

# REGULARIZATION BY PROJECTION IN VARIABLE HILBERT SCALES

PETER MATHÉ AND NADINE SCHÖNE

ABSTRACT. We introduce and analyze projection schemes for solving ill-posed linear operator equations in Hilbert space. Control parameters are shown to yield convergence. Emphasis is on their approximation theoretic content. For smoothness given in terms of general source conditions we establish convergence rates, shown to be optimal in some cases. The study is accomplished with a discussion on condition numbers.

## 1. INTRODUCTION

We shall study the approximate solution of linear ill-posed operator equations in Hilbert space, that is, we are given a linear operator equation

$$(1) \quad y^\delta = Ax^\dagger + \delta\xi$$

where the operator  $A$  maps from some Hilbert space  $X$  *injectively* into some Hilbert space  $Y$ . As usual in regularization we assume that instead of exact data  $y = Ax^\dagger \in Y$  we are given perturbed data  $y^\delta \in Y$ , and the goal is to reconstruct the unknown solution  $x^\dagger$  from these data by projection methods. Within the present context we assume that the noise is bounded and deterministic, hence  $\|\xi\| \leq 1$ , and  $\delta$  is the (known) noise level.

There is vast literature on the stable recovery of  $x^\dagger$  from noisy data  $y^\delta$ , we mention only the monograph [3]. The mathematical analysis of projection schemes started by Natterer in [19]. The authors in [22] also give a detailed account on projection methods. Several quantities similar to the ones occurring here are used in their analysis. However, their approximation theoretic meaning is left unclear.

---

*Date:* Version: December 1, 2007.

*2000 Mathematics Subject Classification.* 47A52, secondary: 65J20.

*Key words and phrases.* regularization, ill-posed problems, projection method, general source condition, condition number.

Here we consider *two-sided discretization* in the form  $Q_n A P_m$ , for projections  $Q_n$  and  $P_m$  in  $Y$  and  $X$ , respectively. The approximate solution is then obtained from a (best approximate) solution to a related system of linear equations. The problem of convergence of projection schemes can be formulated as follows:

- (i) Can we assign the discretization levels  $m = m(\delta)$  and  $n = n(\delta)$  in such a way, that the corresponding approximate solutions converge to the true underlying solution as  $\delta \rightarrow 0$ ?
- (ii) Can we guarantee convergence rates if the true solution  $x^\dagger$  has certain smoothness properties? Which properties of the chosen discretization are responsible?
- (iii) How do the condition numbers of the related systems of linear equations depend on the noise level  $\delta$ ?

We shall give answers to these questions for ill-posed problems in Hilbert spaces, driven by a *compact* operator  $A$ . This restriction allows us to control the convergence behavior of the projection schemes by some parameters, the behavior of which is known in many cases, see § 3. Moreover, in § 4 we analyze the convergence of projection schemes when smoothness is given in terms of *general source conditions*, a subject which became attractive recently, see [14, 15, 18]. For power type source conditions and one-sided discretization we refer to [8]. Here we establish both, an *a priori* parameter choice which provides order optimal regularization for certain smoothness classes, and an *a posteriori* parameter choice. We conclude our study with an analysis of the condition numbers of the linear systems related to the projection schemes in § 5. This analysis complements initial study from [4]. We establish the order of the condition numbers in terms of the noise level  $\delta$ .

## 2. REGULARIZATION BY PROJECTION

We first recall some results from approximation theory of compact operators in Hilbert spaces. Our notation follows [20, Chapt. 11]. If  $A: X \rightarrow Y$ , acting between Hilbert spaces  $X$  and  $Y$  is compact and injective then it admits as (monotonic) Schmidt-representation (singular value decomposition), say

$$(2) \quad Ax = \sum_{j=1}^{\infty} s_j \langle x, v_j \rangle u_j, \quad x \in X,$$

for some orthonormal systems  $\{v_i\}_{i=1}^{\infty} \subset X$  and  $\{u_j\}_{j=1}^{\infty} \subset X$  and a non-increasing sequence  $s_1 \geq s_2 \geq \dots > 0$ , converging to zero,

of singular values of the non-negative self-adjoint mapping  $(A^*A)^{1/2}$ . These singular values coincide with the *approximation numbers*, i.e.,

$$(3) \quad s_j = a_j(A) := \inf \{ \|A - L\|, \quad \text{rank}(L) < j \}, \quad j = 1, 2, \dots$$

In Hilbert space these approximation numbers coincide with the Gelfand- and Kolmogorov numbers, it thus holds true that

$$\begin{aligned} s_j &= a_j(A) = \inf \{ \|A(I - L)\|, \quad \text{rank}(L) < j \} \\ &= \inf \{ \|(I - L')A\|, \quad \text{rank}(L') < j \}, \end{aligned}$$

where the inf is taken over all orthogonal projections  $L$  and  $L'$  in  $X$  and  $Y$ , respectively. In particular we observe that  $s_1 = \|A\|$ .

Projection methods are obtained from specific discretization, and we assume that instead of  $y^\delta \in Y$  we are given a finite collection

$$(4) \quad y_j^\delta = \langle Ax^\dagger, \varphi_j \rangle + \delta \langle \xi, \varphi_j \rangle, \quad j = 1, \dots, n,$$

associated to some system  $\{\varphi_1, \dots, \varphi_n\} \subset Y$ . We shall denote the corresponding projections onto the space  $Y_n := \text{span}(\varphi_1, \dots, \varphi_n) \subset Y$  by  $Q_n$ . We may rewrite the above equation (4) as  $Q_n y^\delta = Q_n A x^\dagger + \delta Q_n \xi$ . In addition to this discretization from the left we may furthermore want the reconstructed solution to be represented through some (finite) system  $\{\psi_1, \dots, \psi_m\} \subset X$ , and we associate the projection  $P_m$  related to the corresponding subspace  $X_m := \text{span}(\psi_1, \dots, \psi_m) \subset X$ . We shall call the chosen systems *designs*, as these reflect the way the data are obtained and the form the approximate solution is sought for. We also assume without further mentioning, that  $\dim(X_m) = m$  and that the spaces  $X_m$  and  $Y_n$  are nested.

*Remark 1.* Of special importance are designs which correspond to the singular system in (2), i.e., when  $\psi_i := v_i$ ,  $i = 1, 2, \dots$  and/or  $\varphi_j := u_j$ ,  $j = 1, 2, \dots$ . Projection schemes with such designs result in specific *spectral cut-off*, a fact which will not be discussed here.

Having chosen, both the projections  $Q_n$  and  $P_m$  we obtain the approximate solution as (the minimal norm) solution, say  $x_{m,n}^\delta$  to the algebraic system corresponding to the equation

$$(5) \quad Q_n y^\delta = Q_n A P_m x_{m,n}^\delta.$$

Using the *Moore-Penrose inverse* of the mapping  $B_{m,n} := Q_n A P_m$ , see [3, Chapt. 2.1] for details, and we are thus led to the following

**Definition 1.** Given orthogonal projections  $Q_n$  in  $Y$  and  $P_m$  in  $X$  we call the Moore-Penrose inverse  $B_{m,n}^+$  a corresponding *projection method*.

The approximate solution to (1) obtained from  $B_{m,n}^+$  is

$$(6) \quad x_{m,n}^\delta := B_{m,n}^+ y^\delta \quad (= B_{m,n}^+ Q_n y^\delta).$$

The determination of  $x_{m,n}^\delta$  from the data  $Q_n y^\delta$  as in (5) leads to solving a system of linear equations. We recall from [3, Prop. 2.3] that  $B_{m,n}^+ B_{m,n} = P_{\ker^\perp(B_{m,n})}$ . Since  $\ker^\perp(B_{m,n}) \subset X_m$ , it will be important to understand when there is equality. This problem was also considered in [22] and we state

**Lemma 1.** *System (5) has a unique solution if and only if  $X_m = \ker^\perp(B_{m,n})$ .*

*Proof.* It is always true that  $\ker^\perp(B_{m,n}) \subset X_m$ . Otherwise suppose that there is  $x \in X_m$  with a decomposition  $x = x_0 + x_1$  with  $x_0 \in \ker^\perp(B_{m,n})$  and  $0 \neq x_1 \in \ker(B_{m,n})$ . Thus  $B_{m,n} x_1 = 0$ , and unique solvability is violated.  $\square$

Therefore we shall assume that this is the case, at least for  $n = n(m)$  large enough.

**Property (S)** (Unique solvability). There are  $m_0 \in \mathbb{N}$  and a mapping  $n := n(m)$ ,  $m = m_0, m_0 + 1, \dots$  such that  $X_m^\perp = \ker(B_{m,n})$ ,  $m \geq m_0$ .

*Remark 2.* Under Property (S)  $B_{m,n}|_{X_m}$  is invertible and its Moore-Penrose inverse equals its inverse  $B_{m,n}^+ = (B_{m,n}|_{X_m})^{-1}$ .

Moreover, it is well known that the projection method  $B_{m,n}^+$  equals the corresponding one related to the *Gauss-symmetrized* version of (5), i.e.,

$$B_{m,n}^* Q_n y^\delta = B_{m,n}^* B_{m,n} x_{m,n}^\delta,$$

precisely we have  $B_{m,n}^+ = (B_{m,n}^* B_{m,n})^+ B_{m,n}^*$ . From this representation it is also clear that  $B_{m,n}^+ y^\delta = B_{m,n}^+ Q_n y^\delta$ , i.e., the projection method uses exactly the given data from (4).

We mention that the projection method as introduced above corresponds to the mapping  $Q_h$ , as used in [19], whereas the mappings  $P_h$  from there correspond to  $P_m$  within the present context of Hilbert spaces. Unique solvability was also assumed in [19].

Within the present context we shall study convergence properties of projection methods, and this requires us to consider families of projection methods, introduced next.

**Definition 2.** A non-increasing mapping  $\delta \mapsto (m(\delta), n(\delta))$ ,  $0 < \delta \leq \delta_0$  is called *parameter choice*.

**Definition 3.** Suppose that we are given designs  $\{\varphi_j, j \in \mathbb{N}\} \subset Y$  and  $\{\psi_i, i \in \mathbb{N}\} \subset X$  with corresponding projections  $Q_n$  and  $P_m$ . Each parameter choice we assign the family  $\left\{B_{m(\delta),n(\delta)}^+\right\}$  of projection methods. This is called *projection scheme*. If the designs have Property (S), then we agree to say that the projection scheme has property (S).

### 3. CONTROLLING THE ERROR

At any instance  $x^\dagger \in X$  the error of a projection method  $B_{m,n}^+$  with approximate solution  $x_{m,n}^\delta$  is given by  $\|x^\dagger - x_{m,n}^\delta\|$ . We are interested in conditions that yield convergent projection schemes, as  $\delta \rightarrow 0$ .

**Definition 4** (convergence). Let  $\delta \mapsto (m(\delta), n(\delta))$ ,  $0 < \delta \leq \delta_0$  be a parameter choice. The corresponding projection scheme  $\left\{B_{m(\delta),n(\delta)}^+\right\}_{\delta>0}$  is called *convergent* if  $\|x^\dagger - x_{m(\delta),n(\delta)}^\delta\| \rightarrow 0$  as  $\delta \rightarrow 0$ .

Of particular interest are projection schemes which are convergent *uniformly* on classes  $M$  of input, where we define the uniform error of any method  $B_{m,n}^+$  as

$$(7) \quad e(B_{m,n}^+, M, \delta) := \sup_{x^\dagger \in M} \sup_{\|y^\delta - Ax^\dagger\| \leq \delta} \|x^\dagger - B_{m,n}^+ y^\delta\|.$$

There is a trivial error decomposition using the triangle inequality as

$$(8) \quad \|x^\dagger - x_{m,n}^\delta\| \leq \|x^\dagger - B_{m,n}^+ Ax^\dagger\| + \|B_{m,n}^+ Ax^\dagger - B_{m,n}^+ y^\delta\|,$$

into the noise-free and pure noise components. In usual regularization theory the analysis of the noise-free component is related to the notion of a *regularization*, and we adopt this to the present context.

**Definition 5** (regularization). If for each  $x^\dagger \in X$  there are  $m_0 \in \mathbb{N}$  and a mapping  $n = n(m)$ ,  $m \geq m_0$  such that  $\|x^\dagger - B_{m,n(m)}^+ Ax^\dagger\| \rightarrow 0$  as  $m \rightarrow \infty$ , then the family  $B_{m,n(m)}^+$  is said to be a *regularization*.

*Remark 3.* The analysis of the noise-free term is of particular interest, and recently, see e.g. [7, 5] this was called *profile function*. In the present context we would obtain (having chosen  $n = n(m)$ ) that any decreasing function  $f(m)$ ,  $m \in \mathbb{N}$  with

$$(9) \quad \|x^\dagger - B_{m,n(m)}^+ Ax^\dagger\| \leq f(m), \quad m \in \mathbb{N}$$

is a profile function for  $(x^\dagger, B_{m,n(m)}^+)$ .

**3.1. Quasi-optimality and robustness.** With the error decomposition (8) the following definitions of quasi-optimality and robustness are natural, as they allow to bound the first and second components, respectively. This has first been recognized in [19].

**Definition 6** (Quasi-optimality). A projection scheme  $\{B_{m(\delta),n(\delta)}^+\}$  is called *quasi-optimal* if

- (1) For each  $x^\dagger \in X$  it holds true that  $P_m x^\dagger \rightarrow x^\dagger$  as  $m \rightarrow \infty$ , and
- (2) there is  $D_Q < \infty$  for which

$$(10) \quad \|B_{m(\delta),n(\delta)}^+ A\| \leq D_Q, \quad \text{for all } 0 < \delta \leq \delta_0.$$

The following is immediate.

**Proposition 1.** *Any quasi-optimal projection scheme with Property (S) constitutes a regularization. Precisely it holds*

$$(11) \quad \|x^\dagger - B_{m(\delta),n(\delta)}^+ A x^\dagger\| \leq (1 + D_Q) \|x^\dagger - P_{m(\delta)} x^\dagger\|, \quad x^\dagger \in X,$$

if  $n = n(m)$  is chosen accordingly.

*Proof.* Under Property (S), and with  $n = n(m)$  as there, we have  $B_{m,n}^+ B_{m,n} = P_m$ ,  $m \geq m_0$ . Thus we can bound

$$(12) \quad \begin{aligned} \|x^\dagger - B_{m,n}^+ A x^\dagger\| &\leq \|x^\dagger - B_{m,n}^+ A P_m x^\dagger\| + \|B_{m,n}^+ A x^\dagger - B_{m,n}^+ A P_m x^\dagger\| \\ &= \|x^\dagger - B_{m,n}^+ Q_n A P_m x^\dagger\| + \|B_{m,n}^+ A (x^\dagger - P_m x^\dagger)\| \\ &\leq (1 + D_Q) \|x^\dagger - P_m x^\dagger\|, \end{aligned}$$

which tends to zero as  $m \rightarrow \infty$ .  $\square$

In particular this result asserts that for quasi-optimal schemes the residual for noise-free data tends to zero as  $m \rightarrow \infty$ , hence such are regularizing. In classical regularization theory this property is strengthened to the concept of qualification, which describes the way how regularizing schemes react on underlying smoothness assumptions. This subject will be discussed in § 4.2, later.

The next concept is used to bound the pure noise component in (8).

**Definition 7** (Robustness). A projection scheme  $\{B_{m(\delta),n(\delta)}^+\}$  is called *robust* if there is a constant  $D_R < \infty$  such that

$$(13) \quad \|B_{m,n}^+ : Y \rightarrow X\| \leq D_R \sigma_m, \quad m \in \mathbb{N},$$

where  $\sigma_m$  is defined as

$$(14) \quad \sigma_m := \sup \left\{ \frac{\|P_m z\|}{\|A P_m z\|}, \quad P_m z \neq 0 \right\}.$$

*Remark 4.* Notice that by definition the mapping  $m \mapsto \sigma_m$  is non-decreasing.

Furthermore, it is well known (see [19]) that under Property (S) necessarily  $\sigma_m \leq \|B_{m,n}^+\|$ , such that  $D_R \geq 1$  in this case.

Robustness allows for an estimate similar to (11).

**Proposition 2.** *For a robust projection scheme the noise-free term can be bounded as*

$$\|x^\dagger - B_{m,n}^+ A x^\dagger\| \leq (1 + D_R \sigma_m \|A(I - P_m)\|) \|x^\dagger - P_m x^\dagger\|.$$

*Proof.* The estimate follows the one from (12), but uses robustness for the second summand and the fact that

$$\|A(x^\dagger - P_m x^\dagger)\| \leq \|A(I - P_m)\| \|x^\dagger - P_m x^\dagger\|.$$

□

We arrive at the basic implication of these properties.

**Theorem 1.** *Given  $x^\dagger \in X$ , we assign*

$$(15) \quad m(\delta) := \sup \left\{ m \geq 1, \quad \frac{d(x^\dagger, X_m)}{\sigma_m} \geq \delta \right\}.$$

*Assume furthermore that Property (S) holds true and  $n(\delta) = n(m(\delta))$  is chosen accordingly. If the corresponding projection scheme  $\{B_{m(\delta),n(\delta)}^+\}$  is quasi-optimal and robust, then it is convergent.*

*Proof.* Under Property (S), quasi-optimality and robustness, and using the bound (11), the error estimate (8) specifies to

$$(16) \quad \|x^\dagger - x_{m,n(m)}^\delta\| \leq (1 + D_Q) \|x^\dagger - P_m x^\dagger\| + \delta D_R \sigma_m.$$

Now we distinct two cases. First, if there is  $M < \infty$  such that  $m(\delta) < M$  as  $\delta \rightarrow 0$ , then  $d(x^\dagger, X_m) = 0$  for  $m \geq M$  and we have  $\|x^\dagger - x_{M,n(M)}^\delta\| \leq D_R \sigma_M \delta$ , which tends to zero as  $\delta \rightarrow 0$ .

Otherwise,  $m(\delta) \rightarrow \infty$  as  $\delta \rightarrow 0$ , and we can continue from (16) to deduce

$$\|x^\dagger - x_{m(\delta),n(m(\delta))}^\delta\| \leq (1 + D_Q + D_R) \|x^\dagger - P_{m(\delta)} x^\dagger\|.$$

The right hand side converges to zero as  $\delta \rightarrow 0$ , by quasi-optimality. □

The above Theorem 1 is actually non-constructive, since  $x^\dagger$  is not known. However, it implies that quasi-optimal and robust schemes admit a convergent parameter choice. One goal of the present study is to establish such choice, *a priori* – based entirely on smoothness properties of  $x^\dagger$  – or *a posteriori* – based on the noise level  $\delta$  and the data  $Q_n y^\delta$ .

**3.2. Control parameters.** The following parameters  $\varrho_m$  and  $\eta_n$ , related to the chosen design systems are crucial. We let

$$(17) \quad \varrho_m := \|A(I - P_m)\|, \quad m \in \mathbb{N} \quad \text{and}$$

$$(18) \quad \eta_n := \|(I - Q_n)A\|, \quad n \in \mathbb{N}.$$

These control parameters are indeed important in understanding the convergence behavior of projection schemes as this is established below. Recall the definition of  $\sigma_m$  from (14).

**Proposition 3.** *Suppose that there are a constant  $0 < \tau < 1$  and a mapping  $n := n(m)$ ,  $m \geq m_0$  for which  $\eta_{n(m)}\sigma_m \leq \tau$ . Then the following estimate holds true for  $m \geq m_0$ :*

$$(19) \quad \|AP_m z\| \leq \frac{1}{1 - \tau} \|Q_{n(m)}AP_m z\|, \quad z \in X.$$

*Proof.* If  $m \geq m_0$  and  $AP_m z \neq 0$  then we estimate

$$(20) \quad \begin{aligned} \|AP_m z\| &\leq \|(I - Q_{n(m)})AP_m z\| + \|Q_{n(m)}AP_m z\| \\ &\leq \eta_{n(m)}\|P_m z\| + \|Q_{n(m)}AP_m z\| \\ &\leq \eta_{n(m)} \cdot \sigma_m \|AP_m z\| + \|Q_{n(m)}AP_m z\| \\ &\leq \tau \|AP_m z\| + \|Q_{n(m)}AP_m z\| \end{aligned}$$

This implies  $(1 - \tau) \|AP_m z\| \leq \|Q_{n(m)}AP_m z\|$ , hence (19).  $\square$

**Theorem 2.** *If  $\{B_{m(\delta), n(\delta)}^+\}$  is a projection scheme which obeys the assumptions of Proposition 3, then*

- (1) *Property (S) is satisfied,*
- (2) *the scheme is robust (with  $D_R = 1/(1 - \tau)$ ).*
- (3) *Suppose that there is a constant  $C < \infty$  for which  $\varrho_m\sigma_m \leq C$ ,  $m \in \mathbb{N}$ . If the projections  $P_m$  converge point-wise as  $m \rightarrow \infty$ , then the projection scheme is quasi-optimal (with  $D_Q = 1 + D_R C$ ).*
- (4) *If we assign  $m = m(\delta)$  according to (15), and  $n = n(m(\delta))$  according to Property (S), then the projection scheme is convergent.*

*Proof.* Since  $A$  is injective we deduce from (19) that  $B_{m, n(m)} z = 0$  yields  $P_m z = 0$ , hence  $\ker(B_{m, n(m)}) \subset X_m^\perp$ , which in turn implies the validity of Property (S) by this mapping  $n = n(m)$ .

Next, by definition of  $\sigma_m$  we have that  $\|P_m z\| \leq \sigma_m \|AP_m z\|$ ,  $P_m z \neq 0$ , thus (with  $n := n(m)$ ) estimate (19) yields

$$(21) \quad \|P_m z\| \leq \frac{\sigma_m}{1 - \tau} \|B_{m, n} z\|.$$

Under Property (S) we have  $B_{m,n}^+ B_{m,n} = P_m$ . Furthermore  $B_{m,n}$  decomposes  $Y$  into  $Y = \mathcal{R}(B_{m,n}) \oplus \mathcal{R}^\perp(B_{m,n})$ . If  $y = y_1 + y_2$  according to this decomposition, then  $B_{m,n}^+ y = B_{m,n}^+ y_1 = B_{m,n}^+ B_{m,n} z$  for some  $z \in X$ , and we deduce

$$\|B_{m,n}^+ y\| = \|B_{m,n}^+ B_{m,n} z\| = \|P_m z\| \leq \frac{\sigma_m}{1-\tau} \|B_{m,n} z\| \leq \frac{\sigma_m}{1-\tau} \|y\|,$$

because  $\|B_{m,n} z\| \leq \|y\|$ , which proves robustness.

To establish quasi-optimality we use the robustness and find

$$\begin{aligned} \|B_{m,n}^+ A\| &\leq \|B_{m,n}^+ Q_n A P_m\| + \|B_{m,n}^+ Q_n A (I - P_m)\| \\ &\leq 1 + \|B_{m,n}^+\| \|Q_n\| \|A(I - P_m)\| \leq 1 + D_R \sigma_m \varrho_m \\ &= 1 + D_R C, \end{aligned}$$

which in conjunction with the other assumptions establishes quasi-optimality, and in the light of Theorem 1 also convergence.  $\square$

*Remark 5.* As seen from the beginning of the proof, the assumption of injectivity for  $A$  can be relaxed to  $\ker(A) \cap X_m = \{0\}$ ,  $m \geq m_0$ . Such assumption is met in [22].

**3.3. Bernstein- and Jackson-type inequalities.** In the previous subsection we indicated the importance of the control parameters  $\varrho_m$  and  $\eta_n$  from (17) and (18) for convergence properties of projection schemes. Now we shall establish natural sufficient conditions on these quantities in order to guarantee the applicability of Theorem 2. Such conditions are usually met in discussions on projection schemes, see e.g. [8, 13]. For instance, one of the crucial assumptions, see e.g., [19, Eq. (3.1)], used to establish quasi-optimality and robustness is as follows: There is a constant  $C$  such that for all  $x$  one can find  $x_m \in X_m$  for which

$$\|x - x_m\| + \sigma_m \|A(x - x_m)\| \leq C \|x\|,$$

with  $\sigma_m$  from (14). In Hilbert space, by letting  $x_m := P_m x$  this reduces to

$$\|(I - P_m)x\| + \sigma_m \|A(I - P_m)x\| \leq C \|x\|,$$

which further reduces to

$$(22) \quad \sigma_m \varrho_m \leq C, \quad m \in \mathbb{N},$$

an assumption which was also made in Theorem 2 (3).

Here we highlight some approximation theoretic content of such requirements. A more detailed discussion of such requirements in the context of variable Hilbert scales is given in [12], and with slightly different notation in [6]. We start with *Jackson-type estimates*, as these

are natural requirements on the approximation power of the underlying design.

**Assumption A.1.** There is a constant  $C_P$  such that

$$\varrho_m \leq C_P \inf \{ \|A(I - L_m)\|, L_m \text{ with } \text{rank}(L_m) \leq m \} = C_P s_{m+1},$$

where the infimum is taken with respect to orthogonal projections  $L_m$  in the space  $X$ .

*Remark 6.* From the discussion in § 2 we know that the infimum on the right hand side above is exactly the  $(m + 1)$ st approximation number of the operator  $A$ . Thus Assumption A.1 asserts, that the design for the reconstruction space must be capable to achieve the best possible order of accuracy. This is known to hold for many systems, as specific finite elements and splines, we refer to [1] and the discussion in [19, § 4].

Notice that such strong assumption is not required for the data design and the corresponding control parameter  $\eta_n$ . All what will be required is  $\eta_n \rightarrow 0$  as  $n \rightarrow \infty$ , which is certainly a minimal requirement.

The other assumption is less obvious. It reflects smoothness properties of the design, which is hidden in its present formulation, see [13] for a discussion of Bernstein- and Jackson-type inequalities for projection schemes.

**Assumption A.2** (Bernstein inequality). There is a constant  $C_B < \infty$  for which

$$(23) \quad \|P_m z\| \leq \frac{C_B}{s_m} \|AP_m z\|, \quad z \in X.$$

*Remark 7.* We stress, that necessarily  $\|AP_m z\|/s_m \leq \|P_m z\|$ , which was discussed in [12]. In particular  $1/s_m \leq \sigma_m$ ,  $m \in \mathbb{N}$ .

The following result asserts that this is fulfilled by projections onto the singular system of  $A$ .

**Lemma 2.** *If the design  $\{v_i\}_{i \in \mathbb{N}}$  corresponds to the singular system of the operator  $A$  as in (2), then A.2 is fulfilled with  $C_B = 1$ .*

*Proof.* For  $z \in X$  we can bound

$$\|P_m z\|^2 = \sum_{i=1}^m |\langle z, v_i \rangle|^2 = \sum_{i=1}^m \frac{1}{s_i^2} s_i^2 |\langle z, v_i \rangle|^2 \leq \frac{1}{s_m^2} \|AP_m z\|^2,$$

which gives (23) with  $C_B = 1$ . □

The following obvious consequence for  $\sigma_m$  from (14) is important.

**Proposition 4.** *If the projections  $P_m$  obey A.2, then  $\sigma_m \leq C_B/s_m$ .*

In terms of the constants from A.1 and A.2 we can indeed guarantee the validity of the assumptions made in Theorem 2.

**Proposition 5.** *Under assumptions A.1 and A.2 the bound (22) holds true for  $C := C_B C_P$ .*

*If  $n = n(m)$  is chosen such that*

$$(24) \quad \eta_{n(m)} \leq \frac{1}{2C_B C_P} \varrho_m$$

*then*

$$(25) \quad \eta_{n(m)} \sigma_m \leq \frac{1}{2}.$$

*Proof.* First, using Proposition 4, we estimate

$$\varrho_m \sigma_m \leq C_P s_{m+1} \cdot \frac{C_B}{s_m} = C_B C_P.$$

Furthermore, if  $n(m)$  is chosen as indicated in (24) then by (22) (with  $C := C_B C_P$ ) we obtain that

$$\eta_{n(m)} \sigma_m \leq \frac{1}{2C_B C_P} \varrho_m \sigma_m \leq \frac{1}{2},$$

which completes the proof.  $\square$

We explicitly mention the following converse result to a bound like (22), under the following restriction to the decay of the singular values of  $A$ .

**Assumption A.3** (Decay of singular values). There is a constant  $0 < \gamma < 1$  such that  $s_{m+1}/s_m \geq \gamma$ ,  $m \in \mathbb{N}$ .

*Remark 8.* Plainly, this assumption is fulfilled for singular values decaying polynomially. But this holds also true for singular values decaying exponentially fast, as e.g.  $s_m = \exp(-\kappa m)$  (with  $\gamma := e^{-\kappa}$ ), whereas super-exponentially decaying singular values are not covered by this approach.

**Proposition 6.** *Suppose that the singular values of  $A$  obey Assumption A.3. If (22) holds true for some constant  $C < \infty$  then both the Assumptions A.1 and A.2 must be satisfied.*

*Proof.* Suppose that (22) is fulfilled but Ass. A.1 is violated, hence along a sub-sequence we have  $\varrho_{m_k}/s_{m_k+1} \rightarrow \infty$  as  $k \rightarrow \infty$ . Then, in the light of Remark 7 we derive that

$$\gamma \leq \frac{s_{m_k+1}}{s_{m_k}} \leq \sigma_{m_k} s_{m_k+1} = \varrho_{m_k} \sigma_{m_k} \frac{s_{m_k+1}}{\varrho_{m_k}} \leq C \frac{s_{m_k+1}}{\varrho_{m_k}} \rightarrow 0,$$

which is a contradiction. In the case of the violation of A.2 the proof is similar and we omit it.  $\square$

**3.4. One-sided discretization.** Many authors restrict the issue of discretization to one side, either from the right or from the left, which are commonly referred to as *least-squares* and *dual least-squares* solutions, see [19] or the more recent monograph [9]. In these cases we turn from the original equation (1) to either  $y^\delta = AP_m x^\dagger + \delta\xi$  or  $Q_n y^\delta = Q_n A x^\dagger + \delta Q_n \xi$ , respectively. Notice, that least-squares regularization is only possible, if we may choose the design of the data space. We stress, that such one-sided discretization is *not* obtained from the two-sided one by letting  $Q = I$  or  $P = I$ ! Instead these correspond to projections  $Q_m := P_{\mathcal{R}(AP_m)}$  and  $P_n = P_{\mathcal{R}(A^*Q_n)}$ .

Below we shall briefly discuss, how the previous assumptions A.1 and A.2, imposed on the design spaces may be fulfilled in these cases.

*Least squares discretization.* In this case we have a one-sided discretization from the right, i.e.,  $B_m = AP_m$ , hence  $Q_m := P_{\mathcal{R}(B_m)}$ . It is well known, that in this case robustness holds true automatically, since

$$\sup_{y \in Y_m} \frac{\|B_m^\perp y\|}{\|y\|} = \sup_{x \in X_m} \frac{\|B_m^\perp B_m x\|}{\|B_m x\|} = \sup_{x \in X_m} \frac{\|x\|}{\|B_m x\|} = \sup_{x \in X_m} \frac{\|x\|}{\|Ax\|} = \sigma_m,$$

thus  $C_R = 1$ . Quasi-optimality follows as in the proof of Theorem 2 (3) and from Proposition 5 under Assumptions A.1 and A.2.

*Dual least squares discretization.* In this case we let  $B_n := Q_n A$ . Again, it is well known that quasi-optimality holds automatically, since

$$B_n^\perp A = B_n^\perp Q_n A = P_n.$$

Robustness for the dual least squares discretization is more subtle. To have the Bernstein-type inequality from A.2 we may argue, see e.g. [8, p. 1528], as follows

$$\begin{aligned} \inf_z \frac{\|AP_m z\|}{\|P_m z\|} &= \inf_v \frac{\|AA^*Q_n v\|}{\|A^*Q_n v\|} \\ &\geq \inf_v \frac{\langle AA^*Q_n v, Q_n v \rangle}{\|A^*Q_n v\| \|Q_n v\|} = \inf_v \frac{\|A^*Q_n v\|^2}{\|A^*Q_n v\| \|Q_n v\|} \\ &= \inf_v \frac{\|A^*Q_n v\|}{\|Q_n v\|}. \end{aligned}$$

Thus, if the data spaces  $Y_n$  obey a Bernstein-type inequality with respect to the operator  $A^*$ , then the Assumption A.2 is also satisfied.

Finally, since  $(I - P_n)A^*Q_n = 0$ , we see that

$$\varrho_n = \|A(I - P_n)\| = \|(I - P_n)A^*\| = \|(I - P_n)A^*(I - Q_n)\| \leq \eta_n.$$

This means, that a Jackson-type inequality, but for the design  $Y_n$  with respect to the operator  $A^*$  implies the one for  $X_n$  and  $A$  as in A.1. Thus we may require that there is  $C_Q$  for which  $\eta_n \leq C_Q s_{n+1}$ ,  $n = 1, 2, \dots$ . Overall, for the dual least squares discretization, both the assumptions of Bernstein and Jackson-type inequalities, but for the design spaces  $Y_n$  and with respect to the operator  $A^*$ , imply the corresponding assumptions for the spaces  $X_n$  and the operator  $A$ . Such assumptions were actually made in [8, Ass. 2&3], and hence the present analysis extends this to two-sided discretization.

#### 4. CONVERGENCE RATES

If we want to obtain convergence rates uniformly for classes  $M$  of instances, see (7) for the error definition, then some form of *smoothness* will help. Here we restrict ourselves to smoothness given in terms of general source conditions as follows, we refer to [15, 14, 2], and to [5] for a more general point of view.

**4.1. Smoothness in terms of general source conditions.** We call a non-negative continuous increasing function  $\varphi: [0, \|A^*A\|] \rightarrow [0, \infty)$  which obeys  $\varphi(0) = 0$ , an *index function*, see [5]. Below we shall denote the operator  $H := A^*A$  for simplicity. To such functions we assign (using spectral calculus) sets

$$(26) \quad H_\varphi := \{x \in X, \quad x = \varphi(H)v, \quad \|v\| \leq 1\}.$$

*Remark 9.* If  $x^\dagger \in H_\varphi$  then we say that smoothness of  $x^\dagger$  is given in terms of a general source condition.

In previous studies where general source conditions were used, the sets  $H_\varphi(R) := RH_\varphi$  were taken, with constant  $0 < R < \infty$ . But this is only a matter of scaling, since  $H_\varphi(R) = H_{\varphi_R}$ , for the scaled index function  $\varphi_R := R\varphi$ .

These sets  $H_\varphi$  are centrally symmetric and convex, they can thus be considered as *unit balls* of some Hilbert space, which we denote by  $X_\varphi^H$ . Notice that  $X_\varphi^H \subset X$ , and that this embedding is compact, since the operator  $A$  is assumed to be compact, see [5, § 2]. Therefore we may think of sets  $H_\varphi$  as smoothness classes (relative to the operator  $A$ ). The spaces  $X_\varphi^H$  belong to *variable Hilbert scales* related to the operator  $H$ . Here we shall not need much of the calculus known for such scales, we refer to [5, 17] for details. The crucial observation is

**Proposition 7.** *Let  $x^\dagger \in H_\varphi$  for an index function  $\varphi$ . If the designs obey assumptions A.1 and A.2, and  $n = n(m)$  is chosen as indicated*

in Proposition 5, then

$$(27) \quad \|x^\dagger - B_{m,n(m)}^+ Ax^\dagger\| \leq (1 + 2C_B C_P) \|I - P_m : X_\varphi^H \rightarrow X\|.$$

*Proof.* As indicated in the proof of Proposition 2, under the assumptions made we have for each  $x^\dagger \in X$  that

$$(28) \quad \|x^\dagger - B_{m,n}^+ Ax^\dagger\| \leq (1 + D_R \sigma_m \varrho_m) \|x^\dagger - P_m x^\dagger\|.$$

Under the specifications  $D_R = 2$ ,  $\varrho_m \sigma_m \leq C = C_B C_P$  made in Proposition 5, the proof can be completed for  $x^\dagger \in H_\varphi$ .  $\square$

**4.2. Qualification of projection methods.** In usual regularization theory the capability of a regularization to take smoothness into account leads to the notion of a *qualification*, we refer to [5] for a recent discussion, where in particular it was shown in [5, Sect. 5] that qualification yields bounds for the noise-free term. For one-sided projection schemes a discussion on this subject can be found in [6, § 5.1].

Within the present context, for quasi-optimal projection schemes this capability is reduced to properties of the projections  $P_m$ ,  $m = m_0, \dots$ , and corresponding spaces  $X_m$ , as can be seen from (11).

**Definition 8.** Let  $\psi$  be an index function. The projections  $P_m$ ,  $m = m_0, \dots$  onto spaces  $X_m$  of dimension  $m$  have *qualification*  $\psi$  if there is a constant  $\gamma$  for which

$$(29) \quad \|(I - P_m)\psi(A^*A)\| \leq \gamma\psi(s_{m+1}^2), \quad m = m_0, m_0 + 1, \dots$$

*Remark 10.* The best possible bound in (29) is for  $\gamma = 1$ , see Equation (3) and the following discussion, applied to the operator  $\psi(A^*A)$  in § 2, and this is provided by the spectral projection. Therefore, spectral projections have arbitrary qualification.

In the light of the above definition we may rephrase the Jackson-type inequality from Assumption A.1. This asserts that we assume that the projections have qualification “one half”, or precisely that the function  $\psi(t) = \sqrt{t}$  is a qualification.

Within previous studies on projection schemes, starting from [19, § 4] such assumptions were made. In particular, the bound from Eq (4.3) *ibid.* asserts that the approximation schemes under consideration have qualification  $\psi(t) = t^s$ .

*Remark 11.* It is interesting to notice that for one-sided discretization from the left, i.e., when  $P_n = P_{\mathcal{R}(A^*Q_n)}$ , then the validity of a Jackson-type inequality for the operator  $A^*$  yields a qualification one, which means with function  $\psi(t) = t$ . This was shown and exploited in [6, Lem. 5].

One important question is how a known qualification allows to deduce properties with respect to other smoothness. Here this question is related to *interpolation*.

**Proposition 8.** *Suppose that the projections  $P_m$ ,  $m = m_0, \dots$  have qualification  $\psi$ . If  $\varphi$  is an index function such that the function*

$$t \mapsto \varphi^2 \left( (\psi^2)^{-1}(t) \right), \quad 0 \leq t \leq \|A^*A\|,$$

*is defined on  $[0, \|A^*A\|]$  and concave there, then the projections have qualification  $\varphi$ .*

*Proof.* This is a straight forward application of [14, Append. A, Cor. 2] for  $T := I - P_m: X \rightarrow X$  and parameter  $\theta := 1$ , which yields that

$$\varphi^{-1} \left( \frac{\|(I - P_m)\varphi(A^*A)\|}{\|I - P_m\|} \right) \leq \psi^{-1} \left( \frac{\|(I - P_m)\psi(A^*A)\|}{\|I - P_m\|} \right).$$

Since  $\|I - P_m\| = 1$  this rewrites, using the qualification  $\psi$ , as

$$\begin{aligned} \|(I - P_m)\varphi(A^*A)\| &\leq \varphi \left( \psi^{-1} (\|(I - P_m)\psi(A^*A)\|) \right) \\ &\leq \varphi \left( \psi^{-1} (\gamma\psi(s_{m+1}^2)) \right) \leq \gamma\varphi \left( \psi^{-1} (\psi(s_{m+1}^2)) \right) \\ &= \gamma\varphi(s_{m+1}^2), \end{aligned}$$

where we used the concavity to move  $\gamma$  to the front.  $\square$

We accompany this result with the case that the smoothness  $\varphi$  is “larger than” the qualification  $\psi$ , i.e., when  $\varphi(t) \leq C\psi(t)$  for some finite constant  $C$ . In this case we have at least the bound

$$\|(I - P_m)\varphi(A^*A)\| = \|(I - P_m)\psi(A^*A) \left( \frac{\varphi}{\psi} \right) (A^*A)\| \leq C\gamma\psi(s_{m+1}^2).$$

The subsequent error analysis easily extends to projections of higher qualification, however we present them for the minimal case covered by Assumption A.1, i.e., when qualification is assumed for  $\psi(t) = \sqrt{t}$ . Then we need to distinguish two cases for the sequence of projections  $P_m$ ,  $m = m_0, \dots$ .

**Case 1:** The index function  $\varphi$  has a concave square, or

**Case 2:** there is  $C_\varphi < \infty$  with  $\varphi(t) \leq C_\varphi\sqrt{t}$ .

The above discussion allows to summarize norm bounds for the projections in both cases.

**Proposition 9.** *Let  $\varphi$  be an index function. Then*

$$(30) \quad \|I - P_m: X_\varphi^H \rightarrow X\| \leq C_P \begin{cases} \varphi(s_{m+1}^2), & \text{in Case 1,} \\ C_\varphi s_{m+1}, & \text{in Case 2.} \end{cases}$$

**4.3. A priori error bounds.** We will provide error estimates valid *uniformly* for  $x^\dagger \in H_\varphi$ , and start with the following lower bound from [15, Cor. 1]. As usual, given the index function  $\varphi$  we assign the related index function

$$(31) \quad \Theta(t) := \sqrt{t}\varphi(t), \quad 0 \leq t \leq \|A^*A\|.$$

The following result in Theorem 3 holds in more general form (see [15, Thm. 1]), however we shall state it under Assumption A.3.

**Theorem 3.** *Suppose that  $x^\dagger \in H_\varphi$  and that the function*

$$t \longrightarrow \varphi^2((\Theta^2)^{-1}(t)), \quad 0 < t \leq \Theta(\|A^*A\|),$$

*is concave. If the operator  $A$  obeys A.3, then there is  $c_\gamma > 0$  such that for any projection scheme  $\{B_{m(\delta),n(\delta)}^+\}$  we have*

$$(32) \quad e(B_{m(\delta),n(\delta)}^+, H_\varphi, \delta) \geq c_\gamma \varphi(\Theta^{-1}(\delta)), \quad 0 < \delta \leq \Theta(\|A^*A\|).$$

We shall use the following bound for the pure noise term, which is an immediate consequence of Propositions 4 and 5.

**Proposition 10.** *If the assumptions A.1 and A.2 are valid, and if we choose  $n = n(m)$  according to (24) then*

$$(33) \quad \|B_{m,n}^+\| \leq 2C_B/s_m.$$

Now we are able to state and prove the main error estimate for projection schemes.

**Theorem 4.** *Let  $x^\dagger \in H_\varphi$  for an index function  $\varphi$ . Assume that the designs obey assumptions A.1 and A.2. Let*

$$(34) \quad m_* = m_*(\delta) := \begin{cases} \sup \{m, \Theta(s_m^2) \geq \delta\}, & \text{in Case 1,} \\ \sup \{m, s_m^2 \geq \delta\}, & \text{in Case 2,} \end{cases}$$

*and  $n_* = n(m_*)$  according to (24). Then*

$$\textbf{Case 1: } e(B_{m_*,n_*}^+, H_\varphi, \delta) \leq 2(1 + 2C_B C_P) C_P \varphi(s_{m_*}^2),$$

$$\textbf{Case 2: } e(B_{m_*,n_*}^+, H_\varphi, \delta) \leq 2(1 + 2C_B C_P) C_P C_\varphi s_{m_*}.$$

*Proof.* From Proposition 10 we obtain that

$$\delta \|B_{m,n}^+\| \leq 2C_B \delta / s_m \leq 2C_B C_P \delta / s_m,$$

and we derive from Proposition 7 that

$$\begin{aligned} \|x^\dagger - x_{m_*,n_*}^\delta\| &\leq (1 + 2C_B C_P) \|I - P_{m_*} : X_\varphi^H \rightarrow X\| + 2C_B C_P \delta / s_{m_*} \\ &\leq (1 + 2C_B C_P) \left( \|I - P_{m_*} : X_\varphi^H \rightarrow X\| + \frac{\delta}{s_{m_*}} \right). \end{aligned}$$

The norm bounds in Proposition 9 yield the desired estimates in both cases.  $\square$

The above error bounds may be given the usual form.

**Corollary 1.** *Under the assumptions of Theorem 4 and with the choices of  $m_*$  and  $n_*$  as there, the following error bounds hold true.*

$$\text{Case 1: } e(B_{m_*, n_*}^+, H_\varphi, \delta) \leq 2(1 + 2C_B C_P) C_P \varphi(\Theta^{-1}(\delta)).$$

$$\text{Case 2: } e(B_{m_*, n_*}^+, H_\varphi, \delta) \leq 2(1 + 2C_B C_P) C_P C_\varphi \sqrt{\delta}.$$

*Proof.* If  $m_*$  is chosen from (34), then  $\Theta(s_{m_*+1}^2) \leq \delta$ . Moreover, the choice from (15) yields

$$\frac{\delta}{s_{m_*}} \leq \frac{\delta}{\sqrt{\Theta^{-1}(\delta)}} = \varphi(\Theta^{-1}(\delta))$$

Thus, in Case 1 we continue

$$\begin{aligned} \|x^\dagger - x_{m_*, n_*}^\delta\| &\leq (1 + 2C_B C_P) C_P (\varphi(s_{m_*+1}^2) + \delta/s_{m_*}) \\ &\leq (1 + 2C_B C_P) C_P 2\varphi(\Theta^{-1}(\delta)). \end{aligned}$$

The proof in Case 2 is similar.  $\square$

*Remark 12.* It is worth-while to notice that in Case 1 we obtain the optimal order of reconstruction, see Theorem 3. Moreover, in the above analysis, based on Proposition 9, there occurs a saturation in the obtained rates, at a level  $\sqrt{\delta}$ . This effect is due to the assumed limited qualification and we refer to the discussion in § 4.2. Under higher qualification this saturation will occur later, and for spectral projections the optimal rates are obtained for arbitrary smoothness.

**4.4. A posteriori parameter choice.** Often the smoothness of the true solution is not known, and the parameter choice of  $m_*$  and  $n_*$ , respectively, must be determined from the available noise level  $\delta$  and the data  $Q_n y^\delta$ . Several such parameter choice strategies are known, and we mention the recent references which deal with smoothness in terms of general source conditions, [14, 16] for using the discrepancy principle under discretization, and [10, 11] for the Lepskiĭ balancing principle. For the discrepancy principle, but in a more restrictive setup, we also mention [8, 22]. Here we restrict ourselves to the Lepskiĭ principle, since it is easily explained and since its use is standard, nowadays. For this principle to work we need a valid bound of the noise term which does not contain any smoothness assumption. Such bound was provided in Proposition 10, if we choose  $n = n(m)$  appropriately.

Thus, to start with, we choose the finest discretization level  $M$  such that  $s_M \asymp 1/\delta$ . Given  $q > 1$  we let  $k := \lfloor \log_q M \rfloor + 1$ . In our approach we restrict discretization to the levels from

$$\Delta_k := \{m_j = \lfloor q^{k-j} \rfloor, \quad j = 1, \dots, k\}.$$

For  $m_j \in \Delta_k$  we assign  $n_j := n(m_j)$  according to (24). We stress that this choice does not depend on smoothness properties of  $x^\dagger$ . It depends only on properties of the design spaces. Finally we determine the approximate solutions by  $x_j := B_{m_j, n_j}^+ Q_{n_j} y^\delta$ ,  $j = 1, \dots, k$ . As can be seen from (28) the following error bound holds true for  $m_j \in \Delta_k$ .

$$\|x^\dagger - x_{m_j, n_j}^\delta\| \leq \frac{1}{2} (2(1 + 2C_B C_P) \|(I - P_{m_j})x^\dagger\| + 4C_B \delta / s_{m_j})$$

In accordance with [10] we assign the non-increasing function  $\Psi(j) := 4C_B / s_{m_j}$ ,  $j = 1, 2, \dots, k$ . In terms of the notion and notation from [10] the functions

$$\Phi(j) := 2(1 + 2C_B C_P) \|(I - P_m)x^\dagger\|, \quad j = 1, \dots, k$$

are *admissible* under Assumption A.2, and if only the technical assumption  $\varrho_M \leq 1/(\varphi(a)C_P)$  is fulfilled by our choice of  $M$ . This guarantees that  $\Phi(1) \leq \Psi(1)$ . Thus if we determine the Lepskiĭ index  $\bar{m}$  as

$$(35) \quad \bar{m} := \max \{m, \quad \|x_l - x_m\| \leq 2\Psi(l), \quad l < m\},$$

then, under Assumptions A.1, A.2, and the additional decay restriction A.3, we have

$$\|x^\dagger - x_{\bar{m}}\| \leq \frac{12}{\gamma} (1 + 2C_B C_P) C_P \begin{cases} \varphi(\Theta^{-1}(\delta)), & \text{in Case 1} \\ C_\varphi \sqrt{\delta}, & \text{in Case 2.} \end{cases}$$

This reproduces the error bounds from Corollary 1 up to a factor  $6/\gamma$ .

## 5. BOUNDING THE CONDITION NUMBER

The solution to equation (5) is stable, as it is finite dimensional. However, if  $m$  and  $n$  increase the conditioning of the resulting system will be worse. The conditioning of projection methods has been discussed in some papers, and we mention [4, 21].

Stability as analyzed here is concerned with the conditioning of the mapping  $B_{m, n}^+$ . If  $\{B_{m(\delta), n(\delta)}^+\}$  is robust and has property (S), then, if  $n = n(m)$  is chosen accordingly, we have observed in Remark 4, that

$$(36) \quad \sigma_m \leq \|B_{m, n(m)}^+\| \leq D_R \sigma_m.$$

and its dependence in  $m$  is important. Thus, in this case, we are interested in the behavior of the *condition numbers*

$$(37) \quad \kappa_m := \|B_{m,n(m)}^+\| \|B_{m,n(m)}\|, \quad m \geq m_0.$$

The goal is to establish its relation to the singular values of the operator  $A$  governing the equation (1). If the designs correspond to the singular system of the operator  $A$ , then we may let  $n(m) = m$  and obtain

$$(38) \quad \kappa_m = \frac{s_1}{s_m}.$$

This is best possible in the following sense.

**Theorem 5.** *Suppose that the projection scheme  $\{B_{m(\delta),n(\delta)}^+\}$  obeys assumptions A.1 and A.2. Let  $n = n(m)$  be chosen to satisfy Property (S). The following bounds hold true.*

$$(39) \quad 1 \leq \liminf_{m \rightarrow \infty} \kappa_m \frac{s_m}{s_1} \leq \limsup_{m \rightarrow \infty} \kappa_m \frac{s_m}{s_1} \leq C_B D_R.$$

*Proof.* We first observe that

$$\left| \|A\|^2 - \|B_{m,n}\|^2 \right| = \left| \|A^*A\| - \|B_{m,n}^*B_{m,n}\| \right| \leq \|A^*A - B_{m,n}^*B_{m,n}\|.$$

A simple estimate (see [14, Prop. 1]) yields

$$\|A^*A - B_{m,n}^*B_{m,n}\| \leq 2\|A^*A\| \varrho_m + \eta_n^2,$$

Furthermore, by Remark 7 and the bounds from (36) we derive

$$\frac{1}{s_m} \leq \sigma_m \leq \|B_{m,n}^+\| \leq D_R \sigma_m \leq C_B D_R \frac{1}{s_m},$$

from which the proof can be completed, since  $\varrho_m$  as well as  $\eta_n$  tend to zero as  $m \rightarrow \infty$ .  $\square$

Of course, in practice one wants to recover the unknown solution at the optimal rate with a projection scheme having a small condition number, depending on the noise level  $\delta$ .

**Corollary 2.** *Let  $A$  obey A.3. Under the assumptions of Theorem 4, if  $m_*$  is chosen according to (34) then*

$$(40) \quad \kappa_*^\delta := \kappa_{m_*(\delta)} \asymp^1 \begin{cases} \frac{\varphi(\Theta^{-1}(\delta))}{\delta}, & \text{in Case 1} \\ 1/\sqrt{\delta}, & \text{in Case 2,} \end{cases} \quad \text{as } \delta \rightarrow 0.$$

---

<sup>1</sup>The symbol  $f(t) \asymp g(t)$  means, that there are constants  $0 < c < C < \infty$  for which  $c \leq f(t)/g(t) \leq C$  as  $t \rightarrow 0$ .

*Proof.* By the choice of  $m_*$  according to (34) we have under Assumption A.3 that  $\Theta(s_{m_*}^2) \asymp \delta$ . By Theorem 5 also  $\kappa_{m_*} \asymp 1/s_{m_*}$ , hence, in Case 1, we obtain

$$\kappa_{m_*} \asymp \frac{1}{\sqrt{\Theta^{-1}(\delta)}} = \frac{\varphi(\Theta^{-1}(\delta))}{\delta},$$

where the latter equality is easily checked. Case 2 is treated similarly, and the proof is omitted.  $\square$

*Remark 13.* This result establishes how the conditioning of the projection scheme depends on the underlying smoothness. For smoothness as treated here, the condition numbers increase at least as  $1/\sqrt{\delta}$ , but one may always achieve at most  $1/\delta$ .

#### ACKNOWLEDGMENT

The author wants to thank an unknown referee for initiating a discussion on qualification of projection methods. This led to the considerations in § 4.2, and may be viewed as a first attempt to tackle this problem.

#### REFERENCES

- [1] Robert A. Adams. *Sobolev spaces*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.
- [2] Albrecht Böttcher, Bernd Hofmann, Ulrich Tautenhahn, and Masahiro Yamamoto. Convergence rates for Tikhonov regularization from different kinds of smoothness conditions. *Appl. Anal.*, 85(5):555–578, 2006.
- [3] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.
- [4] B. Hofmann and G. Fleischer. Stability rates for linear ill-posed problems with compact and non-compact operators. *Z. Anal. Anwendungen*, 18(2):267–286, 1999.
- [5] Bernd Hofmann and Peter Mathé. Analysis of profile functions for general linear regularization methods. *SIAM J. Numer. Anal.*, 45(3):1122–1141 (electronic), 2007.
- [6] Bernd Hofmann, Peter Mathé, and Sergei V. Pereverzev. Regularization by projection: Approximation theoretic aspects and distance functions. preprint 2007-10, TU Chemnitz, Fak. für Math., 2007. to appear *J. Inverse Ill-Posed Probl.*
- [7] Bernd Hofmann and Masahiro Yamamoto. Convergence rates for Tikhonov regularization based on range inclusions. *Inverse Problems*, 21(3):805–820, 2005.
- [8] Barbara Kaltenbacher. Regularization by projection with a posteriori discretization level choice for linear and nonlinear ill-posed problems. *Inverse Problems*, 16(5):1523–1539, 2000.

- [9] Andreas Kirsch. *An introduction to the mathematical theory of inverse problems*, volume 120 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1996.
- [10] Peter Mathé. The Lepskiĭ principle revisited. *Inverse Problems*, 22(3):L11–L15, 2006.
- [11] Peter Mathé. What do we learn from the discrepancy principle? *Z. Anal. Anwend.*, 25(4):411–420, 2006.
- [12] Peter Mathé and Bernd Hofmann. Direct and inverse results in variable Hilbert scales. preprint 2006-27, TU Chemnitz, 2006.
- [13] Peter Mathé and Sergei V. Pereverzev. Optimal discretization of inverse problems in Hilbert scales. Regularization and self-regularization of projection methods. *SIAM J. Numer. Anal.*, 38(6):1999–2021, 2001.
- [14] Peter Mathé and Sergei V. Pereverzev. Discretization strategy for linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, 19(6):1263–1277, 2003.
- [15] Peter Mathé and Sergei V. Pereverzev. Geometry of linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, 19(3):789–803, 2003.
- [16] Peter Mathé and Sergei V. Pereverzev. The discretized discrepancy principle under general source conditions. *J. Complexity*, 22(3):371–381, 2006.
- [17] Peter Mathé and Ulrich Tautenhahn. Interpolation in variable Hilbert scales with application to inverse problems. *Inverse Problems*, 22(6):2271–2297, 2006.
- [18] M. T. Nair, S. V. Pereverzev, and U. Tautenhahn. Regularization in Hilbert scales under general smoothing conditions. *Inverse Problems*, 21(6):1851–1869, 2005.
- [19] Frank Natterer. Regularisierung schlecht gestellter Probleme durch Projektionsverfahren. *Numer. Math.*, 28(3):329–341, 1977.
- [20] Albrecht Pietsch. *Operator ideals*. North-Holland Publishing Co., Amsterdam, 1980. Translated from German by the author.
- [21] Eberhard Schock. What are the proper condition numbers of discretized ill-posed problems? *Linear Algebra Appl.*, 81:129–136, 1986.
- [22] G. M. Vaĭnikko and U. A. Khyamarik. Projection methods and self-regularization in ill-posed problems. *Izv. Vyssh. Uchebn. Zaved. Mat.*, 84(10):3–17, 1985.

WEIERSTRASS INSTITUTE FOR APPLIED ANALYSIS AND STOCHASTICS, MOHREN-  
STRASSE 39, D-10117 BERLIN, GERMANY

*E-mail address:* `mathe@wias-berlin.de`

FREE UNIVERSITY BERLIN, ARNIMALLEE 14, D-14195 BERLIN, GERMANY

*E-mail address:* `inia@mi.fu-berlin.de`