2023年PQCD研讨会 2023年11月04日

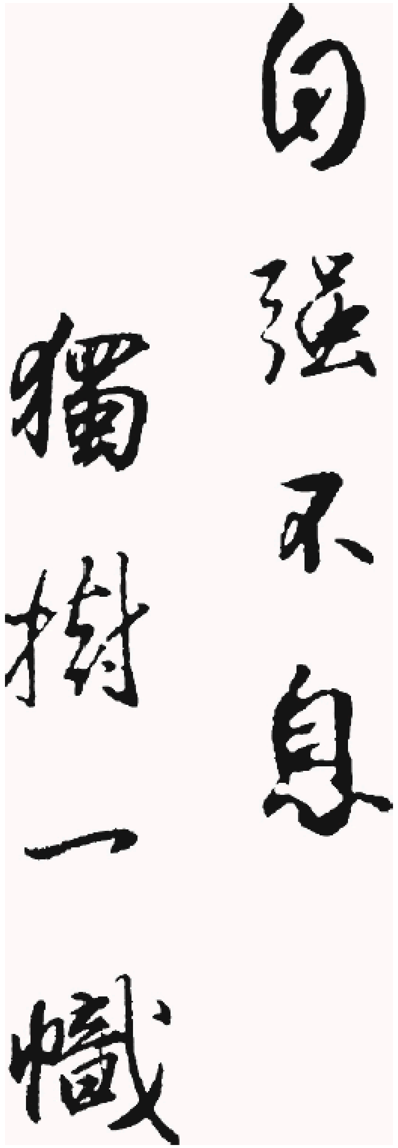
Mathematical foundation of the inverse problem approach

Ao-Sheng Xiong

Lanzhou University

Based on [**ASX**, Ting Wei(魏婷), Fu-Sheng Yu(于福升)]
arXiv: 2211.13753

目录



1 研究动机

重大科学难题：非微扰物理量难以计算，是世纪难题

夸克禁闭和新物理……

现有的非微扰方法各有优缺点

格点QCD(LQCD)：第一性原理的计算方法，但需要超级计算机且对激发态等物理量计算难度大。

其他方法：如QCD求和规则、Dyson-Schwinger方程等都有模型依赖性。唯象方法预言能力有限。

发展新的方法“反问题方法”计算非微扰物理量

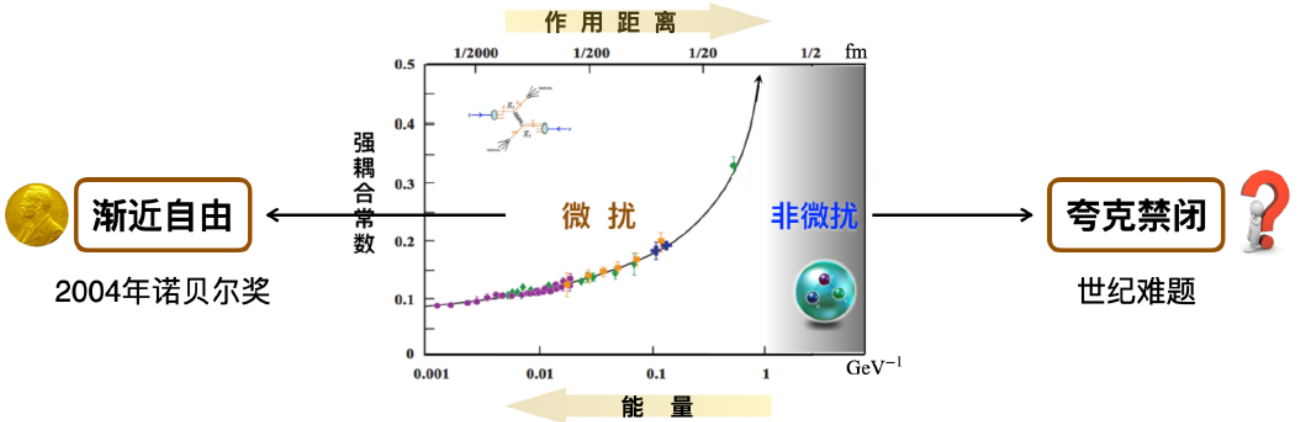
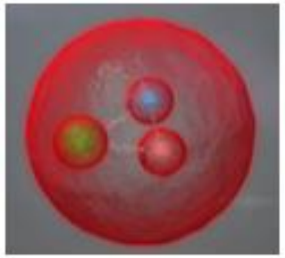
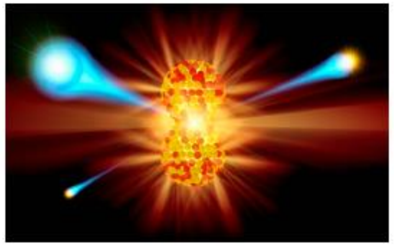
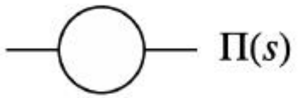


图1：高能微扰与低能非微扰



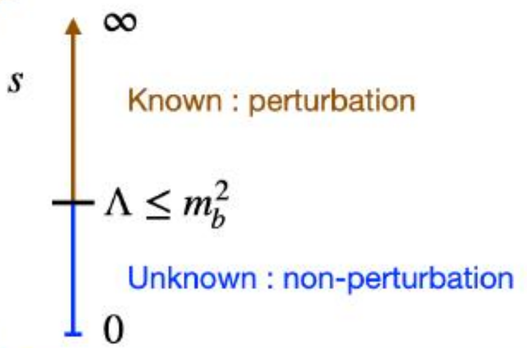


基于量子场论的色散关系，通过已知的微扰计算反解未知的非微扰物理量：



Dispersion Relation:

$$\mathcal{R}e[\Pi(s)] = \frac{1}{\pi}\mathcal{P}\int_0^\infty \frac{\mathcal{I}m[\Pi(s')]}{s-s'}ds'$$



If $s > \Lambda$, $\mathcal{P}\int_0^\Lambda \frac{\mathcal{I}m[\Pi(s')]}{s-s'}ds' = \pi\mathcal{R}e[\Pi(s)] - \mathcal{P}\int_\Lambda^\infty \frac{\mathcal{I}m[\Pi(s')]}{s-s'}ds'$

To be solved

calculable

H.N.Li, H.U, F.R.Xu, **F.S.Yu**, Phys.Lett.B 810(2020)

$$Kf = \int_a^b \frac{f(x)}{y-x}dx = g(y), y \in [c,d], c > b$$

已知积分结果求被积函数需要用到反问题理论。

2 反问题介绍

反问题理论是成熟的数学分支领域：

与反问题对比的是正问题；

1950年左右发展至今；

在众多重要的前沿科学领域中广泛运用；

线性方程组的求解、地质探勘、缪子成像、CT……

If $s > \Lambda$, $$\mathcal{P}\int_0^\Lambda \frac{\mathcal{I}m[\Pi(s')]}{s-s'}ds' = \pi\mathcal{R}e[\Pi(s)] - \mathcal{P}\int_\Lambda^\infty \frac{\mathcal{I}m[\Pi(s')]}{s-s'}ds'$$

To be solved

calculable

$$\int_a^b \frac{f(x)}{y-x}dx = g(y), y \in [c,d], c > b$$

微扰论

+

反问题理论

色散关系
反问题方法

非微扰
物理量

$$\int_a^b \frac{f(x)}{y-x}dx = g(y), y \in [c,d], c > b$$

色散关系的算子K: $F(L^2(a,b)) \to G(L^2(c,d))$

正问题：$Kf = g$

反问题：$f = K^{-1}g$

适定性是研究反问题的基础：

Define: *The operator equation (3.1) is called well-posed if the following holds [8]:*

1.*Existence: For every* $g \in G$ *there is (at least one)* $f \in F$ *such that* $Kf = g$;

2.*Uniqueness: For every* $g \in G$ *there is at most one* $f \in F$ *with* $Kf = g$;

3.*Stability: The solution* f *depends continuously on* g; *that is, for every sequence* $(f_n) \subset F$ *with* $Kf_n \to Kf (n \to \infty)$, *it follows that* $f_n \to f (n \to \infty)$

若上述有一个条件不满足，就称该问题是不适定问题。

存在唯一性等价于逆算子$K^{-1}: G \to F$存在，稳定性等价于逆算子K^{-1}连续（有界）。

反问题的两个简单例子：

- $Kx=10^{-6}x=y, x\in R, y\in R$。其反问题是$x=K^{-1}y=10^{6}y$。

若$y_1=1, y_2=1.01$，其对应的解$x_1=10^6, x_2=10^6+10^4$

- 线性方程组的一般形式为：$A_{n\times n}x_{n\times 1}=b_{n\times 1}$。

上述方程组的解存在当且仅当$r_A=r_{(A,b)}$；当解存在时，解唯一当且仅当$r_A=n$；

$$\begin{cases} 2x_1+3x_2=8 \\ 2x_1+3.00001x_2=8.00002 \end{cases}$$

解为$x_1=1, x_2=2$;

当对方程组的右端加上微小的误差时

$$\begin{cases} 2x_1+3x_2=8 \\ 2x_1+3.00001x_2=8.00003 \end{cases}$$

解却变成 $x_1=-0.5$，$x_2=3$ 。解发生了巨大的变化，其逆矩阵A^{-1}行列式很大。

3 适定性研究

解的存在性：

$$\int_a^b \frac{f(x)}{y-x}dx = g(y), y \in [c,d], c > b$$

R(K) is the analytic functions

$$\begin{aligned} Kf &= \int_a^b \frac{f(x)}{y-x}dx = \int_a^b \frac{1}{y}\frac{f(x)}{1-\frac{x}{y}}dx \\ &= \frac{1}{y}\int_a^b \sum_{k=0}^{\infty}(\frac{x}{y})^k f(x)dx = \sum_{k=0}^{\infty}\frac{1}{y^{k+1}}\int_a^b x^k f(x)dx \end{aligned}$$

where the last equal sign uses the control convergence theorem:

$$\begin{aligned} \sum_{k=0}^{\infty}\frac{1}{y^{k+1}}\int_a^b x^k f(x)dx &\le \sum_{k=0}^{\infty}\frac{1}{y^{k+1}}\int_a^b x^{2k}dx||f(x)||_{L^2(a,b)} \\ &\le \sum_{k=0}^{\infty}\frac{1}{y^{k+1}}b^k\sqrt{b-a}||f(x)||_{L^2(a,b)} \\ &\le \sum_{k=0}^{\infty}\frac{1}{y^{k+1}}(\frac{b}{c})^k||f(x)||_{L^2(a,b)} \le \infty \end{aligned}$$

Thus, the R(K) is the analytic functions $\in [c,d]$

3 适定性研究

解的存在性：☺ 解的唯一性：☺

$$\int_a^b \frac{f(x)}{y-x}dx = g(y), y \in [c,d], c > b$$

Theorem 3.3. *Suppose that $f_1(x)$, $f_2(x) \in L^2(a,b)$. If $Kf_1 = Kf_2 = g(y)$, $y \in [c,d]$, then we have $f_1(x) = f_2(x)$, a. e. $x \in [a,b]$.*

Proof. Since K is a linear operator, we know that $Kf_1 - Kf_2 = K(f_1 - f_2) = 0$. Therefore, in order to prove $f_1(x) = f_2(x)$, a. e. $x \in [a,b]$, we just need to prove that $Kf = 0$ implies $f(x) = 0$, a. e. $x \in [a,b]$.

It is easy to obtain that $Kf = \int_a^b \frac{1}{y-x} f(x)dx = \int_a^b \left(\frac{1}{y}\sum_{k=0}^{\infty}(\frac{x}{y})^k\right) f(x)dx$. Since $x \in [a,b]$, $y \in [c,d]$, $c > b$, we know $|\frac{x}{y}| \le |\frac{b}{c}| < 1$, which implies that $\left|\sum_{k=0}^{\infty}(\frac{x}{y})^k f(x)\right| \le \sum_{k=0}^{\infty}(\frac{b}{c})^k |f(x)|$ for all $x \in [a,b]$. Combined with $\int_a^b |f(x)|dx < +\infty$ and the control convergence theorem, we have

$$y\int_a^b \frac{1}{y-x} f(x)dx = \sum_{k=0}^{\infty} \frac{1}{y^k}\int_a^b x^k f(x)dx = 0, \quad y \in [c,d]. \tag{3.4}$$

If $d = +\infty$, by using (3.4), we have

$$\int_a^b f(x)dx + \frac{1}{y}\int_a^b xf(x)dx + \cdots + \frac{1}{y^k}\int_a^b x^k f(x)dx + \cdots = 0, \quad y \in (c, +\infty). \tag{3.5}$$

Letting $y \to +\infty$ in (3.5), we have $\int_a^b f(x)dx = 0$. Then multiplying y on both sides of (3.5) and letting $y \to +\infty$, we also have $\int_a^b xf(x)dx = 0$. Repeating above process, we can obtain that

$$\int_a^b x^k f(x)dx = 0, \quad k = 0, 1, 2, \cdots. \tag{3.6}$$

By using (3.6), we know that $\int_a^b f(x)Q_n(x)dx = 0$. Combined with the Cauchy inequality, we have

$$\begin{aligned}
\|f\|^2_{L^2(a,b)} &= \int_a^b f^2(x)dx = \int_a^b \left(f^2(x) - f(x)Q_n(x)\right)dx \\
&\le \int_a^b |f(x)| \cdot |f(x) - Q_n(x)|dx \\
&\le \left(\int_a^b f^2(x)dx\right)^{\frac{1}{2}} \left(\int_a^b |f(x) - Q_n(x)|^2 dx\right)^{\frac{1}{2}} \\
&= \|f\|_{L^2(a,b)} \|f - Q_n\|_{L^2(a,b)} \\
&\le (\epsilon + \epsilon\sqrt{b-a})\|f\|_{L^2(a,b)},
\end{aligned}$$

which implies that $\|f\|_{L^2(a,b)} \le \epsilon + \epsilon\sqrt{b-a}$.

Letting $\epsilon \to 0$, we have $\|f\|_{L^2(a,b)} = 0$, i. e. $f(x) = 0$, a. e. $x \in [a,b]$. The proof is completed. □

3 适定性研究

解的存在性： **解的唯一性：** **解的稳定性：**

We show the instability of the inverse problem of dispersion relation by the special case. Taking $a = 0$, $b = 1$, $c = 2$, $d = 3$, $f_2(x) = f_1(x) + \sqrt{n}\cos(n\pi x)$, and $f_{1,2}$ are the solutions of $g_{1,2}$ with $g_i(y) = \int_0^1 \frac{1}{y-x} f_i(x)dx$. As $n \to \infty$, it is obvious that

$$\|f_2 - f_1\|_{L^2(0,1)} = \left(\int_0^1 (\sqrt{n}\cos(n\pi x))^2 dx\right)^{1/2} = \frac{\sqrt{n}}{\sqrt{2}} \to \infty, \tag{3.7}$$

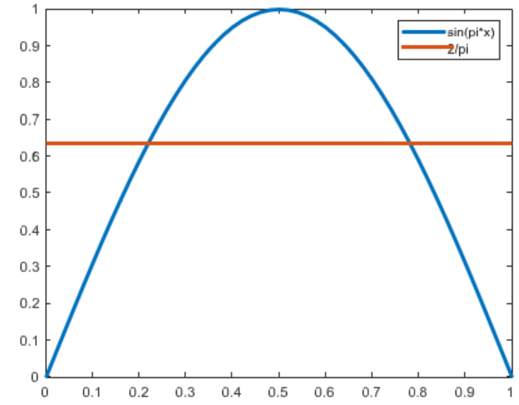
and

$$\|g_2 - g_1\|_{L^2(2,3)} = \frac{1}{\sqrt{n}\pi}\left(\int_2^3 \left(\int_0^1 \left(\frac{1}{y-x}\right)^2 \sin(n\pi x)dx\right)^2 dy\right)^{1/2} \leq \frac{M}{\sqrt{n}\pi} \to 0. \tag{3.8}$$

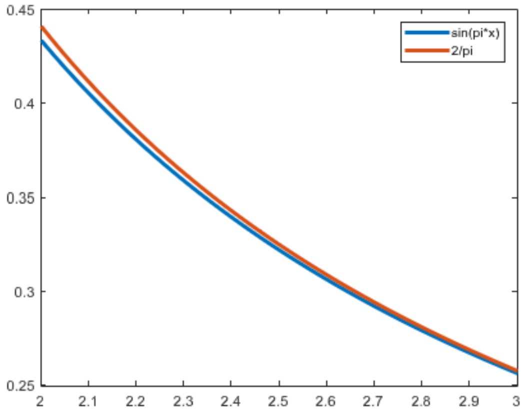
That means the solutions could be changed infinitely even though the noise of the input data is approaching to vanish. So the inverse problem is unstable.

$$\int_a^b \frac{f(x)}{y-x}dx = g(y), y \in [c,d], c > b$$

色散关系的
逆算子无界



$f_{3,4}(x)$



$g_{3,4}(y)$

实际物理中，数据一定存在误差：$\|g^\delta - g\|_{L^2} \leq \delta$

不稳定性是求解该反问题的关键和难题！

4 正则化方法

目标：求解$Kf = g$

数据存在误差,只能求解$Kf^{\delta} = g^{\delta}$，但K^{-1}无界，直接求解会导致$f^{\delta} = K^{-1}g^{\delta}$不收敛到$f$。

$$\|f^{\delta} - f\|_F \le \|K^{-1}g^{\delta} - K^{-1}g\|_F \le \|K^{-1}\| \, \|g^{\delta} - g\|_G = \|K^{-1}\|\delta$$

Andreas Kirsch.2011

正则化算子：

Define: *A regularization strategy is a family of linear and bounded operators $R_\alpha : G \to F, \alpha > 0$, such that $\lim_{\alpha \to 0} R_\alpha K f = f$ for all $f \in F$, where the α is the regularization parameter [8].*

构造一族有界算子R_α去逼近无界算子K^{-1},即可**将不适定问题转化成近似的适定问题。**

$f_\alpha^{\delta} = R_\alpha g^{\delta}$

$$\begin{aligned} \|f_\alpha^{\delta} - f\|_F &\le \|R_\alpha g^{\delta} - R_\alpha g\|_F + \|R_\alpha g - f\|_F \\ &\le \|R_\alpha\| \, \|g^{\delta} - g\|_G + \|R_\alpha K f - f\|_F \\ &\le \boxed{\|R_\alpha\|\delta} + \boxed{\|R_\alpha K f - f\|_F} \end{aligned}$$

$$\delta\|R_\alpha\| \propto \frac{\delta}{\alpha^{M_1}}, M_1 > 0$$

$$\|R_\alpha K f - f\|_F \propto \alpha^{M_2}, M_2 > 0$$

严谨的正则化参数选取方法均可选出合适的 α ！

4 正则化方法

Tikhonov正则化方法：

$$f_\alpha^\delta = argmin\left\{\frac{1}{2}||Kf - g^\delta||_{L^2}^2 + \boxed{\frac{\alpha}{2}||f||_{L^2}^2}\right\}$$

Y. X. Zhang 2017
X. B. Yan 2018
X. B. Yan 2021
M. Asakawa 2001

第一项可看作χ^2拟合，第二项是罚项；

罚项的形式可根据实际问题做改进：H^1、TV、L^1……

$\alpha > 0$是正则化参数，对结果有影响，不能过大或过小。

Tikhonov正则化对应的正则化算子： $\boldsymbol{R_\alpha = (K^*K + \alpha I)^{-1}K^*, f_\alpha^\delta = R_\alpha g^\delta}$

Tikhonov正则化解和真解的估计：

先验信息：$f = K^*v, v \in G, \left||v|\right|_G \le E$ $\quad ||f_\alpha^\delta - f||_F \le \dfrac{\delta}{2\sqrt{\alpha}} + \dfrac{\sqrt{\alpha}E}{2}$

令$\alpha = \delta/E$，则有$\left||f_\alpha^\delta - f|\right|_F \le \sqrt{\delta E} \to 0, as\ \delta \to 0$

1、当误差趋于零时，正则化解一定能收敛到真解；

2、误差可以系统性控制。

4 正则化方法

正则化参数选取方法：

先验选取：结合先验信息$f = K^* v$，理论上清楚展示正则化方法的收敛性；

后验选取：实际计算中方便运用，只知道输入数据或误差即可。

L-curve法：

刘继军 2005

肖庭延 2003

$$\alpha = argmin\{||f_\alpha^\delta||_{L^2}||Kf_\alpha^\delta - g^\delta||_{L^2}\}$$

$\left|\left|f_\alpha^\delta\right|\right|_F$和$\left|\left|Kf_\alpha^\delta - g^\delta\right|\right|$同时达到最小

5 数值结果

$$\int_a^b \frac{f(x)}{y-x}dx = g(y), y \in [c,d], c > b$$

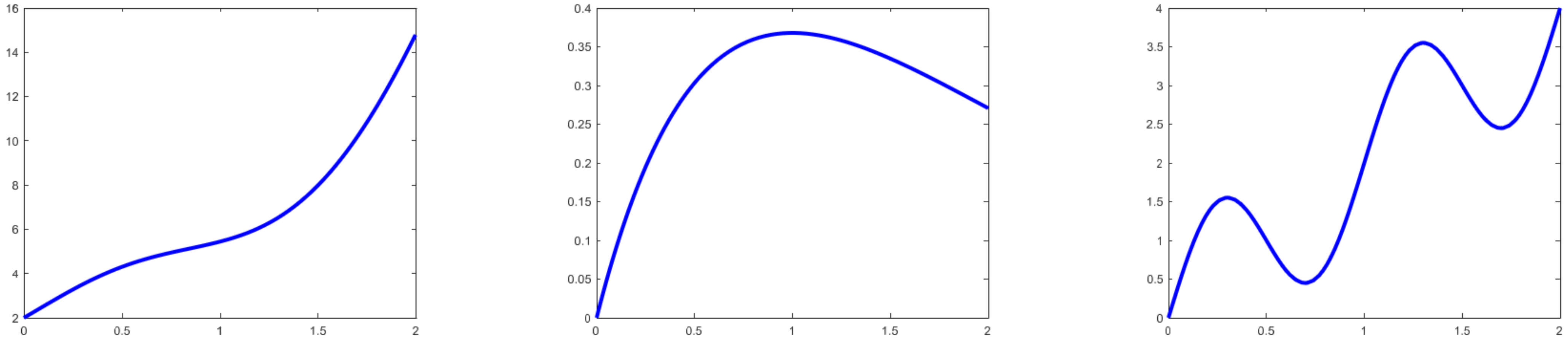
Toy Models:

构造真实解$f(x) = a_1 f_1(x) + a_2 f_2(x)$和对应的真实输入数据$g(y)$，并对真实数据加误差得到误差数据$g^\delta(y)$，以此来测试正则化方法的可靠性。

Model 1: a monotonic function as $f_1(x) = \sin(\pi x)$, $f_2(x) = e^x$;

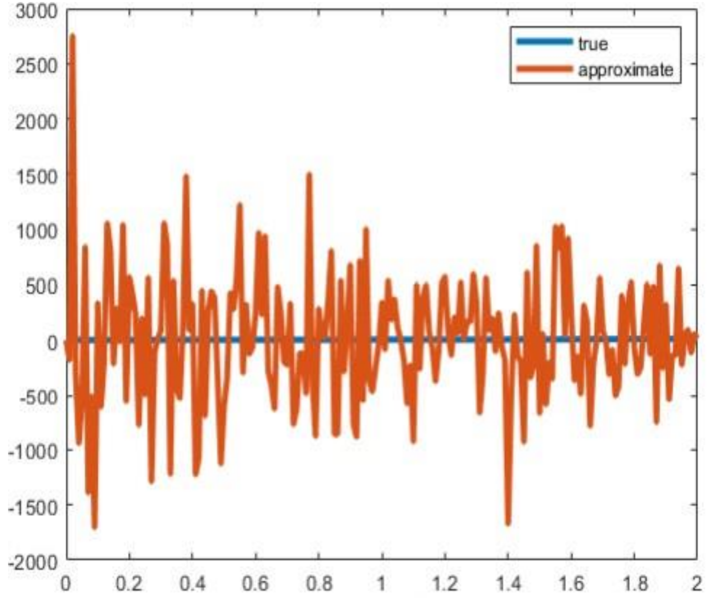
Model 2: a simple non-monotonic function as $f_1(x) = xe^{-x}$, $f_2(x) = 0$;

Model 3: an oscillating function as $f_1(x) = \sin(2\pi x)$, $f_2(x) = x$.

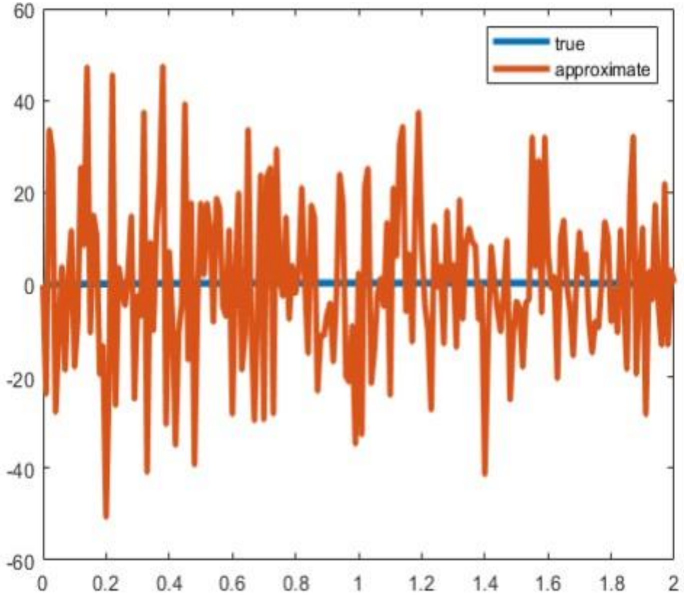


model 1　　model 2　　model 3

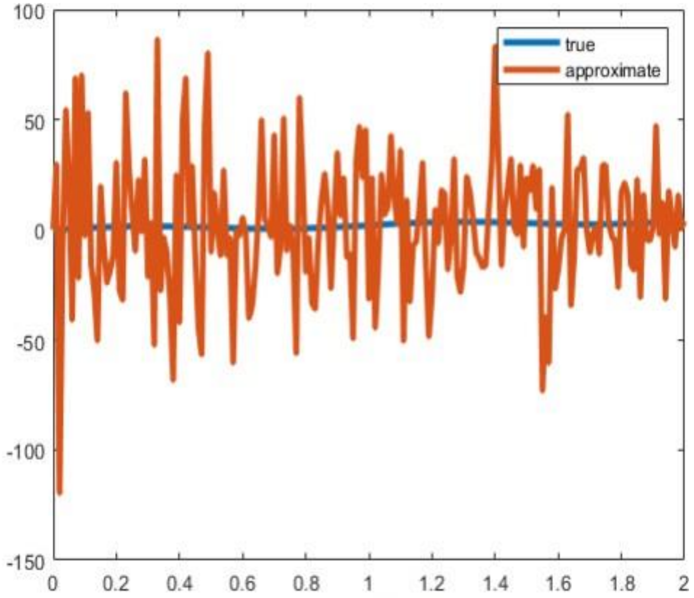
使用经典方法的结果：



model 1



model 2



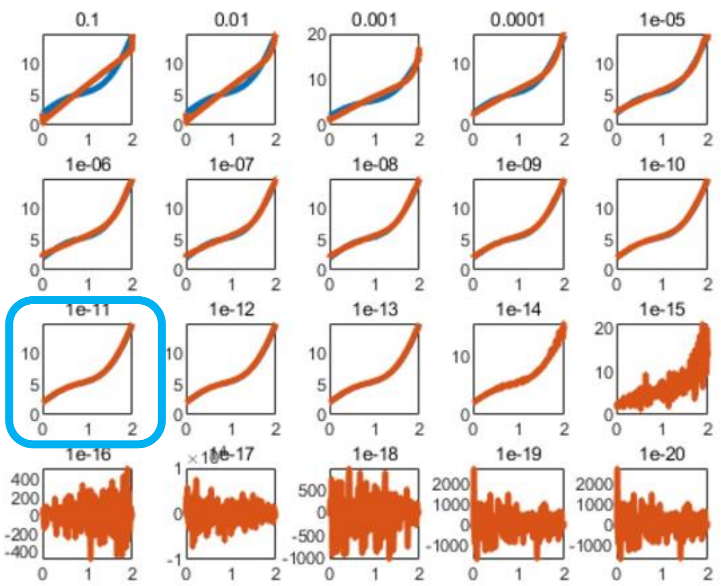
model 3

清楚地验证了该反问题的不稳定性；

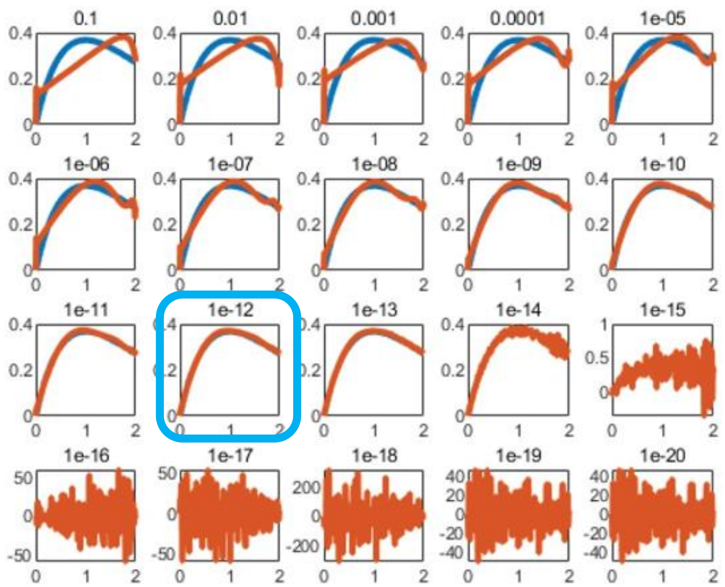
无法使用经典的方法求解，必须要使用正则化方法。

5 数值结果

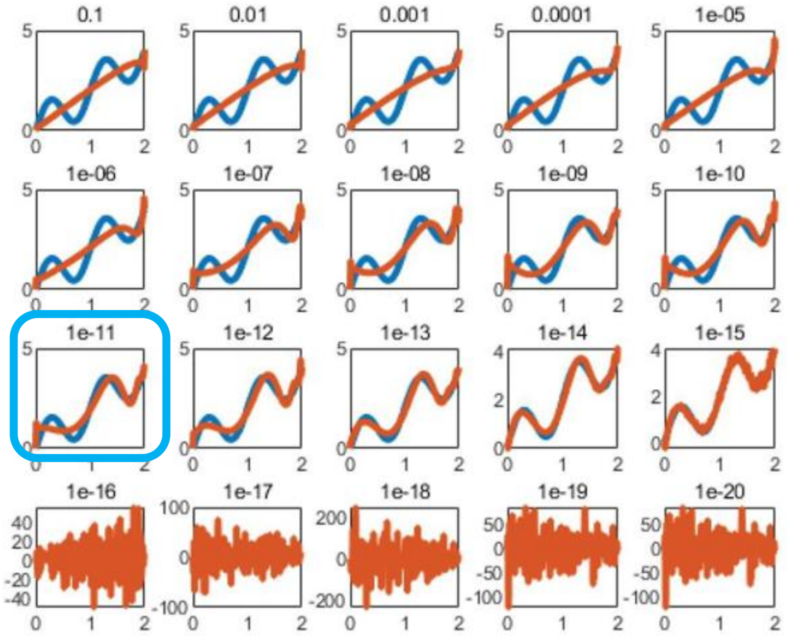
正则化参数对结果的影响：



model 1



model 2



model 3

正则化参数α的选取很重要，符合理论预期；

L-curve曲线法可以较好地选取α

正则化方法可以很好地解决色散关系的反问题。

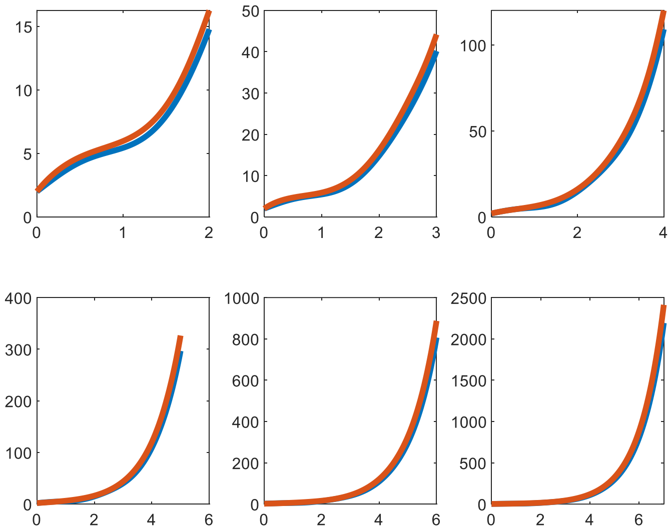
5 数值结果

If $s > \Lambda$, $\mathcal{P}\int_0^\Lambda \frac{\mathcal{I}m[\Pi(s')]}{s-s'}ds' = \pi\mathcal{R}e[\Pi(s)] - \mathcal{P}\int_\Lambda^\infty \frac{\mathcal{I}m[\Pi(s')]}{s-s'}ds'$

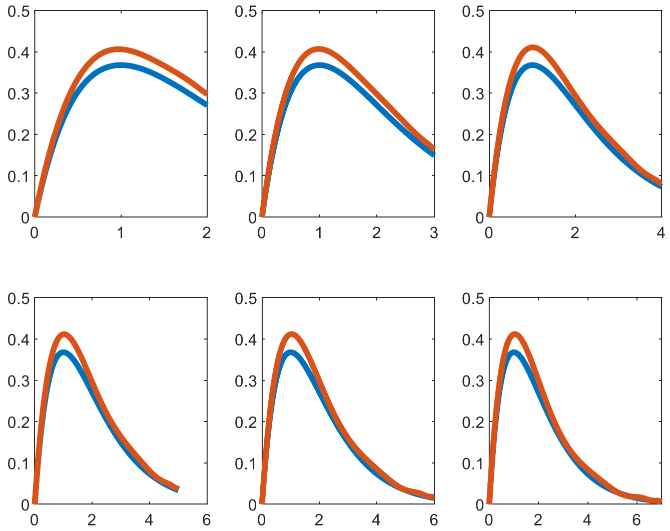
To be solved

calculable

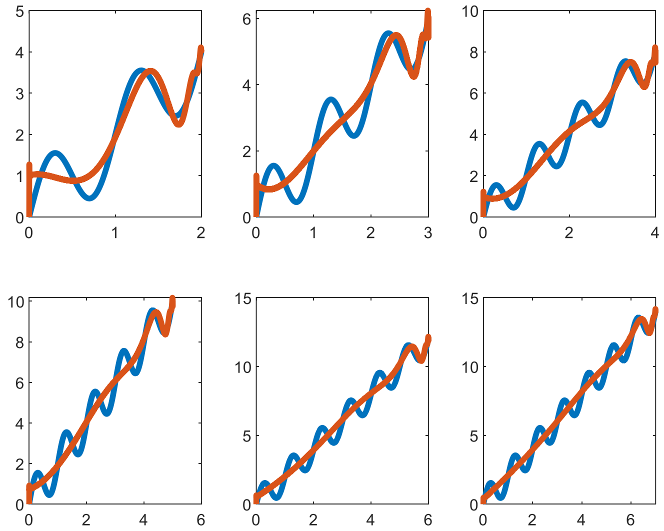
截断$\boldsymbol{\Lambda = b}$的敏感性：$\boldsymbol{\Lambda = 2, \dots, 7}$



model 1



model 2

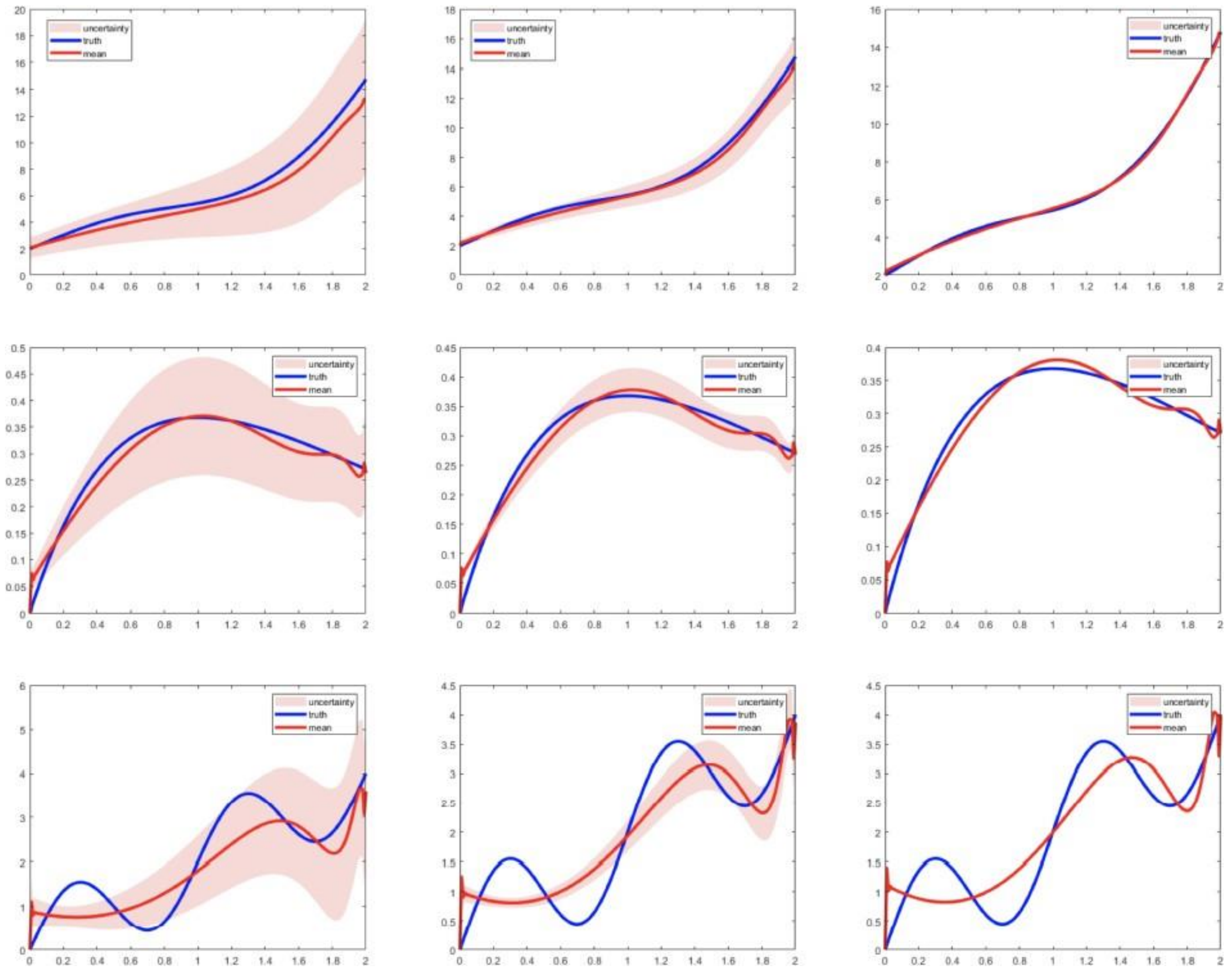


model 3

函数形式简单的函数，截断Λ的影响不大。

5 数值结果

Tikhonov正则化方法和L-curve法的结果：



输入误差分别为30%，10%，1%；

随着输入误差降低，解的误差也逐渐降低，两者误差水平基本一致

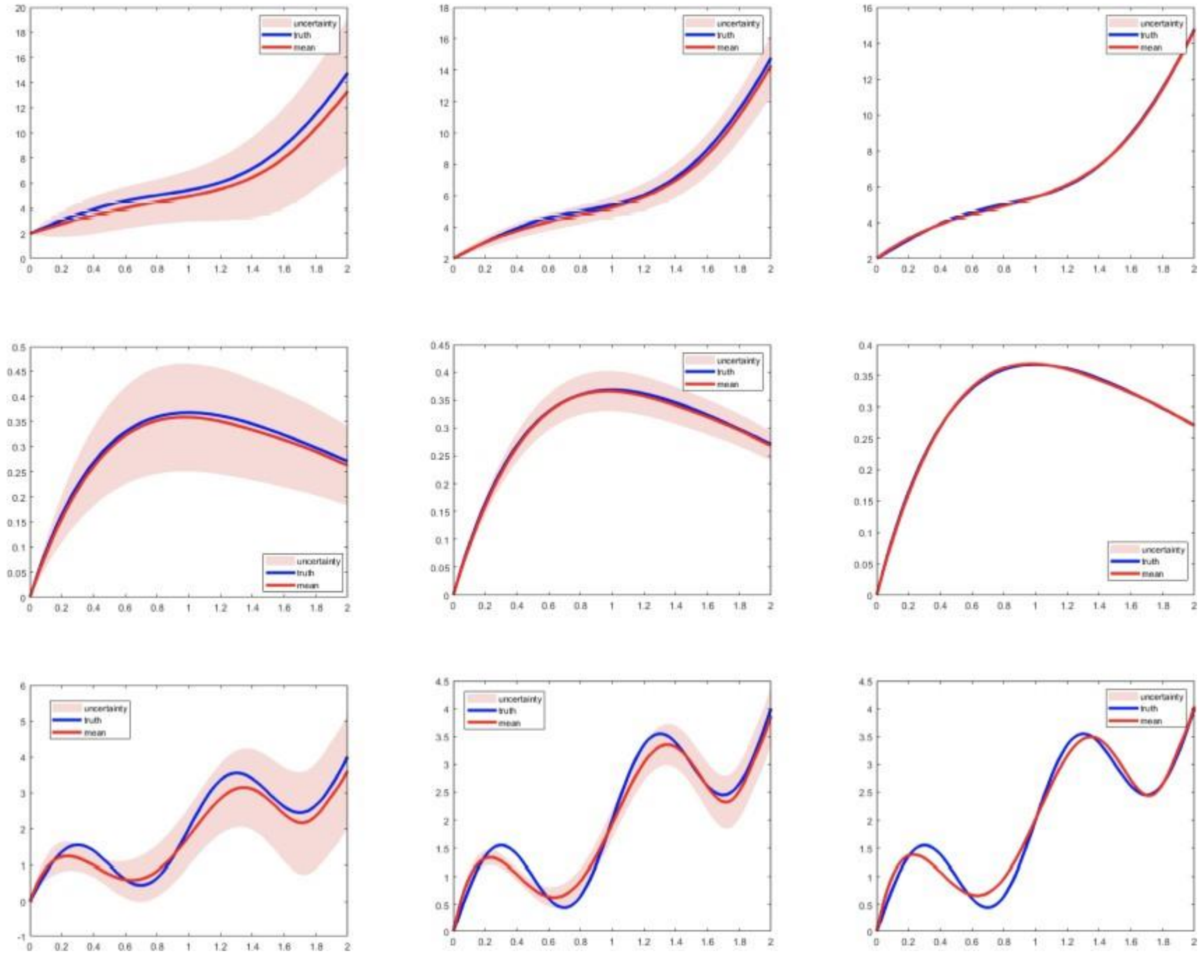
Model 1能够很好地反演，而Model 2和Model 3的反演效果不是特别好；

积分方程将解磨平

可对正则化方法进行改进。

5 数值结果

改进1：限制解的空间： $||f||_{L^2}^2 \rightarrow ||f||_{H^1}^2$



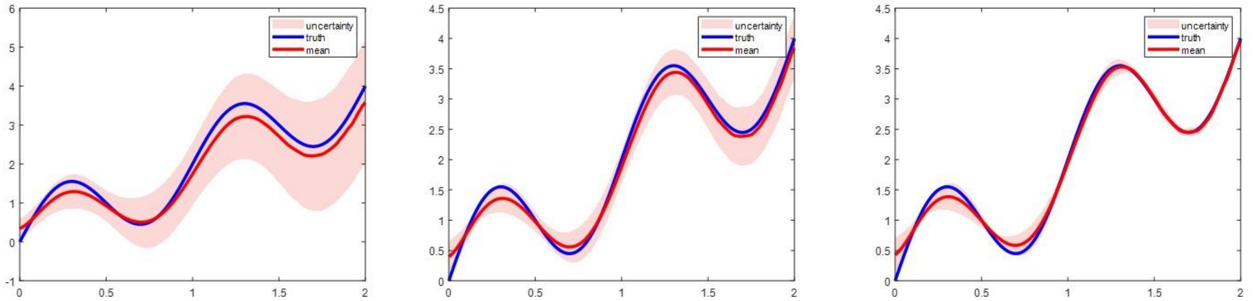
改进2：运用迭代，将每次反演的结果保存并作为下次迭代的先验信息

(1) *Compute* $r_k^\delta = g^\delta - K f_k$

(2) *Solve* $h_k = min\{\frac{1}{2}||Kh - r_k^\delta||_{L^2} + \frac{\alpha_k}{2}||h||_{H^1}\}$

(3) *Update* $f_{k+1} = f_k + h_k$

(4) Stop by the L-curve method

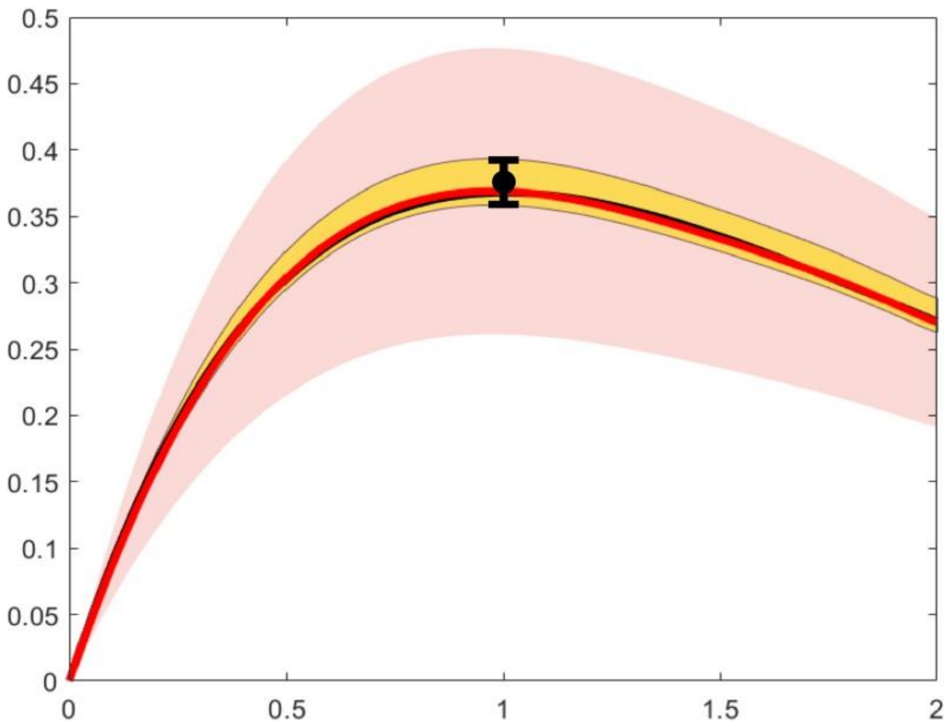


5 数值结果

反问题方法的优势：

1. 基于严谨的数学框架；
2. 能够系统性地控制误差；
3. 不需要过大的计算资源；
4. 能够计算全部非微扰能区、激发态等；
5. 针对具体问题进行改进；
6. 可以与其他方法结合并互补；

与其他非微扰方法结合：



6 总结与展望

总结：

- 计算非微扰量的新方法；
- 严格证明色散关系反问题的不适定性；
- 运用正则化方法将不适定问题转化为适定性问题；
- 数值试验验证正则化方法的有效性。

展望：

- 数学框架的进一步搭建；
- 物理问题的深入应用，解决重大基本问题；
- 寻找反问题方法运用的新方向。

感谢

谢谢大家的聆听！

- Solve $10^{-6}x = y, x, y \in R,\ then\ x = 10^6\, y.$
- If $y = 1, y^\delta = 1.01, \delta = 0.01$, then $x = 10^6,\ \ x^\delta = 10^6 + 10^4$. (compute error is too large)
- Tikhonov regularized solution is

$$x_\alpha^\delta = 10^{-6} y^\delta / (\alpha + 10^{-12})$$

- If take α=10^{-14}, the x_α^δ=10^6 =x (very good)

附录

α	x_α^δ
10^{-6}	1.009998990
10^{-7}	10.099899001
10^{-8}	100.989901010
10^{-9}	1008.991008991
10^{-10}	10000.000000000
10^{-11}	91818.181818182
10^{-12}	505000.000000000
10^{-13}	918181.818181818
10^{-14}	**1000000.000000000**
10^{-15}	1008991.008991009
10^{-16}	1009899.010098990
10^{-17}	1009989.900100999
10^{-18}	1009998.990001010
10^{-19}	1009999.899000010
10^{-20}	1009999.989900000

- How to choose α?

In fact, we can prove that

$$|x_\alpha^\delta - x| \le |x_\alpha^\delta - x_\alpha| + |x_\alpha - x|$$

$$\le \left|\frac{(y^\delta - y)10^{-6}}{\alpha + 10^{-12}}\right| + \left|\frac{y10^{-6}}{\alpha+10^{-12}} - x\right| \le \frac{0.01}{2\sqrt{\alpha}} + \frac{|x|10^6\sqrt{\alpha}}{2}$$

If we know $|x| \le 10^6$, let RHS=min

i.e.

$$\frac{\sqrt{\alpha}}{2}10^{12} = \frac{\delta}{2\sqrt{\alpha}}$$

➔

$$\alpha=\delta/10^{12}=10^{-14}$$

We need to use some information of x which is unknown.