Generalized Local Regularization for Ill-Posed Problems

Patricia K. Lamm

Department of Mathematics
Michigan State University

AIP2009 July 22, 2009

Based on joint work with Cara Brooks, Zhewei Dai, and Xiaoyue Luo
Problem: Solve

\[ Au = f \]

for \( \bar{u} \in \text{dom}(A) \subset X \), where the “ideal” data \( f \in \mathcal{R}(A) \subset X \) is only approximately given.

We assume

- the operator \( A \) may be linear or nonlinear;
- \( X \) is a Hilbert or reflexive Banach space;
- Existence/uniqueness of solution \( \bar{u} \) for \( f \in \mathcal{R}(A) \), which does not depend continuously on \( f \);
- The data is only approximately given by

\[ f^\delta \in X, \text{ with } \| f - f^\delta \|_X < \delta. \]

(Precise conditions on \( A \) to be given later.)
Examples of (related) regularization methods

- Equation for 0th order Tikhonov regularization:
  
  Let $\alpha > 0$ and find $u^\delta_\alpha \in \text{dom}(A) \subset X$ minimizing
  \[
  \|Au - f^\delta\|_X^2 + \alpha \|u\|_X^2.
  \]
  
  If $A$ bounded linear, the regularized solution $u^\delta_\alpha$ satisfies
  \[
  A^*Au + \alpha u = A^*f^\delta.
  \]  

Features (for generic linear case):

Pros:

- Solution of (1) easily handled for $f^\delta \in X$
- “All-purpose” regularizer
- Easy to implement

Cons:

- Stability of $(A^*A)^{-1}$ vs $A^{-1}$
- $A$ nonlinear $\implies$ modify or discard (1)
- Ignores special structure of $A$
- Can be oversmoothing
Examples of (related) regularization methods

- Equation for Lavrent’ev (or simplified) regularization:

In contrast to the Tikhonov regularization equation for linear $A$,

$$A^*Au + \alpha u = A^*f^\delta,$$

the method of Lavrent’ev finds $u^\delta_\alpha \in \text{dom}(A)$ satisfying

$$Au + \alpha u = f^\delta.$$

Features:

Pros:
- Underlying operator is still $A$, so preserves structure of equation;
- Method the same for both linear and nonlinear problems
- Easy to implement

Cons:
- Solvability guaranteed only for certain $A$
- Numerical results not always satisfactory
Diagonal” quantities

Comparisons:

Both Tikhonov (linear case) and Lavrent’ev methods stabilize using the ”diagonal quantity” $\alpha u$.

But local regularization also involves a “diagonal quantity”:

- **Tikhonov and Lavrent’ev regularization:**
  * $\alpha u$ added to the underlying operator, which makes a small shift of singular values of the operator away from zero.

- **Local regularization:**
  * We first subtract a diagonal quantity from the operator . . . use this to construct a diagonal $\sim \alpha u$ to be added back in. $\Rightarrow \alpha u$ more closely linked to $A \Rightarrow$ can preserve structure
  * The subtraction effectively splits the underlying operator (As we’ll see later, the splitting should be related to the sensitivity of the data to local changes in the solution.)
Local regularization

- **The equation for local regularization:**

For each \( \alpha \in (0, \bar{\alpha}] \) and \( X_\alpha \subset X \), let \( u_\alpha^\delta \in \text{dom}(A_\alpha) \subset X_\alpha \) satisfy

\[
A_\alpha u + a_\alpha u = f_\alpha^\delta
\]

for given

- (nonlinear) operator \( A_\alpha : \text{dom}(A_\alpha) \subset X_\alpha \mapsto X \),
- parameter (function/operator) \( a_\alpha(\cdot) \), and for
- \( f_\alpha^\delta \in X_\alpha \) defined via

\[
f_\alpha^\delta = T_\alpha f^\delta,
\]

for a family \( \{ T_\alpha \} \) of bounded linear operators, \( T_\alpha : X \mapsto X_\alpha \),

\[
\| T_\alpha f \|_{X_\alpha} \leq \| f \|_X, \quad \text{uniformly in } f \in X.
\]

[E.g., \( T_\alpha = I \), or \( T_\alpha = \text{a (normalized) mollification operator,} \]

or a number of other choices.]
Compare Tikhonov, Lavrent’ev, and local regularization

Comments about the local regularization equation,

$$A_\alpha u + a_\alpha u = T_\alpha f^\delta$$

- Tikhonov and Lavrent’ev are special cases, with parameter-independent $A_\alpha$ and $X_\alpha = X$ for each:

<table>
<thead>
<tr>
<th>Method</th>
<th>$A_\alpha$</th>
<th>$a_\alpha$</th>
<th>$T_\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tikhonov:</td>
<td>$A^*A$</td>
<td>$\alpha$ (scalar)</td>
<td>$A^*$</td>
</tr>
<tr>
<td>Lavrent’ev:</td>
<td>$A$</td>
<td>$\alpha$ (scalar)</td>
<td>$I$</td>
</tr>
</tbody>
</table>

- In local regularization, $T_\alpha$ plays the role of “consolidating” a local piece of the function $f^\delta$:
  
  Tikhonov: $T_\alpha = A^*$ (not local at all)
  
  Lavrent’ev: $T_\alpha = I$ (may be too local)
Selection of $A_\alpha$, $a_\alpha$, and $T_\alpha$ for local regularization method:

**Motivation:** Look at what’s needed to show stability and convergence to $\bar{u}$ of $u_\delta$, the solution of the (generic) equation

$$A_\alpha u + a_\alpha u = T_\alpha f_\delta.$$  \hspace{1cm} (2)

- **To simplify the motivation:**
  - Take operators $A$, $A_\alpha$ bounded linear, $\alpha > 0$.
  - Take $X_\alpha = X$, $\alpha > 0$.

- **Will make comparisons with:**
  - stability/convergence results for Tikhonov/Lavrent’ev (since these methods also take the form of equation (2)).
Stability of the three methods

(1) **Stability for methods based on** \( \mathcal{A}_\alpha u + a_\alpha u = T_\alpha f^\delta \):

Want: solution \( u^\delta_\alpha \) **unique** + **depends continuously on data** \( f^\delta \).

\[ \implies \text{need } (\mathcal{A}_\alpha + a_\alpha I) \text{ invertible with bounded inverse.} \]

- **This is easy** for Tikhonov regularization:

  \[ \mathcal{A}_\alpha = \mathcal{A}^* \mathcal{A} \text{ automatically nonnegative self-adjoint.} \]

- **Not automatic** for Lavrent'ev or local regularization:

  Need added conditions on \( \mathcal{A}_\alpha \).

  E.g., **assume** \( \mathcal{A}_\alpha \) nonnegative self-adjoint.

**Note difference:**

- For Lavrent'ev \( \implies \) assumption is on \( \mathcal{A}_\alpha = \mathcal{A} \);
- For local reg. \( \implies \) assumption is on \( \mathcal{A}_\alpha \) (we construct).
Convergence for the three methods

So assume we already have stability, or that

\[ \| (A_\alpha + a_\alpha I)^{-1} \| = O \left( \frac{1}{c(\alpha)} \right) \]

where \( c(\alpha) \to 0 \) as \( \alpha \to 0 \); here \( \| \cdot \| \) denotes the operator norm.

(2) Convergence for methods based on \( A_\alpha u + a_\alpha u = T_\alpha f^\delta \):

Must show:

For each \( f \in \mathcal{R}(A) \) there is a choice of \( \alpha = \alpha(\delta, f^\delta) \) for which

\[ \alpha(\delta, f^\delta) \to 0 \quad \text{as} \quad \delta \to 0, \]

and for which

\[ u^\delta_{\alpha(\delta, f^\delta)} \to \bar{u} \quad \text{as} \quad \delta \to 0, \]

for all \( f^\delta \in X \) satisfying \( \| f - f^\delta \|_X \leq \delta \).
Verification of **convergence** of \( u_\delta^\alpha \) to \( \bar{u} \), where

\[
A_\alpha u_\delta^\alpha + a_\alpha u_\delta^\alpha = T_\alpha f^\delta. \tag{3}
\]

with \( A_\alpha, a_\alpha, T_\alpha \) given:

- **First** note \( \bar{u} \in X \) uniquely satisfies

  \[
  A\bar{u} = f, \quad \implies T_\alpha A\bar{u} = T_\alpha f, \quad \implies A_\alpha \bar{u} + a_\alpha \bar{u} = T_\alpha f + (A_\alpha \bar{u} + a_\alpha \bar{u} - T_\alpha A\bar{u}). \tag{4}
  \]

- **So subtract:** equation \((3)-(4)\) to obtain

  \[
  (A_\alpha + a_\alpha I)(u_\delta^\alpha - \bar{u}) = T_\alpha d^\delta - (A_\alpha \bar{u} + a_\alpha \bar{u} - T_\alpha A\bar{u}),
  \]

  where \( d^\delta \equiv (f^\delta - f) \in X \) satisfies \( \|d^\delta\|_X \leq \delta \).
We then have

\[
(u_\delta - \bar{u}) = (A_\alpha + a_\alpha I)^{-1} T_\alpha d_\delta + (A_\alpha + a_\alpha I)^{-1} [(T_\alpha A\bar{u} - A_\alpha \bar{u}) - a_\alpha \bar{u}],
\]

where

\[
\left\| (A_\alpha + a_\alpha I)^{-1} T_\alpha d_\delta \right\|_X = O\left(\frac{\delta}{c(\alpha)}\right).
\]

**Goal:** select \( \alpha = \alpha(\delta, f_\delta) \) so that \( \alpha(\delta, f_\delta) \to 0 \) as \( \delta \to 0 \), and so that both terms above in \( (u_\delta - \bar{u}) \) go to zero as \( \delta \to 0 \). That is,

- **The term due to data error goes to zero**, i.e.,

\[
\frac{\delta}{c(\alpha(\delta, f_\delta))} \to 0 \quad \text{as} \quad \delta \to 0.
\]

- **The term due to method error goes to zero:**

\[
(A_\alpha + a_\alpha I)^{-1} (D_\alpha \bar{u} - a_\alpha \bar{u}) \to 0 \quad \text{as} \quad \delta \to 0.
\]

Here we have defined \( D_\alpha = T_\alpha A - A_\alpha \).
Note:

- There is generally little difficulty in making an *a priori* choice of \( \alpha = \alpha(\delta, f^\delta) \) so that \( \alpha(\delta, f^\delta) \to 0 \) and

\[
\frac{\delta}{c(\alpha(\delta, f^\delta))} \to 0 \quad \text{as} \quad \delta \to 0.
\]

- The problem then is to show that for a selection of \( \alpha \)-values with \( \alpha \to 0 \), the operator \( (A_\alpha + a_\alpha I) \) is continuously invertible, and

\[
(A_\alpha + a_\alpha I)^{-1}(D_\alpha \bar{u} - a_\alpha \bar{u}) \to 0, \quad \text{as} \quad \alpha \to 0.
\]

It is well-known that these last statements are true (for linear \( A \)) in the case of:

- **Tikhonov regularization**, for \( A \) bounded;

- **Lavrent’ev regularization**, for a narrower class of operators \( A \) (e.g. \( A \) nonnegative self-adjoint).
For **local regularization**, we’ll use the desired convergence condition,

\[(A_\alpha + a_\alpha I)^{-1} (D_\alpha \bar{u} - a_\alpha \bar{u}) \to 0 \text{ as } \alpha \to 0, \quad (5)\]

to motivate the selection of \(A_\alpha\) and \(a_\alpha\). (Recall \(D_\alpha \equiv T_\alpha A - A_\alpha\).)

- That is, \(\bar{u}\) satisfies
  \[A\bar{u} = f\]
  \[\implies T_\alpha A\bar{u} = T_\alpha f,\]

  where the operator \(T_\alpha A\) may be split as follows:

  \[T_\alpha A = A_\alpha + (T_\alpha A - A_\alpha)\]

  \[\implies T_\alpha A = A_\alpha + D_\alpha.\]

- Using (5), this splitting of \(T_\alpha A\) needs to be such that (roughly)

  \[D_\alpha \bar{u} \approx a_\alpha \bar{u} \quad \text{for } \alpha \approx 0,\]

  for some choice of \(a_\alpha\), i.e., \(D_\alpha \approx “\text{diagonal}”\) when applied to \(\bar{u}\).
So in **local regularization**:

- **We first select** $T_\alpha$ (selection criteria to be discussed later).
- **We then select** $A_\alpha$ and $a_\alpha$ (and define $D_\alpha \equiv T_\alpha A - A_\alpha$) so that for $\alpha > 0$ we have:
  1. $(A_\alpha + a_\alpha I)$ continuously invertible, and
  2. the operator $(D_\alpha - a_\alpha I)$ satisfies

$$\left\| (A_\alpha + a_\alpha I)^{-1} (D_\alpha - a_\alpha I) \tilde{u} \right\|_X \to 0 \quad \text{as} \quad \alpha \to 0.$$ 

Roughly, we take the modified equation $T_\alpha A u = T_\alpha f$, and split its leading operator via $T_\alpha A = A_\alpha + D_\alpha$, where

- $A_\alpha = \text{the “bigger, nearly continuously invertible” part of } T_\alpha A$,
- $D_\alpha = \text{the “smaller, diagonal-like” part of } T_\alpha A$. 

The above statement of convergence, namely,

\[ \left\|(A_\alpha + a_\alpha I)^{-1} (D_\alpha - a_\alpha I) \bar{u} \right\|_\chi \to 0 \text{ as } \alpha \to 0. \]  \hspace{1cm} (6)

is obtained for Tikhonov & Lavrent’ev, without splitting \( T_\alpha A \):

- **Tikhonov**: \( T_\alpha A = A_\alpha = A^*A, \quad D_\alpha = 0, \quad a_\alpha = \alpha. \)
- **Lavrent’ev**: \( T_\alpha A = A_\alpha = A, \quad D_\alpha = 0, \quad a_\alpha = \alpha. \)

In these cases the convergence as \( \alpha \to 0 \) in (6) takes the form (respectively) of

\[ \left\|(A^*A + \alpha I)^{-1} (-\alpha \bar{u}) \right\|_\chi \to 0, \quad \left\|(A + \alpha I)^{-1} (-\alpha \bar{u}) \right\|_\chi \to 0. \]

**In contrast:** With *local regularization* we can obtain a stronger statement of convergence (than (6)) for many problems.
Stronger statement of convergence: A stronger result than

\[ \left\| (A_\alpha + a_\alpha I)^{-1} (D_\alpha \tilde{u} - a_\alpha \bar{u}) \right\|_X \to 0 \quad \text{as} \quad \alpha \to 0, \]

is obtained if the method satisfies

\[ \left\| D_\alpha \tilde{u} - a_\alpha \bar{u} \right\|_X = o\left( c(\alpha) \right) \quad \text{as} \quad \alpha \to 0 \quad \text{(7)} \]

(recall that \( \left\| (A_\alpha + a_\alpha I)^{-1} \right\| = O \left( \frac{1}{c(\alpha)} \right) \)).

- The stronger result (7) cannot hold for Tikhonov or Lavrent’ev (in general). For these methods the usual case is

\[ \left\| (A_\alpha + \alpha I)^{-1} \right\| = O \left( \frac{1}{\alpha} \right) \quad \text{as} \quad \alpha \to 0, \]

i.e., \( c(\alpha) = \alpha \), but \( \left\| D_\alpha \tilde{u} - a_\alpha \bar{u} \right\|_X = \left\| -\alpha \bar{u} \right\|_X = O(\alpha) \)

\[ \implies \left\| D_\alpha \tilde{u} - a_\alpha \bar{u} \right\|_X \neq o\left( c(\alpha) \right) \quad \text{as} \quad \alpha \to 0. \]
Local regularization: the selection of $T_\alpha$

How to select $T_\alpha$ in the method of local regularization?

- We will select $T_\alpha$ in order to:
  
  - Facilitate a regularization method which is “localized” in an appropriate sense.
  
  - Work toward the goal of finding a method which is easier/faster than classical regularization methods.
  
  - Effectively utilize information about the sensitivity of the data $f$ to (local) changes in $u$ (for $f$, $u$ satisfying the $Au = f$).

We will demonstrate the selection of $T_\alpha$, as well as of $A_\alpha$ and $a_\alpha$, through some examples.
Examples of local regularization

Three motivating examples . . . from where local regularization theory is the most complete, i.e., for Volterra operators.

1 Linear problem on $X = L^p(0, 1), 1 < p \leq \infty$:

$$
\int_0^t k(t, s)u(s) \, ds = f(t), \quad a.e. \quad t \in [0, 1].
$$

2 Nonlinear Hammerstein problem on $X = C[0, 1]$:

$$
\int_0^t k(t, s)g(s, u(s)) \, ds = f(t), \quad a.e. \quad t \in [0, 1],
$$

under (standard) conditions on $g$.

3 Nonlinear autoconvolution problem, $X = L^2(0, 1), C[0, 1]$:

$$
\int_0^t u(t - s)u(s) \, ds = f(t), \quad a.e. \quad t \in [0, 1].
$$
Linear Volterra example

- First consider the **linear** Volterra problem

  \[ A u(t) = \int_0^t k(t, s) u(s) \, ds = f(t), \quad \text{a.e. } t \in [0, 1], \]

  in the special case of \( X = L^2(0, 1) \), and for \( k \in L^2([0, 1] \times [0, 1]) \).

- To determine an appropriate choice for \( T_\alpha \), we first look at **sensitivities** of the data \( f = A u \) to **local parts** of \( u \):

  \( A \) is bounded linear on \( X \) \( \implies \) the Frechet differential of \( A \) at \( u \in X \) with increment \( \Delta u \in X \) is given by

  \[ DA(u; \Delta u) = A \Delta u, \]

  i.e.,

  \[ DA(u; \Delta u)(t) = \int_0^t k(t, s) \Delta u(s) \, ds, \quad \text{a.e. } t \in [0, 1]. \]
Graphs of: (1) $\Delta u =$ “unit impulse” of width .1, (1st & 3rd rows of graphs above), and (2) the response $DA(u; \Delta u)$ (2nd & 4th rows), immediately below each $\Delta u$. (Here the kernel for $A$ is $k(t, s) = 1.$)
Combined graphs of $\mathcal{D}A(u; \Delta u)$ for each “unit impulse” $\Delta u$ of width .1 (As before, the kernel for $A$ is $k(t, s) = 1$.)
Sensitivities of data for linear Volterra problem

Combined graphs of $\mathcal{D}A(u; \Delta u)$ for each “unit impulse” $\Delta u$ of width .1 (Here the kernel for $A$ is $k(t, s) = t - s$.)
Combined graphs of $\mathcal{D}A(u; \Delta u)$ for each “unit impulse” $\Delta u$ of width .1 (Here the kernel for $A$ is $k(t, s) = (t - s)^2$.)
Combined graphs of $\mathcal{D}A(u; \Delta u)$ for each “unit impulse” $\Delta u$ of width $.1$ (Here the kernel for $A$ is $k(t, s) = (t - s)^3$.)
for the linear Volterra problem, the solution $u$ at $t$ most influences data $f$ on the interval $[t, t + \alpha]$, for some $\alpha > 0$ small.

The length $\alpha$ of the interval is the regularization parameter. (Alternatively, use two parameters – one for sensitivity region and another for regularization. E.g., see [PKL & Dai, 2005])

- One possible choice of $T_\alpha$:
  \[
  T_\alpha f(t) = \frac{1}{\alpha} \int_0^\alpha f(t + \rho) \, d\rho, \quad a.e. \ t \in [0, 1 - \alpha],
  \]
  (averaging function on $[t, t + \alpha]$).

- More generally,
  \[
  T_\alpha f(t) = \frac{\int_0^\alpha f(t + \rho) \, d\eta_\alpha(\rho)}{\int_0^\alpha d\eta_\alpha(\rho)}, \quad a.e. \ t \in [0, 1 - \alpha],
  \]
  for suitable measure $\eta_\alpha$ on $[0, \alpha]$. 

For example, we will apply the simple averaging $T_{\alpha}$ to the original linear Volterra equation. Assume $\alpha \in (0, \bar{\alpha}]$, for $\bar{\alpha} > 0$ small:

- Start with the original equation $Au = f$, or

$$\int_0^t k(t, s)u(s) \, ds = f(t), \quad \text{a.e. } t \in [0, 1].$$

- Next compute $T_{\alpha}Au = T_{\alpha}f$, for a.e. $t \in [0, 1 - \alpha]$,

$$\frac{1}{\alpha} \int_0^\alpha \int_0^{t + \rho} k(t + \rho, s)u(s) \, ds \, d\rho = \frac{1}{\alpha} \int_0^\alpha f(t + \rho) \, d\rho.$$

- **Now split**: $T_{\alpha}A = A_{\alpha} + D_{\alpha}$, where $D_{\alpha}$ is nearly "diagonal":

$$T_{\alpha}Au(t) = \frac{1}{\alpha} \int_0^\alpha \int_0^t k(t + \rho, s)u(s) \, ds \, d\rho$$

$$+ \frac{1}{\alpha} \int_0^\alpha \int_t^{t + \rho} k(t + \rho, s)u(s) \, ds \, d\rho.$$
Selection of $A_\alpha$, $a_\alpha$ for linear Volterra problem.

So the splitting of $T_\alpha A$ is thus given by

$$T_\alpha A = A_\alpha + D_\alpha,$$

where, for a.e. $t \in [0, 1 - \alpha]$,

$$A_\alpha u(t) = \int_0^t k_\alpha(t, s)u(s) \, ds,$$

$$D_\alpha u(t) = \frac{1}{\alpha} \int_0^\alpha \int_t^{t+\rho} k(t + \rho, s)u(s) \, ds \, d\rho,$$

and $k_\alpha(t, s) \equiv \frac{1}{\alpha} \int_0^\alpha k(t + \rho, s) \, d\rho$.

$$\implies A_\alpha \text{ is of the same form as } A.$$

- Finally: select $a_\alpha = a_\alpha(t)$ so that $D_\alpha \tilde{u} \approx a_\alpha \tilde{u}$ for $\alpha$ small.
Selection of $a_\alpha$ for linear Volterra problem.

• Need $a_\alpha = a_\alpha(t)$ so that, for a.e. $t \in [0, 1 - \alpha]$,

$$D_\alpha u(t) = \frac{1}{\alpha} \int_0^\alpha \int_t^{t+\rho} k(t + \rho, s) u(s) \, ds \, d\rho \approx a_\alpha(t) u(t),$$

when $u = \bar{u}$.

• If we approximate $u$ on the local interval $[t, t + \alpha]$ by the constant $u(t)$, then we have (formally)

$$D_\alpha u(t) \approx \left( \frac{1}{\alpha} \int_0^\alpha \int_t^{t+\rho} k(t + \rho, s) \, ds \, d\rho \right) u(t) = a_\alpha(t) u(t),$$

for

$$a_\alpha(t) = \frac{1}{\alpha} \int_0^\alpha \int_t^{t+\rho} k(t + \rho, s) \, ds \, d\rho, \quad \text{a.e. } t \in [0, 1 - \alpha].$$
To summarize: The particular construction we have chosen leads to a local regularization equation for the linear Volterra problem given, for noisy data $f_\delta$, by

$$A_\alpha u + a_\alpha u = f_\delta \quad \text{on} \quad X_\alpha = L^2(0, 1 - \alpha),$$

where for a.e. $t \in [0, 1 - \alpha]$,

$$f_\alpha^\delta(t) = T_\alpha f^\delta(t) = \frac{1}{\alpha} \int_0^\alpha f^\delta(t + \rho) \, d\rho,$$

$$A_\alpha u(t) = \int_0^t k_\alpha(t, s)u(s) \, ds$$

$$= \int_0^t \left( \frac{1}{\alpha} \int_0^\alpha k(t + \rho, s) \, d\rho \right) u(s) \, ds,$$

$$a_\alpha(t) = \frac{1}{\alpha} \int_0^\alpha \int_0^{t+\rho} k(t + \rho, s)ds \, d\rho.$$
Result: For a large class of kernels $k$ and for $\alpha$ sufficiently small,

1. The quantity $a_\alpha$,

\[ a_\alpha(t) = \frac{1}{\alpha} \int_0^\alpha \int_t^{t+\rho} k(t+\rho, s) ds \, d\rho, \]

satisfies

\[ 0 < c_1 \, c(\alpha) \leq a_\alpha(t) \leq c_2 \, c(\alpha), \]

a.e. $t \in [0, 1 - \alpha]$, for some $c(\alpha) > 0$ with the property that

\[ c(\alpha) \to 0 \quad \text{as} \quad \alpha \to 0; \quad \text{and,} \]

2. $(A_\alpha + a_\alpha I)$ is invertible with

\[ \| (A_\alpha + a_\alpha I)^{-1} \| = \mathcal{O} \left( \frac{1}{c(\alpha)} \right) \quad \text{as} \quad \alpha \to 0. \]
**Question:** So why does the local regularization method satisfy the stronger convergence result?

Recall that in this case we need to verify

\[ \| (D_\alpha \tilde{u} - a_\alpha \tilde{u}) \|_{X_\alpha} = o \left( c(\alpha) \right) \quad \text{as} \quad \alpha \to 0. \]

- In contrast to Tikhonov and Lavrent’ev, where \( X_\alpha = X \) and

\[ \| (D_\alpha \tilde{u} - \alpha \tilde{u}) \|_{X_\alpha} = \| 0 - \alpha \tilde{u} \|_{X_\alpha} \neq o(\alpha), \]

- the method of local regularization gives

\[
D_\alpha \tilde{u}(t) - a_\alpha(t)\tilde{u}(t) = \frac{1}{\alpha} \int_0^\alpha \int_t^{t+\rho} k(t+\rho, s) \tilde{u}(s) \, ds \, d\rho - \left( \frac{1}{\alpha} \int_0^\alpha \int_t^{t+\rho} k(t + \rho, s) \, ds \, d\rho \right) \tilde{u}(t).
\]
So for **local regularization** we have

\[
D_\alpha \bar{u}(t) - a_\alpha(t)\bar{u}(t) = \frac{1}{\alpha} \int_0^\alpha \int_t^{t+\rho} k(t + \rho, s) (\bar{u}(s) - \bar{u}(t)) \, ds \, d\rho,
\]
a.e. \( t \in [0, 1 - \alpha) \), and thus as \( \alpha \to 0 \),

\[
\| D_\alpha \bar{u} - a_\alpha \bar{u} \|_{X_\alpha} = \mathcal{O} \left( \| a_\alpha(\cdot) \|_{L_\infty} \right) \cdot \| S_\alpha \bar{u} - \bar{u} \|_{X_\alpha} = \mathcal{O} \left( c(\alpha) \right) \cdot \| S_\alpha \bar{u} - \bar{u} \|_{X_\alpha}
\]

where

\[
S_\alpha \bar{u}(t) = \frac{1}{\alpha} \int_0^\alpha \bar{u}(t + s) \, ds.
\]

Note that \( S_\alpha \bar{u}(t) \) is a special case of the Hardy-Littlewood maximal function for \( \bar{u} \) at \( t \).
So we have

$$
\| D_\alpha \bar{u} - a_\alpha \bar{u} \|_{X_\alpha} = O \left( c(\alpha) \right) \cdot \| S_\alpha \bar{u} - \bar{u} \|_{X_\alpha},
$$

and it follows from the Lebesgue differentiation theorem that

$$
\| S_\alpha \bar{u} - \bar{u} \|_{X_\alpha} \to 0 \quad \text{as} \quad \alpha \to 0,
$$

for any $\bar{u} \in X$. Thus

$$
\| D_\alpha \bar{u} - a_\alpha \bar{u} \|_{X_\alpha} = o \left( c(\alpha) \right) \quad \text{as} \quad \alpha \to 0,
$$

and the stronger convergence result is obtained.

For smoother $\bar{u}$, it’s obvious that from this same estimate we can obtain a rate of convergence.
More general results for the linear Volterra problem


- $f \in \mathcal{R}(A) \subset X = L^p[0, 1], \ f^\delta \in L^p[0, 1], \ 1 < p \leq \infty$.
- For $\nu$-smoothing problems, with a large class of measures (depending on $\nu$) used to define suitable operators $T_\alpha$.
- Convergence as $\delta \to 0$ of $u^\delta_\alpha$ to $\bar{u}$ in $X_\alpha = L^p[0, 1-\alpha]$ for a priori choices of $\alpha = \alpha(\delta, f^\delta)$.
- Optimal rates of convergence for the $\nu$-smoothing problem.
- A discrepancy principle of the form

$$d(\alpha) = \tau \delta,$$

where

$$d(\alpha) \equiv \left( \alpha^m \frac{\|f^\delta\|_X}{\|T_\alpha f^\delta\|_{X_\alpha}} \right) \|A_\alpha u^\delta_\alpha - T_\alpha f^\delta\|_{X_\alpha}.$$
Nonlinear Volterra problems

Nonlinear Hammerstein problem:

Start with the usual local regularization equation

\[ A_\alpha u + a_\alpha u = T_\alpha f^\delta, \]

except now \( A_\alpha \) and the term \( a_\alpha u \) are both nonlinear in \( u \).

• How to handle the nonlinearity?
  Use the local regularization interval to facilitate linearization of the nonlinear problem \([\text{Luo thesis } '07; \text{ Brooks, PKL, Luo } '09]\)
  i.e., for \( t > 0 \), let \( \tau = \min\{t, \alpha\} \), and linearize \( a_\alpha = a_\alpha(t, u(t)) \) about the prior \( u \)-value \( u(t - \tau) \).

• Numerical implementation of the linearized problem:
  \[ u = \left( T_\alpha f^\delta - A_\alpha u \right)/a_\alpha(\cdot), \quad t > 0 \]
  \( \implies \) solve a linear problem for \( u(t) \) sequentially, since \( A_\alpha u \) and (the linearized) \( a_\alpha \) only involve values of \( u \) prior to \( t \).
Nonlinear autoconvolution problem:

For the **autoconvolution problem** (as with Hammerstein), the usual local regularization equation

\[ A_\alpha u + a_\alpha u = T_\alpha f^\delta \]  

(8)

involves *nonlinearities* in both \( A_\alpha \) and the term \( a_\alpha u \).

- **First:** Solve (8) for \( u(t) \), \( t \in [0, \alpha] \) \( (\alpha > 0 \text{ small}) \):
  * Nonlinear equation (8) not too difficult to solve on \([0, \alpha]\);
  * Need only a reasonable approximation to \( u \) on \([0, \alpha]\) for convergence

- **Then:** Solve (8) for \( u(t) \), \( t \in (\alpha, 1] \) using **linear** methods:
  * The (approximate) solution \( u \) on \((\alpha, 1]\) can be found **sequentially** from (8) using **linear** techniques.

- Thus, linearization of (8) not required in practice for the autoconvolution problem (but used as tool in convergence theory).

[Dai thesis, 2006; Dai & PKL, 2008; Brooks, Dai, & PKL, 2009]
• **Numerical results for nonlinear Volterra problems:**

  In general,
  - Qualitatively equivalent to Tikhonov solution, except not oversmoothed; considerably better than Lavrent’ev solution.
  - Many times faster than a nonlinear Tikhonov (iterative) approach in practice.

  (See above references.)

• **Other non-Volterra work:**

  - **2-D image deblurring**
    Older work: [Cui thesis], [Cui, Lamm, Scofield, ’07].
    Newer results (to appear) are based on generalized local regularization approach \(\Rightarrow\) less computationally intensive.

  - **General \(n\)-dimensional problems governed by a nonlinear monotone operator** (to appear; from generalized theory).
Sample numerical results

- Tikhonov: 68% rel error in u
- Lavrentiev: 78% rel error in u
- Local reg. $\alpha=0.06$: 6% rel error in u

- Tikhonov: 40% rel error in u
- Lavrentiev: 66% rel error in u
- Local reg. $\alpha=0.05$: 4% rel error in u

- Tikhonov: 23% rel error in u
- Lavrentiev: 53% rel error in u
- Local reg. $\alpha=0.04$: 3% rel error in u

- Tikhonov: 14% rel error in u
- Lavrentiev: 88% rel error in u
- Local reg. $\alpha=0.08$: 10% rel error in u

- 16% rel data error
- 8% rel data error
- 4% rel data error
- 2% rel data error