Physics-based models for measurement correlations. Application to an inverse Sturm-Liouville problem.

Guillaume Bal* Kui Ren†

February 15, 2009

Abstract

In many inverse problems, the measurement operator, which maps objects of interest to available measurements, is a smoothing (regularizing) operator. Its inverse is therefore unbounded and as a consequence, only the low frequency component of the object of interest is accessible from inevitably noisy measurements. In many inverse problems however, the neglected high frequency component may significantly affect the measured data. Using simple scaling arguments, we characterize the influence of the high frequency component. We then consider situations where the correlation function of such an influence may be estimated by asymptotic expansions, for instance as a random corrector in homogenization theory. This allows us to consistently eliminate the high frequency component and derive a closed-form, more accurate, inverse problem for the low frequency component of the object of interest. We present the asymptotic expression of the correlation matrix of the eigenvalues in a Sturm-Liouville problem with unknown potential. We propose an iterative algorithm for the reconstruction of the potential from knowledge of the eigenvalues and show that using the approximate correlation matrix significantly improves the reconstructions.

keywords: inverse problem, measurement correlations, Sturm-Liouville, random fluctuations, central limit correction to homogenization.

1 Introduction

Consider a general inverse problem of the form

$$y = \mathfrak{M}(x) + n, \tag{1}$$

where x is the unknown quantity, $\mathfrak{M}$ the measurement operator, n some noise term, and y the available noisy measurements. We assume that the linearization of $\mathfrak{M}$ is a

*Department of Applied Physics and Applied Mathematics, Columbia University, New York NY, 10027; gb2030@columbia.edu

†Department of Mathematics, University of Texas at Austin, Austin, TX 78712; ren@math.utexas.edu

compact, hence smoothing, operator. We also assume that x and y are represented in given bases, which may be chosen because the linear approximation of $\mathfrak{M}$ is sparse for these bases (for instance as the bases in the singular value decomposition of the linearization of $\mathfrak{M}$) or because objects of interest are sparsely represented with such a choice.

The effect of the noise then typically implies that the low-frequency component of x may be reconstructed relatively accurately while the high-frequency component of x is not accessible. Such a high-frequency component is then usually eliminated from the reconstruction by choosing an appropriate regularization $\mathfrak{R}_c$ of the inverse map $\mathfrak{M}^{-1}$, even though its effects on the measured data may not be negligible.

We propose a framework that allows us to account for errors in the measurements generated by the high frequency component of x. More specifically, let us define x_0 as the low frequency part of x and $x_n = x - x_0$ its high frequency part. Here, low frequency refers to the first coordinates of x in the chosen basis that we believe we can reconstruct, whereas high frequency refers to the rest of x. Upon linearizing (1) in the vicinity of x_0, we obtain the new (possibly) non-linear inverse problem for x_0:

$$y_0 = \mathfrak{M}(x_0) + \mathfrak{M}'(x_0)x_n + n. \tag{2}$$

Here, y_0 stands for the low frequency part of y. We may set $y_0 = y$ if all the data are retained or y_0 as a low-frequency projection of y if we believe that some components of the measurements are too noisy to be of any use in the reconstruction. The matrix $\mathfrak{M}' = \mathfrak{M}'(x_0)$ models the coupling between the high frequency part of x and the low frequency part of y. It may vanish in certain situations, for instance when $\mathfrak{M}$ is linear and x and y are decomposed in the bases used in the singular decomposition of $\mathfrak{M}$. When $\mathfrak{M}'$ does not vanish, the non-recoverable component of x implicitly increases the noise component in y_0.

Since x_n is not known, $\mathfrak{M}'x_n$ is difficult to compute or even estimate. It may however be modeled statistically, i.e., as the realization of a random distribution, in which case $\mathfrak{M}'x_n$ may at least be estimated statistically. Unfortunately, in many practical situations, x_n is a high-dimensional object. Its modeling and the estimation of the parameters involved in the modeling are therefore rather delicate and often outright impractical. Fortunately, such a complex modeling and estimation problem may not be necessary in practical situations where sufficient *statistical averaging* takes place.

This leads us to the main regime of interest in this paper. We assume that x_n is a spatially highly oscillatory object and that y_0 is a set of measurements of a differential equation involving the object x_n. Moreover, we assume the existence of a macroscopic theory that provides a *simple* expression (hence, easily parameterizable) for the law of the vector y_0 in terms of that of x_n. The macroscopic model we have in mind here comes from homogenization theory, and more precisely as an analysis of the random corrector to homogenization. Explicit asymptotic expressions then allow us to approximately characterize the law of y_0 as a function of the law of x_n.

Under such (admittedly restrictive) assumptions, a strategy for the inverse problem may be formulated as follows. From an initial guess for x_0, we obtain the approximate law for $\mathfrak{M}'(x_0)x_n$. This allows us to obtain an updated reconstruction of x_0, for instance by minimizing its variance using the appropriate approximation for y_0. The two-step procedure is applied iteratively until convergence.

What we gain from this is a relatively inexpensive, physics-based, asymptotically optimal mitigation of the measurement noise induced by fluctuations of the un-recoverable high frequency components of a parameter of interest. The resulting correlation matrix for such measurements is typically far from being diagonal and thus may provide far superior reconstructions for x_0 than when the influence of x_n on the measurements is ignored.

In section 3 below, we apply the methodology to the reconstruction of the potential in a one dimensional Schrödinger (Sturm-Liouville) equation from the measurement of two sets of eigenvalues corresponding to different boundary conditions. Because of inaccuracies in the measurements of large eigenvalues, the small scale structure of the potential is not accessible. It is then replaced by a highly oscillatory random process. Recent results on central limit corrections to homogenization [1] then allow us to approximately characterize the influence of the randomness on the measured eigenvalues, which form a highly correlated vector. The resulting correlation matrix for the measured eigenvalues depends on one unknown scaling parameter, which measures the strength of the oscillatory random process. We will see that in some situations, the proposed method significantly reduces the variance of the reconstructed potential at the cost of estimating one unknown parameter modeling the influence of the unrecoverable high frequency parameters.

The rest of the paper is structured as follows. Section 2 analyzes the scaling properties of a simplified version of (2). We then describe the physical setting and proposed reconstruction method for the one dimensional inverse spectral problem in section 3. Numerical simulations that allow us to quantify the interest of the method are presented in section 4.

2 Scaling and regularization

We now consider the influence of the various noise contributions in the following extremely simplified yet plausible scenario. We refer the reader to e.g. [3] for extensions to other regularization methods.

Linearization and high frequency cut-off. We consider a linearization of the nonlinear problem (1) of the form

$$y = Ax + n. \tag{3}$$

Let us assume that x and y are functions of a spatial variable t in $\mathbb{R}^d$ and that A is diagonal in the Fourier domain and has for symbol (Fourier multiplier)

$$\hat{A}(\xi) = \langle \xi \rangle^{-\alpha}, \qquad \langle \xi \rangle = \sqrt{1 + |\xi|^2}, \quad \alpha > 0. \tag{4}$$

This means that $Ax = \mathcal{F}^{-1}(\hat{A}(\xi)\hat{x}(\xi))$, where $\hat{x}(\xi)$ is the Fourier transform of $x(t)$ and $\mathcal{F}^{-1}$ is the inverse Fourier transform. Let us assume moreover that $x \in H^\beta$ for some $\beta > 0$, where H^β with norm $\| \cdot \|_\beta$ is the Hilbert space of functions with β square integrable derivatives, or equivalently such that $\int_{\mathbb{R}^d} \langle \xi \rangle^{2\beta} |\hat{x}(\xi)|^2 d\xi < \infty$.

Then the linear inverse problem (3) may be inverted as follows. Denote by $\delta = \|n\|$ the norm of the noise term, where $\| \cdot \|$ is the L^2 norm. We can define R_c with symbol

$\hat{R}_c(\xi)$ as

$$\hat{R}_c(\xi) = \begin{cases} \langle \xi \rangle^{\alpha}, & |\xi| < \xi_c \\ 0 & |\xi| > \xi_c, \end{cases} , \tag{5}$$

as an approximate inverse for A and $x_0 := R_c y$ as an approximate solution to (3). Then, classically [3], we obtain the following error estimate

$$\|x - x_0\| \leq \|R_c\|\delta + \|(R_c A - I)x\| \leq \langle \xi_c \rangle^{\alpha}\delta + \langle \xi_c \rangle^{-\beta}\|x\|_{\beta},$$

where we also denote by $\|\cdot\|$ the norm of linear bounded operators in $\mathcal{L}(L^2)$. It remains to choose $\langle \xi_c \rangle = (\delta^{-1}\|x\|_{\beta})^{\frac{1}{\alpha+\beta}}$ to obtain that the error is bounded by

$$\|x - x_0\| = \|x_n\| \leq \|x\|_{\beta}^{\frac{\alpha}{\alpha+\beta}} \delta^{\frac{\beta}{\alpha+\beta}}. \tag{6}$$

The above result provides us with a spatial resolution $\varepsilon = \langle \xi_c \rangle^{-1} \sim \delta^{\frac{1}{\alpha+\beta}}$.

This extremely simplified example displays the main features of the inverse problem of interest in this paper. The oscillations of x at a larger scale than ε are reconstructed in x_0 while the oscillations of x at a smaller scale than ε are not reconstructed because of the presence of noise n in (3).

High frequency low frequency coupling. Let us now come back to (1), which we replace by (2) assuming that x_n is small as was shown above. We further linearize (2) about a guessed value x_0 and obtain the following equation for δx_0

$$\delta y_0 := y_0 - \mathfrak{M}(x_0) = A_0 \delta x_0 + \mathfrak{M}'(x_0)x_n + n. \tag{7}$$

Here, we have replaced $\mathfrak{M}(x_0 + \delta x_0)$ by its linearization $\mathfrak{M}(x_0) + A_0 \delta x_0$.

Now (7) has the form (3) with a noise term given by $\mathfrak{M}'(x_0)x_n + n$. A worst-case scenario estimate for the new noise contribution is

$$\|\mathfrak{M}'(x_0)x_n\| \leq \|\mathfrak{M}'(x_0)\|\|x_n\| = \|\mathfrak{M}'(x_0)\|\|x\|_{\beta}^{\frac{\alpha}{\alpha+\beta}} \delta^{\frac{\beta}{\alpha+\beta}}. \tag{8}$$

This estimate shows a new contribution potentially significantly larger than $\delta = \|n\|$. Such an estimate is quite pessimistic in two ways. First of all, it is really $R_c\mathfrak{M}'(x_0)x_n$ that should be estimated, where R_c is an approximation to the unbounded operator A_0^{-1} assuming that A_0 is of the same form as A in (4) above, i.e., is a smoothing operator of order $-\alpha$ (it smoothes by $\alpha > 0$ anti-derivatives). We know that the error term $R_c n$ is of order $\delta^{\frac{\beta}{\alpha+\beta}}$ as was shown in (6). Assuming that $\|R_c\mathfrak{M}'(x_0)\|$ is of order $\|R_c\|\|\mathfrak{M}'(x_0)\|$, the worst-case-scenario bound for $R_c\mathfrak{M}'(x_0)x_n$ corresponding to (8) is

$$\|R_c\mathfrak{M}'(x_0)x_n\| \leq \|R_c\|\|\mathfrak{M}'(x_0)\|\|x_n\| \leq \|\mathfrak{M}'(x_0)\|\|x\|_{\beta}^{\frac{\alpha}{\alpha+\beta}} \delta^{\frac{\beta-\alpha}{\beta+\alpha}}$$

This may be pessimistic because $\|R_c\mathfrak{M}'(x_0)\|$ might in some applications be significantly smaller than $\|R_c\|\|\mathfrak{M}'(x_0)\|$, for instance when $\mathfrak{M}'(x_0)$ is also a smoothing operator. The above bound assumes that $\mathfrak{M}'(x_0)x_n$ is more or less constant on the whole domain $0 \leq |\xi| < \xi_c$ of definition of R_c. This is a realistic assumption in the inverse spectral application in light of (16) below, where we observe that the norm of the corrector is independent of the spectral parameter (the parameter ξ in this section and m in (16)).

Of more fundamental interest in this paper, the estimate in (8) is also quite pessimistic in many practical settings because x_n is highly oscillatory as it oscillates at the scale $\varepsilon \sim \delta^{\frac{1}{\alpha+\beta}}$. Because x_n is unknown and will remain unknown to us assuming that ε has been chosen correctly, it makes sense to model x_n as the realization of a random field. Let $t \in \mathbb{R}^d$ denote the spatial variable. Since x_n oscillates at the scale ε, it is then reasonable to model x_n as

$$x_n(t) = \mathfrak{x}(\frac{t}{\varepsilon}), \tag{9}$$

where $\mathfrak{x}(t)$ is a random field. As an element in a relatively general class of random processes, we assume for concreteness that $\mathfrak{x}(t)$ is a bounded, mean-zero, stationary process with integrable correlation function:

$$\infty > \sigma^2 = \hat{\mathsf{R}}(0) = \int_{\mathbb{R}^d} \mathsf{R}(t)dt, \qquad \mathsf{R}(t) = \mathbb{E}\{\mathfrak{x}(t+u)\mathfrak{x}(u)\}. \tag{10}$$

To be consistent with our regularity assumptions on x_n, we may assume that $\sigma = \|x\|_\beta^{\frac{\alpha}{\alpha+\beta}} \delta^{\frac{\beta}{\alpha+\beta}}$. A simple calculation then shows that

$$\mathbb{E}\|x_n\|^2 = \int_{\mathbb{R}^d} \mathsf{R}\Big(\frac{t}{\varepsilon}\Big)dt = \varepsilon^d \sigma^2. \tag{11}$$

In other words, the root mean square norm of x_n, i.e., the square root of the ensemble average of $\|x_n\|^2$, is $\varepsilon^{\frac{d}{2}}$ times smaller than its maximal value so that under plausible assumptions, we find, using formal scaling arguments, the more realistic expression

$$\big(\mathbb{E}\{\|\mathfrak{M}'(x_0)x_n\|^2\}\big)^{\frac{1}{2}} \le \|\mathfrak{M}'(x_0)\|\sigma\varepsilon^{\frac{d}{2}}\delta^{\frac{\beta}{\beta+\alpha}} \sim \delta^{\mathfrak{g}}, \qquad \mathfrak{g} = \frac{\beta+\frac{d}{2}}{\beta+\alpha}. \tag{12}$$

Such a formal estimate has to be justified for each specific application of interest. The error on δx_0 obtained by solving (7) is now of the form

$$\big(\mathbb{E}\{\|R_c(\mathfrak{M}'(x_0)x_n + n)\|^2\}\big)^{\frac{1}{2}} \le \big(\|\mathfrak{M}'(x_0)\|\sigma\delta^{\frac{\beta+\frac{d}{2}-\alpha}{\beta+\alpha}} + \|x\|_\beta^{\frac{\alpha}{\alpha+\beta}}\delta^{\frac{\beta}{\beta+\alpha}}\big). \tag{13}$$

When $d > 2\alpha$, then the new noise contribution is smaller than the original noise n and the proposed correction is asymptotically negligible. This is the case when the linear operator A_0 does not damp high frequencies too strongly when mapping the unknown vector x to the measurement data y so that α is small. When α is large so that $d < 2\alpha$, then we are in a situation where the noise generated by the unaccessible high frequency component x_n of x dominates the measurement error. In such situations, it is important to have access to a good model for the correlation function of the measurements y if one wants to accurately reconstruct the low frequency component x_0 of x.

3 Inverse spectral problem

We now consider the reconstruction of the potential in a one-dimensional Schrödinger equation from knowledge of two sets of eigenvalues. For the practical, theoretical, and

numerical aspects of the inverse spectral problem, we refer the reader to e.g. [2, 4, 5, 6, 8, 9, 10, 11, 12]. Let the potential $q(t)$ be decomposed as

$$q(t) = q_0(t) + q_\varepsilon(t),$$

where we assume that $x_0 \equiv q_0(t)$ is the identifiable low frequency component and $x_n \equiv q_\varepsilon(t)$ the non-identifiable high frequency part of the potential. We assume that $q_\varepsilon(t) = Q(\frac{t}{\varepsilon})$, where Q is a mean-zero, stationary, bounded random process with integrable correlation function as in (10). The Schrödinger (Sturm Liouville) equation on the interval $(0, 1)$ is given by

$$\begin{aligned} &-\frac{d^2u_m^\varepsilon}{dt^2} + \big(q_0(t) + q_\varepsilon(t)\big)u_m^\varepsilon = \lambda_m^\varepsilon u_m^\varepsilon \\ &u_m^\varepsilon(0) = 0, \qquad \frac{du_m^\varepsilon}{dt}(1) + Hu_m^\varepsilon(1) = 0. \end{aligned} \tag{14}$$

We impose Dirichlet conditions at $t = 0$ to simplify and denote by $(\lambda_m^\varepsilon, u_m^\varepsilon)$ and $(\mu_m^\varepsilon, v_m^\varepsilon)$ the eigenvalues and eigenvectors obtained by setting $H = H_1$ and $H = H_2$, respectively, with $H_1 \neq H_2$.

To conform with the notation of the preceding section, we note that

$$y = (\lambda_m^\varepsilon, \mu_m^\varepsilon) = \mathfrak{M}(q_0(t) + q_\varepsilon(t)) + n. \tag{15}$$

Asymptotic law of physical random fluctuations. We denote by u_m and λ_m the solutions of the unperturbed spectral equation (14) where q_ε is set to zero. With the above decomposition into slowly and rapidly oscillatory components of the potential, we can use the results obtained in [1] to show that under reasonably general assumptions on the random process $Q(t)$, we have

$$(\lambda_m^\varepsilon)^{-1} = (\lambda_m)^{-1} + \varepsilon^{\frac{1}{2}}\sigma \int_0^1 u_m^2(t)dW_t, \tag{16}$$

plus lower-order contributions (in the sense of distributions), where dW_t is standard Wiener measure (spatial white noise). In other words, the above central limit result shows that the measured eigenvalues are approximately deterministic plus a normal contribution with a variance of order ε. We have a similar expression for μ_m^ε, and consequently an explicit expression for the correlation matrix of the random measurements y. The latter expression, however, depends on the unknown eigenvectors u_m, which are nonlinear functionals of the low-frequency component q_0.

Up to lower-order contributions, the expansion (16) may be used to find that

$$\lambda_m^\varepsilon = \lambda_m - \lambda_m^2\varepsilon^{\frac{1}{2}}\sigma \int_0^1 u_m^2(t)dW_t. \tag{17}$$

Note that the Wiener measure dW_t is the same for all values of m and for both spectral λ_m and μ_m. The above formula thus allows us to calculate the approximate covariance matrix of the eigenvalues and obtain that

$$\tilde{\Sigma}_{mn}(q_0) := \mathbb{E}\{(\lambda_m^\varepsilon - \lambda_m)(\lambda_n^\varepsilon - \lambda_n)\} = \varepsilon\sigma^2\lambda_m^2\lambda_n^2 \int_0^1 u_m^2(t)u_n^2(t)dt, \tag{18}$$

and similar expressions for the covariances involving the eigenvalues μ_m^ε. To relate the above expansion with the asymptotic expansion (2), we may define $\mathfrak{M}(q_0) := (\lambda_m, \mu_m)_{m\geq 1}$ the unperturbed measurements for the low frequency component of the potential and identify the fluctuations as

$$\mathfrak{M}'(q_0)q_\varepsilon(t) := \Big(-\lambda_m^2 \varepsilon^{\frac{1}{2}} \sigma \int_0^1 u_m^2(t) dW_t \, , \, -\mu_m^2 \varepsilon^{\frac{1}{2}} \sigma \int_0^1 v_m^2(t) dW_t \Big)_{m\geq 1},$$

where the high frequency component of the potential $q_\varepsilon(t)$ now asymptotically appears only in the stochastic integral and the constant σ. Note that $\mathfrak{M}'(q_0)$ does indeed depend on q_0 via λ_m and u_m. This allows us to recast the inverse problem as

$$\Lambda = \mathfrak{M}(q_0) + \mathcal{N}_\varepsilon(q_0) + n, \tag{19}$$

where $\mathcal{N}_\varepsilon + n = \mathcal{N}_\varepsilon(q_0) + n$ is a discrete Gaussian process with explicit covariance matrix $\Sigma = \Sigma(q_0)$ provided that n is Gaussian. When n is independent of the high frequency component $q_\varepsilon(t)$, then $\Sigma(q_0) = \tilde{\Sigma}(q_0) + \Sigma_n$, where $\tilde{\Sigma}(q_0)$ is the covariance matrix in (18) and Σ_n that of the process n.

Scaling and noise terms. Consider the Dirichlet problem with $H = \infty$. We know that for bounded potentials, (see e.g. [8] and references therein)

$$\lambda_m = m^2\pi^2 + \int_0^1 q(t)dt - \Big(1 + O(\frac{1}{m})\Big) \int_0^1 q(t)\cos(2\pi m t)dt.$$

Let us assume that noise $n = n(m)$ grows like δm^α for a fixed constant δ, which is rather optimistic for $\alpha < 2$ as noise may be considered as proportional to λ_m, i.e., proportional to m^2. Then we see that the mth even Fourier coefficient $\hat{q}_m$ of $q(t)$ may be reconstructed provided that its influence is larger than the noise level. Assuming that $\hat{q}_m$ decays like $m^{-\beta}$, the cut-off frequency is thus $m_c^{-\beta} = \delta m_c^\alpha$, i.e., as in the preceding section, $m_c = \delta^{-\frac{1}{\alpha+\beta}} = \frac{1}{\varepsilon}$. As in the preceding section, α quantifies how high frequencies of the object of interest are damped in the measurements and β quantifies a priori regularity on the object of interest.

We thus define q_0 as the sum over all Fourier modes with indices below the critical value m_c and $q_\varepsilon(t)$ as the rest of the potential, which may indeed be written as a function of $\frac{t}{\varepsilon}$. Based on the asymptotic expansion available in (17), we thus obtain that the high frequency component of the potential generates a noise of order $\varepsilon^{\frac{1}{2}} = \delta^{\frac{1}{2(\alpha+\beta)}}$. Such a noise contribution will dominate the standard noise n with norm $\|n\| = \delta$ as soon as $2(\alpha+\beta) > 1$ and so certainly as soon as $2\alpha > 1$ independently of the value of β.

Inverse problem and MAP estimator. It remains to devise an algorithm that solves the inverse problem (19) for q_0. Since we have constructed a probabilistic framework for the rapidly oscillating component q_ε, we might as well assume that q_0 is drawn from a prior distribution, which here we assume follows a Gaussian distribution with diagonal covariance matrix γI with γ a small (non-negative) constant [7]. The typical role of γ is to construct biased estimators of q_0 that overall reduce the average error on the reconstruction of q_0.

Now that we have constructed a prior model for q_0 and that we have our likelihood function in (19), we may use Bayes' theorem to infer that the posterior distribution $\Pi(q_0)$ based on the available data is given by

$$\Pi(q_0) \propto \exp\Big(-\frac{1}{2}(\Lambda - \mathfrak{M}(q_0)) \cdot \Sigma^{-1}(q_0)(\Lambda - \mathfrak{M}(q_0))\Big) \exp\big(-\frac{\gamma}{2}|q_0|^2\big). \tag{20}$$

In what follows, we present the potential q_0 that maximizes the above functional (the maximal a posteriori, or MAP, solution) and so do not need to calculate the normalizing constant that makes the right-hand side in (20) a probability density.

To simplify the presentation, we assume here and below that $\Sigma_n = 0$ so that the noise contribution in the measured eigenvalues is generated by the random part $q_\varepsilon(t)$ of the potential. Because the asymptotic correlation matrix $\Sigma = \Sigma(q_0)$ defined in (18) depends on the unknown parameter q_0, the reconstruction is necessarily iterative. Let N (resp. $N+1$) be the number of available measurements for the μ (resp. λ) spectrum so that the total vector of measurements y is represented by a $M = 2N+1$ vector. Asymptotic expansions show that for large values of m, we have

$$\lambda_m \approx (m\pi)^2 + \int_0^1 q_0(t)dt. \tag{21}$$

Assuming a constant initial guess q_0^0, we estimate it as $q_0^0 = \lambda_N - (N\pi)^2$. Assuming that q_0^k has been constructed, we estimate the variance $\Sigma(q_0^k)$ by using the explicit formula (17), which requires solving an eigenvalue problem. We then calculate q_0^{k+1} by maximizing

$$\exp\Big(-\frac{1}{2}(\Lambda - \mathfrak{M}(q_0^{k+1})) \cdot \Sigma^{-1}(q_0^k)(\Lambda - \mathfrak{M}(q_0^{k+1}))\Big) \exp\big(-\frac{\gamma}{2}|q_0^{k+1}|^2\big). \tag{22}$$

The solution to this classical minimization problem is obtained by using a Newton method [13]. The iterative method converged quite rapidly and robustly in the numerical simulations that were considered.

Note that the approximate correlation matrix Σ of the eigenvalues defined in (18) depends on the eigenvectors u^m, which are estimated iteratively in the above algorithm, and on the strength parameter $\varepsilon^2\sigma$. We observe that, asymptotically, Σ is proportional to $\varepsilon\sigma^2$. We also observe that the maximization in (20) is not modified when both $\Sigma^{-1}(q_0)$ and γ are multiplied by a constant. In other words, maximizing (20) is equivalent to maximizing

$$\exp\Big(-\frac{1}{2}(\Lambda - \mathfrak{M}(q_0)) \cdot (\varepsilon^2\sigma\Sigma^{-1}(q_0))(\Lambda - \mathfrak{M}(q_0))\Big) \exp\big(-\frac{\varepsilon^2\sigma\gamma}{2}|q_0|^2\big).$$

It is therefore $\varepsilon^2\sigma\gamma$ that should be estimated rather than $\varepsilon^2\sigma$ and γ independently. When that parameter is unknown, a discrepancy principle, such as the Morozov principle, may then be used to figure out which parameter best fits available data [3]. In this paper, we assume that $\varepsilon^2\sigma\gamma$ is known a priori and in our numerical experiments, we choose that parameter so that $\mathrm{Tr}(\varepsilon^2\sigma\Sigma^{-1}(q_0)) = \mathrm{Tr}(\varepsilon^2\sigma\gamma I)$, i.e., the influences from the noise $\varepsilon^2\sigma$ and from the regularization γ are comparable.

The solution to the above algorithm very much depends on the structure of the extra-diagonal terms in $\varepsilon^2\sigma\Sigma^{-1}(q_0)$. The asymptotic formula (17) is useful precisely in that

it provides an approximate model for the extra-diagonal components of the correlation matrix, which are difficult to estimate in practice and are often neglected, at the price of possibly severe inaccuracies in the reconstruction as the next section demonstrates.

4 Numerical simulations

In what follows, we assume that the Gaussian noise n is set to 0. This may be justified by assuming that noise is overwhelming for large eigenvalues, so that N need be finite, and that noise is relatively mild for the eigenvalues that are measured. This simplifying assumption allows us to concentrate on the effects of the spatial random fluctuations, which is the main object of interest in this paper. We also assume that the regularization parameter γ and the strength of the "nonlinear" noise $\varepsilon^2\sigma$ are such that $\mathrm{Tr}(\varepsilon^2\sigma\Sigma^{-1}(q_0)) = \mathrm{Tr}(\varepsilon^2\sigma\gamma I)$ as was indicated in the preceding section. We consider six different reconstructions, all based on maximizing the posterior distribution (20), but with different choices of the correlation matrix Σ.

In each simulation, q_0 is deterministic and chosen of the form

$$q_0(t) = \sum_{k=-K}^{K} c_k e^{-i2\pi kt}, \tag{23}$$

with $K = 2$ and $K = 10$ depending on the simulation so that $2K + 1$ parameters need be reconstructed. Since q_0 is real-valued, we have $\overline{c_{-k}} = c_k$. The two spectra λ and μ are obtained by setting $H_1 = \infty$ and $H_2 = 0$. The number of measured eigenvalues M is equal to 5, 21, or 41 in the simulations below. One more eigenvalue comes from the λ spectrum than from the μ spectrum. The practically relevant cases are $M = 2K + 1 = 5$ and $M = 2K + 1 = 21$ where the number of measurements equals the number of coefficients we believe we can reconstruct. We also consider over-determined reconstructions with $M > 2K + 1$.

The accuracy of the reconstruction $\tilde{q}_0$ is measured in the relative L^2 norm

$$\epsilon_0 = \frac{\mathbb{E}\Big(\sum\limits_{|k|<K} |c_k - \tilde{c}_k|^2\Big)^{\frac{1}{2}}}{\Big(\sum\limits_{|k|<K} |c_k|^2\Big)^{\frac{1}{2}}}, \tag{24}$$

In our simulations, ϵ_0 is estimated by averaging over 200 realizations of the noise $q_\varepsilon(t)$; their standard deviation was found to be extremely small (less than 5% of the mean), which implies that the relative L^2 norm is very stable statistically.

The random coefficient $q_\varepsilon(t)$ is modeled as a superposition of coefficients as in (23), where $|k|$ now runs from 11 to 50 and where the coefficients c_k are chosen at random (so that q_ε is real-valued) with a flat power spectrum (i.e., the variance of c_k is independent of k). The tail of the power spectrum is in fact not very important since the influence of each mode c_k is roughly of order $\varepsilon_k^{\frac{1}{2}} \approx k^{-\frac{1}{2}}$ by application of the central limit result in (16). When $K = 10$, there is no "spectral gap" between the last mode of q_0, which is considered deterministic, and the first mode of q_ε, which is considered random. This is

the realistic situation physically. When $K = 2$, we then have a spectral gap between the deterministic and random parts, which may only happen under quite restrictive physical assumptions.

In what follows, we call "$\mathbf{\Sigma}$" the iterative solution obtained by applying the iterative algorithm described in (22). We compare this solution to five solutions obtained with different correlation matrices. The first three correlation matrices are not observable a priori: they are given by the exact correlation matrix $\mathbf{\Sigma_a}$ obtained by solving the eigenvalue problems a large number of times (five hundred times in our simulations) with the (known in synthetic experiments but unknown in practice) exact statistics for q_ε and exact low frequency component q_0. The correlation matrix $\mathbf{\Sigma_b}$ is obtained by applying the asymptotic formula (16) with the exact low frequency component q_0. The correlation matrix $\mathbf{\Sigma_d}$ is given by the diagonal of Σ_a. Such a matrix thus completely misses the cross-correlations between the measured eigenvalues.

The last two correlation matrices are accessible experimentally. The first matrix $\mathbf{\Sigma_c}$ is given by the asymptotic formula (16) by using the first constant guess q_0^0, for which the eigenvectors are given by explicit sin and cos functions by the method of separation of variables. This intermediate matrix does not require us to iterate in (22) although it still requires an asymptotic model of the form (16). Finally, the last correlation matrix $\mathbf{\Sigma_e}$ is simply given by identity.

ϵ_0	$M = 5$	$M = 21$		$M = 41$	
Covariance	$\mathbf{2K+1=5}$	$2K+1=5$	$\mathbf{2K+1=21}$	$2K+1=5$	$2K+1=21$
Σ	**9.3 %**	7.6%	**8.5 %**	6.2 %	6.7%
Σ_a	**4.7 %**	4.0 %	**4.4 %**	3.3 %	4.0%
Σ_b	**11.9 %**	10.0 %	**12.3 %**	7.9 %	11.0%
Σ_c	**19.5 %**	13.8 %	**16.7 %**	8.9 %	15.8%
Σ_d	**28.5 %**	22.6 %	**19.8 %**	17.4 %	21.9%
Σ_e	**36.7 %**	26.6 %	**29.4 %**	19.5 %	26.9%

Table 1: Relative L^2 error ϵ_0 (in percentages) on the reconstruction of the potential q_0 using the different correlation matrices. In bold are the L^2 norms of the errors in the columns corresponding to $M = 2K + 1$. Underlined are the results obtained from the *accessible* (observable) correlations Σ, Σ_c, and Σ_e.

The errors we obtained on the reconstruction of q_0 in the scenarios described above are collected in Tab.1. Typical realizations of the reconstructed potentials using the correlations Σ, Σ_a, and Σ_e, are presented in Fig. 1, Fig. 2, and Fig. 3 for $M = 5$, $M = 21$, and $M = 41$, respectively. We observe that the asymptotic correlation Σ provides reconstructions that are visually of similar quality to those obtained using the unaccessible exact correlation Σ_a. The use of the diagonal correlation Σ_e, which is the best one can do in the absence of any prior knowledge for the correlations or any asymptotic expansion such as (16), generates significantly worse reconstructions.

As expected, the minimum variance in the reconstruction is obtained when the exact statistics Σ_a are used to model the measurement noise. The error made is less than 5% for all such reconstructions. This is the minimal error that can be made as it assumes full knowledge of the statistics of the measured eigenvalues (λ_m, μ_m).

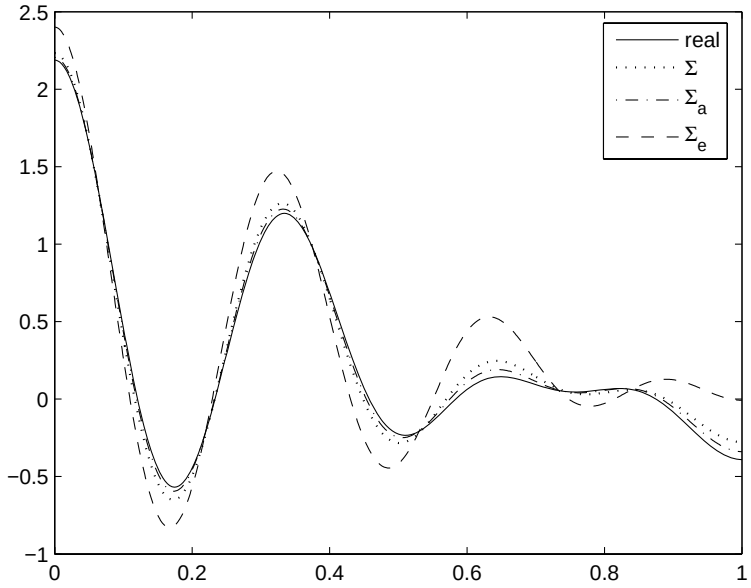

Figure 1: Reconstructions of the potential $q_0(x)$ (solid) with covariance matrices Σ_a (dot-dashed), Σ_e (dashed) and Σ (dotted). The potential q_0 consists of $2K+1=5$ Fourier modes and $M=5$ eigenvalues are used in each reconstruction.

The iterative reconstruction based on the asymptotic correlation Σ provides solutions of similar quality with errors on the order of $8-9\%$. Whereas the construction of Σ_a requires full knowledge of the statistics of the measurements, no information other than the measurements of the M eigenvalues is necessary in the construction of Σ. We have therefore a parameter-free reconstruction method that performs almost as well (it roughly doubles the error) as a method requiring knowledge of the full statistical description of the measurements.

A significantly larger error, of order $15-20\%$ is obtained by using the correlation Σ_c, which is constructed by using the asymptotic expression (16) with a constant potential q_0^0. This shows that the asymptotic correlations generated by q_0 are significantly different from those obtained by the constant potential q_0^0. It is therefore necessary to iterate as specified in (22) to obtain a decent approximation of the correlation matrix Σ_a.

The reconstructions based on Σ_b are somewhat less accurate than those based on Σ. By insisting that the correlation be based on the asymptotic expansion of the true (unknown) q_0, we obtain a larger variance than by letting the correlation adapt iteratively to the optimal potential q_0. That Σ_b performs significantly less accurately than Σ_a shows that the asymptotic expansion (16) is not extraordinarily accurate. This is to be expected since ε in our simulations is rather large as only the modes corresponding to $|k| \geq 11$ (as opposed to $|k| \geq 101$, say) are supposed random. We expect reconstructions based on Σ to perform between the optimal reconstructions based on Σ_a and those based on Σ_b and the asymptotic formula (16). This has been verified in all of our numerical experiments.

These errors should be contrasted with the solutions obtained by assuming that the correlation is proportional to identity or is diagonal, which in the absence of any physical model, may be the best available option. For both models based on (quite different)

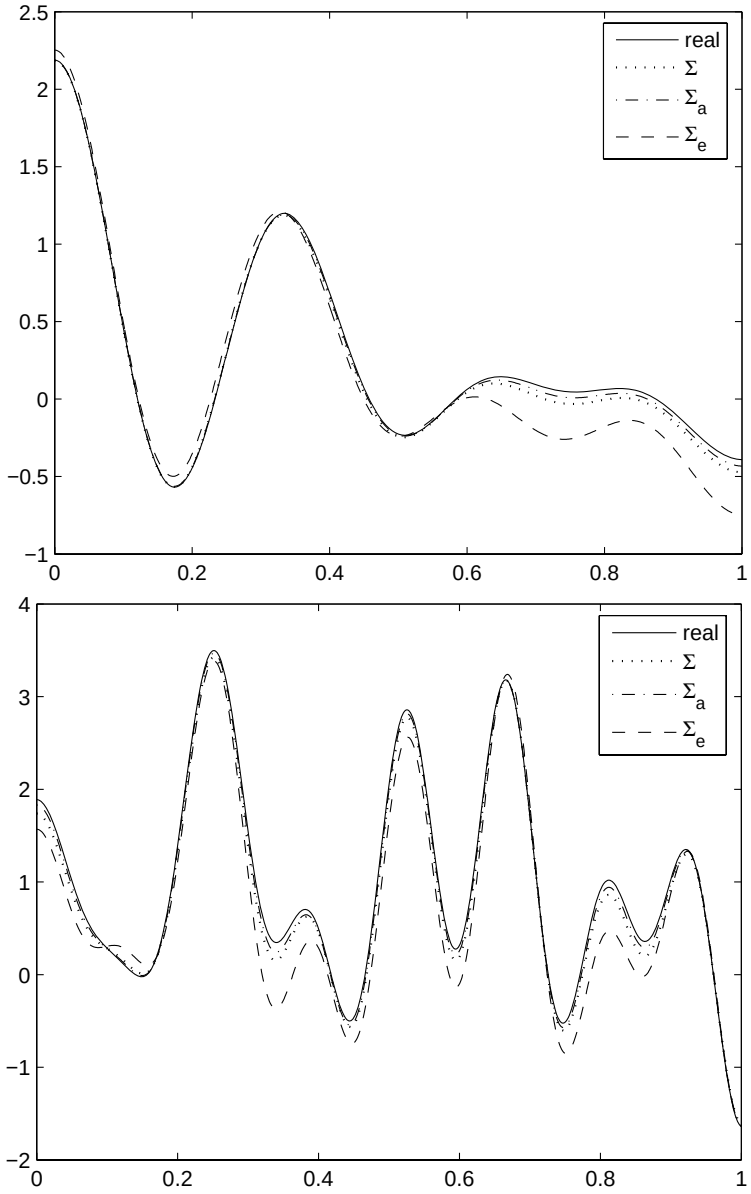

Figure 2: Reconstructions of the potential $q_0(x)$ (solid) with covariance matrices Σ_a (dot-dashed), Σ_e (dashed) and Σ (dotted). Top: q_0 consists of $2K+1=5$ Fourier modes; Bottom: q_0 consists of $2K+1=21$ Fourier modes. $M=21$ eigenvalues are used in each reconstruction.

diagonal assumptions, namely Σ_d and Σ_e, the errors in the reconstructions, on the order of 30%, are significantly larger. This shows the importance of modeling the off-diagonal component of the correlation matrix in a reasonably accurate manner.

Finally, we observe that over-determined measurements (e.g., $M=41$) only somewhat marginally improve the reconstructions. This is to be expected since we assume that the eigenvalues for all values of M are fairly accurately measured. Our noise contribution is exclusively coming from the randomness in the potential q_ε. The reconstructions corresponding to $2K+1=5$ and based on a spectral gap (since the coefficients c_k for $|k|$ between 3 and 10 are set to 0) also perform marginally better than the more phys-

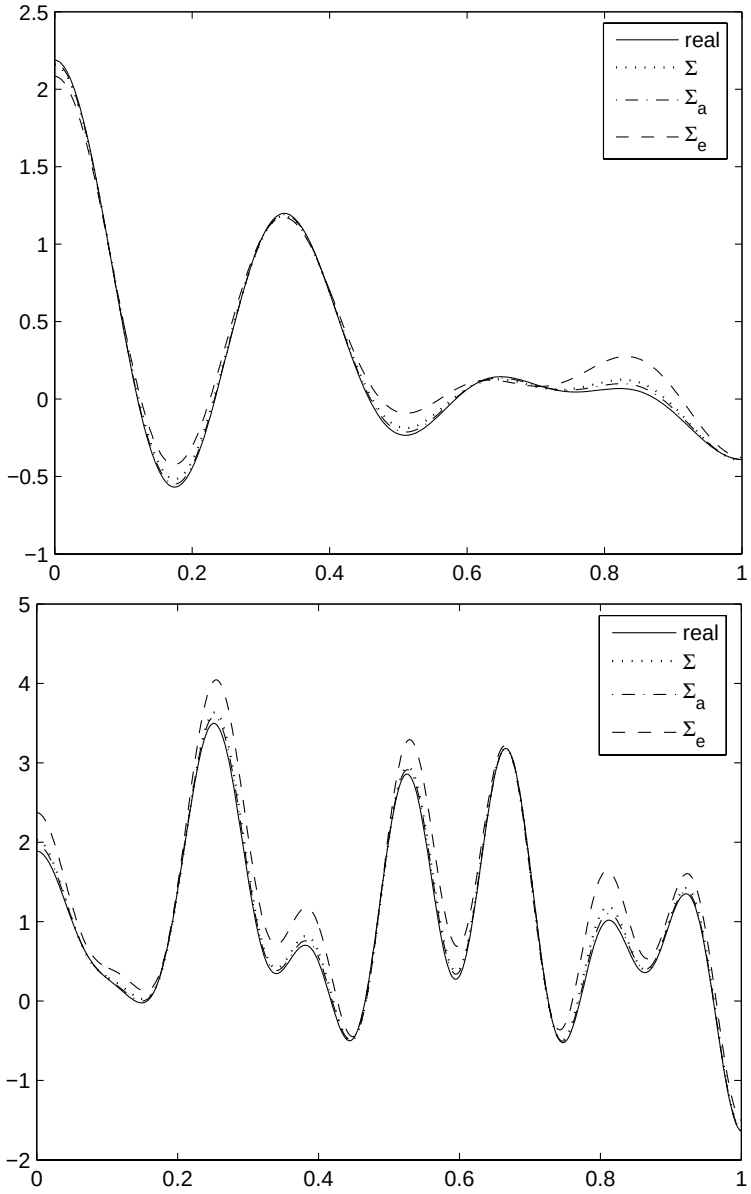

Figure 3: Reconstructions of the potential $q_0(x)$ (solid) with covariance matrices Σ_a (dot-dashed), Σ_e (dashed) and Σ (dotted). Top: q_0 consists of $2K+1=5$ Fourier modes; Bottom: q_0 consists of $2K+1=21$ Fourier modes. $M=41$ eigenvalues are used in each reconstruction.

ical reconstructions based on $2K+1=21$ for over-determined measurements $M=21$ and $M=41$.

5 Summary

Inverse problems are characterized by the degree of smoothness of the measurement operator mapping the object $x=x_0+x_n$ we are interested in reconstructing to what we may measure. When that operator is significantly smoothing (α above is positive),

then the (properly defined) high frequency component x_n of the object of interest is invisible to the inevitably noisy available data. We are concerned with situations in which the measurement operator efficiently couples the high frequency component x_n to the available data so that neglecting to account for the presence of x_n may generate sizeable errors in the reconstruction. This is for instance the case in the Sturm Liouville problem considered in this paper.

Since x_n is not accessible, it has to be modeled a priori. Statistical descriptions are then quite satisfactory in the generic setting where the detailed structure of x_n may not be guessed. It remains to infer the parameters used in the statistical description either from prior knowledge or from the available data. Fortunately, such a description is not necessary when self-averaging mechanisms (e.g. of central limit type) simplify the influence of the random field x_n on the available data. In the Sturm Liouville problem and in other problems with a physical origin, it turns out that the influence of x_n on the measured data may be approximated by an explicit Gaussian law. Moreover, that Gaussian law is essentially modeled by one unknown scaling parameter σ.

Theoretical estimates for the cut-off frequency ξ_c separating x_0 from x_n have been presented in section 2 based on the smoothing properties of the measurement operator and the prior regularity in H^β imposed on x. Once the cut-off frequency has been chosen or estimated (a difficult and problem-dependent question that was addressed in an ad hoc manner in our numerical inversion of the Sturm Liouville problem), two terms contribute to errors in the reconstruction of x_0. The first term in (6) is simply the error made by neglecting x_n in the reconstruction (this term was omitted in our numerical simulations) while the second term in (13) quantifies the influence of x_n on the available measurement. The formal estimates in (12) and (13) are the main theoretical result of the paper. Although they have to be justified for each problem of interest, they provide a reasonable estimate for the influence of the invisible part x_n on the reconstruction of x_0. That the statistical description of x_n asymptotically reduces to one parameter as an application of the central limit theorem is exemplified by the inverse spectral problem considered in section 3 and more concretely by the expression (17) describing available measurements.

The frequency cut-off in our numerical simulations of the inverse spectral problem, for instance with $M = 21$ measured eigenvalues, is rather arbitrary and reflects our belief that only those $M = 21$ first eigenvalues may be measured accurately. Once this cut-off is chosen, we apply the classical MAP algorithm and the iterative scheme (22) to solve the inverse spectral problem. A common feature of this and many other inversion algorithms (such as e.g. least square algorithms), is the importance of the correlation matrix Σ, which weighs the measured data according to the confidence we have in them. The main advantage of modeling the influence of x_n on the measured data is precisely that it allows one to obtain a more accurate description of the correlation matrix Σ than when x_n is simply treated as uncorrelated noise.

In the absence of any model for x_n, the only choice for Σ is to assume that it is proportional to identity. This lack of understanding of the correlations in the available measurements generated a reconstruction of x_0 with an average of **29.4**% error in the L^2 sense according to Tab.1. Prior knowledge of the exact correlation Σ_a provides a much more accurate reconstructions, with an average error dropping to **4.4**%. Such prior knowledge is unrealistic in many settings. Reconstructions performed using the

central limit-based asymptotic formula (17) provide quite accurate reconstructions, with an average error equal to **8.5**% and *do not* require the estimation (or prior knowledge) of any additional parameters. In the configuration considered in this paper, the parameter-free asymptotic modeling of x_n allows us to obtain much more accurate reconstructions of the low frequency component of the potential than methods that do not model x_n or treat it as uncorrelated "white" noise.

Acknowledgment

This work was supported in part by NSF Grants DMS-0239097 and DMS-0804696. We would like to thank the reviewers and the editorial board of Inverse Problems for comments and suggestions that helped us improve the presentation of our main ideas.

References

[1] G. Bal, *Central limits and homogenization in random media*, Multiscale Model. Simul., 7(2) (2008), pp. 677–702.

[2] K. Chadan, D. Colton, L. Päivärinta, and W. Rundell, *An Introduction to Inverse Scattering and Inverse Spectral Problems*, SIAM, Philadelphia, 1997.

[3] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, 1996.

[4] G. Freiling and V. Yurko, *Inverse Sturm-Liouville Problems and Their Applications*, Nova Science Publishers, Huntington, NY, 2001.

[5] I. M. Gelfand and B. M. Levitan, *On the determination of a differential equation from its spectral function*, Trans. Amer. Math. Soc., 1 (1951), pp. 253–304.

[6] H. Hochstadt, *The inverse Sturm-Liouville problem*, Comm. Pure Appl. Math., 26 (1973), pp. 715–729.

[7] J. P. Kaipio and E. Somersalo, *Statistical and Computational inverse problems*, Springer Verlag, New York, 2004.

[8] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, Springer-Verlag, New York, 1996.

[9] N. Levinson, *The inverse Sturm-Liouville problem*, Math. Tidsskr. B, 25 (1949), pp. 25–30.

[10] V. A. Marchenko, *Sturm-Liouville Operators and Applications*, Birkhäuser, Berl, 1986.

[11] J. R. McLaughlin and W. Rundell, *A uniqueness theorem for an inverse Sturm-Liouville problem*, J. Math. Phys., 28 (1987), pp. 1471–1472.

[12] W. Rundell and P. E. Sacks, *Reconstruction techniques for classical inverse Sturm-Liouville problems*, Math. Comp., 58 (1992), pp. 161–183.

[13] C. R. Vogel, *Computational Methods for Inverse Problems*, Frontiers Appl. Math., SIAM, Philadelphia, 2002.