On Weakly Bounded Noise in Ill-Posed Problems

P.P.B. Eggermont[†], V.N. LaRiccia[†], M.Z. Nashed[‡]

[†]University of Delaware and [‡]University of Central Florida
August 13, 2008

Abstract

We study compact operator equations with noisy data in Hilbert space. Instead of assuming that the error in the data converges strongly to 0, we only assume weak convergence. Under the usual source conditions, we derive optimal convergence rates for convexly constrained Phillips-Tikhonov regularization. We also discuss a discrepancy principle and prove its asymptotic behavior. As an example, we discuss compact integral equations in $L^2(0,1)$ with data perturbed by white noise, as well as the discrete case.

Keywords: Ill-posed problems, Tikhonov regularization, weakly bounded noise, white noise, random discrepancy principle, reproducing kernel Hilbert spaces
2000 Math. Subject Classification: 47A52, 45Q05, 45R05, 46E25, 60G35

1. Introduction

Let X and Y be Hilbert spaces, and let $K : X \to Y$ be linear and compact, with nonclosed range. Suppose that one observes the data $y \in Y$ such that

$$y = K x_o + \delta \ , \tag{1.1}$$

for some unknown $x_o \in X$, and with the "noise" δ. The goal is to recover x_o. The problem (1.1) is ill-posed in the sense of Hadamard [18]: The solution of (1.1) will not depend continuously on y in the implied topology, even assuming its existence and uniqueness, because K has no bounded (generalized) inverse.

There are two views of the noise in ill-posed problems. In the classical approach to the analysis of (1.1), one assumes that $\| \delta \|_Y$ is "small", and one investigates what happens when

$$\| \delta \|_Y \longrightarrow 0 \; . \tag{1.2}$$

The hope is that this will allow one to make inferences regarding the small noise case. This approach originated with TIKHONOV [41], and started seriously with PHILLIPS [37], TIKHONOV [42] and TWOMEY [44]. See GROETSCH [17], TIKHONOV *et al.* [43] and ENGL *et al.* [12].

There are many cases where the assumption (1.2) is not realistic. In signal processing, one considers high frequency noise, and one designs low-pass filters to combat the problem. The size of the noise is largely irrelevant here. See, e.g., DUAN *et al.* [7] and WANG *et al.* [45]. In image processing, see, e.g., SCHERES *et al.* [39], signal-to-noise ratios of 10% or less are not unheard of. Of course, there the *randomness* of the noise saves the day. This brings us to the second approach to (1.1), based on a probabilistic point of view, where one assumes that the noise is random. Moreover, one typically assumes that only a finite number of observations are available, e.g., in the form

$$y_i = \langle \, \ell_i \, , \, K x_o \, \rangle_Y + \varepsilon_i \; , \quad i = 1, 2, \cdots, n \; , \tag{1.3}$$

where $\ell_1, \ell_2, \cdots, \ell_n$ are given (known) elements of Y, and $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ are independent Normal$(0 \, , \sigma^2)$ random variables. In this approach, one investigates what happens when more and more information becomes available, i.e., when $n \to \infty$, but the size of the noise (σ) stays the same. This approach originated (very early) with SUDAKOV and KHALFIN [40]. Some recent references are MAIR and RUYMGAART [27], BISSANTZ *et al.* [2] and HOHAGE and PRICOP [21]. The above approach constitutes the *frequentist* view of statistical problems. There is a different approach, based on the *Bayesian* view of the world. Here, oversimplifying somewhat, the unknown x_o is considered to be random with a known prior distribution, and one assumes that x_o is drawn at random from this distribution. Then, one considers (1.3) conditional on x_o, and one determines the posterior distribution of x_o. The estimator of x_o is then the posterior mean or posterior mode. See, e.g., FRANKLIN [13] and KAIPIO and SOMERSALO [22]. We shall not consider the Bayesian view further.

It is not unusual to study random noise in ill-posed problems in the form

$$y = Kx_o + \eta\,\xi \;, \tag{1.4}$$

where ξ is standard white noise and η is a scale parameter, see, e.g., Cavalier *et al.* [6] and Mathé and Pereverzev [30], and references therein. (When estimating Kx_o instead of x_o, the problems (1.3) and (1.4) with $\eta = \sigma/\sqrt{n}$, are asymptotically equivalent in a precise sense. See, Brown and Low [3] and references therein.) It should be noted that (1.4) is no longer set in the Hilbert space Y, since $\|\,\xi\,\|_Y = +\infty$. It has to be interpreted in the weak sense, i.e., (1.4) is equivalent to

$$\langle\, \ell\,,\, y\,\rangle_Y = \langle\, \ell\,,\, Kx_o\,\rangle_Y + \eta\,\langle\, \ell\,,\, \xi\,\rangle_Y \quad \text{for all} \quad \ell \in Y \;, \tag{1.5}$$

where $\langle\, \ell\,,\, \xi\,\rangle_Y$ is a normal random variable with mean 0 and variance $\|\,\ell\,\|_Y^2$. This is very close to saying that $\eta\,\xi \to 0$ weakly in Y as $\eta \to 0$, even though $\xi \notin Y$.

As a bridge between the classical and probabilistic points of view, we propose here a semi-classical approach and assume that the noise δ in (1.1) is small in the weak sense, and study what happens if instead of (1.2),

$$\delta \longrightarrow 0 \quad \text{weakly in } Y \;, \tag{1.6}$$

while $\|\,\delta\,\|_Y$ is not necessarily small (denoted as $\|\,\delta\,\|_Y \asymp 1$). The precise assumptions are that for some operator

$$T \,:\, Y \longrightarrow Y \quad \text{linear, compact, Hermitian, positive definite} \;, \tag{1.7}$$

we have that

$$K(X) \subset T^m(Y) \quad \text{for some} \quad m \geqslant 1 \;, \quad \text{and} \tag{1.8}$$

$$\eta^2 \;\stackrel{\text{def}}{=}\; \langle\, \delta\,,\, T\delta\,\rangle_Y \longrightarrow 0 \tag{1.9}$$

(m need not be an integer). The last condition may be replaced by

$$\widetilde{\eta}^{\,2} \stackrel{\text{def}}{=} \sup_{0<\beta\leqslant 1} \beta\,\langle\, \delta\,,\, (\,I + \beta^2\,T^{-2}\,)^{-1}\delta\,\rangle_X \longrightarrow 0 \;. \tag{1.10}$$

Thus, T (or rather T^m) measures the smoothing effect of the operator K, and T must be able to filter the noise in an asymptotic sense. Of course, the compactness of T guarantees that $\eta \to 0$, but (1.8)-(1.9) attempts to express that T is ideally suited to (1.1)-(1.3). One may refer to η as the weak bound on the noise, and speak of *weakly bounded* noise if (1.7)-(1.9) hold. Thus, in the semi-classical approach, we assume that η is "small", and inverstigate what happens when $\eta \to 0$. Note that (1.9) does not imply a bound on $\| \delta \|_Y$. In fact, it is possible that $\langle \delta , T\delta \rangle_Y \to 0$ but $\| \delta \|_Y \to \infty$.

REMARK 1. There does not appear to be an easy way to transform the weakly bounded version of (1.1) into a strongly bounded version without repercussions. One could consider the model, obtained from (1.1) after multiplying by $T^{1/2}$,

$$T^{1/2} y = T^{1/2} K x_o + T^{1/2} \delta \ ,$$

where now the noise $T^{1/1}\delta$ converges stronly to 0, but the disadvantage is that roughly speaking, now one has to "invert" $T^{1/2}K$, which is harder than "inverting" just K. However, a viable alternative is to pre-filter the data as in KLANN *et al.* [23], e.g., by way of the generalized spline smoothing problem

$$\text{minimize} \quad \| z - y \|_Y^2 + \gamma \, \| T^{-k} x \|_X^2 \ ,$$

for suitable γ and k, and then to consider the model

$$z = K x_o + \varepsilon \ ,$$

with $\| \varepsilon \|_X \to 0$. Of course, this must be shown, presumably along the lines of this paper.

REMARK 2. The spaces $T^m(Y)$, $m > 0$, with the norms $\| T^{-m} y \|_Y$ for $y \in T^m(Y)$, form a (compact) *Hilbert scale*, see KREIN and PETUNIN [24]. (Later on, we refer to this is a Hilbert scale in the narrow sense.) It is unclear whether they meant it as the definition or as an example of a Hilbert scale, having formally defined a *Banach scale* as a collection of Banach spaces, E_α, $\alpha_o \leqslant \alpha \leqslant \beta_o$, with norms $\| \cdot \|_\alpha$, such that

(a) E_β is densely embedded in E_α, when $\beta > \alpha$, and for a constant $c_{\alpha,\beta}$,

$$\| x \|_\alpha \leqslant c_{\alpha,\beta} \, \| x \|_\beta \quad \text{for all} \quad x \in E_\beta \ ;$$

(b) (Interpolation) For all α, β, γ with $\alpha_o \leqslant \alpha < \gamma < \beta \leqslant \beta_o$, there exists a finite constant $c_{\alpha,\beta,\gamma}$ such that with $\theta = (\beta - \gamma)/(\beta - \alpha)$,

$$\| x \|_\gamma \leqslant c_{\alpha,\beta,\gamma} \, \| x \|_\alpha^\theta \, \| x \|_\beta^{1-\theta} \quad \text{for all} \quad x \in E_\beta \ .$$

If Krein and Petunin [24] really meant their definition of a Hilbert scale, then we arrive at the odd situation where, e.g., the usual scale of Sobolev spaces of functions on an open, bounded subset of $\mathbb{R}^n$ with smooth boundary, which are Hilbert spaces, would be a Banach scale, but as shown by Neubauer [33], not a Hilbert scale. For Sobolev spaces on (all of) $\mathbb{R}^n$, the two notions coincide, as shown by Krein and Petunin [24]. (We are tempted to speak of a Hilbert scale in the wide sense if the collection of Hilbert spaces forms a Banach scale.) Be that as it may, Hilbert scales in the narrow sense have been used extensively in ill-posed problems, dating back to Natterer [32]. see, e.g., Mair and Ruymgaart [27], , Liu and Nashed [25], Mathé and Pereverzev [28], Egger [9], Hohage and Pricop [21]. Here, for "general" source conditions, Hegland [20] is of interest.

In this paper, we limit ourselves to the investigation of Phillips-Tikhonov regularization of (1.1) (but with convex constraints) for weakly bounded noise under the general source conditions of Groetsch [17] and Neubauer [34], and following Gfrerer [15], [16], formulate a weak discrepancy principle for the selection of the regularization parameter (in the unconstrained case). The application to compact integral equations in $L^2(0,1)$ with data perturbed by white noise as well as discrete versions are considered also.

We hasten to add that much more general regularization schemes, e.g., as in Burger and Osher [5], and iterative regularization methods, such as (accelerated) Landweber iteration, Hanke [19], conjugate gradients, Eicke *et al.* [11], Frommer and Maass [14], or semi-iterative methods, Egger [9], under "general" source conditions, e.g., as in Hegland [20], Mathé and Pereverzev [28], [29], can and should be investigated in the context of weakly bounded noise, but we leave that for later.

2. Tikhonov regularization

In this section, we discuss Tikhonov regularization of linear ill-posed operator equations with weakly bounded noise in Hilbert space.

Let X, Y be Hilbert spaces and let $K : X \longrightarrow Y$ be linear and compact with nonclosed range. We observe the data $y \in Y$ according to the model

$$y = K\,x_o + \delta \tag{2.1}$$

for some unknown $x_o \in X$, and where δ represents weakly bounded noise in the sense of (1.7)-(1.9). The goal is to recover x_o . As the recovery scheme, we use Phillips-Tikhonov regularization, i.e., the unknown x_o is approximated by the solution $x = x^{\alpha,\delta}$ of the penalized least-squares problem,

$$\text{(2.2)} \qquad \begin{array}{ll} \text{minimize} & \mathrm{LS}(\, x \,|\, y \,) \stackrel{\text{def}}{=} \| \, Kx - y \, \|_Y^{\,2} + \alpha \, \| \, x \, \|_X^{\,2} \\ \text{subject to} & x \in C \ . \end{array}$$

Here, C is a closed, convex subset of X, whose purpose it is to capture any qualitative and quantitative *a priori* information one might possess about x_o . In (2.2), the parameter α is positive, and must be chosen appropriately. The actual goal is to give asymptotic bounds on the error $\| \, x^{\alpha,\delta} - x_o \, \|_{\alpha,X}$, in terms of η , see the definition (1.9), as $\eta \to 0$. Here, for all $x \in X$,

$$\text{(2.3)} \qquad \| \, x \, \|_{\alpha,X}^{\,2} \stackrel{\text{def}}{=} \| \, Kx \, \|_Y^{\,2} + \alpha \, \| \, x \, \|_X^{\,2}$$

defines a useful norm on X . Note that $\alpha \, \| \, x \, \|_X \leqslant \| \, x \, \|_{\alpha,X}$ for all $x \in X$. However, it is well-known that the $\| \, K(x^{\alpha,\delta} - x_o) \, \|_Y -$ and $\| \, x^{\alpha,\delta} - x_o \, \|_X -$ parts of the error behave "differently", Groetsch [17], Neubauer [34].

We assume the usual source condition,

$$\text{(2.4)} \qquad x_o = (K^* K)^\nu \, z_o \quad \text{for some} \quad 0 < \nu \leqslant 1 \ ,$$

and $z_o \in X$. Initially, we pretend that we know the value of ν , and choose the regularization parameter α depending on ν . For the discrepancy principle, we need $\nu = 1$.

Remark 3. The "proper" source condition is given by Neubauer [34] and Gfrerer [15], [16], and involves the spectral properties of K . Thus, let $\{ \, v_i \, , \, w_i \, \}_i$ be a complete biorthonormal system of singular functions of the operator $K : X \to Y$ with singular values $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant 0$, such that

$$K \, v_i = \sigma_i \, w_i \ , \quad i = 1, 2, \cdots \ .$$

Then, the proper source condition is

$$\sum_{\sigma_i \leqslant \alpha} \big| \, \langle \, x_o \, , \, v_i \, \rangle_X \, \big|^2 = \mathcal{O}\big(\, \alpha^{2\nu} \, \big) \ .$$

This source condition, and also (2.4), is "independent" of the assumptions about the noise, be they classical or semi-classical. So, in the absence of constraints, the observations of Groetsch [17] and Neubauer [34] about the necessity of the source conditions for optimal convergence rates in the case of noiseless data would be in effect.

Main Theorem. *For weakly bounded noise in the sense of* (1.7)-(1.9), *under the source condition* (2.4),

$$\| x^{\alpha,\delta} - x_o \|_X \leqslant c\,\alpha^{-1/2} \left\{ \alpha^{-1/4m}\, \eta + \alpha^{\nu+1/2} \right\} ,$$

and

$$\| K(x^{\alpha,\delta} - x_o) \|_Y \leqslant c \left\{ \alpha^{-1/4m}\, \eta + \alpha^{\mu} \right\} ,$$

where $\mu = \min(\nu + \frac{1}{2} , 1)$.

Main Corollary. *Under the conditions of the Main Theorem,*

$$\| x^{\alpha,\delta} - x_o \|_X = \mathcal{O}\big(\eta^{4m\nu/(4m\nu+2m+1)} \big) ,$$

provided $\alpha \asymp \eta^{4m/(4m\nu+2m+1)}$. *Likewise,*

$$\| K(x^{\alpha,\delta} - x_o) \|_Y = \mathcal{O}\big(\eta^{4m\mu/(4m\mu+1)} \big) ,$$

provided $\alpha \asymp \eta^{4m/(4m\mu+1)}$, *with* μ *as in the Main Theorem. (The statement* $\alpha \asymp \eta^p$ *for a fixed* p *means that*

$$0 < \liminf_{\eta\to 0}\ \alpha\,\eta^{-p} \leqslant \limsup_{\eta\to 0}\ \alpha\,\eta^{-p} < \infty\ .\ )$$

Remark 4. Note that for $m \to \infty$, one gets $\| x^{\alpha,\delta} - x_o \|_X = \mathcal{O}\big(\eta^{2\nu/(2\nu+1)} \big)$, and $\| K(x^{\alpha,\delta} - x_o) \|_Y = \mathcal{O}\big(\eta \big)$, which are the best possible rates in this scheme. □

We now outline the proof of the Main Theorem by way of some basic results that are proved in later sections. The principal tool goes back at least as far as Ribière [38]. Of course, the constraints require attention. We also need to establish a continuity result : How does the solution of (2.2) depend on the data y . For precision, we introduce the notation

(2.5) $$x_\alpha(y) = x^{\alpha,\delta} \quad \text{where} \quad x^{\alpha,\delta} \quad \text{solves} \quad (2.2) ,$$

and define $x^{\alpha,o}$ as the solution of (2.2) in the noiseless constrained case,

$$x^{\alpha,o} = x_\alpha(Kx_o) \ . \tag{2.6}$$

Finally, let $x = \theta^{\alpha,\delta}$ be the *unconstrained* solution of (2.2)

$$\theta^{\alpha,\delta} = (\, K^*K + \alpha\, I\,)^{-1}\, K^*y \ , \tag{2.7}$$

and let $\theta^{\alpha,o}$ be the unconstrained noiseless solution, i.e., given by (2.7) with $y = Kx_o$.

The first set of results says that we can ignore the constraints. Of course, if $x^{\alpha,\delta}$ is much more accurate than $\theta^{\alpha,\delta}$, e.g., if the set C is "small" in some sense (the most extreme example being $C = \{\, x_o \,\}$, but classes of analytic functions come to mind) then this is of not much use. The result is an elementary application of convex optimization.

DEFINITION 1. *Let $C \subset X$ be closed and convex. The Phillips-Tikhonov projection of $x \in X$ onto C is defined as the solution of the problem*

$$\text{minimize} \quad \|\, x - b \,\|_{\alpha,X} \quad \text{subject to} \quad b \in C \ .$$

THEOREM 2. *$x^{\alpha,\delta}$ is the Phillips-Tikhonov projection of $\theta^{\alpha,\delta}$ onto C.*

COROLLARY 3. $\quad \|\, x^{\alpha,\delta} - x_o \,\|_{\alpha,X} \leqslant \|\, \theta^{\alpha,\delta} - x_o \,\|_{\alpha,X}$.

Of course, we need to worry about $\|\, \theta^{\alpha,\delta} - x_o \,\|_{\alpha,X}$, especially in view of the source condition (2.4). More about this later. The second set of results expresses the (lack of) continuity of $x_\alpha(y)$ as function of y.

THEOREM 4. (RIBIÈRE [38]) *For y and $z \in Y$,*

$$\|\, x_\alpha(y) - x_\alpha(z) \,\|^2_{\alpha,X} \leqslant \langle\, y - z \,,\, K\big(x_\alpha(y) - x_\alpha(z)\big)\,\rangle_Y \ .$$

The following Corollary 5 represents the decomposition of $x^{\alpha,\delta}$ into a part corresponding to noiseless data, and a part due to pure noise data. (In the absence of constraints, the decomposition would be exact.)

COROLLARY 5. (RIBIÈRE [38]) $\quad \|\, x^{\alpha,\delta} - x^{\alpha,o} \,\|^2_{\alpha,X} \leqslant \langle\, \delta \,,\, K(x^{\alpha,\delta} - x^{\alpha,o})\,\rangle_Y$.

In the next section, we discuss the contribution of the noise, culminating in the following theorem.

THEOREM 6. *Under the assumptions* (1.7)-(1.9) *on the weakly bounded noise* δ *, for all* $x \in X$ *,*

$$\left| \left\langle \delta \, , \, Kx \right\rangle_Y \right| \leqslant \alpha^{-1/4m} \, \eta \, \| \, x \, \|_{\alpha,X} \; .$$

We must now deal with $x^{\alpha,o} - x_o$. First, Theorem 2 gets rid of the constraints in (2.2).

COROLLARY 7. $\quad \| \, x^{\alpha,o} - x_o \, \|_{\alpha,X} \leqslant \| \, \theta^{\alpha,o} - x_o \, \|_{\alpha,X}$.

Now, it is easy to show that $\| \, \theta^{\alpha,o} - x_o \, \|^2_{\alpha,X} = \mathcal{O}\big(\, \alpha^{2\nu+1} \, \big)$ for $0 < \nu \leqslant \frac{1}{2}$, but unfortunately, for $\nu > \frac{1}{2}$, it fails: We are stuck at the bound $\mathcal{O}\big(\, \alpha^2 \, \big)$. Some cheap tricks must be applied to get around this: Perturb x_o slightly, i.e., define w_o and $w^{\alpha,o}$ by

$$w_o = \big(\, (K^*K)^2 + \alpha^2 \, I \, \big)^{\nu/2} \, z_o \quad , \quad w^{\alpha,o} = x_\alpha(K w_o) \; , \tag{2.8}$$

with z_o as in the source condition (2.4). From Theorem 4, we get

COROLLARY 8. $\quad \| \, w^{\alpha,o} - x^{\alpha,o} \, \|_{\alpha,X} \leqslant \| \, K(w_o - x_o) \, \|_Y$.

Finally, we need the following two results.

LEMMA 9. *Under the source condition* (2.4)*, for* $0 < \nu \leqslant 1$ *,*

$$\| \, w_o - x_o \, \|_{\alpha,X} = \mathcal{O}\big(\, \alpha^{\nu+1/2} \, \big) \; .$$

LEMMA 10. *Under the source condition* (2.4)*, for* $0 < \nu \leqslant 1$ *,*

$$\| \, w^{\alpha,o} - w_o \, \|_X = \mathcal{O}\big(\, \alpha^\nu \, \big)$$

It should be noted that for $\nu > \frac{1}{2}$, we are stuck again,

$$\| \, K(w^{\alpha,o} - w_o) \, \|_Y = \mathcal{O}\big(\, \alpha \, \big) \; ,$$

but this time it does not matter.

PROOF OF THE MAIN THEOREM. We obviously have

$$\| \, x^{\alpha,\delta} - x_o \, \|_X \leqslant \| \, x^{\alpha,\delta} - x^{\alpha,o} \, \|_X + \| \, x^{\alpha,o} - w^{\alpha,o} \, \|_X + \| \, w^{\alpha,o} - w_o \, \|_X + \| \, w_o - x_o \, \|_X \; .$$

Now, Corollary 5 and Theorem 6 take care of the first term on the right: $\| x^{\alpha,\delta} - x^{\alpha,o} \|_{\alpha,X} \leqslant \alpha^{-1/(4m)} \eta$, so that

$$\| x^{\alpha,\delta} - x^{\alpha,o} \|_{X} \leqslant \alpha^{-1/2-1/4m} \eta = \alpha^{-(2m+1)/4m} \eta \ .$$

Next, the second term succumbs to Corollary 8 and Lemma 9. The required bound on the third term is provided by Lemma 10, and the fourth term by Lemma 9. So,

$$\| x^{\alpha,\delta} - x_o \|_{X} = \mathcal{O}\big(\alpha^{-(2m+1)/4m} \eta + \alpha^{\nu} \big) \ ,$$

and for $\alpha \asymp \eta^{4m/(4m\nu+2m)}$, the bound of the Main Theorem is proved. Q.e.d.

3. Weakly bounded noise

In this section, we provide bounds on $\langle \delta \, , \, Kx \rangle_{Y}$, for all $x \in X$, thus proving Theorem 6. However, neither the operator K nor the weakly bounded noise δ enter into the picture until the very end.

Recall the magical operator T from the conditions (1.7)-(1.9) on the weakly bounded noise. Let $\{ \lambda_i \, , \, u_i \}_{i=1}^{\infty}$ be the complete orthonormal system of eigenvectors and eigenvalues of T, i.e.,

(3.1) $$T u_i = \lambda_i u_i \ , \quad i = 1, 2, \cdots \ ,$$

where, without loss of generality, $1 = \lambda_1 \geqslant \lambda_2 \geqslant \cdots > 0$. Then, every $y \in Y$ may be represented as

(3.2) $$y = \sum_{i=1}^{\infty} \langle y \, , \, u_i \rangle_{Y} u_i \qquad \text{(strong convergence in } Y\text{)} \ .$$

For $m \geqslant 1$ and $0 < \beta \leqslant 1$, it is useful to introduce the inner products $\langle \cdot \, , \cdot \rangle_{m,\beta}$ on $T^m(Y)$ by

(3.3) $$\langle y \, , \, z \rangle_{m,\beta} = \langle y \, , \, z \rangle_{Y} + \beta^{2m} \langle T^{-m} y \, , \, T^{-m} z \rangle_{Y} \ .$$

Define the norms $\| \cdot \|_{m,\beta}$ in the usual way.

The growth (decay?) of these norms with increasing m is under control.

LEMMA 11. *For* $1 \leqslant m < n$, *for all* $0 < \beta \leqslant 1$, *and for all* $y \in T^n(Y)$,

$$\| \, y \, \|_{m,\beta} \leqslant 2 \, \| \, y \, \|_{n,\beta} \; .$$

PROOF. A simple computation does the trick. Using the representation (3.2) for $y \in Y$, we get

$$\begin{aligned} \| \, y \, \|_{m,\beta}^2 &= \sum_{i=1}^{\infty} \big(\, 1 + (\beta/\lambda_i)^{2m} \, \big) \, | \, \langle \, y \, , \, u_i \, \rangle_Y \, |^2 \\ &= \sum_{i=1}^{\infty} a_{n,m,i} \, \big(\, 1 + (\beta \, \lambda_i)^{2n} \, \big) \, | \, \langle \, y \, , \, u_i \, \rangle_Y \, |^2 \; , \end{aligned}$$

where

$$a_{n,m,i} = \frac{1 + (\beta/\lambda_i)^{2m}}{1 + (\beta/\lambda_i)^{2n}} \leqslant \sup_{t>0} \frac{1 + \, t^{\,2m}}{1 + \, t^{\,2n}} \; .$$

Now, for $0 < \; t \; \leqslant 1$, the expression under consideration is bounded by 2 (the numerator is at most 2, the denominator is at least 1), whereas for $t \geqslant 1$, it is bounded by 1 ($m < n$). Thus, $a_{n,m,i} \leqslant 2$ for all relevant n, m and i . Q.e.d.

The task at hand is to bound $\langle \, \delta \, , \, y \, \rangle_Y$ in terms of $\| \, y \, \|_{1,\beta}$. This amounts to computing the dual norm.

LEMMA 12. *For all* $z \in Y$ *and* $y \in T(Y)$.

$$| \, \langle \, z \, , \, y \, \rangle_Y \, |^2 \leqslant \| \, y \, \|_{1,\beta}^2 \, \langle \, z \, , \, (\, I + \beta^2 \, T^{-2} \,)^{-1} \, z \, \rangle_Y \; .$$

PROOF. Note that the operator $I + \beta^2 \, T^{-2}$ is unbounded but has a bounded inverse, and then so does the Hermitian square root, $(\, I + \beta^2 \, T^{-2} \,)^{1/2}$. Now, for $y \in T(Y)$, we have that

$$\begin{aligned} | \, \langle \, z \, , \, y \, \rangle_Y \, |^2 &= | \, \langle \, (\, I + \beta^2 \, T^{-2} \,)^{-1/2} \, z \, , \, (\, I + \beta^2 \, T^{-2} \,)^{1/2} \, y \, \rangle_Y \, |^2 \\ &\leqslant \| \, (\, I + \beta^2 \, T^{-2} \,)^{-1/2} \, z \, \|_Y^2 \, \| \, (\, I + \beta^2 \, T^{-2} \,)^{1/2} \, y \, \|_Y^2 \; , \end{aligned}$$

by Cauchy-Schwarz, and this equals $\langle \, z \, , \, (\, I + \beta^2 \, T^{-2} \,) \, z \, \rangle_Y \, \| \, y \, \|_{1,\beta}^2$, and that does it. Q.e.d.

The above lemma tells us how to bound $\langle \, z \, , \, y \, \rangle_Y$. We need to bound this further in terms of $\langle \, z \, , \, Tz \, \rangle_Y$.

Lemma 13. *For all* $z \in Y$,

$$\left\langle z\,,\,(\,I+\beta^2\,T^{-2}\,)^{-1}\,z\,\right\rangle_Y \leqslant (\,2\beta)^{-1}\left\langle z\,,\,Tz\,\right\rangle_Y \;.$$

Proof. We have

$$\left\langle z\,,\,(\,I+\beta^2\,T^{-2}\,)^{-1}\,z\,\right\rangle_Y = \left\langle z\,,\,(\,T+\beta^2\,T^{-1}\,)^{-1}\,T\,z\,\right\rangle_Y \leqslant r\left\langle z\,,\,Tz\,\right\rangle_Y \;,$$

where r is the spectral radius of $(\,T+\beta^2\,T^{-1}\,)^{-1}$. Then,

$$\begin{aligned} r &\leqslant \sup_{t>0}\;(\,t\,+\beta^2\,t^{\,-1}\,)^{-1} = \sup_{t>0}\;t\,(\,t^{\,2}\,+\beta^2\,)^{-1} \\ &= \sup_{\tau>0}\;\beta\,\tau(\,\beta^2\,\tau^2\,+\beta^2\,)^{-1} = \beta^{-1}\sup_{\tau>0}\;\tau(\,\tau^2\,+1\,)^{-1}\;. \end{aligned}$$

so that $r \leqslant (2\beta)^{-1}$. Q.e.d.

All the pieces are now in place.

Proof of Theorem 6. Let $m \geqslant 1$ and let δ be weakly bounded noise in the sense of (1.7)-(1.9). Let $x \in X$. Then $Kx \in T^m(Y)$, so surely, $Kx \in T(Y)$. Then Lemmas 12 and 13 imply that for $0 < \beta \leqslant 1$,

$$\begin{aligned} \left|\left\langle\,\delta\,,\,Kx\,\right\rangle_Y\right|^2 &\leqslant \|\,Kx\,\|_{1,\beta}^2\,\left\langle\,\delta\,,\,(\,I+\beta^2\,T^2\,)^{-1}\,\delta\,\right\rangle_Y \\ &\leqslant (2\beta)^{-1}\,\|\,Kx\,\|_{1,\beta}^2\,\left\langle\,\delta\,,\,T\delta\,\right\rangle_Y \\ &\leqslant (2\beta)^{-1}\,\eta^2\,\|\,Kx\,\|_{1,\beta}^2 \leqslant \beta^{-1}\,\eta^2\,\|\,Kx\,\|_{m,\beta}^2 \;, \end{aligned}$$

the last line by virtue of Lemma 11. Finally, by (1.7),

$$\|\,Kx\,\|_{m,\beta}^2 = \|\,Kx\,\|_Y^2 + \beta^{2m}\,\|\,T^{-m}Kx\,\|_Y^2 \leqslant \|\,Kx\,\|_Y^2 + c\,\beta^{2m}\,\|\,x\,\|_X^2 \;.$$

Now, take $\beta = \alpha^{1/2m}$, so that $\|\,Kx\,\|_{m,\beta} = \|\,x\,\|_{\alpha,X}$, and we are done. Q.e.d.

4. Continuity and constraints

In this section, we prove the characterization of the constrained regularized solution, Theorem 2, and show the continuity of the regularized solution on the data. We also provide the required bounds on the noiseless regularized solutions.

First,we show that the constrained regularized solution is the Phillips-Tikhonov projection of the unconstrained solution onto the constraint set, and then the obvious relation between the constrained and unconstrained error.

Proof of Theorem 2. Let $y \in Y$ be arbitrary, and let θ be the unconstrained Phillips-Tikhonov regularized solution, i.e.,

$$\theta = (\, K^* K + \alpha\, I \,)^{-1}\, K^* y \ .$$

Consider the functional $\mathrm{LS}(\, x \,|\, y \,)$ of (2.2). Its Fréchet derivative with respect to x is given by $\nabla\mathrm{LS}(\, x \,|\, y \,) = 2\, K^*(Kx - y) + 2\,\alpha\, x$, so that $\nabla\mathrm{LS}(\, \theta \,|\, y \,) = 0$. Now, quadratic Taylor expansion shows that for all $x \in X$,

$$\begin{aligned} \mathrm{LS}(\, x \,|\, y \,) &= \mathrm{LS}(\, \theta \,|\, y \,) + \big\langle\, \nabla\mathrm{LS}(\, \theta \,|\, y \,)\,,\, x - \theta \,\big\rangle_X + \|\, x - \theta \,\|_{\alpha,X}^2 \\ &= \mathrm{LS}(\, \theta \,|\, y \,) + \|\, x - \theta \,\|_{\alpha,X}^2 \ . \end{aligned}$$

Thus, the minimizers of $\mathrm{LS}(\, x \,|\, y \,)$ and $\|\, x - \theta \,\|_{\alpha,X}$ over $x \in C$ are one and the same. Q.e.d.

Proof of Corollary 3. Again by Quadratic Taylor expansion,

$$\|\, \theta - x_o \,\|_{\alpha,X}^2 = \|\, x^{\alpha,o} - x_o \,\|_{\alpha,X}^2 + \|\, \theta - x^{\alpha,o} \,\|_{\alpha,X}^2 + 2\, q \ ,$$

with $q = \big\langle\, K(x^{\alpha,o} - x_o)\,,\, K(\theta - x^{\alpha,o}) \,\big\rangle_Y + \alpha\, \big\langle\, x^{\alpha,o} - x_o \,,\, \theta - x^{\alpha,o} \,\big\rangle_X$. Now,

$$\begin{aligned} q &= \big\langle\, x^{\alpha,o} - x_o \,,\, K^*K(\theta - x^{\alpha,o}) + \alpha\,(\theta - x^{\alpha,o}) \,\big\rangle_X \\ &= \big\langle\, x_o - x^{\alpha,o} \,,\, K^*K(x^{\alpha,o} - \theta) + \alpha\,(x^{\alpha,o} - \theta) \,\big\rangle_X \\ &= \tfrac{1}{2}\, \big\langle\, x_o - x^{\alpha,o} \,,\, \nabla\mathrm{LS}(\, x^{\alpha,o} \,|\, K\theta \,) \,\big\rangle_X \geqslant 0 \ , \end{aligned}$$

the last inequality by the necessary and sufficient conditions for a minimum of $\mathrm{LS}(\, x \,|\, K\theta \,)$ over $x \in C$. Q.e.d.

We now show the continuity of the regularized solution, which implies the bound on the error due to the noise.

Proof of Theorem 4. Let $y\,,\ z \in Y$. Let $x = x_\alpha(y)$ and $w = x_\alpha(v)$. Then, by the necessary and sufficient conditions for the minimum of a convex differentiable minimization problem, $\left\langle \nabla\mathrm{LS}(\, x \,|\, y \,) \,,\, w - x \,\right\rangle_X \geqslant 0$. Then, by quadratic Taylor expansion,

$$\mathrm{LS}(\, w \,|\, y \,) - \mathrm{LS}(\, x \,|\, y \,) = \left\langle \nabla\mathrm{LS}(\, x \,|\, y \,) \,,\, w - x \,\right\rangle_X + \|\, w - x \,\|_{\alpha,X}^2 \geqslant \|\, w - x \,\|_{\alpha,X}^2 \;.$$

Likewise, one obtains $\mathrm{LS}(\, x \,|\, z \,) - \mathrm{LS}(\, w \,|\, z \,) \geqslant \|\, x - w \,\|_X^2$, and then upon adding these two inequalities and simplifying gives

$$2 \,\|\, w - x \,\|_{\alpha,X}^2 \leqslant \|\, Kw - y \,\|_Y^2 - \|\, Kx - y \,\|_Y^2 + \|\, Kx - z \,\|_Y^2 - \|\, Kw - z \,\|_Y^2 \;.$$

The right hand side is equal to $2 \left\langle \, y - z \,,\, K(x - w) \,\right\rangle_Y$. Q.e.d.

The Corollaries 5 and 7 are obvious, as is Corollary 8. We proceed with the proofs of Lemmas 9 and 10.

Proof of Lemma 9. Since

$$w_o - x_o = \left\{ \left((\, K^* K \,)^2 + \alpha^2 I \right)^{\nu/2} - (K^* K)^{\nu} \right\} z_o \;,$$

then

$$\|\, w_o - x_o \,\|_X \leqslant \varrho \,\|\, z_o \,\|_X \;,$$

where ϱ is the spectral radius of $(\, (K^* K)^2 + \alpha^2 I \,)^{\nu/2} - (K^* K)^{\nu}$. Thus,

$$\begin{aligned} \varrho &\leqslant \sup_{t>0} \; (\, t^2 + \alpha^2 \,)^{\nu/2} - t^{\nu} = \sup_{\tau>0} \; (\, \alpha^2 \tau^2 + \alpha^2 \,)^{\nu/2} - (\alpha\, \tau \,)^{\nu} \\ &= \alpha^{\nu} \sup_{\tau>0} \; (\, \tau^2 + 1 \,)^{\nu/2} - \tau^{\nu} \;, \end{aligned}$$

and the supremum is finite for $0 < \nu \leqslant 1$.

In the same vein, $\|\, K(w_o - x_o) \,\|_Y \leqslant r \,\|\, z_o \,\|_X$, where r is the spectral radius of $K \left\{ \left((K^* K)^2 + \alpha^2 I \right)^{\nu/2} - (K^* K)^{\nu} \right\}$. Then, similar to the above,

$$\begin{aligned} r &\leqslant \sup_{t>0} \;\; t \;\left\{ (\, t^4 + \alpha^2)^{\nu/2} - \, t^{2\nu} \right\} \\ &= \alpha^{\nu+1/2} \sup_{\tau>0} \;\; \tau \left\{ \, (\tau^4 + 1)^{\nu/2} - \tau^{2\nu} \right\} \;, \end{aligned}$$

and the supremum is finite for $0 < \nu \leqslant 1$ (for $\nu \leqslant \frac{3}{2}$, in fact). Q.e.d.

Proof of Lemma 9. We have

$$\begin{aligned} w^{\alpha,o} - w_o &= \big((K^*K + \alpha I)^{-1} K^*K - I \big)\, w_o \\ &= -\alpha\, (K^*K + \alpha I)^{-1}\, w_o \\ &= -\alpha\, (K^*K + \alpha I)^{-1} \big((K^*K)^2 + \alpha^2 I \big)^{\nu/2} z_o \;, \end{aligned}$$

so that

$$\| \, w^{\alpha,o} - w_o \, \|_X \leqslant R \, \| \, z_o \, \|_X \;,$$

with R the spectral radius of $-\alpha\, (K^*K + \alpha I)^{-1} \big((K^*K)^2 + \alpha^2 I \big)^{\nu/2}$. Consequently,

$$\begin{aligned} R &\leqslant \alpha \sup_{t>0} \; (\, t + \alpha)^{-1} (\, t^2 + \alpha^2 \,)^{\nu/2} \\ &= \alpha \sup_{\tau>0} \; (\alpha\, \tau + \alpha)^{-1} (\alpha^2 \tau^2 + \alpha^2 \,)^{\nu/2} \\ &= \alpha^{\nu} \sup_{\tau>0} \; (\, \tau + 1 \,)^{-1} (\, \tau^2 + 1 \,)^{\nu/2} \;, \end{aligned}$$

and for $\nu \leqslant 1$, the supremum is finite. Q.e.d.

Remark 5. Computing a bound on $\| \, K(w^{\alpha,o} - w_o) \, \|_Y$, as in the previous proof gives $\| \, K(w^{\alpha,o} - w_o) \, \|_Y \leqslant \mathfrak{R} \, \| \, z_o \, \|$, with $\mathfrak{R}$ the spectral radius of $-\alpha\, K(K^*K + \alpha I)^{-1} \big((K^*K)^2 + \alpha^2 I \big)^{\nu/2}$, and so

$$\begin{aligned} \mathfrak{R} &\leqslant \alpha \sup_{t>0} \;\; t\, (\, t^2 + \alpha)^{-1} (\, t^4 + \alpha^2 \,)^{\nu/2} \\ &= \alpha \sup_{\tau>0} \;\; \alpha^{1/2} \tau (\alpha \tau^2 + \alpha)^{-1} (\alpha^2 \tau^4 + \alpha^2 \,)^{\nu/2} \\ &= \alpha^{\nu+1/2} \sup_{\tau>0} \;\; \tau (\tau^2 + 1)^{-1} (\tau^4 + 1 \,)^{\nu/2} \;, \end{aligned}$$

but now, unfortunately, the supremum is finite for $\nu \leqslant \frac{1}{2}$ and infinite for $\nu > \frac{1}{2}$. So, one has to be more precise in the range of t,

$$\mathfrak{R} = \alpha \sup_{0 < t \leqslant k} \;\; t \; (\, t^2 + \alpha)^{-1} (\, t^4 + \alpha^2 \,)^{\nu/2} \;,$$

where k is the largest singular value of K. The upshot of it is that one only gets $\| \, K(w^{\alpha,o} - w_o) \, \|_Y = \mathcal{O}\big(\alpha^{\mu} \big)$ with $\mu = \min(\, 1 \,, \, \nu + \frac{1}{2} \,)$. □

5. A weak discrepancy principle

In this section, we discuss a weak discrepancy principle for choosing the regularization parameter α, in the unconstrained problem, i.e., (2.2) with $\mathcal{C} = X$, under the conditions (1.7)-(1.9) on the weakly bounded noise and the source condition

$$x_o = K^* K\, w_o \quad \text{for some} \quad w_o \in X\,,\ w_o \neq 0\,. \tag{5.1}$$

i.e., (2.4) with $\boldsymbol{\nu = 1}$. Our discrepancy principle is essentially the one of Gfrerer [15], [16] based on iterated Tikhonov regularization, but with some obvious, necessary modifications.

Recall that in the classical setting, where the noise is assumed to converge strongly to 0,

$$\| \delta \|_Y \longrightarrow 0\,, \tag{5.2}$$

the discrepancy principle of Gfrerer [15], [16] states that α should be chosen such that for some $n \in \mathbb{N}$,

$$\alpha^{2n+1} \left\langle y\,, (\, KK^* + \alpha I\,)^{-2n-1}\, y \right\rangle_Y = C\, \| \delta \|_Y^2 \tag{5.3}$$

for some constant $C \geqslant 1$. Under the source condition (5.1) for $0 < \nu \leqslant n$, Gfrerer [15], [16] proves that this "works" for all $n \in \mathbb{N}$, but here, we are considering the case $n = 1$ only.

It is clear that in the setting of weakly bounded noise, any reference to $\| \delta \|_Y$ is not very informative, since asymptotically, it is just a positive constant. (In fact, it may even be $+\infty$, as in the case of white noise.) Now, the left hand side on (5.3) involves the term (for $n = 1$)

$$\alpha^3 \left\langle \delta\,, (\, KK^* + \alpha I\,)^{-3}\, \delta \right\rangle_Y\,,$$

about which in the context of weakly bounded noise, it is hard to say anything useful. All of these problems go away if we replace one of the y in (5.3) by the smoothed version $Kx^{\alpha,\delta}$. Then, (5.3) takes the form

$$\alpha^3 \left\langle K^*y\,, (\, KK^* + \alpha I\,)^{-4}\, K^*y \right\rangle_Y = C\, \alpha^{-1/2m}\, \eta^2 \tag{5.4}$$

for an appropriately chosen constant C. This may be rewritten in various ways. First, the left hand side of (5.4) equals $\alpha^3 \| (\, K^*K + \alpha I\,)^{-1} x^{\alpha,\delta} \|^2$, Now, using the second Tikhonov iterate, see Gfrerer [15], [16],

$$z^{\alpha,\delta} = 2\, x^{\alpha,\delta} - (\, K^*K + \alpha I\,)^{-1}\, K^*K x^{\alpha,\delta} \tag{5.5}$$

(and $z^{\alpha,o}$ being the noise-free version), one notes that

$$z^{\alpha,\delta} - x^{\alpha,\delta} = \alpha\,(\,K^*K + \alpha\,I\,)^{-1}\,x^{\alpha,\delta} = \alpha\,(\,K^*K + \alpha\,I\,)^{-2}\,K^*y\ ,$$

and so

$$\alpha\,\|\,z^{\alpha,\delta} - x^{\alpha,\delta}\,\|_X^{\,2} = \alpha^3\,\big\langle\,K^*y\,,\,(\,K^*K + \alpha\,I\,)^{-4}\,K^*y\,\big\rangle_Y\ . \tag{5.6}$$

REMARK 6. Following the development of § 4 regarding the spectral radii of various operators, it is easy to show that under the source condition (2.4),

$$\|\,z^{\alpha,o} - x_o\,\|_X = \mathcal{O}\big(\,\alpha^\nu\,\big)\ ,$$

but now for all ν with $0 < \nu \leqslant 2$! See GFRERER [15], [16].

The import of the source condition may now be stated.

LEMMA 14. *Under the source condition* (5.1), *for* $\alpha \to 0$,

$$\|\,x^{\alpha,o} - x_o\,\|_X = \alpha\,\|\,w_o\,\|_X + o(\,\alpha\,) \quad \textit{and} \quad \|\,z^{\alpha,o} - x^{\alpha,o}\,\|_X = \alpha\,\|\,w_o\,\|_X + o(\,\alpha\,)\ .$$

PROOF. One shows that

$$\begin{aligned} x^{\alpha,\delta} - x_o &= -\alpha\,(\,K^*K + \alpha\,I\,)^{-1}\,x_o = -\alpha\,(\,K^*K + \alpha\,I\,)^{-1}\,K^*Kw_o \\ &= -\alpha\,w_o + \alpha^2\,(\,K^*K + \alpha\,I\,)^{-1}\,w_o\ , \end{aligned}$$

and that $\alpha\,\|\,(\,K^*K + \alpha\,I\,)^{-1}\,w_o\,\|_X \to 0$ as $\alpha \to 0$. Thus,

$$\|\,x^{\alpha,o} - x_o\,\|_X = \alpha\,\|\,w_o\,\|_X + o(\,\alpha\,)\ .$$

Similarly, from (5.5),

$$\begin{aligned} z^{\alpha,o} - x^{\alpha,o} &= \big(\,I - (\,K^*K + \alpha\,I\,)^{-1}K^*K\,\big)\,x^{\alpha,o} \\ &= \alpha\,(\,K^*K + \alpha\,I\,)^{-2}K^*K\,x_o \\ &= \alpha\,(\,K^*K + \alpha\,I\,)^{-1}\,x_o - \alpha^2\,(\,K^*K + \alpha\,I\,)^{-2}\,x_o \\ &= \alpha\,w_o + \varepsilon_o\ , \end{aligned}$$

with

$$\varepsilon_o = -2\,\alpha^2\,(\,K^*K + \alpha\,I\,)^{-2}\,w_o + \alpha^3\,(\,K^*K + \alpha\,I\,)^{-3}\,w_o\ .$$

Since $\alpha\,\|\,(\,K^*K + \alpha\,I\,)^{-1}\,\|_X \leqslant 1$ it follows as above that $\|\,\varepsilon_o\,\|_X = o(\alpha)$. Thus,

$$\|\,z^{\alpha,o} - x^{\alpha,o}\,\|_X = \alpha\,\|\,w_o\,\|_X + o(\,\alpha\,)\ .$$

Since $\| w_o \|_X \neq 0$, the lemma follows. Q.e.d.

We are now ready for the discrepancy principle.

A WEAK DISCREPANCY PRINCIPLE. Under the conditions (1.7)-(1.9) on the weakly bounded noise, choose α as the smallest solution of

$$\alpha^{1/2} \| z^{\alpha,\delta} - x^{\alpha,\delta} \|_X = C\, \alpha^{-1/4m}\, \eta \; . \tag{5.7}$$

Here, C is a large enough constant.

REMARK 7. Ideally, one would want to choose the constant $C = C(\alpha)$ to depend on α, such that

$$C(\alpha)\, \alpha^{-1/4m}\, \eta = \| x^{\alpha,\delta} - x^{\alpha,o} \|_X = \| (K^*K + \alpha I)^{-1} K^* \delta \|_X \; .$$

Under certain circumstances, the right hand side might be estimable.

THEOREM 15. (RATES FOR THE DISCREPANCY PRINCIPLE) *Under the source condition* (5.1) *and the conditions* (1.7)-(1.9) *on the weakly bounded noise, the discrepancy equation* (5.7) *(with C large enough) implies that*

$$\| x^{\alpha,\delta} - x_o \|_X = \mathcal{O}\big(\eta^{4m/(6m+1)} \big) \; .$$

We break the proof up into a number of manageable pieces. The first one is to rewrite the weak discrepancy principle as a minimum principle. One should contrast this with the approach of GFRERER [15], [16].

LEMMA 16. (MINIMUM PRINCIPLE) *The Weak Discrepancy Principle* (5.7) *is equivalent to finding stationary points of*

$$\begin{aligned} &\text{minimize} \quad F(\gamma,\eta) \stackrel{def}{=} \langle\, y\, ,\, (I + \gamma\; KK^*)^{-3}\, y \,\rangle_Y \;+\; 6\, m\, C^2\, \gamma^{1/2m}\, \eta^2 \\ &\text{subject to} \quad 0 < \gamma < \infty \;, \end{aligned}$$

where $\gamma \equiv 1/\alpha$ *.*

PROOF. Using (5.6), one sees that

$$\alpha \, \| z^{\alpha,\delta} - x^{\alpha,\delta} \|_X^2 = \tfrac{1}{3}\, \alpha\, \frac{d}{d\alpha} \Big\{ \alpha^3 \langle\, y\, ,\, (KK^* + \alpha I)^{-3}\, y \,\rangle_Y \Big\} \; ,$$

and so the Discrepancy Principle equation (5.7) is equivalent to

$$\frac{d}{d\alpha}\Big\{\, \alpha^3 \big\langle\, y\,,\,(\,KK^* + \alpha\, I\,)^{-3}\, y\,\big\rangle_Y \,+ 6\, m\, C^2\, \alpha^{-1/2m}\, \eta^2 \,\Big\} = 0\ .$$

Thus, we are finding critical points of the one-dimensional minimization problem,

$$\alpha^3 \big\langle\, y\,,\,(\,KK^* + \alpha\, I\,)^{-3}\, y\,\big\rangle_Y + C_1\, \alpha^{-1/2m}\, \eta^2\ .$$

The transformation $\gamma = 1/\alpha$ then yields the advertised minimization problem (but it seems obvious that we want (local) minima). Q.e.d.

LEMMA 17. *Let $y \in Y$ and the constant C be fixed. Then, there exists a positive constant ϑ_o such that the global minimizer $\gamma = \gamma(\vartheta)$ of $F(\gamma,\vartheta)$ exists.*

REMARK 8. Of course, we would have liked to apply the implicit function theorem to show that $\gamma(\vartheta)$ is a continuous, decreasing function of ϑ. However, we are unable to show that $\partial^2 F/\partial\gamma^2 > 0$ at the minimizer $\gamma(\vartheta)$.

PROOF OF LEMMA 17. It is clear that F is infinitely many times differentiable in both arguments. First, we show that $F(\gamma,\vartheta)$ attains a global minimum away from 0, for ϑ small enough. Observe that $F(0,\vartheta) = 0$ and that $F(\gamma,\vartheta) \to \infty$ for $\gamma \to \infty$. Next, one observes that

$$\frac{d}{d\gamma}\big\langle\, y\,,\,(\,I + \gamma\, KK^*\,)^{-3} y\,\big\rangle_Y = -3\big\langle\, y\,,\,(\,I + \gamma\, KK^*\,)^{-4} KK^* y\,\big\rangle_Y < 0\ ,$$

so that $\big\langle\, y\,,\,(\,I + \gamma\, KK^*\,)^{-3} y\,\big\rangle_Y$ is strictly decreasing. Thus, e.g.,

$$\big\langle\, y\,,\,(\,I + KK^*\,)^{-3} y\,\big\rangle_Y < \|\, y\,\|_Y^2\ .$$

Then,

$$F(1,\vartheta) = \big\langle\, y\,,\,(\,I + KK^*\,)^{-3} y\,\big\rangle_Y + 6\, m\, C^2\, \vartheta^2 < \|\, y\,\|_Y^2 = F(0,\vartheta)$$

for

$$\vartheta^2 < \vartheta_o^2 \stackrel{\text{def}}{=} (\,6\, m\, C^2\,)^{-1} \big\{\, \|\, y\,\|_Y^2 - \big\langle\, y\,,\,(\,I + KK^*\,)^{-3} y\,\big\rangle_Y \,\big\}\ .$$

So, for $\vartheta < \vartheta_o$, there is a global minimum away from 0. Q.e.d.

Lemma 18. *Under the provisions of Theorem* 15*, the discrepancy principle equation* (5.7) *(with* C *large enough) is asymptotically equivalent with*

$$\alpha^{1/2} \| z^{\alpha,\delta} - x^{\alpha,o} \|_X \asymp \alpha^{-1/4m} \eta \ .$$

Proof. Let $\varepsilon = z^{\alpha,\delta} - x^{\alpha,\delta} - (z^{\alpha,o} - x^{\alpha,o})$. A simple calculation,

$$\begin{aligned} \varepsilon &= \alpha (K^*K + \alpha I)^{-2} K^* \delta \\ &= -\alpha (K^*K + \alpha I)^{-1} (x^{\alpha,\delta} - x_o) \ , \end{aligned}$$

shows that

$$\| \varepsilon \|_X \leqslant \| x^{\alpha,\delta} - x_o \|_X \leqslant C_1 \alpha^{-(1/2)-1/4m} \eta \ ,$$

for a suitable constant C_1, the last bound by Corollary 5 and Theorem 6. It follows from the triangle inequality, that

$$\| x^{2\alpha,o} - x^{\alpha,o} \|_X - \| \varepsilon \|_X \leqslant \| x^{2\alpha,\delta} - x^{\alpha,\delta} \|_X \leqslant \| x^{2\alpha,o} - x^{\alpha,o} \|_X + \| \varepsilon \|_X \ ,$$

and so

$$(5.8) \qquad (C - C_1) \alpha^{-1/4m} \eta \leqslant \alpha^{1/2} \| x^{2\alpha,o} - x^{\alpha,o} \|_X \leqslant (C + C_1) \alpha^{-1/4m} \eta \ .$$

So, if C is large enough, the lemma follows. Q.e.d.

Corollary 19. *Under the provisions of Theorem 15, the regularization parameter* α *chosen according to* (5.7) *satisfies*

$$\alpha \asymp \eta^{4m/(6m+1)} \ .$$

Proof. Follows from Lemmas 14 and 18. Q.e.d.

Proof of Theorem 15. This now follows from Lemma 14 and Corollary 19. Q.e.d.

We finish this section with an interesting extension of Lemma 14. Unfortunately, it does not enable us to prove the above results for the more general source condition (2.4) for all $0 < \nu < 1$.

Theorem 20. *Let* $0 < \nu \leqslant 1$. *Let* $x_o \in X$ *be arbitrary. Then,*

$$\| z^{\alpha,o} - x^{\alpha,o} \|_X = \mathcal{O}(\alpha^\nu) \qquad (\alpha \to 0)$$

if and only if

$$\| x^{2\alpha,o} - x^{\alpha,o} \|_X = \mathcal{O}(\alpha^\nu) \qquad (\alpha \to 0)$$

if and only if

$$\| x^{\alpha,o} - x_o \|_X = \mathcal{O}(\alpha^\nu) \qquad (\alpha \to 0) \ .$$

PROOF. The first equivalence is fairly straightforward. As a preliminary observation, note that the operator $L : X \to X$, defined for $\alpha > 0$ as

$$L = (K^*K + \alpha I)(K^*K + 2\alpha I)(K^*K + \alpha I)^{-2} ,$$

is bounded and has a bounded inverse, with $\| L \|_X \leqslant 2$ and $\| L^{-1} \|_X \leqslant 1$. Since $z^{\alpha,o} - x^{\alpha,o} = \alpha (K^*K + \alpha)^{-2} K^*K x_o$, and

$$\alpha (K^*K + \alpha I)^{-1} (K^*K + 2\alpha I)^{-1} = (K^*K + \alpha I)^{-1} - (K^*K + 2\alpha I)^{-1} ,$$

then $z^{\alpha,o} - x^{\alpha,o} = L (x^{2\alpha,o} - x^{\alpha,o})$, which implies that

$$\| x^{2\alpha,o} - x^{\alpha,o} \|_X \leqslant \| z^{\alpha,o} - x^{\alpha,o} \|_X \leqslant 2 \| x^{2\alpha,o} - x^{\alpha,o} \|_X , \tag{5.9}$$

and you could not ask for more.

The final equivalence is now straight forward. The if-part is obvious. We introduce the "abbreviation" $x(\alpha) = x^{\alpha,o}$. Next, note that $x(\alpha) \to x_o$ strongly in X, so that

$$\| x(2^{-i}\alpha) - x_o \|_X^2 \longrightarrow 0 \qquad , \qquad i \to \infty ,$$

uniformly in α, $0 < \alpha \leqslant 1$. Then,

$$x^{\alpha,o} - x_o = \sum_{i=0}^{\infty} \left\{ x(2^{-i}\alpha) - x(2^{-i-1}\alpha) \right\} ,$$

and the series converges strongly in X, resulting in/as shown by

$$\begin{aligned} \| x^{\alpha,o} - x_o \|_X &\leqslant \sum_{i=0}^{\infty} \| x(2^{-i}\alpha) - x(2^{-i-1}\alpha) \|_X \\ &\leqslant \sum_{i=0}^{\infty} c\,(2^{-i}\alpha)^{\nu} = \mathcal{O}\big(\alpha^{\nu} \big) . \end{aligned}$$

This shows the only-if-part. Q.e.d.

6. White noise

In this section, we discuss Fredholm integral equations of the first kind on $L^2(0,1)$ with data perturbed by white noise. We assume that the integral operator maps into a Sobolev space of a suitable order. This permits a treatment similar to that of the previous sections, with two exceptions. The first difference is that the set-up of (1.7)-(1.8) does not work, since the Sobolev scale is not a Hilbert scale, see Remark 2, but this is more than offset by the fact that they are reproducing kernel Hilbert spaces. The second difference is that white noise is not square integrable, so the least-squares formulation of Tikhonov regularization must be modified. (We do not consider constraints.)

Sobolev spaces. For technical convenience, we consider Sobolev spaces of integer order only. Let $m \geqslant 1$ be an integer, and let $H^m(0,1)$ be the Sobolev space of functions on $(0,1)$ whose m-the derivative is square integrable. On $H^m(0,1)$, the standard inner product is defined as

$$\langle u, w \rangle_m = \langle u, w \rangle + \langle u^{(m)}, w^{(m)} \rangle , \tag{6.1}$$

with the induced norm $\| \cdot \|_m$. Here, $\langle \cdot, \cdot \rangle$ is the usual $L^2(0,1)$ inner product (and later, $\| \cdot \|$ is the usual $L^2(0,1)$ norm). Actually, it makes sense to define inner products, for $\beta > 0$,

$$\langle u, w \rangle_{m,\beta} = \langle u, w \rangle + \beta^{2m} \langle u^{(m)}, w^{(m)} \rangle , \tag{6.2}$$

with the induced norm $\| \cdot \|_{m,\beta}$. It is well-known that $H^m(0,1)$ for $m \geqslant 1$ ($m > \frac{1}{2}$ if one includes noninteger orders) is a reproducing kernel Hilbert space, thanks to the Sobolev embedding theorem, which implies that for all $t \in [0,1]$ and all $u \in H^m(0,1)$,

$$| u(t) | \leqslant c_1 \| u \|_1 \leqslant c_m \| u \|_m , \tag{6.3}$$

for suitable constants c_m. See, e.g., Adams and Fournier [1]. Then a simple scaling argument shows that for all β, $0 < \beta \leqslant \frac{1}{2}$, and all $t \in [0,1]$ and $u \in H^m(0,1)$,

$$| u(t) | \leqslant c_1 \beta^{-1/2} \| u \|_{1,\beta} \leqslant c_m \beta^{-1/2} \| u \|_{m,\beta} \tag{6.4}$$

with the same c_m. Thus, the point evaluation functionals $t \longmapsto u(t)$ are continuous, and may be represented as $\langle \cdot, \cdot \rangle_{m,\beta}$ inner products with the reproducing kernel, denoted by $\mathfrak{R}_{m,\beta}(t,s)$. So, for all $u \in H^m(0,1)$,

$$u(t) = \langle \mathfrak{R}_{m,\beta}(t, \cdot), u \rangle_{m,\beta} , \quad t \in [0,1] , \tag{6.5}$$

and then, with the same c_m as before, for all $t \in [0,1]$,

$$\| \mathfrak{R}_{m,\beta}(t, \cdot) \|_{m,\beta} \leqslant c_m \beta^{-1/2} . \tag{6.6}$$

We note that $\mathfrak{R}_{m,\beta}$ is the Green's function for the boundary value problem

$$(6.7)\qquad \begin{aligned} &(-\beta^2)^m\, u^{(2m)} + u = f \quad \text{on} \quad (\,0\,,\,1\,)\;, \\ &u^{(k)}(0) = u^{(k)}(1) = 0\;, \quad k = m, \cdots, 2m-1\;. \end{aligned}$$

Finally, for $m = 1$, the complete system of orthogonal eigenfunctions for

$$(6.8)\qquad \begin{aligned} &-\,\beta^2\, u^{(2)} = \lambda\, u \qquad \text{on} \quad (\,0\,,\,1\,) \\ &u'(0) = u'(1) = 0\;, \end{aligned}$$

are given by

$$(6.9)\qquad u_k(\,t\,) = \cos(\,\pi k\, t\,) \quad , \quad \lambda_k = (\pi k)^2 \quad , \quad k = 0, 1, 2, \cdots$$

In particular, this gives rise to the expansion

$$(6.10)\quad \mathfrak{R}_{1,\beta}(\,t\,,\,s\,) = 1 + 2\sum_{k=1}^{\infty}(\,1 + (\pi\beta k)^2\,)^{-1}\, u_k(\,t\,)\, u_k(\,s\,)\;, \quad t\,,\,s \in [\,0\,,\,1\,]\;.$$

For more useful properties of the reproducing kernels, see Eggermont and LaRiccia [10]. For the asymptotics of the eigenvalues and eigenfunctions for the boundary value problem (6.7), see Duistermaat [8] by way of Oudshoorn [35].

The integral operator. Let $K : L^2(0,1) \to H^m(0,1)$ be a linear integral operator, given by

$$(6.11)\qquad Kx(\,t\,) = \int_0^1 k(\,t\,,s)\, x(s)\, ds\;, \quad t \in [\,0\,,\,1\,]\;,$$

where k is m times differentiable with respect to its first argument, and $k^{(m)}$, the m-th derivative of k with respect to its first argument, is integrable in the sense that

$$(6.12)\qquad \sup_{t\in[\,0\,,\,1\,]} \int_0^1 \big\{\, |\, k^{(m)}(\,t\,,s)\,| + |\, k^{(m)}(s,\,t\,)\,|\,\big\}\, dt < \infty\;.$$

Then, K is indeed a bounded linear operator mapping $L^2(0,1)$ into $H^m(0,1)$.

Data perturbed by white noise. Let $x_o \in L^2(0,1)$. Consider the model

$$(6.13)\qquad y = K\, x_o + \eta\, W' \qquad \text{on} \quad [\,0\,,\,1\,]\;,$$

where $\eta > 0$, and $W'(t)$ is white noise, i.e., formally it is the derivative of a standard Wiener process $W(t)$, with $W(0) = 0$. The goal is to estimate x_o, and to obtain asymptotic bounds on the error in terms of η as $\eta \to 0$. Of course, white noise is not square integrable, formally

$$\| W' \| = +\infty \ , \tag{6.14}$$

so (6.13) does not make sense in $L^2(0,1)$. However, white noise may be integrated against functions in $H^1(0\,,1)$ say, so that the equation $K\,x = y$ makes sense in $H^{-1}(0\,,1)$, the dual of $H^1(0\,,1)$ with respect to the L^2 inner product. Note that the process

$$v(s) = [\,K^* dW\,](\,s\,) = \int_0^1 k(\,t\,, s)\, dW(\,t\,)\ , \quad s \in [\,0\,,\,1\,]\ , \tag{6.15}$$

is completely characterized by the property that for all $n \in \mathbb{N}$ and all $s_1\,, s_2\,, \cdots, s_n \in [\,0\,,1\,]$, the vector $\big(v(\,s_1), v(\,s_2)\,, \cdots, v(\,s_n)\big)$ is a multivariate normal random variable with mean 0 and variance V_n given by

$$[\,V_n\,]_{i,j} = \mathbb{E}[\,v(\,s_i)\,y(\,s_j)\,] = \ell(\,s\,,\,t\,)\ , \tag{6.16}$$

where $\ell(\,s\,,\,t\,) = \big\langle\, k(\,\cdot\,,\,s\,)\,, k(\,\cdot\,,\,t\,)\,\big\rangle$, $s\,,t \in [\,0\,,1\,]$, is the kernel of the integral operator K^*K, the *covariance operator* of the process (6.15). Now, let $\{\,\varpi_n\,, \kappa_n^2\,\}_{n=1}^{\infty}$ be the eigenfunction-eigenvalue expansion of the compact and Hermitian operator K^*K, then

$$v(s) = \sum_{n=1}^{\infty} \kappa_n\, W_n\, \varpi_n(\,s\,)\ , \quad s \in [\,0\,,\,1\,]\ , \tag{6.17}$$

where

$$W_n = \int_0^1 \varpi_n(\,t\,)\, dW(\,t\,)\ , \quad n = 1, 2, \cdots\ ,$$

are independent standard normal random variables. The infinite series in (6.17) converges in quadratic mean, i.e.,

$$\lim_{n\to\infty}\ \mathbb{E}\left[\ \Big\|\, v - \sum_{i=1}^{n} \kappa_i\, W_i\, \varpi_i\,\Big\|^2\,\right] = 0\ , \tag{6.18}$$

and $v \in L^2(0,1)$ with probability 1. Applying the above development to the Green's function operator $\mathfrak{R}_{1,\beta}$, we get, see (6.10),

$$[\,\mathfrak{R}_{1,\beta}\, dW\,](\,s\,) = Z_0 + 2\sum_{k=1}^{\infty} \big(1 + (\pi\beta k)^2\,\big)^{-1} Z_k\, \cos(\pi k\, s\,)\ , \tag{6.19}$$

for $s \in [0, 1]$, where $Z_0 = W(1) - W(0)$, and

$$Z_k = \int_0^1 u_k(t)\, dW(t)\ , \quad k = 1, 2, \cdots\ ,$$

Then, $Z_0, Z_1, Z_2, \cdots$ are independent standard normal random variables. One verifies that then

$$\| \mathfrak{R}_{m,\beta}\, dW \|_{m,\beta}^2 = \sum_{k=0}^{\infty} (1 + (\pi\beta k)^2)^{-1} Z_k^2\ . \tag{6.20}$$

Remark 9. The expansion (6.17) in terms of the eigenfunctions of the covariance operator goes under the name of Karhunen-Loève expansion. See, e.g., Papoulis [36] for the engineering point of view, and Loève [26] for the authorized probabilistic treatment.

Tikhonov regularization. After this set-up, we are ready: The goal is to recover x_o given the data y following the model (6.13), and of course, we wish to use Tikhonov regularization and the discrepancy principle. It is clear that a least-squares formulation as in (2.2), with $X = Y = L^2(0,1)$, will not work, since y is not square integrable. However, if y were to belong to $L^2(0,1)$, then

$$\begin{aligned} \| Kx - y \|^2 &= \| Kx \|^2 - 2 \langle Kx, y \rangle + \| y \|^2 \\ &= \| Kx \|^2 - 2 \langle x, K^* y \rangle + \| y \|^2\ . \end{aligned}$$

Observe that $K^* y$ belongs to $L^2(0,1)$ with probability 1, see (6.17), and then $\langle Kx, y \rangle = \langle x, K^* y \rangle$ is well defined. Now, the term $\| y \|^2$ is not defined if $y \notin L^2(0,1)$, but we do not need it: We may formulate the regularized least-squared problem as

$$\begin{array}{ll} \text{minimize} & \mathfrak{L}(x \,|\, y) \stackrel{\text{def}}{=} \| Kx \|^2 - 2 \langle x, K^* y \rangle + \alpha \| x \|^2 \\ \text{subject to} & x \in L^2(0,1)\ , \end{array} \tag{6.21}$$

One verifies that the solution is

$$x^{\alpha,\eta} = (K^*K + \alpha I)^{-1} K^* y\ . \tag{6.22}$$

To complete the set-up, we must address the source condition and the conditions for weakly bounded noise. The source conditions cause no problems: We assume (2.4) or the condition of Remark 3. Regarding the conditions (1.7)-(1.9) on the weakly bounded noise, their only use was in proving the bound of Theorem 2.6 on $\langle \delta, Kx \rangle$. Here, we prove the analogue of Theorem 2.6 for white noise.

LEMMA 21. *Under the assumptions* (6.11)-(6.12) *on the operator* K *, for all* β *,* $0 < \beta \leqslant \frac{1}{2}$ *.*

$$\left| \left\langle W' , Kx \right\rangle \right| \leqslant \beta^{-1/2}\, w_{\beta} \, \| \, Kx \, \|_{m,\beta} \ ,$$

where $\mathbb{E}[\, w_{\beta}^{2} \,] \leqslant c_1$ *, with* c_1 *as in* (6.4)*, with* $m = 1$ *.*

PROOF. Let $x \in L^2(0,1)$. Then, $Kx \in H^m(0,1)$, so certainly in $H^1(\,0\,,\,1\,)$, and by the reproducing kernel property (6.5),

$$[\,Kx\,](\,t\,) = \left\langle \mathfrak{R}_{1,\beta}(\,t\,,\,\cdot\,)\,,\,Kx \,\right\rangle_{1.\beta}\ , \qquad t \in [\,0\,,\,1\,]\ .$$

By Fubini's theorem, we then get

$$\left\langle \, W' \,,\, Kx \,\right\rangle = \int_0^1 \left\langle \, \mathfrak{R}_{1,\beta}(\,t\,,\,\cdot\,)\,,\,Kx \,\right\rangle_{1.\beta} dW(\,t\,) = \left\langle \, \mathfrak{R}_{1,\beta}\, dW\,,\,Kx \,\right\rangle_{1.\beta}$$

where

$$[\,\mathfrak{R}_{1,\beta}\, dW\,](\,s\,) = \int_0^1 \mathfrak{R}_{1,\beta}(\,t\,,s)\, dW(\,t\,)\ , \quad s \in [\,0\,,\,1\,]\ . \tag{6.23}$$

Since $\mathfrak{R}_{1,\beta}\, dW \in L^2(0,1)$ with probability 1, then by Cauchy-Schwarz,

$$\left| \left\langle \, W' \,,\, Kx \,\right\rangle \right| \leqslant \| \, \mathfrak{R}_{1,\beta}\, dW \, \|_{1,\beta} \ \| \, Kx \, \|_{1,\beta} \leqslant C \, \| \, \mathfrak{R}_{1,\beta}\, dW \, \|_{1,\beta} \ \| \, Kx \, \|_{m,\beta} \ .$$

Finally, following (6.20),

$$\mathbb{E}\big[\, \| \, \mathfrak{R}_{1,\beta}\, dW \, \|_{1,\beta}^{2} \,\big] = \sum_{k=0}^{\infty} (\, 1 + (\pi\beta k)^2 \,)^{-1} \leqslant c_1\, \beta^{-1} \ . \tag{6.24}$$

Thus, the choice $w_{\beta} = \beta^{1/2} \, \| \, \mathfrak{R}_{1,\beta}\, dW \, \|_{1,\beta}$ fits the bill. Q.e.d.

Convergence rates on $x^{\alpha,\eta} - x_o$ then follow as in the Main Theorem.

THEOREM 22. *Let* $m \geqslant 1$ *. Under the source condition* (2.4) *on* x_o *, and the conditions* (6.11)-(6.12) *on the operator* K*, the solution* $x^{\alpha,\eta}$ *of* (6.17) *under the model* (6.13) *satisfies*

$$\alpha \, \| \, x^{\alpha,\eta} - x_o \, \|^{\,2} \leqslant \eta^{\,2} \, \| \, \mathfrak{R}_{1,\beta}\, dW \, \|_{1,\beta}^{2} + c \, \alpha^{2\nu+1} \ ,$$

where $\beta \equiv \alpha^{1/2m}$ *, so that for deterministic* α *,*

$$\mathbb{E}\big[\, \| \, x^{\alpha,\eta} - x_o \, \|^{\,2} \,\big] = \mathcal{O}\big(\, \alpha^{-1-1/2m}\, \eta^{\,2} + \alpha^{2\nu} \,\big) \ .$$

For $\alpha \asymp \eta^{\,4m/(4m\nu+2m+1)}$ *, this gives*

$$\mathbb{E}\big[\, \| \, x^{\alpha,\eta} - x_o \, \|^{\,2} \,\big] = \mathcal{O}\big(\, \eta^{\,8m\nu/(4m\nu+2m+1)} \,\big) \ .$$

The random discrepancy principle. Things get considerably more complicated when choosing the regularization parameter by the discrepancy principle, because the resulting α is *random*. The (random) discrepancy principle is really as before: Choose α as the largest solution of

$$\alpha \, \| \, z^{\alpha,\eta} - x^{\alpha,\eta} \, \|^{\,2} = C \, \alpha^{-1/2m} \, \eta^2 \ , \tag{6.25}$$

for a large enough constant C. For emphasis, we denote the random solution as $\alpha = A$. Let $B = A^{1/2m}$. The basic problem is saying something about $\mathbb{E}\big[\, \| \, \mathfrak{R}_{1,B} \, dW \, \|_{1,B} \, \big]$, which obviously requires some knowledge about B.

The best would be to obtain precise bounds on $(\beta/B) - 1$ for a deterministic β, but this we are unable to do. It appears that the only alternative is to determine bounds on $\| \, \mathfrak{R}_{1,\beta} \, dW \, \|_{1,\beta}$, uniformly in β, and then substitute a random B for β. Below, we prove that certain random quantities are bounded in probability. The way we prove this is to show that the expectations of the squared quantities are bounded, and then use the Markov inequality: Recall that if Y is a random variable, then

$$\mathbb{P}[\, | \, Y \, | > \, t \,] \leqslant t^{-2} \, \mathbb{E}[\, Y^2 \,] \ ,$$

and that if $Y = Y(\gamma)$ then $Y = \mathcal{O}\big(\, \gamma \, \big)$ in probability as $\gamma \to 0$ means that

$$\lim_{K \to \infty} \ \lim_{\gamma \to 0} \ \mathbb{P} \, [\, | \, Y(\gamma) \, | \leqslant K \, \gamma \,] = 1 \ .$$

See, e.g., Breiman [4].

Lemma 23. *For* $0 < \beta \leqslant \gamma$, *we have* $\| \, \mathfrak{R}_{1,\gamma} \, dW \, \|_{1,\gamma} \leqslant \| \, \mathfrak{R}_{1,\beta} \, dW \, \|_{1,\beta}$.

Proof. Using the representation (6.20), a simple calculation shows that

$$\begin{aligned} \| \, \mathfrak{R}_{1,\gamma} \, dW \, \|_{1,\gamma}^{2} &= \sum_{k=0}^{\infty} \, (\, 1 + (\pi \gamma k)^2 \,)^{-1} \, Z_k^2 \\ &\leqslant \sum_{k=0}^{\infty} \, (\, 1 + (\pi \beta k)^2 \,)^{-1} \, Z_k^2 = \| \, \mathfrak{R}_{1,\beta} \, dW \, \|_{1,\beta}^{2} \ , \end{aligned}$$

since $\beta \leqslant \gamma$. Q.e.d.

Theorem 24. *There exists a constant* C *such that for* $\gamma \to 0$,

$$\sup_{0 < \beta \leqslant \gamma} \ \beta \, \| \, \mathfrak{R}_{1,\beta} \, dW \, \|_{1,\beta}^{2} \ \leqslant C + \mathcal{O}\big(\sqrt{\gamma} \, \big) \quad \textit{in probability} \ .$$

COROLLARY 25. *If the random variable B satisfies $\varepsilon \leqslant B \leqslant \gamma$ in probability, with $\varepsilon \to 0$ and $\gamma \to 0$, then*

$$\| \mathfrak{R}_{1,B}\, dW \|_{1,B} \leqslant C\, \varepsilon^{-1/2} + \mathcal{O}\big(\sqrt{\gamma\, \varepsilon^{-1}}\,\big) \qquad \textit{in probability}\ .$$

PROOF. Theorem 24 implies that

$$\| \mathfrak{R}_{1,B}\, dW \|_{1,B} = C\, B^{-1/2} + \mathcal{O}\big(B^{-1/2}\, \gamma^{1/2} \big)\ ,$$

in probability. Q.e.d.

PROOF OF THEOREM 24. Let $\beta_i = 2^{-i}\,\gamma$, $i = 0, 1, 2, \cdots$, and define

$$\mathcal{F}(\beta) = \| \mathfrak{R}_{1,\beta}\, dW \|_{1,\beta}^{2}\ .$$

The quantity of interest is $\sup_{0<\beta\leqslant\gamma}\ \beta\, F(\beta)$. Now,

$$\begin{aligned}
\sup_{0<\beta\leqslant\gamma}\ \beta\,\mathcal{F}(\beta) &\leqslant \sup_{1\leqslant i<\infty}\ \sup\big\{\, \beta\,\mathcal{F}(\beta)\ :\ \beta_i \leqslant \beta \leqslant \beta_{i-1} \,\big\} \\
&\leqslant \sup_{1\leqslant i<\infty}\ \beta_{i-1}\,\mathcal{F}(\beta_i) \qquad \text{(by Lemma 23)} \\
&\leqslant 2\ \sup_{1\leqslant i<\infty}\ \beta_i\,\mathcal{F}(\beta_i) \\
&\leqslant 2\,\Big(\sup_{1\leqslant i<\infty}\ \beta_i\,\mathbb{E}[\,\mathcal{F}(\beta_i)\,] + \sup_{1\leqslant i<\infty}\ \beta_i\,\big\{\, \mathcal{F}(\beta_i) - \mathbb{E}[\,\mathcal{F}(\beta_i)\,] \,\big\} \Big)\ .
\end{aligned}$$

By (6.15) we get $C \overset{\text{def}}{=} \sup_{0\leqslant\beta\leqslant1}\ \beta\,\mathbb{E}[\,\mathcal{F}(\beta)\,] < \infty$.

The other term requires more work. First,

$$\begin{aligned}
\mathbb{E}\Big[\sup_{1\leqslant i<\infty}\ \beta_i\,\big\{\, \mathcal{F}(\beta_i) - \mathbb{E}[\,\mathcal{F}(\beta_i)\,] \,\big\} \Big] &\leqslant \mathbb{E}\Big[\sum_{i=1}^{\infty}\ \beta_i\,\big|\, \mathcal{F}(\beta_i) - \mathbb{E}[\,\mathcal{F}(\beta_i)\,] \,\big| \Big] \\
&\leqslant \sum_{i=1}^{\infty}\ \beta_i\,\mathbb{E}\Big[\, \big|\, \mathcal{F}(\beta_i) - \mathbb{E}[\,\mathcal{F}(\beta_i)\,] \,\big| \,\Big] \\
&\leqslant \sum_{i=1}^{\infty}\ \beta_i\,\Big\{ \mathbb{E}\Big[\, \big|\, \mathcal{F}(\beta_i) - \mathbb{E}[\,\mathcal{F}(\beta_i)\,] \,\big|^2 \,\Big] \Big\}^{1/2}\ . \qquad (6.26)
\end{aligned}$$

Finally, using the representation (6.20),

$$\begin{aligned}
\mathbb{E}\Big[\, \big|\, \mathcal{F}(\beta) - \mathbb{E}[\,\mathcal{F}(\beta)\,] \,\big|^2 \,\Big] &= \mathbb{E}\Big[\, \Big|\, \sum_{k=0}^{\infty}\, (\, 1 + (\pi\beta k)^2\,)^{-1}\, (\, Z_k^2 - 1\,)\, \Big|^2\, \Big] \\
&= \sum_{k=0}^{\infty}\, (\, 1 + (\pi\beta k)^2\,)^{-2}\, \mathbb{E}[\, (\, Z_k^2 - 1\,)^2\,] \\
&= 2\, \sum_{k=0}^{\infty}\, (\, 1 + (\pi\beta k)^2\,)^{-2} \asymp \beta^{-1}\ ,
\end{aligned}$$

the last asymptotic equivalence by comparing the series with the corresponding integral. It follows that the expression in (6.26) is bounded by

$$2 \sum_{i=1}^{\infty} \beta_i^{1/2} = 2\,\gamma \sum_{i=1}^{\infty} 2^{-i/2} = 2\,\gamma\,/(\,\sqrt{2}-1\,)\ .$$

Thus,

$$\sup_{0<\beta\leqslant\gamma} \ \beta\,\mathcal{F}(\beta) \leqslant C + \mathcal{O}\big(\sqrt{\gamma}\,\big) \quad \text{in probability}\ ,$$

and the theorem follows. Q.e.d.

Now we are ready to say something about the random α chosen by the discrepancy principle.

LEMMA 26. *Under the source condition* (5.1) *for white noise, the random regularization parameter* $\alpha = A$ *, chosen according to the discrepancy principle* (6.25) *(with* C *large enough), satisfies*

$$A \asymp \eta^{4m/(6m+1)} \qquad \textit{in probability}\ .$$

PROOF. Again, it is obvious that $A \to 0$ in probability. So there is a deterministic γ, with $\gamma \to 0$, and $A \leqslant \gamma^{1/2m}$ in probability, in the sense that $\mathbb{P}[\,A \leqslant \gamma^{1/2m}\,] \longrightarrow 1$.

Let $\varepsilon = z^{\alpha,\eta} - x^{\alpha,\eta} - (z^{\alpha,o} - x^{\alpha,o})$. Then, $\varepsilon = \alpha\,\eta\,(\,K^*K + \alpha\,I\,)^{-2}\,K^* dW$, and so, since $\alpha\,\|\,(\,K^*K + \alpha\,I\,)^{-1}\,\| \leqslant 1$,

$$\|\,\varepsilon\,\| \leqslant \eta\,\|\,(\,K^*K + \alpha\,I\,)^{-1}\,K^* dW\,\| = \|\,x^{\alpha,\eta} - x^{\alpha,o}\,\|\ .$$

From Theorem 22, it follows with $\beta \equiv \alpha^{1/2m}$, that

$$\begin{aligned}
\alpha^{1/2}\,\|\,x^{\alpha,\eta} - x^{\alpha,o}\,\| &\leqslant c_m\,\eta\,\|\,\mathfrak{R}_{1,\beta} dW\,\|_{1,\beta} \\
&\leqslant c_m\,\alpha^{-1/4m}\,\eta\,\sup_{\beta\leqslant\gamma}\ \beta^{1/2}\,\|\,\mathfrak{R}_{1,\beta} dW\,\|_{1,\beta} \\
&\leqslant c_m\,\alpha^{-1/4m}\,\eta\,(\,C + \mathcal{O}\big(\,\gamma^{1/2}\,\big)\,)
\end{aligned}$$

in probability. Since the random discrepancy principle (6.25) implies that

$$\alpha^{1/2}\,\|\,z^{\alpha,o} - x^{\alpha,o}\,\| - \alpha^{1/2}\|\,\varepsilon\,\| \leqslant C\,\alpha^{-1/4m}\,\eta \leqslant \alpha^{1/2}\,\|\,z^{\alpha,o} - x^{\alpha,o}\,\| + \alpha^{1/2}\|\,\varepsilon\,\|\ ,$$

we obtain that for $\alpha = A$,

$$\alpha^{1/2}\,\|\,z^{\alpha,o} - x^{\alpha,o}\,\| \asymp \alpha^{-1/4m}\,\eta$$

in probability. Together with Lemma 14, the lemma follows (if C is large enough). Q.e.d.

THEOREM 27. (RATES FOR THE RANDOM DISCREPANCY PRINCIPLE) *Under the source condition* (5.1), *and with* $\alpha = A$ *determined by the random discrepancy principle* (6.25), *there exists a constant* C_1 *such that*

$$\| x^{\alpha,\eta} - x_o \| \leqslant C_1 \, \eta^{2m/(6m+1)} \big(1 + o(1) \big) \qquad \textit{in probability} .$$

PROOF. Follows from Lemmas 14 and 26. Q.e.d.

7. Discrete noise

Finally, we briefly discuss the case of moment problems for Fredholm integral equations with discrete noise. We show that the analysis proceeds almost exactly as in the white noise case.

The model under consideration is

$$y_{in} = [\, K x_o](\, t_{in}) + d_{in} \ , \quad i = 1, 2, \cdots , n \ , \tag{7.1}$$

where $0 \leqslant t_1 < t_2 < \cdots < t_n \leqslant 1$ are more or less uniformly distributed over $(\, 0 \, , \, 1 \,)$, e.g.,

$$t_{in} = \frac{i-1}{n-1} \ , \quad i = 1, 2, \cdots , n \ . \tag{7.2}$$

In fact, all that is required is that there exists a constant c such that for all u with integrable derivative,

$$\Big| \tfrac{1}{n} \sum_{i=1}^{n} u(\, t_{in}) - \int_0^1 u(\, t \,) \, dt \, \Big| \leqslant c \, n^{-1} \, \| \, u' \, \|_{L^1(\, 0 \, , \, 1 \,)} \ . \tag{7.3}$$

It is straightforward to show that (7.3) holds for the choice (7.2). Regarding the noise, we assume that

$$\begin{array}{l} d_{1,n}, d_{2,n}, \cdots , d_{n,n} \quad \text{are uncorrelated random variables, with} \\ \mathbb{E}[\, d_{in} \,] = 0 \quad , \quad \mathbb{E}[\, d_{in}^2 \,] = \sigma^2 \quad , \quad \mathbb{E}[\, d_{in} \, d_{jn} \,] = 0 \quad \text{for} \quad i \neq j \ . \end{array} \tag{7.4}$$

Now, Tikhonov regularization takes the least-squares form,

$$\begin{array}{ll} \text{minimize} & \tfrac{1}{n} \sum_{i=1}^{n} \big| \, [\, K x \,](\, t_{in}) - y_{in} \, \big|^2 + \alpha \, \| \, x \, \|^2 \\ \text{subject to} & x \in L^2(0,1) \ , \end{array}$$

and it seems we have to start all over again. However, note that

$$\tfrac{1}{n}\sum_{i=1}^{n} \bigl|\,[\,Kx\,](\,t_{in}) - y_{in}\,\bigr|^2 = \tfrac{1}{n}\sum_{i=1}^{n} \bigl|\,[\,K(x-x_o)\,](\,t_{in})\,\bigr|^2 - \tfrac{2}{n}\sum_{i=1}^{n} d_{in}\,[\,K(x-x_o)\,](\,t_{in}\,)\ ,$$

apart from the sum of squared d_{in}. Thus, the above least-squares problem is equivalent to

$$\text{(7.5)} \qquad \text{minimize} \quad \mathfrak{D}_n(\,x\,|\,d_n\,) \quad \text{subject to} \quad x \in L^2(0,1)\ ,$$

where $d_n = (\,d_{1,n}, d_{2,n}, \cdots, d_{n,n}\,)^{\mathrm{T}}$, and $\mathfrak{D}_n$ is the "discrete" functional,

$$\text{(7.6)} \quad \mathfrak{D}_n(\,x\,|\,d_n\,) \stackrel{\text{def}}{=} \tfrac{1}{n}\sum_{i=1}^{n} \bigl|\,[\,K(x-x_o)\,](\,t_{in})\,\bigr|^2 + \alpha\,\|\,x\,\|^2 - \tfrac{2}{n}\sum_{i=1}^{n} d_{in}\,[\,K(x-x_o)\,](\,t_{in}\,)\ .$$

Now, the solution of (7.5) turns out to be remarkably close to the solution of the *partially* continuous version of (7.5),

$$\text{(7.7)} \qquad \text{minimize} \quad \mathfrak{P}_n(\,x\,|\,d_n\,) \quad \text{subject to} \quad x \in L^2(0,1)\ ,$$

with

$$\text{(7.8)} \quad \mathfrak{P}_n(\,x\,|\,d_n\,) \stackrel{\text{def}}{=} \|\,K(x-x_o)\,\|^2 + \alpha\,\|\,x\,\|^2 - \tfrac{2}{n}\sum_{i=1}^{n} d_{in}\,[\,K(x-x_o)\,](\,t_{in}\,)\ .$$

The next claim is that (7.7) is a slightly different reincarnation of the white noise case of (6.21). This may be seen as follows. First, we may write

$$\text{(7.9)} \qquad \tfrac{1}{n}\sum_{i=1}^{n} d_{in}\,[\,K(x-x_o)\,](\,t_{in}\,) = \bigl\langle\, x - x_o\,,\, K^* dW_n\,\bigr\rangle\ ,$$

where W_n is a step function with jumps at the t_{in} of size d_{in}, cf. (6.23),

$$\text{(7.10)} \qquad W_n(\,s\,) = \sum_{t_{in} \leqslant s} d_{in}\ .$$

So, we may rewrite $\mathfrak{P}_n$ as

$$(7.11)\qquad \mathfrak{P}_n(\,x\,|\,d_n\,) = \|\,K(x-x_o)\,\|^{\,2} + \alpha\,\|\,x\,\|^{\,2} - 2\,\big\langle\, x-x_o\,,\,K^*dW_n\,\big\rangle\ .$$

Now, we rewrite the white noise functional $\mathfrak{L}$ of (6.21). Using

$$\|\,Kx\,\|^{\,2} = \|\,K(x-x_o)\,\|^{\,2} + 2\,\big\langle\, x-x_o\,,\,K^*Kx_o\,\big\rangle + \|\,Kx_o\,\|^{\,2}$$

(and ignoring the constant term $\|\,Kx_o\,\|^{\,2}$), and using

$$\big\langle\, x\,,\,K^*y\,\big\rangle = \big\langle\, x-x_o\,,\,K^*y\,\big\rangle + \big\langle\, x_o\,,\,K^*y\,\big\rangle$$

(and ignoring the constant term $\big\langle\, x_o\,,\,K^*y\,\big\rangle$), we may rewrite $\mathfrak{L}$ as

$$(7.12)\qquad \mathfrak{L}(\,x\,|\,y\,) = \|\,K(x-x_o)\,\|^{\,2} + \alpha\,\|\,x\,\|^{\,2} - 2\,\eta\,\big\langle\, x-x_o\,,\,K^*dW\,\big\rangle\ .$$

So, the difference between the discrete noise version (7.5) and the white noise case (6.21) is that we have replaced the white noise process $\eta\,dW(s)$ by $dW_n(s)$.

Before showing the closeness of the solutions of (7.5) and (7.7), we give error bounds on the solution of (7.7). Comparison with the white noise case of Theorem 22 shows that W_n behaves as $\eta\,W$ with $\eta \approx \sigma\,n^{-1/2}$.

Theorem 28. *The solution* $u^{\alpha,n}$ *of* (7.7) *exists and is unique, and under the source condition* (2.4) *and the assumptions* (7.3)-(7.4) *satisfies*

$$\mathbb{E}[\,\|\,u^{\alpha,n}-x_o\,\|^{\,2}\,] = \mathcal{O}\big(\,\alpha^{-1}\,(\alpha^{1/2m}\,n)^{-1} + \alpha^{2\nu}\,\big)$$

provided $\alpha \to 0$ *and* $\alpha^{1/2m}\,n \to \infty$. *For* $\alpha \asymp n^{-2m/(4m\nu+2m+1)}$, *this gives*

$$\mathbb{E}[\,\|\,u^{\alpha,n}-x_o\,\|^{\,2}\,] = \mathcal{O}\big(\,n^{-4m\nu/(4m\nu+2m+1)}\,\big)\ .$$

Proof. Let $\varepsilon = u^{\alpha,n} - x_o$. With the interpretation (7.11), it follows as in Theorem 22 with $\eta\,dW$ replaced by dW_n that

$$\alpha\,\|\,\varepsilon\,\|^{\,2} \leqslant \|\,\mathfrak{R}_{m,\beta}\,dW_n\,\|_{m,\beta}^{\,2} + c\,\alpha^{2\nu+1}\ ,$$

where $\beta \equiv \alpha^{1/2m}$. The required bound then follows from the next lemma. Q.e.d.

Lemma 29. *Under the assumption* (7.4),

$$\mathbb{E}\big[\,\|\,\mathfrak{R}_{m,\beta}\,dW_n\,\|^2_{m,\beta}\,\big] = \mathcal{O}\big(\,(n\beta)^{-1}\,\big)\ .$$

Proof. One computes

$$\begin{aligned}\mathbb{E}\big[\,\|\,\mathfrak{R}_{m,\beta}\,dW_n\,\|^2_{m,\beta}\,\big] &= \mathbb{E}\Big[\,\|\,\tfrac{1}{n}\sum_{i=1}^{n} d_{in}\,\mathfrak{R}_{m,\beta}(\,t_{in},\,\cdot\,)\,\|^2_{m.\beta}\,\Big]\\ &= \mathbb{E}\Big[\,n^{-2}\sum_{i,j=1}^{n} d_{in}\,d_{jn}\,\big\langle\,\mathfrak{R}_{m,\beta}(\,t_{in}\,,\,\cdot\,)\,,\,\mathfrak{R}_{m,\beta}(\,t_{jn}\,,\,\cdot\,)\,\big\rangle_{m,\beta}\Big]\\ &= \sigma^2\,n^{-2}\sum_{i=1}^{n}\|\,\mathfrak{R}_{m,\beta}(\,t_{in},\,\cdot\,)\,\|^2_{m,\beta} = \mathcal{O}\big(\,(n\beta)^{-1}\,\big)\ ,\end{aligned}$$

where we used (7.4) and (6.6). Q.e.d.

We now show that the error due to the replacement of (7.5) by (7.7) is negligible. Roughly speaking, the next theorem says that

$$\alpha^{1/2}\,\|\,x^{\alpha,n} - u^{\alpha,n}\,\| \approx \big(\,\alpha^{1/2}\,\|\,u^{\alpha,n} - x_o\,\|\,\big)^2\ ,$$

and one may draw one's conclusions about $\alpha^{1/2}\,\|\,x^{\alpha,n} - x_o\,\|$.

Theorem 30. *The solution* $x^{\alpha,n}$ *of* (7.5) *and the solution* $u^{\alpha,n}$ *of* (7.7) *satisfy*

$$\mathbb{E}[\,\|\,x^{\alpha,n} - u^{\alpha,n}\,\|^2\,] = o\big(\,\alpha^{-1}\,(\,\alpha^{1/2m}\,n\,)^{-2}\,\big)\ ,$$

provided $\alpha \to 0$ *and* $\alpha^{1/2m}\,n \to \infty$.

Proof. Let $\varepsilon = x^{\alpha,n} - u^{\alpha,n}$. One shows analogous to the proof of Theorem 2.4 that

$$\|\,K\varepsilon\,\|^2 + \alpha\,\|\,\varepsilon\,\|^2 \leqslant \mathfrak{P}_n(\,x^{\alpha,n}\,|\,d_n\,) - \mathfrak{P}_n(\,u^{\alpha,n}\,|\,d_n\,)\ , \qquad \text{and}$$

$$\tfrac{1}{n}\sum_{i=1}^{n}|\,[\,K\varepsilon\,](\,t_{in}\,)\,|^2 + \alpha\,\|\,\varepsilon\,\|^2 \leqslant \mathfrak{D}_n(\,u^{\alpha,n}\,|\,d_n\,) - \mathfrak{D}_n(\,x^{\alpha,n}\,|\,d_n\,)\ .$$

Adding these two inequalities yields

$$\|\,K\varepsilon\,\|^2 + \tfrac{1}{n}\sum_{i=1}^{n}|\,\varepsilon(\,t_{in})\,|^2 + 2\,\alpha\,\|\,\varepsilon\,\|^2 \leqslant RHS\ , \tag{7.13}$$

with

$$RHS = \| K(x^{\alpha,n} - x_o) \|^2 - \| K(u^{\alpha,n} - x_o) \|^2 + \frac{1}{n} \sum_{i=1}^{n} \{ | [K(u^{\alpha,n} - x_o)](t_{in}) |^2 - | [K(x^{\alpha,n} - x_o)](t_{in}) |^2 \} .$$

The right hand side RHS is actually a quadrature error, the function under integration being $f = | K(x^{\alpha,n} - x_o) |^2 - | K(u^{\alpha,n} - x_o) |^2$. Thus, by (7.3),

$$RHS \leqslant c\, n^{-1} \| f' \|_{L^1(0,1)} . \tag{7.14}$$

Now write f as

$$f = K(x^{\alpha,n} + u^{\alpha,n} - 2\, x_o)\; K\varepsilon = K\varphi\, K\varepsilon ,$$

with $\varphi = \varepsilon + 2\,(u^{\alpha,n} - x_o)$. Then, we get with Cauchy-Schwarz, for all $0 < \beta \leqslant \frac{1}{2}$,

$$\begin{aligned} \| f' \|_{L^1(0,1)} &\leqslant \| (K\varphi)' \| \, \| K\varepsilon \| + \| K\varphi \| \, \| (K\varepsilon)' \| \\ &\leqslant \beta^{-1} \{ \| K\varphi \| + \beta \| (K\varphi)' \| \} \{ \| K\varepsilon \| + \beta \| (K\varepsilon)' \| \} \\ &\leqslant 2\, \beta^{-1} \| K\varphi \|_{1,\beta} \| K\varepsilon \|_{1,\beta} \leqslant c\, \beta^{-1} \| K\varphi \|_{m,\beta} \| K\varepsilon \|_{m,\beta} , \end{aligned}$$

for another suitable constant c, the last inequality by (6.4). Let $\| \cdot \|_{\alpha,X}$ be as in (2.3) with $X = L^2(0,1)$. Setting $\beta = \alpha^{1/2m}$, the above shows that for yet another constant c,

$$\| f' \|_{L^1(0,1)} \leqslant c\, \alpha^{1/2m} \| \varphi \|_{\alpha,X} \| \varepsilon \|_{\alpha,X} ,$$

where we used that $K : L^2(0,1) \longrightarrow H^m(0,1)$ is continuous. Now,

$$\| \varphi \|_{\alpha,X} \leqslant \| \varepsilon \|_{\alpha,X} + 2\, \| u^{\alpha,n} - x_o \|_{\alpha,X} ,$$

and it follows that

$$\| f' \|_{L^1(0,1)} \leqslant c\, \alpha^{-1/2m} \{ \| \varepsilon \|^2_{\alpha,X} + 2\, \| \varepsilon \|_{\alpha,X} \| u^{\alpha,n} - x_o \|_{\alpha,X} \} .$$

Substituting *this* into (7.14) and substituting *that* into (7.13) gives

$$\| \varepsilon \|^2_{\alpha,X} \leqslant C\, (\alpha^{1/2m}\, n)^{-1} \| \varepsilon \|^2_{\alpha,X} + 2\, C\, (\alpha^{1/2m}\, n)^{-1} \| \varepsilon \|_{\alpha,X} \| u^{\alpha,n} - x_o \|_{\alpha,X} ,$$

or

$$\big(1 - C\,(\alpha^{1/2m}\, n)^{-1} \big)\, \|\, \varepsilon \,\|_{\alpha,X}^2 \leqslant 2\, C\,(\alpha^{1/2m}\, n)^{-1}\, \|\, \varepsilon \,\|_{\alpha,X} \|\, u^{\alpha,n} - x_o \,\|_{\alpha,X} \;.$$

If $\alpha^{1/2m}\, n \to \infty$, then the factor on the left is (asymptotically) equal to 1. So then,

$$\|\, \varepsilon \,\|_{\alpha,X} \leqslant C_1\, (\, \alpha^{1/2m}\, n)^{-1}\, \|\, u^{\alpha,n} - x_o \,\|_{\alpha,X} \;,$$

from which one obtains

$$\begin{aligned} \mathbb{E}[\, \|\, \varepsilon \,\|_{\alpha,X}^2 \,] &\leqslant C_1^2\, (\, \alpha^{1/2m}\, n)^{-2}\, \mathbb{E}[\, \|\, u^{\alpha,n} - x_o \,\|_{\alpha,X}^2 \,] \\ &= o\big((\, \alpha^{1/2m}\, n)^{-2} \,\big) \;, \end{aligned}$$

where we used a weak form of Theorem 30. The lemma follows. Q.e.d.

References

[1] Adams, R.A., Fournier, J.J.F.; *Sobolev spaces*, Second edition. Amsterdam: Academic Press (2003).

[2] Bissantz, N.; Hohage, T.; Munk, A.; *Consistency and rates of convergence of nonlinear Tikhonov regularization with random noise.* Inverse Problems **20** (2004) 1773–1789.

[4] Breiman, L.; *Probability*, Reading, Mass.: Addison-Wesley (1968). Reprinted, SIAM (1992).

[3] Brown, L.D.; Low, M,; *Asymptotic equivalence of nonparametric regression and white noise,* Ann. Statist. **24** (1996) 2384–2398.

[5] Burger, M.; Osher, S.; *Convergence rates for convex variational regularization.* Inverse problems **20** (2004) 1411–1421.

[6] Cavalier, L.; Golubev, G.K.; Picard, D.; Tsybakov, A.B.; *Oracle inequalities for inverse problems.* Ann. Statist. **30**, (2002) 843–874.

[7] Duan, H.; Ng, B.P.; See, C.M.S.; Fang, J.; *Broadband interference suppression performance of minimum redundancy arrays.* IEEE Trans. Signal Process. **56** (2008) 2406–2416.

[8] Duistermaat, J.J.; *The Sturm-Liouville problem for the operator* $(-d^2/dx^2)^m$ *with Neumann or Dirichlet boundary conditions.* Technical Report 899, Department of Mathematics. Utrecht University (1995).

[9] Egger, H.; *Semi-iterative regularization in Hilbert scales.* SIAM J. Numer. Anal. **44** (2006) 66–81.

[10] Eggermont, P.P.B.; LaRiccia, V.N.; *Equivalent kernels for smoothing splines*, J. Integral Equations Appl. **18**, (2006) 97–225.

[11] Eicke, B.; Louis, A.K.; Plato, R.; *The instability of some gradient methods for ill-posed problems.* Numer. Math. **58** (1990) 129–134.

[12] Engl, H.W.; Hanke, M.; Neubauer, A.; *Regularization of inverse problems.* Dordrecht: Kluwer (1996).

[13] Franklin, J.N.; *Well-posed stochastic extensions of ill-posed linear problems.* J. Math. Anal. Appl. **31** (1970) 682–716.

[14] Frommer, A.; Maass, P.; *Fast CG-based methods for Tikhonov-Phillips regularization.* SIAM J. Sci. Comput. **20**, (1999) 1831–1850.

[15] Gfrerer, H.; *An a posteriori parameter choice for ordinary and iterated Tikhonov regularization of ill-posed problems leading to optimal convergence rates.* Math. Comp. **49** (1987) 523–542.

[16] Gfrerer, H.; *Supplement to An a posteriori parameter choice for ordinary and iterated Tikhonov regularization of ill-posed problems leading to optimal convergence rates.* Math. Comp. **49** (1987) S5–S12.

[17] Groetsch, C.W.; *The theory of Tikhonov regularization for Fredholm equations of the first kind.* Boston: Pitman (1984).

[18] Hadamard, J.; *Sur les problèmes aux derivées partièlles et leur signification physique.* Princeton University Bulletin (1902) 49–52.

[19] Hanke, M.; *Accelerated Landweber iterations for the solution of ill-posed problems.* Numer. Math. **60**, (1991) 341–373.

[20] Hegland, M.; *Variable Hilbert scales and their interpolation inequalities with applications to Tikhonov regularization.* Appl. Anal. **59**, 207–223 (1995).

[21] Hohage, T.; Pricop, M.; *Nonlinear Tikhonov regularization in Hilbert scales for inverse boundary value problems with random noise.* Inverse Problems and Imaging **2** (2008) 271–290.

[22] Kaipio, J.; Somersalo, E.; *Statistical and computational inverse problems.* Berlin: Springer (2005).

[23] Klann, E.; Maass, P.; Ramlau, R.; *Two-step regularization methods for linear inverse problems.* J. Inverse Ill-Posed Probl. **14** (2006) 583–607.

[24] Krein, S.G.; Petunin, Yu.I.; *Scales of Banach spaces.* Russian Math. Surveys

21, No. 2 (1966) 85–159.

[25] Liu, F.; Nashed, M.Z.; *Tikhonov regularization of nonlinear ill-posed problems with closed operators in Hilbert scales.* J. Inverse Ill-Posed Probl. **5** (1997), 363–376

[26] Loève, M.; *Probability theory.* New York: Van Nostrand (1955)

[27] Mair, B.A.; Ruymgaart, F.H.; *Statistical inverse estimation in Hilbert scales.* SIAM J. Appl. Math. **56** (1996) 1424–1444.

[28] Mathé, P.; Pereverzev, S.V.; *Geometry of linear ill-posed problems in variable Hilbert scales.* Inverse Problems **19** (2003) 789–803.

[29] Mathé, P.; Pereverzev, S.V.; *The discretized discrepancy principle under general source conditions.* J. Complexity **22** (2006) 371–381.

[30] Mathé, P.; Pereverzev, S.V.; *Regularization of some linear ill-posed problems with discretized random noisy data.* Math. Comp. **75** (2006) 1913–1929.

[31] Morozov, V.A.; *Methods for solving incorrectly posed problems*, New York: Springer-Verlag (1984).

[32] Natterer, F.; *Error bounds for Tikhonov regularization in Hilbert scales*, Appl. Anal. **18** (1984) 29–37.

[33] Neubauer, A.; *When do Sobolev spaces form a Hilbert scale?* Proc. Amer. Math. Soc. **103** (1988) 557–562.

[34] Neubauer, A. *On converse and saturation results for Tikhonov regularization of linear ill-posed problems.* SIAM J. Numer. Anal. **34** (1997) 517–527.

[35] Oudshoorn, C.G.M; *Asymptotically minimax estimation of a function with jumps.* Bernoulli **4** (1998) 15–33.

[36] Papoulis, A.; *Probability, random variables and stochastic processes.* New York: McGraw-Hill (1965).

[37] Phillips, B.L.; *A technique for the numerical solution of certain integral equations of the first kind.* J. Assoc. Comput. Mach. **9** (1962) 84–97.

[38] Ribière, G.; *Régularisation d'opérateurs.* Rev. Informat. Recherche Opérationnelle **1** (1967) 57–79.

[39] Scheres, S.H.W.; Núñez-Ramírez, R.; Gómez-Llorente, Y.; San Martín, C.; Eggermont, P.P.B.; Carazo, J.M.; *Modeling experimental image formation for maximum-likelihood classification of electron microscopy data.* Structure **15** (2007) 1167–1177.

[40] Sudakov, V.N.; Khalfin, L.A.; *A statistical approach to the correctness of the problems of –athematical physics.* Dokl. Akad. Nauk SSSR **157** (1964) 1058–1060.

[41] Tikhonov, A.N.; *On the stability of inverse problems.* Dokl. Akad. Nauk. SSSR **39** (1943) 195–198. (in Russian)

[42] Tikhonov, A.N.; *On the solution of incorrectly formulated problems and the regularization method.* Dokl. Akad. Nauk. SSSR **151** (1963) 501–504.

[43] Tikhonov, A.N.; Goncharsky, A.V.; Stepanov, V.V.; Yagola, A.G.; *Numerical methods for the solution of ill-posed problems.* Dordrecht: Kluwer (1995).

[44] Twomey, S.; *On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature.* J. Assoc. Comput. Mach. **10** (1963) 97–101.

Food and Resource Economics
University of Delaware
Newark, DE 19717
{eggermon,lariccia}@udel.edu

Department of Mathematics
University of Central Florida
Orlando, FL 32143
znashed@mail.ucf.edu