumption that  $S_f(i)$  are the eigenvalues of the circulant autocorrelation matrix  $\mathbf{R}_f$ .

In Appendix A we show that (24) is the LMMSE filter, and in our experiments we demonstrate that this filter (24) always yields lower mean-square error (MSE) than the one in (29).

An algorithm for iterative image covariance estimation in the context of restoration from random blur was also proposed in Ref. 4. The direct conversion of this algorithm into the DFT domain yields the following power-spectrum update:

$$S_f^{(n+1)}(i) = \frac{1}{N} |F_{lmmse}^{(n)}(i)|^2. \quad (30)$$

By comparing the estimators in (26) and (30), one can see that the estimator in (30) does not include the term  $S_{f/g}^{(n)}(i)$  given in (25) as suggested by the EM algorithm's estimator in (26). In section IV the convergence properties of estimators in (26) and (30) are examined.

### III.B EM Algorithm for SAR Model

The estimation of the full-blown power spectrum ( $N$  coefficients for the full  $\mathbf{R}_f$  model) is a formidable task especially when a limited amount of data is available. This is the case in our restoration problem since only one degraded image is available and there are  $N + 2$  unknown parameters to estimate. Furthermore, in general the total number of relevant data points for power spectrum estimation is limited, since a degraded image usually contains relatively few high-frequency components. If the number of unknown parameters is greater than the number of relevant data points unreliable identification results should be expected <sup>22</sup>.

To reduce the number of unknown parameters in the identification/restoration problem we propose to use a SAR image model in (5). This image model is based on the single unknown parameter  $\alpha$ . Now, there are only three unknown parameters to estimate, i.e.,  $\Psi = \{\alpha, \beta, \gamma\}$ . Following a similar procedure to that in the previous section it can be shown that the E-step of the

EM algorithm for a SAR model reduces to calculating the following functional:

$$\begin{aligned}
F(\Psi; \Psi^{(n)}) &= \log[\det(\alpha^{-1}\mathbf{I})] + \text{tr}[\alpha\mathbf{Q}^t\mathbf{Q}\mathbf{R}_{f|g}^{(n)}] + \mathbf{m}_{f|g}^{(n)t}\alpha\mathbf{Q}^t\mathbf{Q}\mathbf{m}_{f|g}^{(n)} \\
&+ \log[\det(\mathbf{R}_T(\mathbf{f}))] + \text{tr}[\bar{\mathbf{H}}^t\mathbf{R}_T^{-1}(\mathbf{f})\bar{\mathbf{H}}\mathbf{R}_{f|g}^{(n)}] + (\bar{\mathbf{H}}\mathbf{m}_{f|g}^{(n)} - \mathbf{g})^t\mathbf{R}_T^{-1}(\mathbf{f})(\bar{\mathbf{H}}\mathbf{m}_{f|g}^{(n)} - \mathbf{g}),
\end{aligned} \tag{31}$$

in which

$$\mathbf{m}_{f|g}^{(n)} = (\alpha^{(n)}\mathbf{Q}^t\mathbf{Q})^{-1}\bar{\mathbf{H}}^t \left( \bar{\mathbf{H}}(\alpha^{(n)}\mathbf{Q}^t\mathbf{Q})^{-1}\bar{\mathbf{H}}^t + \mathbf{R}_T^{(n)}(\mathbf{f}) \right)^{-1} \mathbf{g} \tag{32}$$

and

$$\mathbf{R}_{f|g}^{(n)} = (\alpha^{(n)}\mathbf{Q}^t\mathbf{Q})^{-1} - (\alpha^{(n)}\mathbf{Q}^t\mathbf{Q})^{-1}\bar{\mathbf{H}}^t \left( \bar{\mathbf{H}}(\alpha^{(n)}\mathbf{Q}^t\mathbf{Q})^{-1}\bar{\mathbf{H}}^t + \mathbf{R}_T^{(n)}(\mathbf{f}) \right)^{-1} \bar{\mathbf{H}}^t(\alpha^{(n)}\mathbf{Q}^t\mathbf{Q})^{-1}. \tag{33}$$

It is easy to show using similar steps to those used to derive Eqs. (23), (24), and (25) that the DFT representation of (31), (32), and (33) is given by

$$\begin{aligned}
F(\Psi; \Psi^{(n)}) &= \sum_{i=0}^{N-1} -\log[\alpha] + \alpha|Q(i)|^2(S_{f|g}^{(n)}(i) + \frac{1}{N}|M_{f|g}^{(n)}(i)|^2) \\
&+ \log(N\beta\frac{1}{\alpha|Q(i)|^2} + \gamma) + \frac{1}{N\beta\frac{1}{\alpha|Q(i)|^2} + \gamma} \left\{ |\bar{H}(i)|^2 S_{f|g}^{(n)}(i) + \frac{1}{N}|\bar{H}(i)M_{f|g}^{(n)}(i) - G(i)|^2 \right\},
\end{aligned} \tag{34}$$

$$M_{f|g}^{(n)}(i) = \frac{\bar{H}^*(i)\frac{1}{\alpha^{(n)}}}{|\bar{H}(i)|^2\frac{1}{\alpha^{(n)}} + N\beta^{(n)}\frac{1}{\alpha^{(n)}} + \gamma^{(n)}|Q(i)|^2} G(i), \tag{35}$$

and

$$S_{f|g}^{(n)}(i) = \frac{\frac{1}{\alpha^{(n)}|Q(i)|^2} \left[ \frac{N\beta^{(n)}}{\alpha^{(n)}} + \gamma^{(n)}|Q(i)|^2 \right]}{\frac{|\bar{H}(i)|^2}{\alpha^{(n)}} + \frac{N\beta^{(n)}}{\alpha^{(n)}} + \gamma^{(n)}|Q(i)|^2}, \tag{36}$$

in which  $Q(i)$  are the eigenvalues of the circulant matrix  $\mathbf{Q}$ , and other quantities were defined in (23). Equations (34), (35), and (36) are the DFT-domain expressions for the E-step of the EM algorithm for SAR model.

The estimates of  $\alpha$ ,  $\beta$ , and  $\gamma$  are obtained in the M-step of the algorithm according to the following equations (see Appendix B):

$$\frac{1}{\alpha^{(n+1)}} = \frac{1}{N} \sum_{i=0}^{N-1} |Q(i)|^2 \left( S_{f|g}^{(n)}(i) + \frac{1}{N}|M_{f|g}^{(n)}(i)|^2 \right) \tag{37}$$

$$\beta^{(n+1)} = \beta^{(n)} \frac{v(\beta^{(n)})}{u(\beta^{(n)})} \tag{38}$$

$$\gamma^{(n+1)} = \gamma^{(n)} \frac{v(\gamma^{(n)})}{u(\gamma^{(n)})}, \tag{39}$$

where  $u(\beta^{(n)})$ ,  $v(\beta^{(n)})$ ,  $u(\gamma^{(n)})$ , and  $v(\gamma^{(n)})$  are given in Appendix B, in (B-18), (B-19), (B-20), and (B-21), respectively. Thus, the M-step of the EM algorithm for SAR model reduces to calculating (37), (38), and (39) at each iteration, in which  $M_{f|g}^{(n)}(i)$  and  $S_{f|g}^{(n)}(i)$  are computed by (35) and (36). In implementing these equation we let  $\epsilon = 0$ . Thus,  $\mathbf{Q}$  becomes the Laplacian operator. Notice that  $Q(0) = 0$  does not cause any problems in our calculations because although  $S_{f|g}^{(n)}(0)$  in (36) appears to be infinite in reality in all the formulas for the SAR model the product  $|Q(i)|^2 S_{f|g}^{(n)}(i)$  appears which for  $i = 0$  is finite.

## IV Convergence Analysis of the EM Power Spectrum Estimators

It is a well-known fact that the EM algorithm converges to a stationary point of the likelihood functional (see Ref. 14). In this section we examine the convergence properties of the estimator in (26) not in the likelihood values, but, rather in terms of the power-spectrum coefficient values. More specifically, we identify the fixed points of the iteration in (26) to see where the power-spectrum coefficients converge. For this purpose we assume that  $\beta$  and  $\gamma$  are known and restrict our observations only to the update of the power-spectrum coefficients.

Dropping the frequency index  $i$  for notational simplicity and substituting (24) and (25) into (26) we obtain

$$S_f^{(n+1)} = \frac{S_f^{(n)} [N\beta S_f^{(n)} + \gamma]}{|\bar{H}|^2 S_f^{(n)} + N\beta S_f^{(n)} + \gamma} + \frac{S_f^{(n)2} |\bar{H}|^2 \frac{1}{N} |G|^2}{[|\bar{H}|^2 S_f^{(n)} + N\beta S_f^{(n)} + \gamma]^2}. \quad (40)$$

We assume that the periodogram estimate of the degraded image power spectrum  $\frac{1}{N}|G|^2$  is equal to the ensemble power spectrum, i.e.,

$$\frac{1}{N}|G|^2 = |\bar{H}|^2 S_f + N\beta S_f + \gamma, \quad (41)$$

where  $S_f$  is the power spectrum of the original image. For the iterative process to be useful, the sequence  $\{S_f^{(1)}, S_f^{(2)}, \dots\}$  must converge to some value  $\tilde{S}_f$  which is a fixed point of the update

equation, i.e.,

$$\tilde{S}_f = S_f^{(n+1)} = S_f^{(n)}. \quad (42)$$

In other words, the update formula of (40) maps  $S_f^{(n)}$  into itself. To conclude convergence it suffices to show that the sequence in (40) is monotone and bounded. This is shown in Appendix C. To find where the sequence of estimates converges we let  $\tilde{S}_f = S_f^{(n+1)} = S_f^{(n)}$  in (40). Solving for the fixed points of the iteration (40) it is straightforward to see that  $\tilde{S}_f$  converges to the power spectrum coefficients  $S_f$ . Thus, we conclude that the proposed EM-based algorithm yields the power-spectrum estimate of the source image in the sense of Eq. (41), under the assumption that the statistics of the PSF and the observation errors are known.

Based on the preceding discussion and on the fact that the estimator in (30) neglects the conditional covariance term in (25), it is clear that (30) cannot converge to the power spectrum of the source image. It is easy to see where this estimator converges. Dropping the frequency index  $i$  for notational simplicity and solving for the fixed points of the iteration (30), we obtain three solutions, given by

$$\tilde{S}_f^2(|\bar{H}|^2 + N\beta)^2 + \tilde{S}_f \left[ 2(|\bar{H}|^2 + N\beta)\gamma - |\bar{H}|^2 \frac{1}{N}|G|^2 \right] + \gamma^2 = 0 \quad \text{and} \quad \tilde{S}_f = 0. \quad (43)$$

Substituting (41) into (43) it is straightforward to show that  $\tilde{S}_f$  converges to  $S_f$  only if  $\beta = \gamma = 0$ . If  $\beta \neq 0$  or  $\gamma \neq 0$ ,  $\tilde{S}_f$  will never converge to the power spectrum  $S_f$ .

## V Numerical Experiments

In this section numerical experiments are reported that test the proposed EM algorithms (for full  $\mathbf{R}_f$  and SAR models), and to compare them with the LMMSE approach in Ref. 4. All three algorithms are also compared to the “ideal” LMMSE restoration filter in (24) assuming the knowledge of all parameters. Since the true image power spectrum (ensemble statistic) is unavailable, for the “ideal” LMMSE filter we used the periodogram estimate of the power spectrum assuming the knowledge of the original image.

Although the proposed EM-based algorithms do not attempt to minimize the mean square error (MSE) of the restored image (it is the likelihood functional that is being optimized), for comparison purposes with the LMMSE filter in Ref. 4 and the “ideal” LMMSE filter we choose the MSE criterion as an objective measure of performance in all our experiments. The (per pixel) MSE is defined as

$$\text{MSE} = \frac{1}{N} \|\mathbf{f} - \hat{\mathbf{f}}\|_2^2, \quad (44)$$

where  $\mathbf{f}$  and  $\hat{\mathbf{f}}$  are the original and the restored (upon convergence) images, respectively. For the EM-based algorithms the convergence was tested in the likelihood, while for the LMMSE approach in Ref. 4 the convergence was tested on two subsequent power-spectrum estimates. To obtain statistically meaningful results we performed Monte-Carlo simulations in which the MSE was averaged over five different noise realizations. We experimentally observed that more than five noise realizations did not change the nature of the MSE results that we present.

The MSE is a function of two noise parameters:  $\beta$  and  $\gamma$ . To avoid plotting the 3-D plot of the MSE versus both noise parameters and to enhance clarity and visibility of the results we plot two representative 2-D MSE plots: (constant- $\gamma$ ): For a fixed  $\text{SNR}_g = 30\text{dB}$  we plot MSE versus  $\text{SNR}_h$  by varying  $\beta$ , and (constant- $\beta$ ): For a fixed  $\text{SNR}_h = 20\text{dB}$  we plot MSE versus  $\text{SNR}_g$  by varying  $\gamma$ . In those plots the noise parameters are expressed in terms of the signal-to-noise ratios (SNR), i.e.,

$$\text{SNR}_h = \frac{\|\bar{\mathbf{h}}\|^2}{N\beta}, \quad (45)$$

where  $\|\bar{\mathbf{h}}\|^2$  is the energy of the known part of the PSF, and

$$\text{SNR}_g = \frac{\|\mathbf{f}\|^2}{N\gamma}, \quad (46)$$

where  $\|\mathbf{f}\|^2$  is the energy of the original image. In what follows we present three experiments in which the  $256 \times 256$  “Lena” test image was used as the source image.

In all experiments presented in this paper Gaussian-shaped PSF given below was used for blurring:

$$h(i, j) = c \cdot \exp \left\{ -\frac{i^2 + j^2}{2 \cdot 3^2} \right\}, \quad \text{for } i, j = -15, -14, \dots, -1, 0, 1, \dots, 14, 15, \quad (47)$$

where  $c$  is a constant chosen so that  $\sum_{i,j} h(i,j) = 1$ . The same kernel as in (47) with the additive white-noise component of variance  $\beta$  was used for restoration. The blurred data was further degraded with additive white observation noise of variance  $\gamma$ . We also performed the experiments where the “smooth” PSF from (47) was used for restoration, while the noisy one ((47) plus the additive PSF noise) was used in the blurring process. Due to the lack of space those experiments are not presented. Similar results in both cases are obtained.

In every iteration of the EM algorithm the function  $F(\Psi; \Psi^{(n)})$  was monitored. The algorithm is terminated based on the convergence of  $F(\Psi; \Psi^{(n)})$ . Furthermore, since the exact minimum of  $F(\Psi; \Psi^{(n)})$  cannot be found in the M-step, decreasing values of  $F(\Psi; \Psi^{(n)})$  guarantee convergence based on the theory of the generalized EM (GEM) algorithm<sup>14</sup>. During all our experiments we observed that the proposed updates for  $S_f$ ,  $\alpha$ ,  $\beta$  and  $\gamma$  reduced  $F(\Psi; \Psi^{(n)})$  at each iteration.

### Experiment 1

In the first part of this experiment we compare the LMMSE restoration filter in Ref. 4 (Eq. (29) in this paper) and the proposed EM-based conditional-mean estimator in (24). For comparison purposes we assume that noise parameters  $\beta$  and  $\gamma$  are known. To examine the lower MSE bound of these restoration filters we supply to both filters the periodogram estimate of  $S_f(i)$  obtained from the source image. Note that in this case the EM-based conditional mean estimator for the full  $\mathbf{R}_f$  prior is equivalent to the “ideal” LMMSE estimator. The constant- $\gamma$  MSE plot is shown in Fig. 1(a), and the constant- $\beta$  MSE plot is shown in Fig. 1(b). It is interesting to notice in Fig. 1(b) that the MSE of the LMMSE in Ref. 4 increases as the variance of the additive noise  $\gamma$  decreases. This counter-intuitive behavior can be explained by the fact that the LMMSE filter in Eq. (29) is underregularized due to the lack of the factor  $N$  that in the term  $\beta$ . Thus, when the additive noise  $\gamma$  is small the underregularization is more severe, therefore, the LMMSE filter in Eq. (29) is less stable which increases the MSE.

In the second part of this experiment in Figs. 2(a) and 2(b) we compare the image power spectrum estimators in (26) (full  $\mathbf{R}_f$  model case) and (37) (SAR model case) with the image power-

spectrum estimator in Ref. 4 (Eq. (30) in this paper) when  $\beta$  and  $\gamma$  are known. The constant- $\gamma$  MSE plot is shown in Fig. 2(a), and the constant- $\beta$  MSE plot is shown in Fig. 2(b). The sample images corresponding to the presented plots are in Fig. 2(b) are shown in Fig. 3. It is interesting to notice in Figs. 2(a) and 2(b) that the “ideal” LMMSE filter, the one where the periodogram estimate using the source image is used for the power spectrum  $S_f$ , yields slightly worse performance, for high SNRs, than the LMMSE where the power spectrum is estimated by the EM algorithm. An explanation for this is that the EM algorithm for high SNRs provides better estimates for the power spectrum of the source signal than the periodogram.

In both parts of this experiment all filters converged within 30 iterations. The presented experiments verify our previous claim that the derived LMMSE filter expression in (24) is the correct one as opposed to the one in (29) from Ref. 4 (without proper scaling of the autocorrelation to compensate for the lack of the N factor, see Appendix A), since it results in lower MSE. Based on the presented experiments and on the convergence analysis in Section IV it is clear that (in both perfectly-known and partially-known blur cases) the EM-based algorithms have the capability to simultaneously identify the image power spectrum and restore the degraded image. In other words, the EM algorithms for both image models attain (or almost attain, at very low SNR’s) the lower MSE bound defined by the “ideal” LMMSE filter.

## Experiment 2

In the first part of this experiment we test the proposed EM algorithms in estimating one of the noise parameters along with the image power spectrum while simultaneously restoring the degraded image. The other noise parameter is assumed known and is held fixed. Since no provision for estimating noise parameters was made in the LMMSE filter in Ref. 4, the filter could not be included in the comparison. The constant- $\gamma$  MSE plot is shown in Fig. 4, and the constant- $\beta$  MSE plot is shown in Fig. 5(a). The corresponding sample images are shown in Fig. 6.

In the second part of this experiment we assume that the PSF is perfectly known. In other words  $\beta = 0$  and we test the EM algorithm for both image models in estimating the observation

noise parameter  $\gamma$  along with the image power spectrum. In this case the proposed EM algorithm with the full  $\mathbf{R}_f$  model becomes equivalent to the iterative Wiener filter described in Ref. 21 (Ch. 6, pp. 158). The MSE versus  $\text{SNR}_g$  plot is shown in Fig. 5(b). In both parts of this experiment we initialized the estimates of the power spectrum of the source image for all filters using the degraded image.

In both parts of the experiment the algorithms converged within 20-30 iterations. Based on these experiments we make the following observations: (i) The EM algorithm (in both perfectly-known and partially-known blur cases) for the SAR model attained (or almost attained at very low SNRs) the “ideal” LMMSE lower-bound. At high SNR’s, in the partially-known blur case, the proposed algorithm actually slightly outperformed the “ideal” LMMSE filter, which only confirms that the periodogram estimate of the power spectrum assuming the knowledge of the original image is inferior to the true (ensemble) statistics. The algorithm also identified the unknown noise parameters with very high accuracy. (ii) The EM algorithm (in both perfectly-known and partially-known blur cases) for the full  $\mathbf{R}_f$  model failed to attain the “ideal” LMMSE lower bound for very low SNR’s. The reason for this could be the lack of information in the data to identify simultaneously  $N + 1$  parameters (the full power-spectrum and the unknown noise parameter). Additional prior knowledge of the unknown parameters (or more data) may be needed.

## VI Conclusions and Future Work

In this paper we derived two image-restoration techniques based on the Expectation-Maximization (EM) algorithm in Ref. 14 which allow for inexact knowledge of the point-spread function (PSF). In the first algorithm the full  $\mathbf{R}_f$  model for the source image was used, while in the second algorithm the SAR model was utilized. We showed that the linear minimum mean square error (LMMSE) algorithm of Refs. 4 and 6 is a special case of our first EM algorithm. Furthermore, we showed and verified experimentally that the correct expression of the LMMSE filter for this problem, when the PSE noise is white, contains a factor  $N$  which multiplies the PSF noise variance. This factor was

omitted in the derivation of the LMMSE filter in Ref. 6 and 4.

Convergence analysis of the EM algorithm for the full  $\mathbf{R}_f$  image model indicates that if good estimates of the statistics of both observation and PSF noises are available, then the algorithm can overcome the lack of knowledge of the signal statistics and produce excellent restorations. The SAR image model -based EM algorithm showed the capability to accurately identify both the power spectrum of the source image and one noise parameter simultaneously while producing excellent restorations.

When the variances of both noises are unknown it is not possible to estimate them accurately. The total noise term  $N\beta S_f(i) + \gamma$  for the Gaussian model or  $\frac{N\beta}{\alpha|Q(i)|^2} + \gamma$  for the SAR model contains  $\beta$  and  $\gamma$  in a complementary fashion and they cannot be distinguished from the data only. We have observed that in this case the iterative estimation algorithm is very ill-behaved and the estimate depends on the initialization. In order to correct this problem knowledge in the form of a prior distribution for  $\beta$  and  $\gamma$  is necessary in order to simultaneously restore the image and estimate all the unknown parameters in this problem. Work along those lines is currently under way (see Refs. 13 Ch. 5 and 19).

Finally, we would like to point out that the algorithms in this paper could be derived without the Gaussian assumption using fixed-point analysis and additive correction in a manner similar to the work in Refs. 23, and 24 for the classical restoration problem where the PSF is perfectly known. However, for this problem this approach is very tedious and somewhat ad-hoc.

## VII Acknowledgment

This work was supported by NSF grant MIP-9309910.

## VIII Appendix A: Derivation of the Eigenvalues of the Noise Autocorrelation Matrix $\mathbf{R}_T(\mathbf{f})$ .

First, we calculate  $\mathbf{R}_T(\mathbf{f})$  in DFT domain using the circulant assumption for  $\Delta H$  and  $\mathbf{R}_f$ . We first rewrite (8) as follows:

$$\begin{aligned}\mathbf{R}_T(\mathbf{f}) &= E_{\Delta h} \left\{ \Delta \mathbf{H} \mathbf{R}_f \Delta \mathbf{H}^t \right\} + \mathbf{R}_{\Delta g} = \left[ E_{\Delta h} \left\{ \Delta \mathbf{H}^t \Delta \mathbf{H} \right\} \right] \mathbf{R}_f + \mathbf{R}_{\Delta g} \\ &= \left[ E_{\Delta h} \left\{ \sum_{i=0}^{N-1} (\Delta \mathbf{H})_i (\Delta \mathbf{H}^t)_i \right\} \right] \mathbf{R}_f + \mathbf{R}_{\Delta g},\end{aligned}\tag{A-1}$$

where  $(\Delta \mathbf{H})_i$  is the  $i^{\text{th}}$  column of matrix  $\Delta \mathbf{H}$ , whose  $n^{\text{th}}$  element is equal to:

$$(\Delta \mathbf{H})_i(n) = \Delta \mathbf{h}(n - i)_{\text{mod } N}.\tag{A-2}$$

Let  $\mathbf{W}$  denote the  $N \times N$  DFT matrix,  $(\cdot)^{-1}$  and  $(\cdot)^H$  denote the inverse and the Hermitian operations, respectively, and  $\mathbf{W}^{-1} = \frac{1}{N} \mathbf{W}^H$ . It is easy to see that the DFT coefficient corresponding to (A-2) is given by

$$[\mathbf{W}(\Delta \mathbf{H})_i](n) = [\mathbf{W}\Delta \mathbf{h}](n) \exp(-j \frac{2\pi}{N} n i),\tag{A-3}$$

where  $j$  is the imaginary unit. Then, the DFT of  $(\Delta \mathbf{H})_i$  can be written as

$$\mathbf{W}(\Delta \mathbf{H})_i = \text{diag} \left[ 1, \exp(-j \frac{2\pi}{N} i), \dots, \exp(-j \frac{2\pi}{N} (N-1)i) \right] [\mathbf{W}\Delta \mathbf{h}].\tag{A-4}$$

Further, we note that

$$\mathbf{W} E_{\Delta h} \left\{ \sum_{i=0}^{N-1} (\Delta \mathbf{H})_i (\Delta \mathbf{H}^H)_i \right\} \mathbf{W}^{-1} = \sum_{i=0}^{N-1} E_{\Delta h} \left\{ \mathbf{W}(\Delta \mathbf{H})_i (\Delta \mathbf{H}^H)_i \mathbf{W}^H \frac{1}{N} \right\}\tag{A-5}$$

can be rewritten using (A-4) as follows:

$$\begin{aligned}\mathbf{W} E_{\Delta h} \left\{ \sum_{i=0}^{N-1} (\Delta \mathbf{H})_i (\Delta \mathbf{H}^H)_i \right\} \mathbf{W}^{-1} &= \sum_{i=0}^{N-1} E_{\Delta h} \left\{ [\mathbf{W}\Delta \mathbf{h}] [\mathbf{W}\Delta \mathbf{h}]^H \frac{1}{N} \right\} \\ &= \sum_{i=0}^{N-1} \mathbf{W} E_{\Delta h} \left\{ \Delta \mathbf{h} \Delta \mathbf{h}^H \right\} \mathbf{W}^{-1} = \sum_{i=0}^{N-1} \mathbf{W} \mathbf{R}_{\Delta h} \mathbf{W}^{-1} \\ &= \sum_{i=0}^{N-1} \text{diag} [\beta, \dots, \beta] = N \text{diag} [\beta, \dots, \beta],\end{aligned}\tag{A-6}$$

where  $\beta$  is the variance of the PSF noise (or the eigenvalues of  $\mathbf{R}_{\Delta h}$ ). Finally, using the result of (A-6), it is straightforward to obtain the DFT of (A-1), i. e.,

$$\mathbf{W} \mathbf{R}_T(\mathbf{f}) \mathbf{W}^{-1} = \text{diag} [N\beta S_f(0) + \gamma, N\beta S_f(1) + \gamma, \dots, N\beta S_f(N-1) + \gamma],\tag{A-7}$$

where  $S_f(i)$  are the power spectrum coefficients of the source image (or the eigenvalues of  $\mathbf{R}_f$ ) and  $\gamma$  denotes the observation-noise power-spectrum (or the eigenvalues of  $\mathbf{R}_{\Delta g}$ ). Thus, the eigenvalues of  $\mathbf{R}_T(\mathbf{f})$  are equal to

$$N\beta S_f(i) + \gamma. \quad (\text{A-8})$$

A different formula for calculating  $\mathbf{R}_T(\mathbf{f})$  was derived in Ref. 4; in the white-noise case this formula becomes (Ref. 4, Eq. (18a), pp. 1256):

$$\mathbf{R}_T(\mathbf{f}) = \beta\mathbf{R}_f + \gamma\mathbf{I}. \quad (\text{A-9})$$

It is easy to see that the eigenvalues of  $\mathbf{R}_T(\mathbf{f})$  in (A-9) are given by

$$\beta S_f(i) + \gamma. \quad (\text{A-10})$$

Comparing the expressions for the eigenvalues of  $\mathbf{R}_T(\mathbf{f})$  in (A-10) and (A-8) we note that (A-10) lacks a factor of  $N$ . Proper scaling of the autocorrelation of  $\mathbf{R}_f$  in the implementation of the LMMSE filter could render this problem non-existent. However, lack of this  $N$  factor compromises the value of this filter as demonstrated in our experiments.

## IX Appendix B: Derivation of the M-step of the EM algorithm

### IX.A Full $\mathbf{R}_f$ Model

From (23) it is clear that its minimization with respect to the  $i^{\text{th}}$  frequency coefficient is independent from other frequencies. Therefore, taking the partial derivative of (23) with respect to  $S_f(i)$ , and setting it equal to zero yields the following nonlinear equation with respect to  $S_f^{(n+1)}(i)$ :

$$\begin{aligned} & \frac{1}{S_f^{(n+1)}(i)} - \frac{1}{[S_f^{(n+1)}(i)]^2} \left\{ S_{f|g}^{(n)}(i) + \frac{1}{N} |M_{f|g}^{(n)}(i)|^2 \right\} + \frac{N\beta^{(n)}}{N\beta^{(n)}S_f^{(n+1)}(i) + \gamma^{(n)}} \\ & - \frac{N\beta^{(n)}}{[N\beta^{(n)}S_f^{(n+1)}(i) + \gamma^{(n)}]^2} \left\{ |\bar{H}(i)|^2 S_{f|g}^{(n)}(i) + \frac{1}{N} |\bar{H} M_{f|g}^{(n)}(i) - G(i)|^2 \right\} = 0. \end{aligned} \quad (\text{B-1})$$

The minimization of (B-1) can be greatly simplified by observing that the dependency of the last two terms of (B-1) with respect to  $S_f(i)$  is very weak compared to the first two terms. This is due to the multiplication of  $S_f(i)$  by a factor of  $N\beta$  in the last two terms of (B-1). This multiplication factor is small (typically less than  $10^{-2}$  for  $\text{SNR}_h = 20\text{dB}$ ) and therefore, the location of the minimum of (23) with respect to  $S_f(i)$  is primarily determined by the first two terms. We verified this in computer simulations in which we minimized (B-1) with respect to  $S_f(i)$  with and without the last two terms: the obtained minimizers were practically identical. We also analyzed the plots of (B-1) as a function of  $S_f(i)$ , and we observed that the curvature of the functional in (B-1) is captured by the first two terms only. Therefore, for the purposes of numerical computations of (B-1) we can neglect the last two terms of (B-1). Now, the minimizer of (B-1) can be obtained in the closed form, i.e.,

$$S_f^{(n+1)}(i) = S_{f|g}^{(n)}(i) + \frac{1}{N} |M_{f|g}^{(n)}(i)|^2. \quad (\text{B-2})$$

Taking the partial derivative of (23) with respect to  $\beta$  and setting it equal to zero we obtain

$$\sum_{i=0}^{N-1} \frac{NS_f^{(n)}(i)}{N\beta^{(n+1)}S_{f|g}^{(n)}(i) + \gamma^{(n)}} = \sum_{i=0}^{N-1} \frac{NS_f^{(n)}(i)}{[N\beta^{(n+1)}S_{f|g}^{(n)}(i) + \gamma^{(n)}]^2} \left\{ |\bar{H}(i)|^2 S_{f|g}^{(n)}(i) + \frac{1}{N} |\bar{H}(i)M_{f|g}^{(n)}(i) - G(i)|^2 \right\}. \quad (\text{B-3})$$

Let

$$u(\beta) = \sum_{i=0}^{N-1} \frac{NS_f^{(n)}(i)}{N\beta S_{f|g}^{(n)}(i) + \gamma^{(n)}} \quad (\text{B-4})$$

and

$$v(\beta) = \sum_{i=0}^{N-1} \frac{NS_f^{(n)}(i)}{[N\beta S_{f|g}^{(n)}(i) + \gamma^{(n)}]^2} \left\{ |\bar{H}(i)|^2 S_{f|g}^{(n)}(i) + \frac{1}{N} |\bar{H}(i)M_{f|g}^{(n)}(i) - G(i)|^2 \right\}. \quad (\text{B-5})$$

Then, the solution of (B-3) is attained when

$$u(\beta) = v(\beta) \quad (\text{B-6})$$

for some  $\beta = \beta^*$ . The functions  $u(\beta)$  and  $v(\beta)$  have linear and quadratic rates of decrease in  $\beta$ , respectively. Furthermore, it can be shown that  $v(0) > u(0)$ . Therefore, (B-6) has a unique solution on the positive  $\beta$  axis. Thus, the solution of (B-6) can be obtained using a simple iterative

procedure:

$$\beta^{(n+1)} = \beta^{(n)} \frac{v(\beta^{(n)})}{u(\beta^{(n)})}, \quad (\text{B-7})$$

where  $u(\beta^{(n)})$  and  $v(\beta^{(n)})$  are Eqs. (B-4) and (B-5) calculated at iteration ( $n$ ). It is straightforward to verify that  $\beta^{(n+1)}$  converges to  $\beta^*$ . To see this, assume that  $\beta^{(n)} < \beta^*$ . This implies  $\frac{v(\beta^{(n)})}{u(\beta^{(n)})} > 1$ , and then  $\beta^{(n+1)} > \beta^{(n)}$ . In this case  $\beta^{(n+1)}$  will converge to  $\beta^*$  from below. Conversely, if  $\beta^{(n)} > \beta^*$ , the ratio  $\frac{v(\beta^{(n)})}{u(\beta^{(n)})} < 1$ , and then  $\beta^{(n+1)} < \beta^{(n)}$  which guarantees convergence from above.

Taking the partial derivative of the last two terms in (23) with respect to  $\gamma$  and setting it equal to zero we obtain:

$$\sum_{i=0}^{N-1} \frac{1}{N\beta^{(n)}S_f^{(n)}(i) + \gamma^{(n+1)}} = \sum_{i=0}^{N-1} \frac{1}{[N\beta^{(n)}S_f^{(n)}(i) + \gamma^{(n+1)}]^2} \left\{ |\bar{H}(i)|^2 S_{f|g}^{(n)}(i) + \frac{1}{N} |\bar{H}(i)M_{f|g}^{(n)}(i) - G(i)|^2 \right\}. \quad (\text{B-8})$$

Following similar steps as in the case of solving (B-3), an iterative procedure for finding the solution of (B-8) can be obtained as follows:

$$\gamma^{(n+1)} = \gamma^{(n)} \frac{v(\gamma^{(n)})}{u(\gamma^{(n)})}, \quad (\text{B-9})$$

where

$$u(\gamma^{(n)}) = \sum_{i=0}^{N-1} \frac{1}{N\beta^{(n)}S_f^{(n)}(i) + \gamma^{(n)}} \quad (\text{B-10})$$

and

$$v(\gamma^{(n)}) = \sum_{i=0}^{N-1} \frac{1}{[N\beta^{(n)}S_f^{(n)}(i) + \gamma^{(n)}]^2} \left\{ |\bar{H}(i)|^2 S_{f|g}^{(n)}(i) + \frac{1}{N} |\bar{H}(i)M_{f|g}^{(n)}(i) - G(i)|^2 \right\}. \quad (\text{B-11})$$

## IX.B SAR Model

Taking the derivative with respect to  $\alpha$  of  $F(\Psi, \Psi^{(n)})$  in Eq. (34), the M-step of the EM algorithm for the SAR model yields the following non-linear equation:

$$\begin{aligned} & -\frac{N}{\alpha^{(n+1)}} + \sum_{i=0}^{N-1} |Q(i)|^2 \left( S_{f|g}^{(n)}(i) + \frac{1}{N} |M_{f|g}^{(n)}(i)|^2 \right) - \sum_{i=0}^{N-1} \frac{N\beta^{(n)}}{N\alpha^{(n+1)}\beta^{(n)} + (\alpha^2)^{(n+1)} |Q(i)|^2 \gamma^{(n)}} \\ & + \sum_{i=0}^{N-1} \frac{N\beta^{(n)}|Q(i)|^2}{(\alpha^2)^{(n+1)} \left[ \frac{N\beta^{(n)}}{\alpha^{(n+1)}} + \gamma^{(n)} \right] |Q(i)|^2} \left\{ |\bar{H}(i)|^2 S_{f|g}^{(n)}(i) + \frac{1}{N} |\bar{H}(i)M_{f|g}^{(n)}(i) - G(i)|^2 \right\} = 0. \end{aligned} \quad (\text{B-12})$$

To simplify this minimization we neglect the last two terms in (B-12), as in the full  $\mathbf{R}_f$  model, since those terms for the same reasons as before, do not significantly influence the minimum of the functional. Then, from (B-12) the closed form solution for  $\alpha$  reads:

$$\frac{1}{\alpha^{(n+1)}} = \frac{1}{N} \sum_{i=0}^{N-1} |Q(i)|^2 \left( S_{f|g}^{(n)}(i) + \frac{1}{N} |M_{f|g}^{(n)}(i)|^2 \right). \quad (\text{B-13})$$

In a manner similar to the Gaussian case, the estimates for  $\beta$  and  $\gamma$  can be obtained respectively from the following equations:

$$\sum_{i=0}^{N-1} \frac{\frac{N}{\alpha^{(n)}}}{\frac{N\beta^{(n+1)}}{\alpha^{(n)}} + \gamma^{(n)} |Q(i)|^2} = \sum_{i=0}^{N-1} \frac{\frac{N|Q(i)|^2}{\alpha^{(n)}}}{\left[ \frac{N\beta^{(n+1)}}{\alpha^{(n)}} + \gamma^{(n)} |Q(i)|^2 \right]^2} \left\{ |\bar{H}(i)|^2 S_{f|g}^{(n)}(i) + \frac{1}{N} |\bar{H}(i) M_{f|g}^{(n)}(i) - G(i)|^2 \right\} \quad (\text{B-14})$$

and

$$\sum_{i=0}^{N-1} \frac{|Q(i)|^2}{\frac{N\beta^{(n)}}{\alpha^{(n)}} + \gamma^{(n+1)} |Q(i)|^2} = \sum_{i=0}^{N-1} \frac{|Q(i)|^4}{\left[ \frac{N\beta^{(n)}}{\alpha^{(n)}} + \gamma^{(n+1)} |Q(i)|^2 \right]^2} \left\{ |\bar{H}(i)|^2 S_{f|g}^{(n)}(i) + \frac{1}{N} |\bar{H}(i) M_{f|g}^{(n)}(i) - G(i)|^2 \right\}. \quad (\text{B-15})$$

By following the same approach as in the Gaussian case the iterative solutions for (B-14) and (B-15), respectively can be obtained as follows:

$$\beta^{(n+1)} = \beta^{(n)} \frac{v(\beta^{(n)})}{u(\beta^{(n)})} \quad (\text{B-16})$$

and

$$\gamma^{(n+1)} = \gamma^{(n)} \frac{v(\gamma^{(n)})}{u(\gamma^{(n)})}, \quad (\text{B-17})$$

where

$$u(\beta^{(n)}) = \sum_{i=0}^{N-1} \frac{\frac{N}{\alpha^{(n)}}}{\frac{N\beta^{(n)}}{\alpha^{(n)}} + \gamma^{(n)} |Q(i)|^2} \quad (\text{B-18})$$

$$v(\beta^{(n)}) = \sum_{i=0}^{N-1} \frac{\frac{N|Q(i)|^2}{\alpha^{(n)}}}{\left[ \frac{N\beta^{(n)}}{\alpha^{(n)}} + \gamma^{(n)} |Q(i)|^2 \right]^2} \left\{ |\bar{H}(i)|^2 S_{f|g}^{(n)}(i) + \frac{1}{N} |\bar{H}(i) M_{f|g}^{(n)}(i) - G(i)|^2 \right\} \quad (\text{B-19})$$

$$u(\gamma^{(n)}) = \sum_{i=0}^{N-1} \frac{|Q(i)|^2}{\frac{N\beta^{(n)}}{\alpha^{(n)}} + \gamma^{(n)} |Q(i)|^2} \quad (\text{B-20})$$

$$v(\gamma^{(n)}) = \sum_{i=0}^{N-1} \frac{|Q(i)|^4}{\left[ \frac{N\beta^{(n)}}{\alpha^{(n)}} + \gamma^{(n)} |Q(i)|^2 \right]^2} \left\{ |\bar{H}(i)|^2 S_{f|g}^{(n)}(i) + \frac{1}{N} |\bar{H}(i) M_{f|g}^{(n)}(i) - G(i)|^2 \right\}. \quad (\text{B-21})$$

## X Appendix C: Convergence proof for the iterative power-spectrum estimator

Let us first examine the first derivative of the iteration sequence in (40). It is straightforward to show that

$$\frac{\partial S_f^{(n+1)}}{\partial S_f^{(n)}} = \frac{2|\bar{H}|^2 \frac{1}{N}|G|^2 S_f^{(n)2}}{[|\bar{H}|^2 S_f^{(n)} + N\beta S_f^{(n)} + \gamma]^3 \gamma} + \frac{\gamma(2N\beta S_f^{(n)} + \gamma) + N\beta S_f^{(n)2}}{[|\bar{H}|^2 S_f^{(n)} + N\beta S_f^{(n)} + \gamma]^2} \geq 0, \quad (\text{C-1})$$

for all  $S_f^{(n)} \geq 0$ , which proves the monotonicity of the sequence. To show boundedness, we rewrite (40) as follows

$$S_f^{(n+1)} = \frac{S_f^{(n)}}{(|\bar{H}|^2 + N\beta)S_f^{(n)} + \gamma} \left[ \frac{\frac{1}{N}|G|^2}{(|\bar{H}|^2 + N\beta)S_f^{(n)} + \gamma} |\bar{H}|^2 S_f^{(n)} + N\beta S_f^{(n)} + \gamma \right], \quad (\text{C-2})$$

where  $\frac{1}{N}|G|^2 = (|\bar{H}|^2 + N\beta)S_f + \gamma$ , and  $S_f$  are the true image power-spectrum coefficients. If  $S_f^{(n)} \geq S_f$ , then

$$(|\bar{H}|^2 + N\beta)S_f^{(n)} + \gamma \geq \frac{1}{N}|G|^2 \quad (\text{C-3})$$

and (C-2) implies

$$S_f^{(n+1)} \leq S_f^{(n)}. \quad (\text{C-4})$$

Thus, the sequence in (40) is bounded from above. Similarly, if

$$(|\bar{H}|^2 + N\beta)S_f^{(n)} + \gamma \leq \frac{1}{N}|G|^2 \quad (\text{C-5})$$

then (C-2) implies

$$S_f^{(n+1)} \geq S_f^{(n)}, \quad (\text{C-6})$$

and the sequence is bounded from below. (The sequence is trivially bounded from below by zero since all quantities in (40) are positive).

## References

- [1] D. Kundur and D. Hatzinakos, "Blind Image Deconvolution," *IEEE Sig. Proc. Mag.*, 4, 43-64 (1996).
- [2] D. Kundur and D. Hatzinakos, "Blind Deconvolution Revisited," *IEEE Sig. Proc. Mag.*, 6, 61-63 (1996).
- [3] D. Slepian, "Linear least squares filtering of distorted images" *J. Opt. Soc. Am.* 7, 918-922 (1967).
- [4] R. K. Ward and B. E. A. Saleh, "Restoration of Images Distorted by Systems of Random Impulse Response", *J. Opt.Soc. Am. A* 8, 1254-1259 (1985).
- [5] R. K. Ward and B. E. A. Saleh, "Deblurring Random Blur", *IEEE Trans. Acoust., Speech and Signal Processing* 10, 1494-1498 (1987).
- [6] L. Guan and R. K. Ward, "Deblurring Random Time-Varying Blur", *J. Opt.Soc. Am. A* 11, 1727-1737 (1989).
- [7] L. Guan and R. K. Ward, "Restoration of Randomly Blurred Images by the Wiener Filter", *IEEE Tran. on Acoust. Speech and Signal Processing* 10, 589-592 (1989).
- [8] P. L. Combettes, and H. J. Trussell, "Methods for Digital Restoration of Signals Degraded by Stochastic Impulse Response," *IEEE Tran. on Acoust. Speech and Signal Processing* 3, pp. 393-401, March 1989.
- [9] V. Z. Mesarović, N. P. Galatsanos, and A. K. Katsaggelos, "Regularized Constrained Total Least Squares Image Restoration," *IEEE Trans. on Image Processing* 8, 1096-1108 (1995).
- [10] V. Z. Mesarović, N. Galatsanos, and M. N. Wernick, "Restoration from partially-known blur using an expectation-maximization algorithm for tomographic reconstruction," 1995 Conference Record, *IEEE Nuclear Science Symposium & Medical Imaging Conference*, pp. 1257-1261, October 1995.
- [11] E. J. Hoffman and M. E. Phelps, *Positron Emission Tomography and Autoradiography: Principles and Applications for the Brain and Heart*, (Raven Press, New York 1986).
- [12] V. Z. Mesarović, N. P. Galatsanos, and M. Wernick, "Restoration from Partially-Known Blur Using an Expectation-Maximization Algorithm," In *Proceedings of Thirtieth ASILOMAR conference*, (Pacific Grove 1996) p. 95
- [13] V. Z. Mesarović, *Image Restoration Under Point Spread Function Uncertainties*, Ph.D. Dissertation, (Illinois Institute of Technology, Chicago, IL 1997).
- [14] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data",

J. Roy. Statist. Soc. B 39, 1-38 (1977).

[15] R. Molina, "On the Hierarchical Bayesian Approach to Image Restoration: Applications to Astronomical Images", IEEE Trans. on Pat. Anal. and Mac. Int. 11, 1122-1188 (1994).

[16] R. Molina, A. K. Katsaggelos, and J. Mateos, "Bayesian and Regularization Methods for Hyperparameter Estimation in Image Restoration," IEEE Trans. on Image Processing 2, 231-245 (1999).

[17] S. M. Kay, Fundamentals of Statistical Signal Processing, Englewood Cliffs, (NJ: Prentice-Hall, 1993).

[18] H. Andrews and B. Hunt, Digital Image Restoration, (Englewood Cliffs, NJ: Prentice-Hall, 1977).

[19] N. Galatsanos, V. Mesarovic, R. Molina, and A. Katsaggelos, "Hyperparameter Estimation Using Hyperpriors for Hierarchical Bayesian Image Restoration from Partially-Known Blurs," In Proceedings of SPIE on Bayesian Inference for Inverse Problems, 3459, (San Diego, July 1998). p. 337.

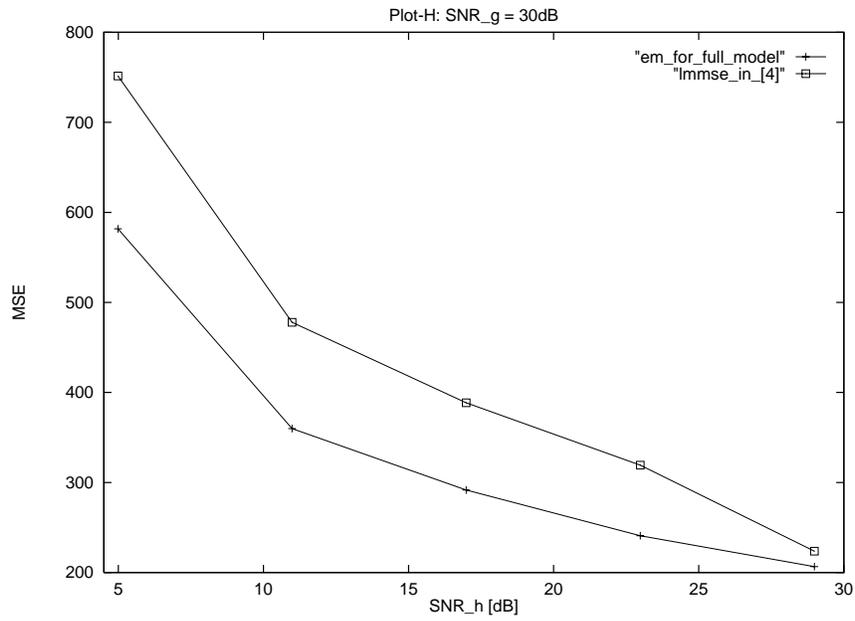
[20] N. P. Galatsanos and A. K. Katsaggelos, "Methods for Choosing the Regularization Parameter and Estimating the Noise Variance in Image Restoration and Their Relation", IEEE Trans. Image Processing 3, 322-336 (1992).

[21] A. K. Katsaggelos and K. T. Lay, "Maximum Likelihood Identification and Restoration of Images Using the Expectation-Maximization Algorithm", in Digital Image Restoration, A. K. Katsaggelos, editor, (Springer Verlag, 1991) pp. 143-176.

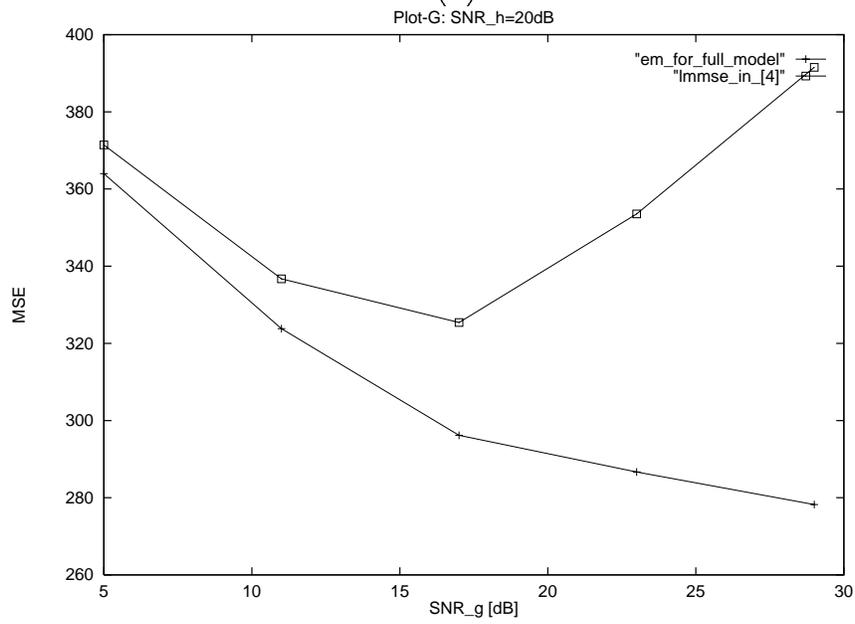
[22] R. L. Lagendijk, J. Biemond, and D. E. Boeke, "Identification and Restoration of Noisy Blurred Images Using the Expectation-Maximization Algorithm", IEEE Trans. Acoust. Speech and Signal Processing 7, 1180-1191 (1990).

[23] A. Hillery and R. T. Chin, "Iterative Wiener Filters for Image Restoration," IEEE Trans. on Signal Processing 8, (1991).

[24] A. Hillery, "Parameter Estimation for Image Restoration", Ph.D. Dissertation, (University of Wisconsin-Madison, Madison 1991.)

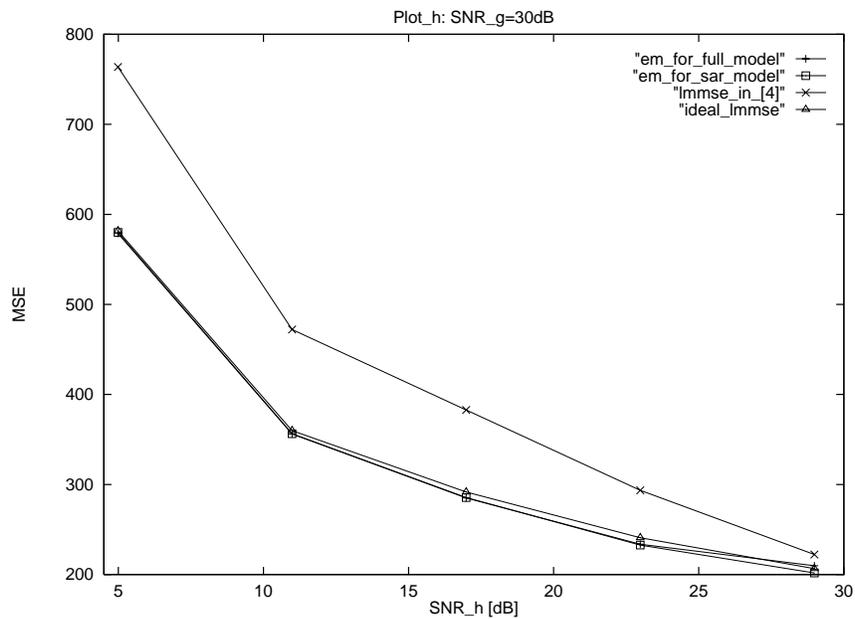


(a)

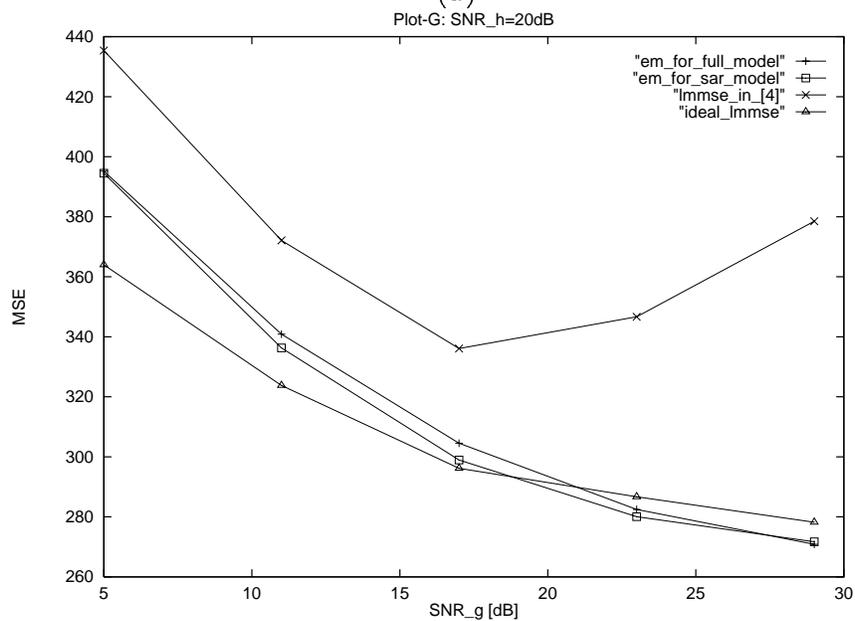


(b)

Figure 1: Experiment 1: Restoration with all statistics assumed known. (a) Constant- $\gamma$  MSE plot. (b) Constant- $\beta$  MSE plot.



(a)



(b)

Figure 2: Experiment 1: Restoration and simultaneous estimation of the image statistics. (a) Constant- $\beta$  MSE plot. (b) Constant- $\gamma$  MSE plot.



(a)



(b)



(c)



(d)



(e)

Figure 3: Experiment 1: Restoration and simultaneous estimation of the image statistics: (a) Degraded image by  $\text{SNR}_h = 20\text{dB}$ ,  $\text{SNR}_g = 11\text{dB}$ . (b) LMMSE in [4], (c) EM with full  $\mathbf{R}_f$  model, (d) EM with SAR model, (e) "ideal" LMMSE.

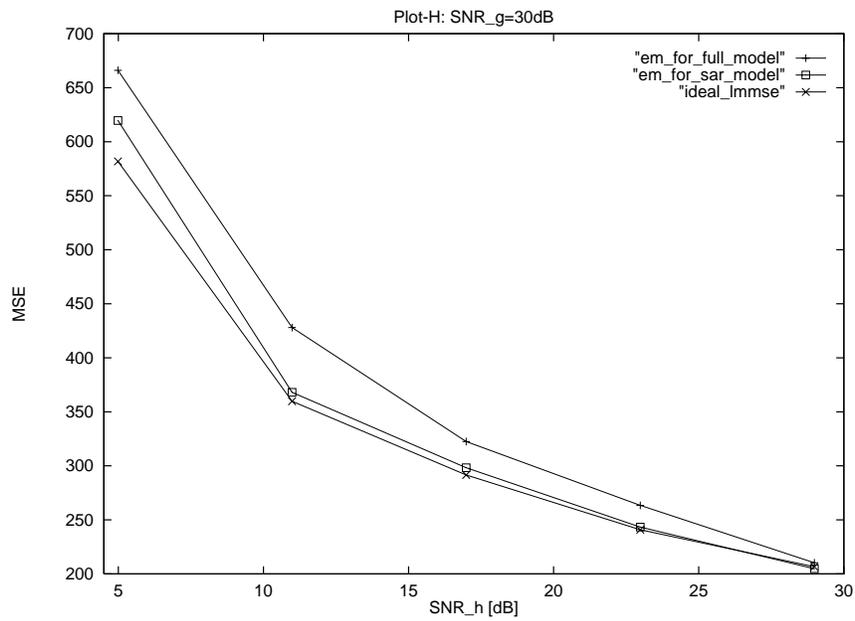
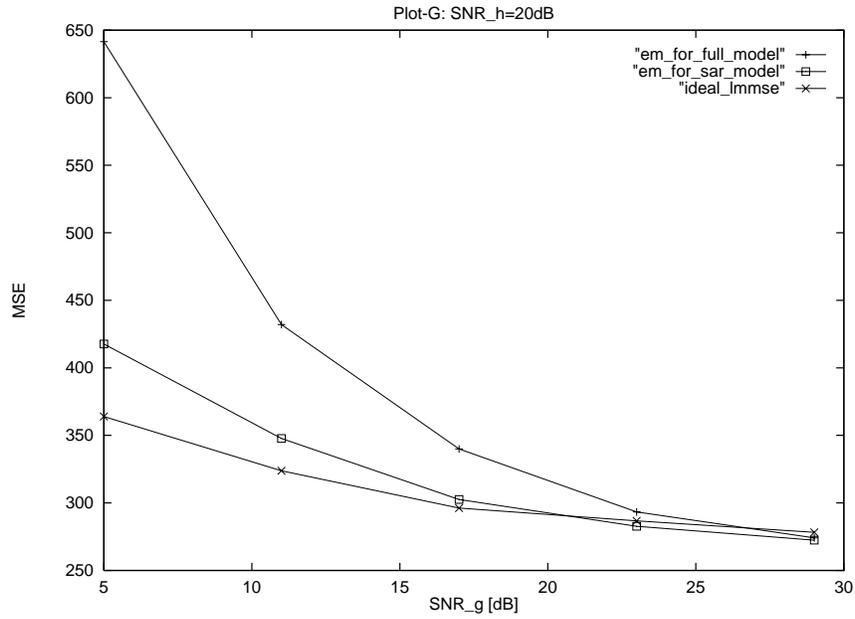
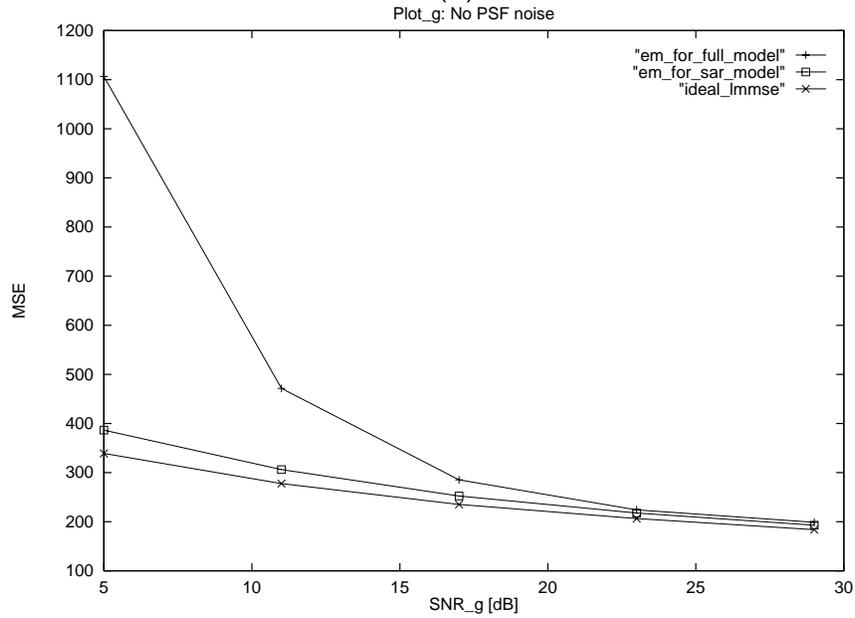


Figure 4: Experiment 2: Restoration and simultaneous estimation of the PSF noise and the image statistics. Constant- $\gamma$  MSE plot.



(a)



(b)

Figure 5: Experiment 2: Restoration and simultaneous estimation of the additive noise and image statistics. (a) Constant- $\beta$  MSE plot for partially known blur case. (b) Constant- $\beta$  MSE plot for perfectly known blur case ( $\beta = 0$ ).



(a)



(b)



(c)



(d)

Figure 6: Experiment 2: Restoration and simultaneous estimation of the additive noise and the image statistics. (a) Degraded image by  $\text{SNR}_h = 20dB$ ,  $\text{SNR}_g = 11dB$ , (b) EM with full  $\mathbf{R}_f$  model, (c) EM with SAR model, (d) “ideal” LMMSE.

## LIST OF FIGURES

Figure 1: Experiment 1: Restoration with all statistics assumed known. (a) Constant- $\gamma$  MSE plot. (b) Constant- $\beta$  MSE plot.

Figure 2: Experiment 1: Restoration and simultaneous estimation of the image statistics. (a) Constant- $\beta$  MSE plot. (b) Constant- $\gamma$  MSE plot.

Figure 3: Experiment 1: Restoration and simultaneous estimation of the image statistics: (a) Degraded image by  $\text{SNR}_h = 20dB$ ,  $\text{SNR}_g = 11dB$ . (b) LMMSE in Ref. 4., (c) EM with full  $\mathbf{R}_f$  model, (d) EM with SAR model, (e) “ideal” LMMSE.

Figure 4: Experiment 2: Restoration and simultaneous estimation of the PSF noise and the image statistics. Constant- $\gamma$  MSE plot.

Figure 5: Experiment 2: Restoration and simultaneous estimation of the additive noise and image statistics. (a) Constant- $\beta$  MSE plot for partially known blur case. (a) Constant- $\beta$  MSE plot for perfectly known blur case ( $\beta = 0$ )

Figure 6: Experiment 2: Restoration and simultaneous estimation of the additive noise and the image statistics. (a) Degraded image by  $\text{SNR}_h = 20dB$ ,  $\text{SNR}_g = 11dB$ , (b) EM with full  $\mathbf{R}_f$  model, (c) EM with SAR model, (d) “ideal” LMMSE.