Math 76 HW4, Fall 2025

Name: Jack Chou, Consulted with Benjamin and Mark

For all plots, make sure to include a title, x-axis label, and y-axis label.

Problem 1

Derive the expressions on p. 64 of the textbook coming from the statistical aspects of the Tikhonov solution given additive Gaussian white noise. Specifically derive the equations for the covariance matrix and the expectation of the solution which introduces bias.

Response: *(it may be easier to do this on paper and submit it alongside the notebook)*

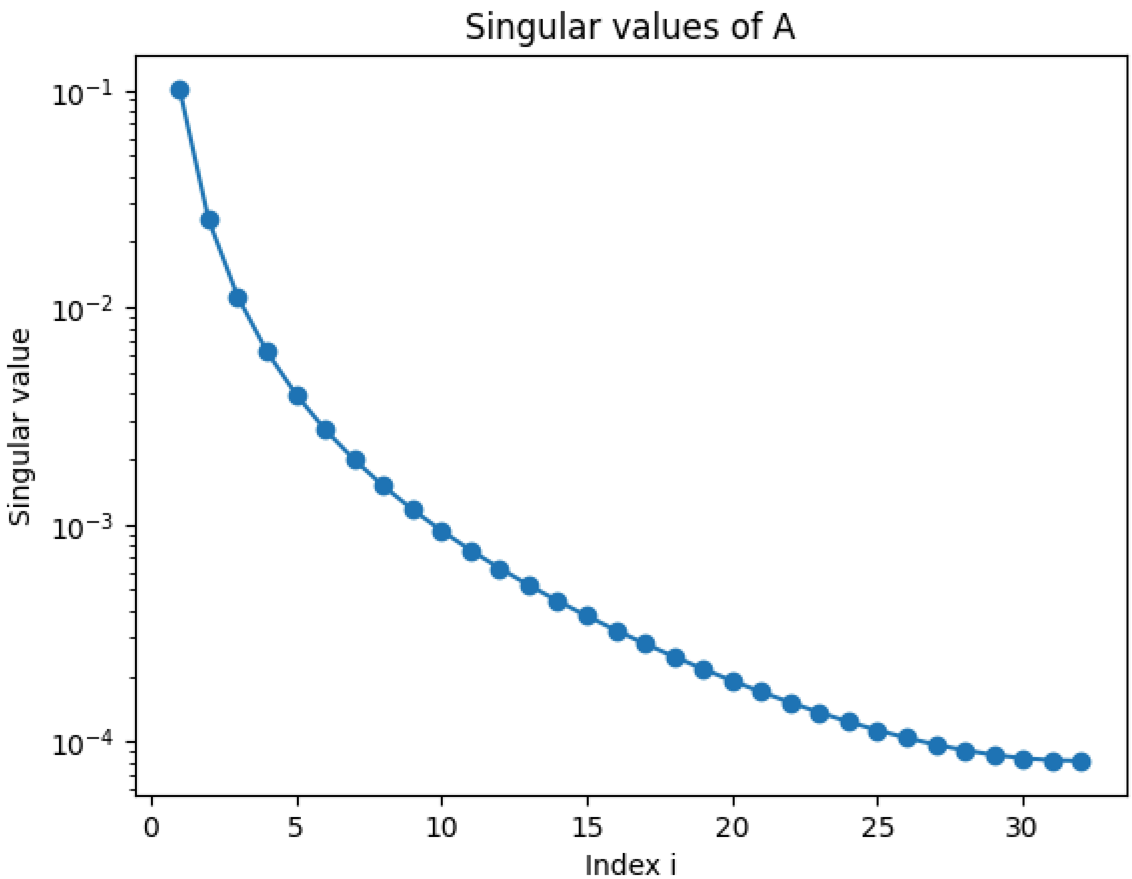
Problem 4.4 (From Oversmoothing to Undersmoothing)

Part A

Use the `deriv2` function to generate the test problem (set $n = 32$). Then use the function `csvd` to compute the SVD of A, and inspect the singular singular values.

```
(32, 32)
(32, 32)
A shape: (32, 32)
Rank: 32
Num of signular values returned: 32
```

```
findfont: Font family ['cmsy10'] not found. Falling back to DejaVu Sans.
findfont: Font family ['cmr10'] not found. Falling back to DejaVu Sans.
findfont: Font family ['cmtt10'] not found. Falling back to DejaVu Sans.
findfont: Font family ['cmmi10'] not found. Falling back to DejaVu Sans.
findfont: Font family ['cmb10'] not found. Falling back to DejaVu Sans.
findfont: Font family ['cmss10'] not found. Falling back to DejaVu Sans.
findfont: Font family ['cmex10'] not found. Falling back to DejaVu Sans.
```



Part B

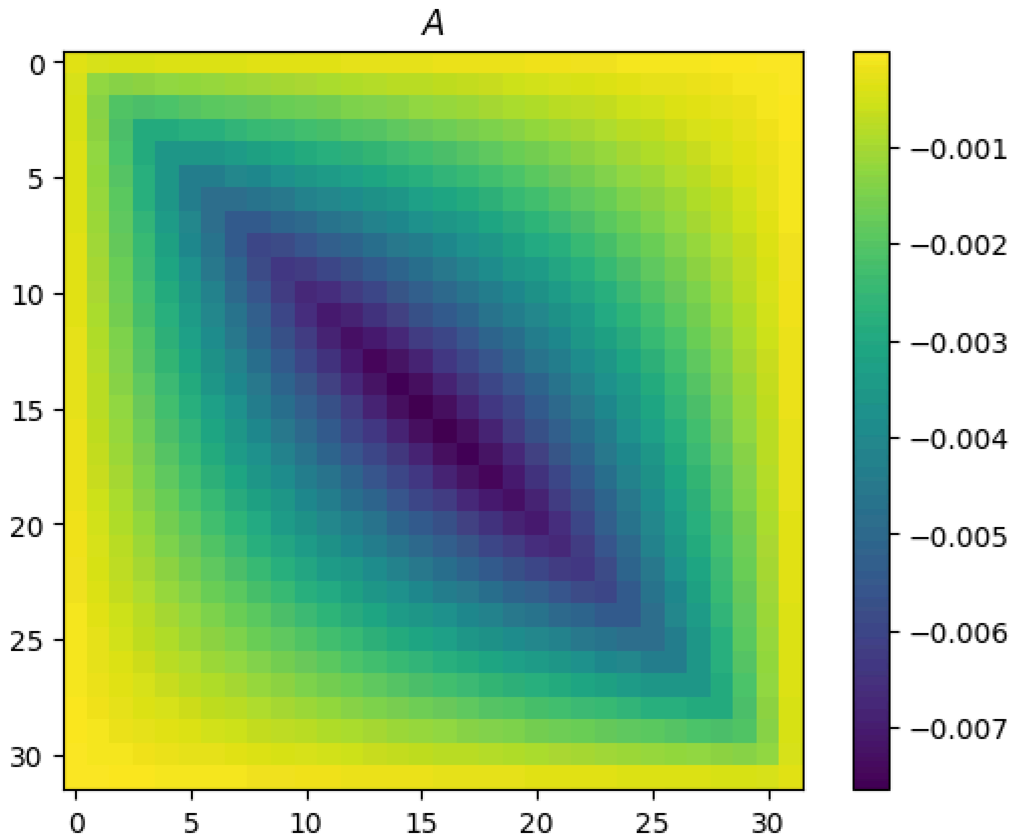
Add a small amount of noise to the right hand side, e.g.,

```
e = 1e-3*\texttt{np.random.normal(size=len(b))}
```

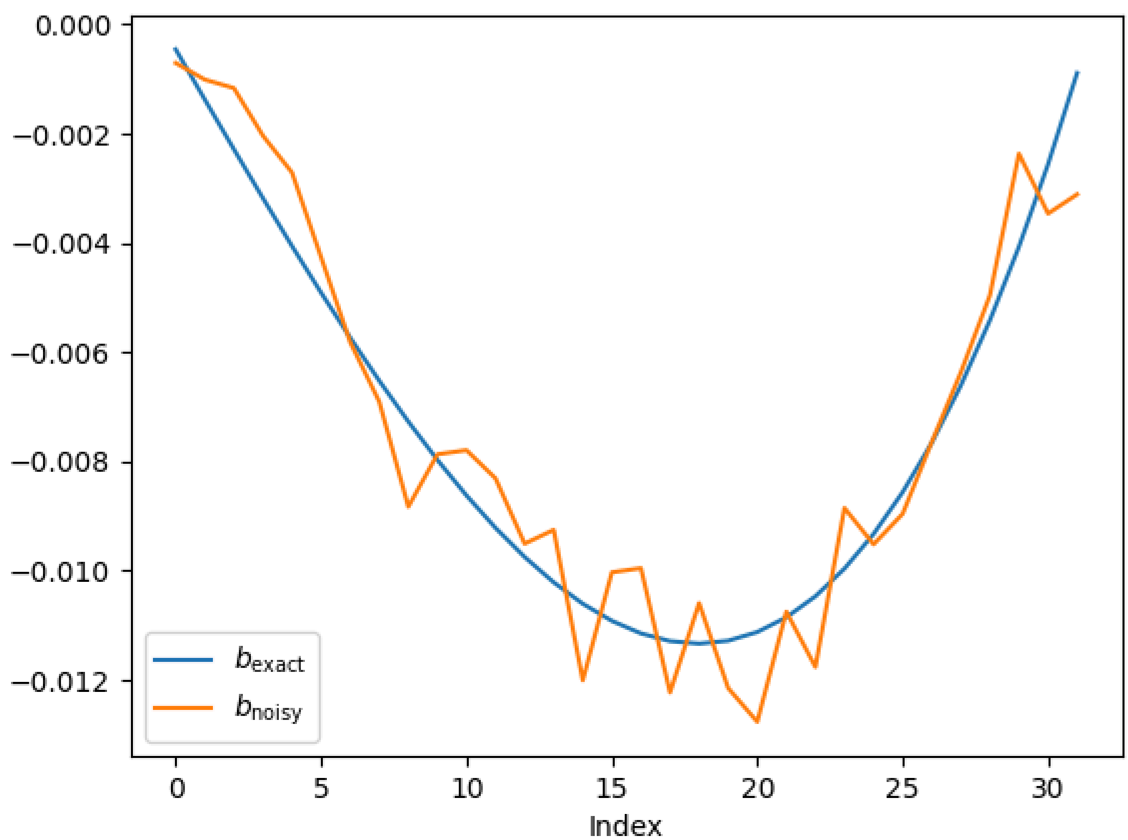
This noise is certainly not visible when plotting the right-hand side vector, but it is very significant with respect to the regularization. For a number of different regularization parameters λ in the range 10^{-3} to 1, compute the corresponding filter factors $\varphi_i^{[\lambda]}$ using the function `fil_fac`, as well as the corresponding Tikhonov solution x_λ by means of

```
X = tikhonov(U, s, V, b, lambdah).
```

For each λ, plot both the filter factors and the solution, and comment on your results. Use a logarithmic distribution of λ-values using `matplotlib`'s `semilogy()` function.

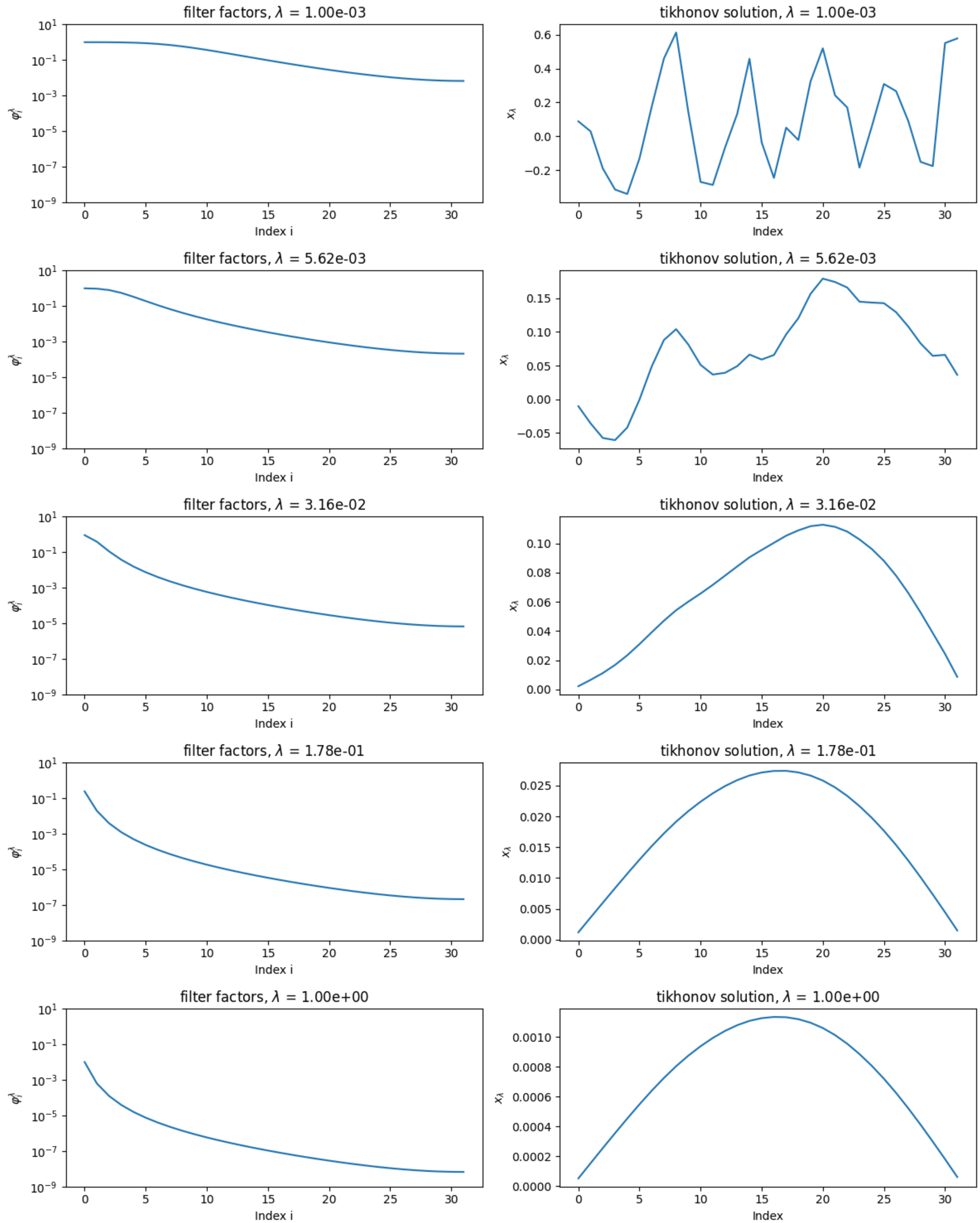


A
0
5
10
15
20
25
30
−0.001
−0.002
−0.003
−0.004
−0.005
−0.006
−0.007



0.000
−0.002
−0.004
−0.006
−0.008
−0.010
−0.012
0
5
10
15
20
25
30
Index
b_{exact}
b_{noisy}

```
<>:13: SyntaxWarning: invalid escape sequence '\l'
<>:13: SyntaxWarning: invalid escape sequence '\l'
C:\Users\sharp\AppData\Local\Temp\ipykernel_30460\836922243.py:13: SyntaxWarning: invalid esc
ape sequence '\l'
  axs[j,0].set_ylabel("$\\varphi_i^{\lambda}$")
```



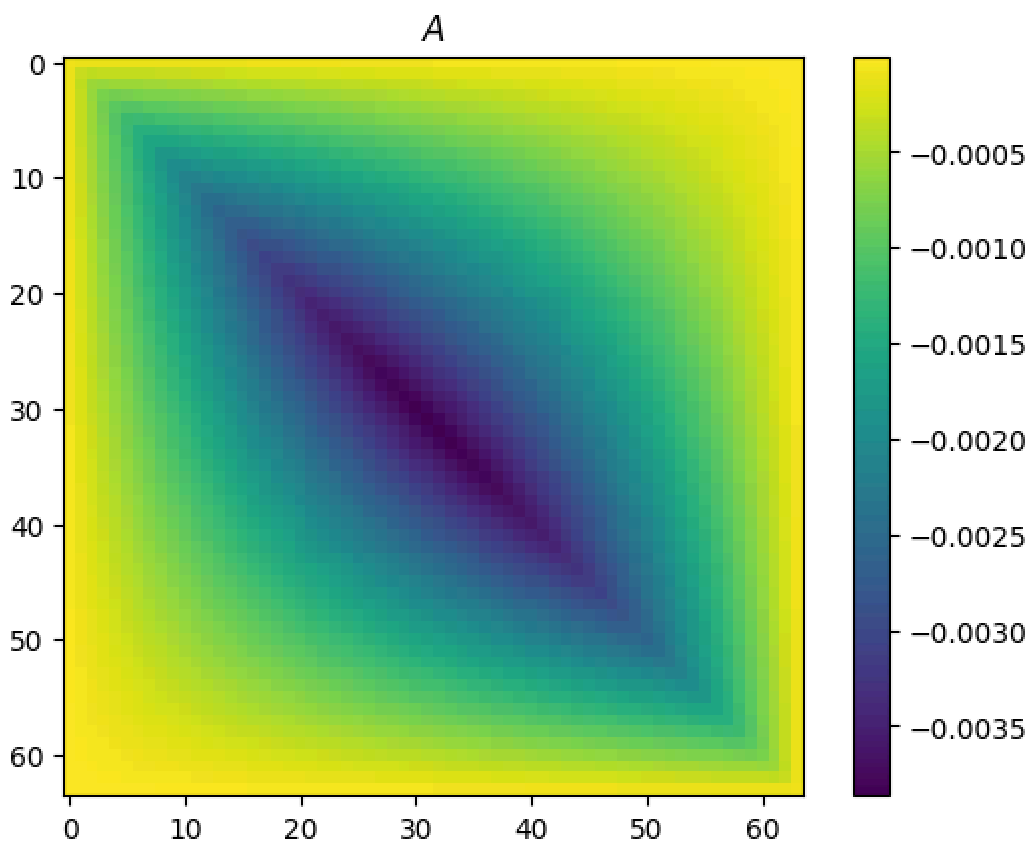
Response: As λ increase, the filter factors suppress more high-index modes and the solution transitions from noisy to smooth/biased. The best λ is the one that damps the onset of noise without removing the essential structure.

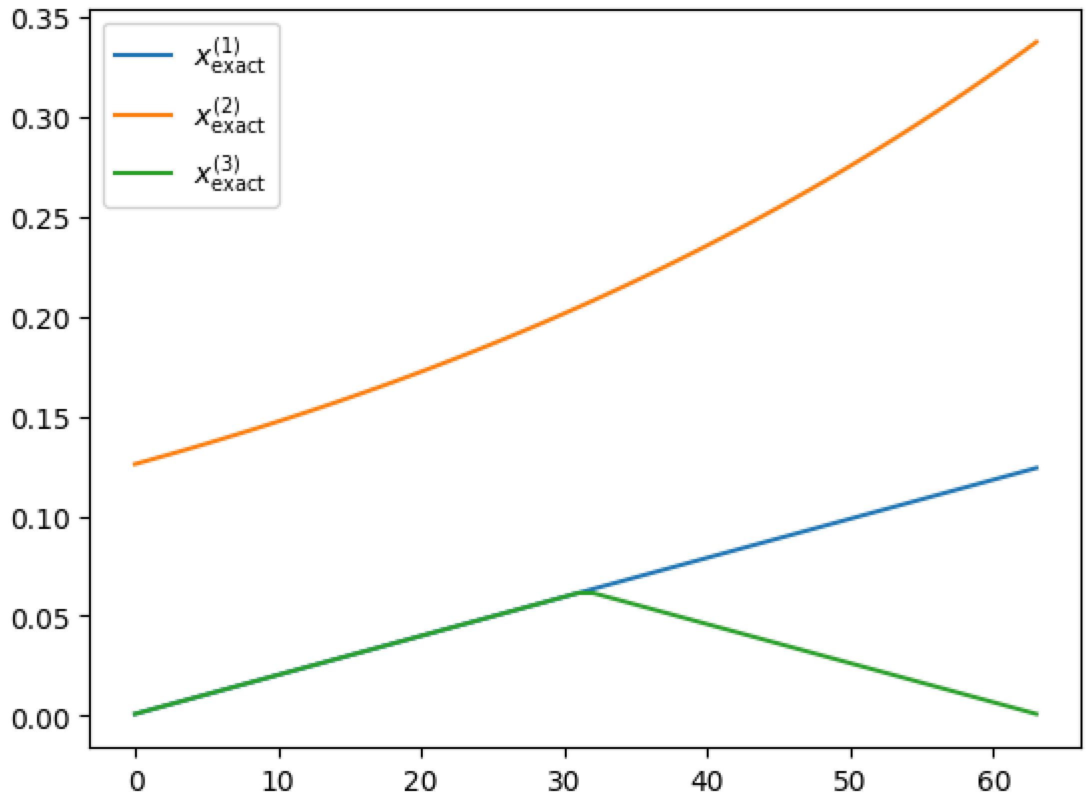
Problem 4.6 (The L-curve)

This exercise illustrates the typical behavior of the L-curve for a discrete ill-posed problem, using the second-derivative test problem from Excercise 2.3.

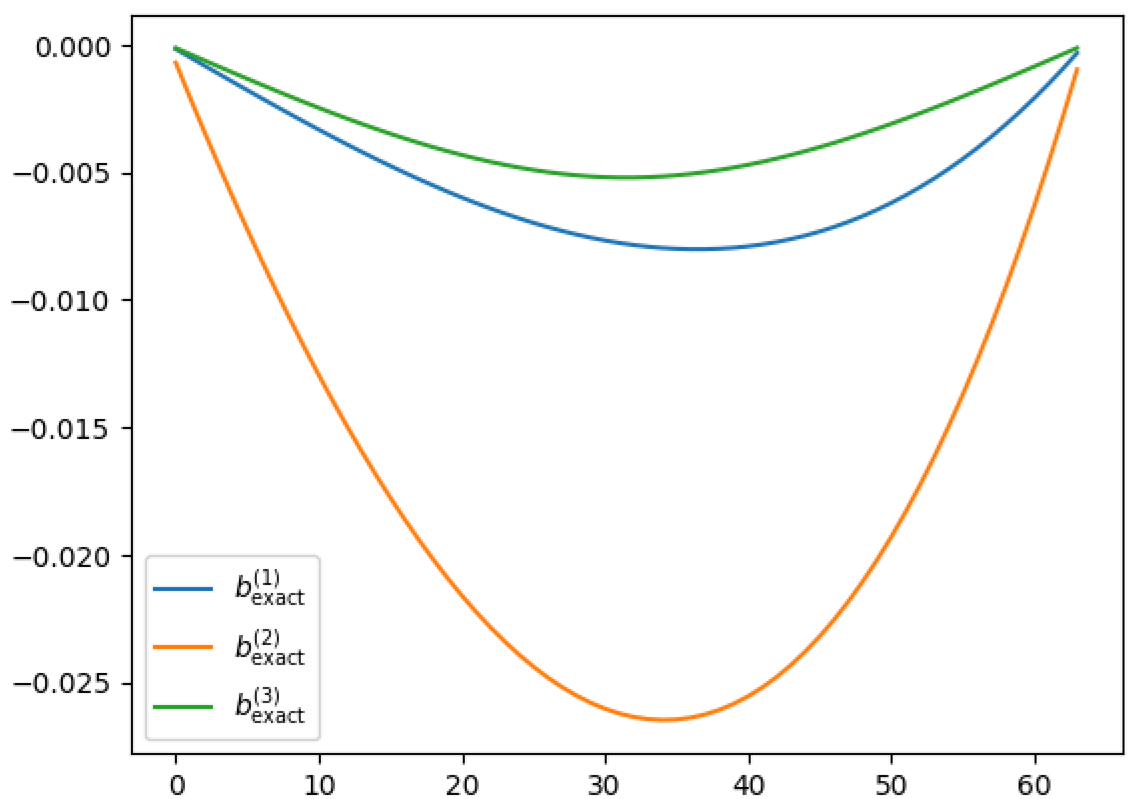
Part A

Generate the test problem `deriv2` with $n = 64$, and add Gaussian white noise scaled such that $\|e\|_2/\|b_{\mathrm{exact}}\|_2 = 10^{-2}$. Then use `l_curve` to plot the L-curves corresponding to the three different right-hand sides b_{exact}, e, and $b_{\mathrm{exact}} + e$. What happens to the corner if you switch to lin-lin scale?

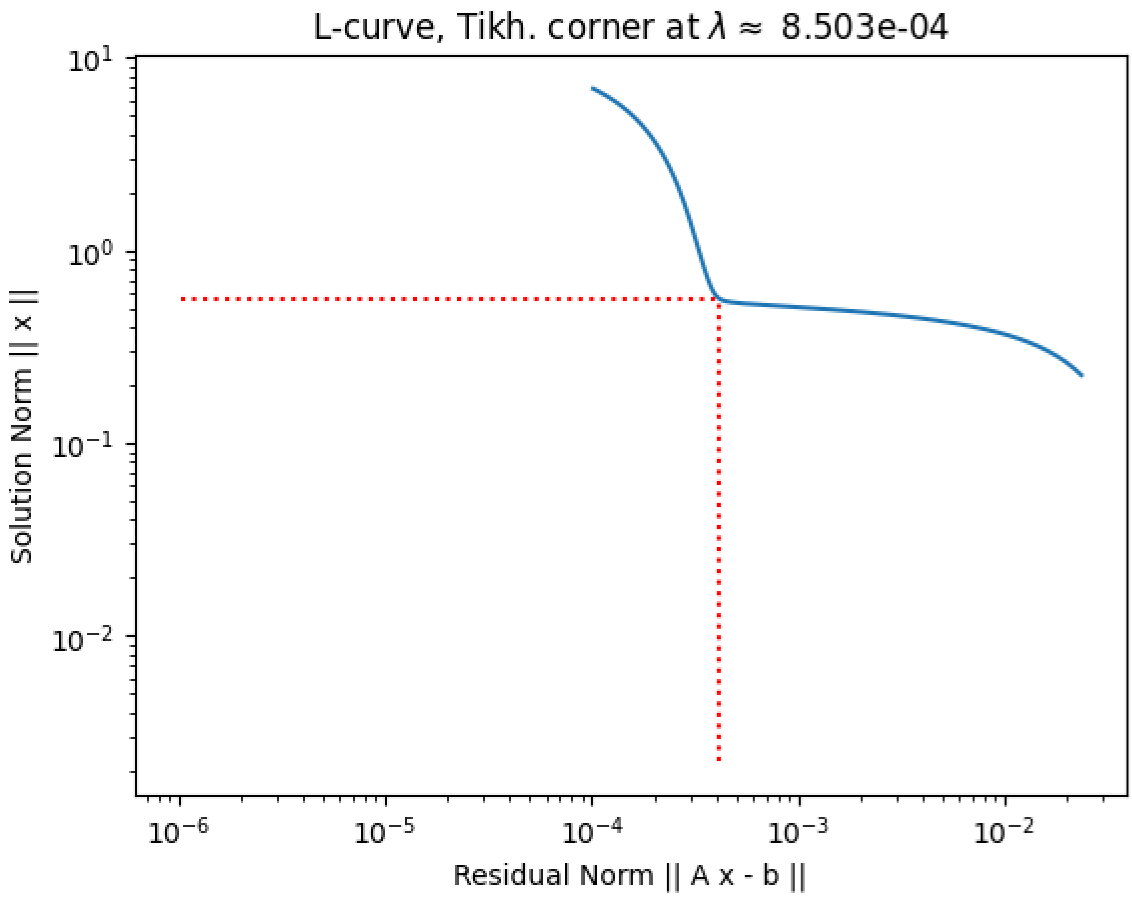


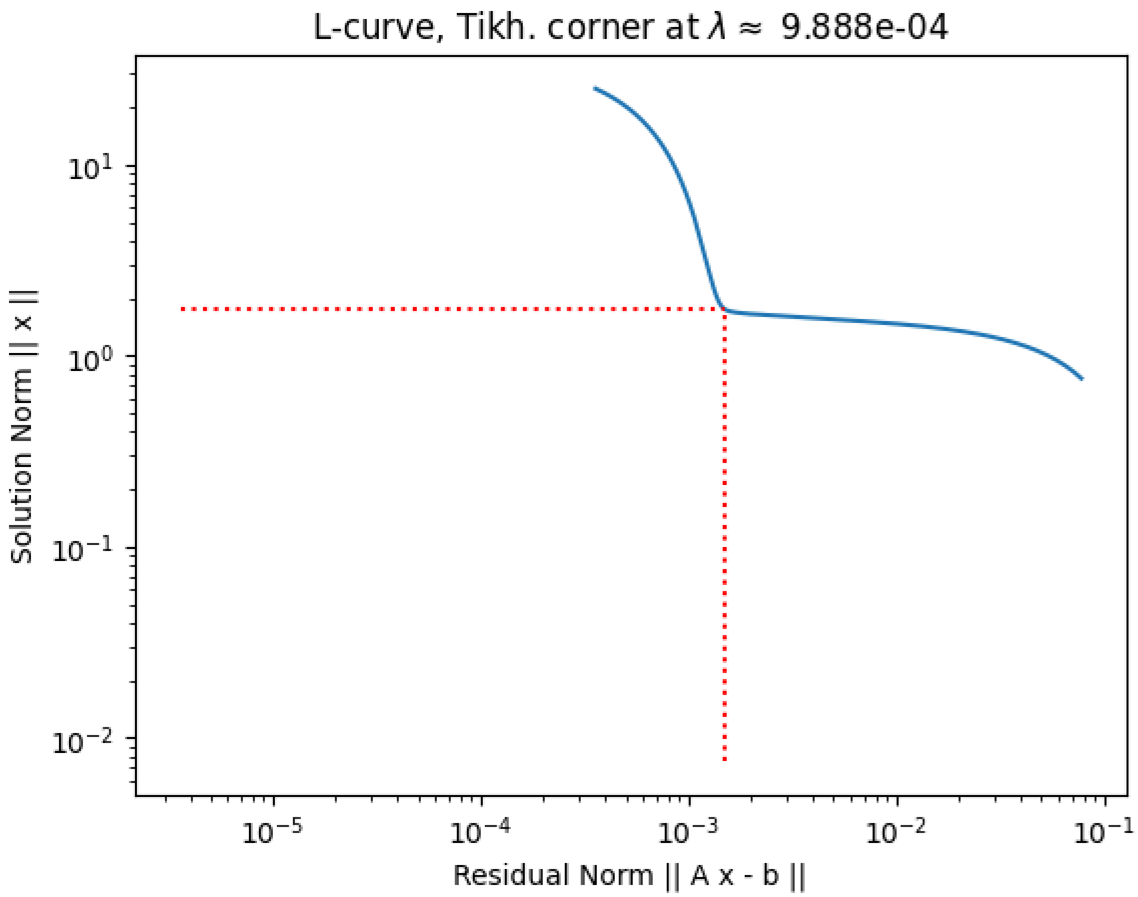


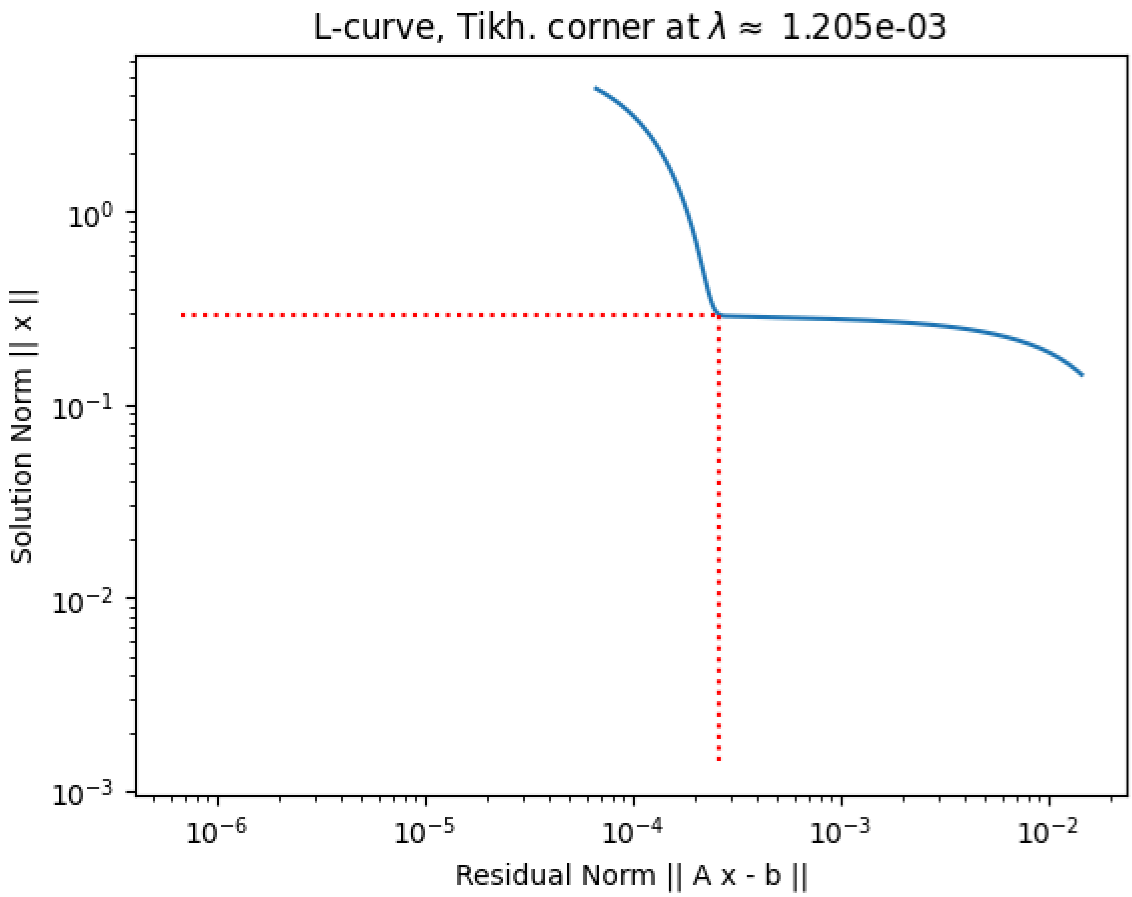
$x^{(1)}_{\text{exact}}$
$x^{(2)}_{\text{exact}}$
$x^{(3)}_{\text{exact}}$
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
10
20
30
40
50
60



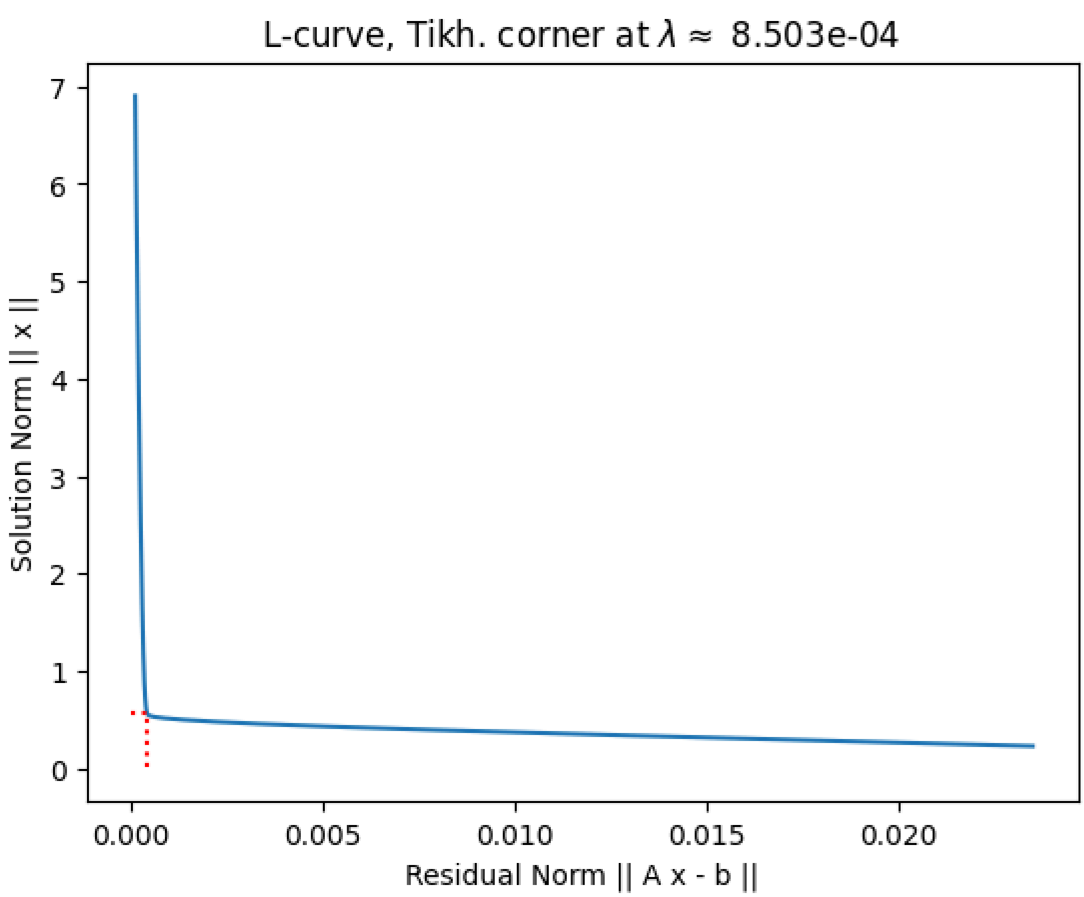
0.000
−0.005
−0.010
−0.015
−0.020
−0.025
$b^{(1)}_{\text{exact}}$
$b^{(2)}_{\text{exact}}$
$b^{(3)}_{\text{exact}}$
0
10
20
30
40
50
60


L-curve, Tikh. corner at $\lambda \approx$ 8.503e-04
10^{1}
10^{0}
10^{-1}
10^{-2}
Solution Norm || x ||
10^{-6}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
Residual Norm || A x - b ||

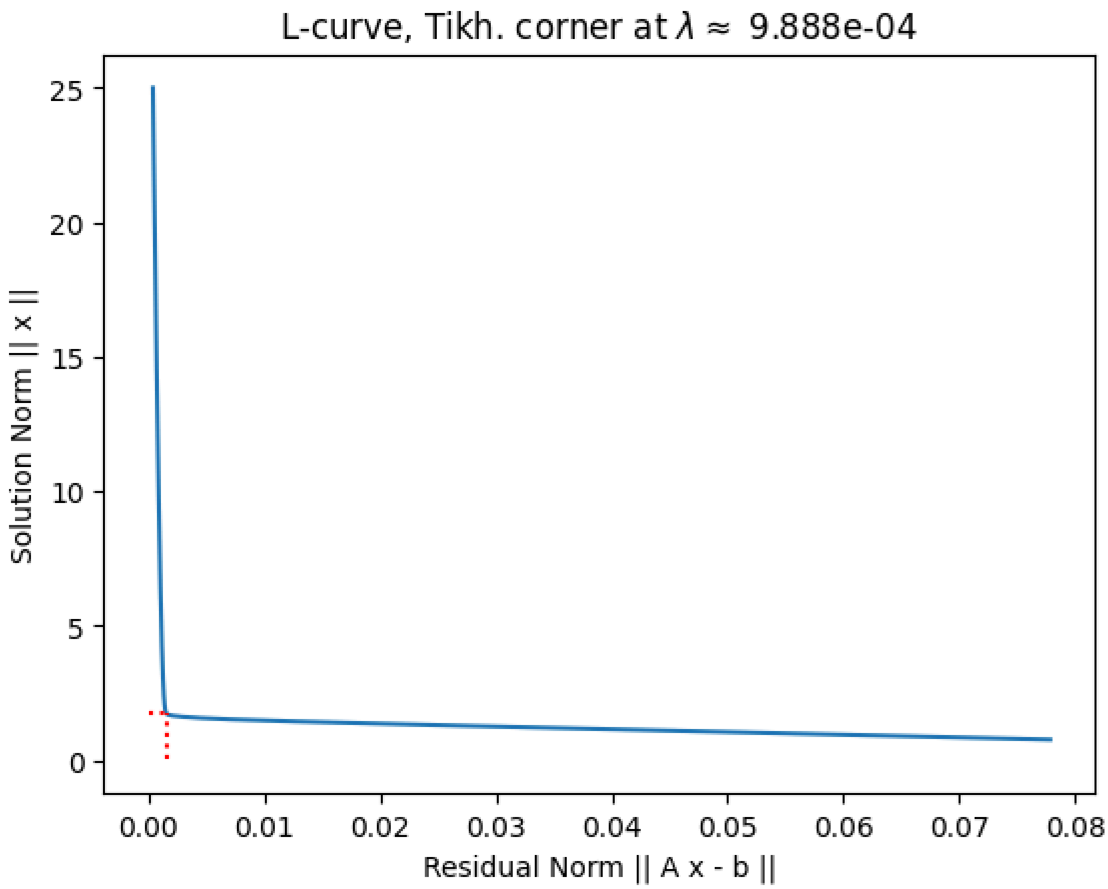

L-curve, Tikh. corner at $\lambda \approx$ 9.888e-04
Solution Norm || x ||
Residual Norm || A x - b ||
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}

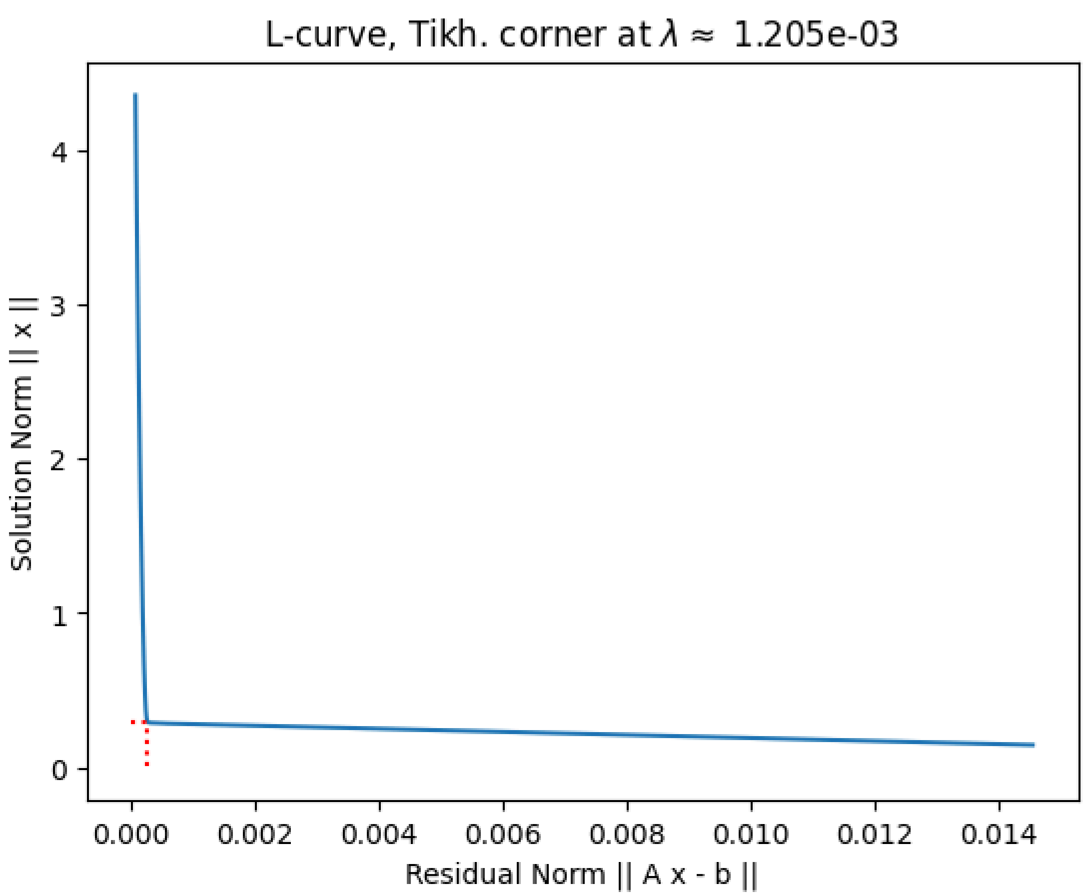


L-curve, Tikh. corner at $\lambda \approx$ 1.205e-03
Solution Norm || x ||
Residual Norm || A x - b ||
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-6}
10^{-5}
10^{-4}
10^{-3}
10^{-2}



L-curve, Tikh. corner at $\lambda \approx$ 8.503e-04
Solution Norm || x ||
Residual Norm || A x - b ||
0
1
2
3
4
5
6
7
0.000
0.005
0.010
0.015
0.020

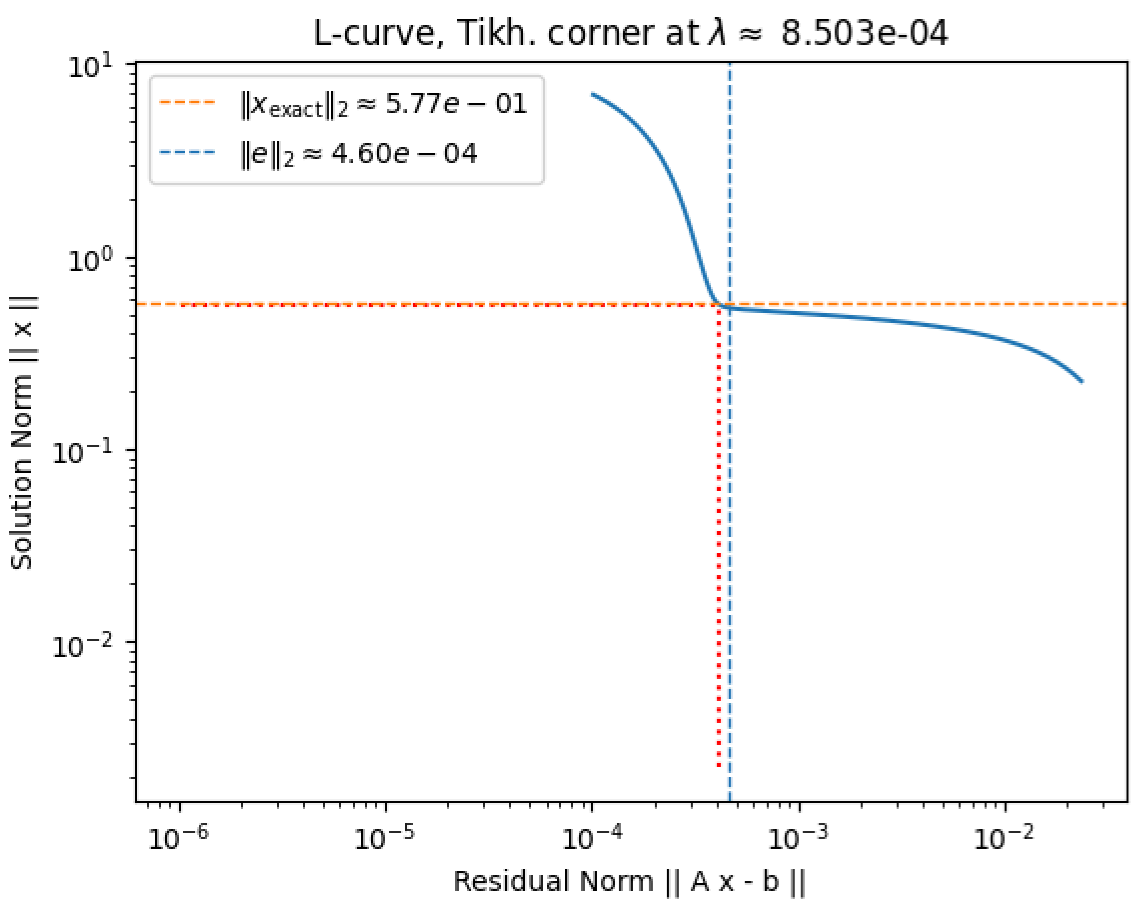

L-curve, Tikh. corner at $\lambda \approx$ 9.888e-04
Solution Norm || x ||
Residual Norm || A x - b ||
0
5
10
15
20
25
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08

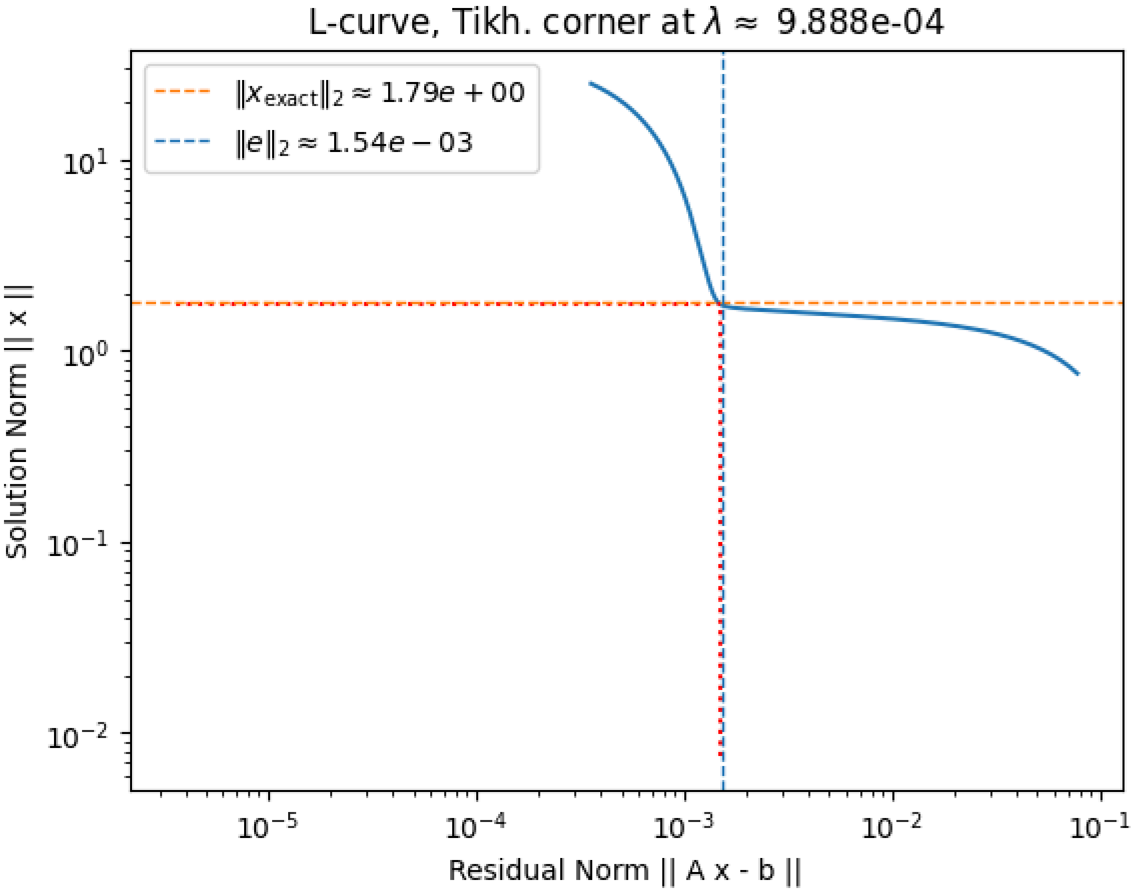

L-curve, Tikh. corner at $\lambda \approx$ 1.205e-03
Solution Norm || x ||
Residual Norm || A x - b ||
0
1
2
3
4
0.000
0.002
0.004
0.006
0.008
0.010
0.012
0.014

Response: Since $||Ax - b||$ and $||x||$ span several orders of magnitude, log-log axes spread the small values and compress the large ones. Log-log is also scale invariant, whereas on lin-lin most points collapse near the axes and the corner becomes flattened and unreliable.

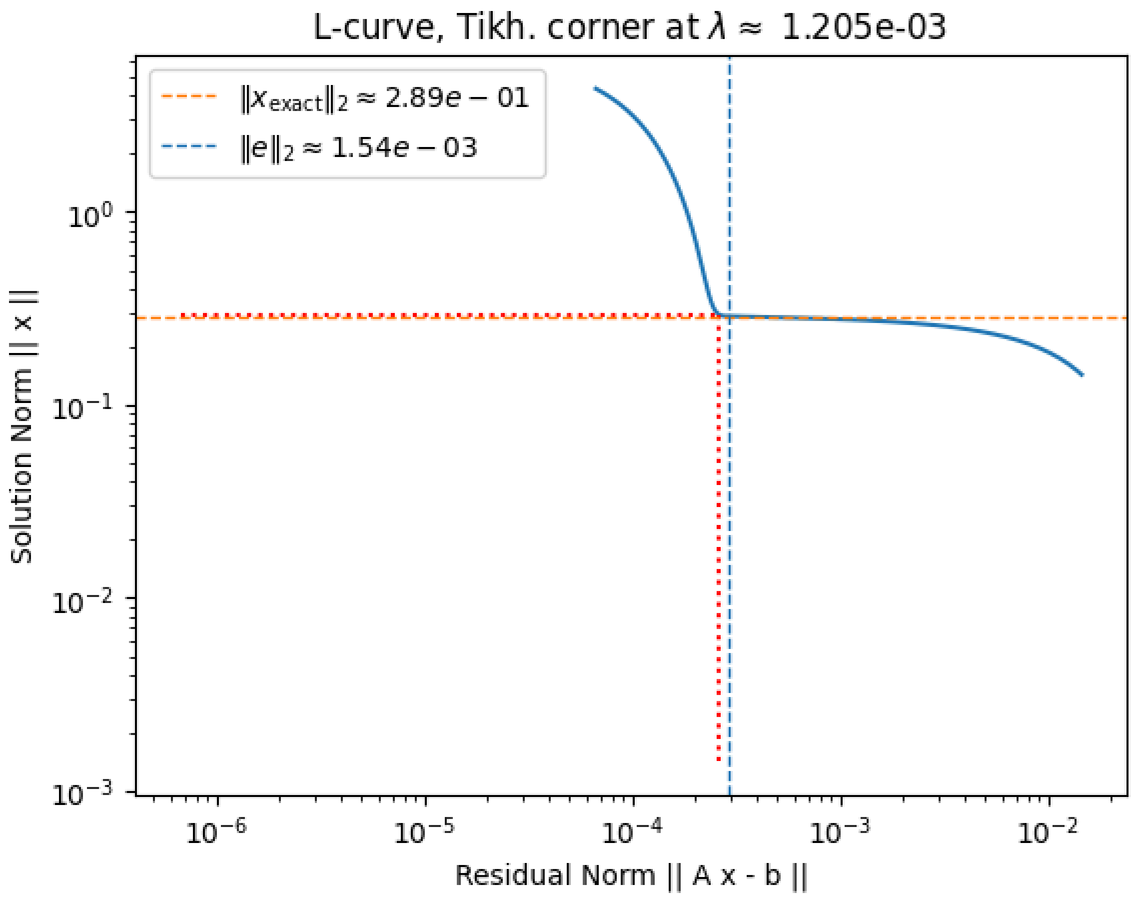
Part B

Switch back to log-log scale and add a horizontal line at $||x_{\text{exact}}||_2$, the norm of the exact solution, and a vertical line at $||e||_2$, the norm of the perturbation. Relate the positions of these lines to the different parts of the L-curve.





L-curve, Tikh. corner at $\lambda \approx$ 9.888e-04
$\|x_{\text{exact}}\|_2 \approx 1.79e+00$
$\|e\|_2 \approx 1.54e-03$
Solution Norm || x ||
Residual Norm || A x - b ||
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}

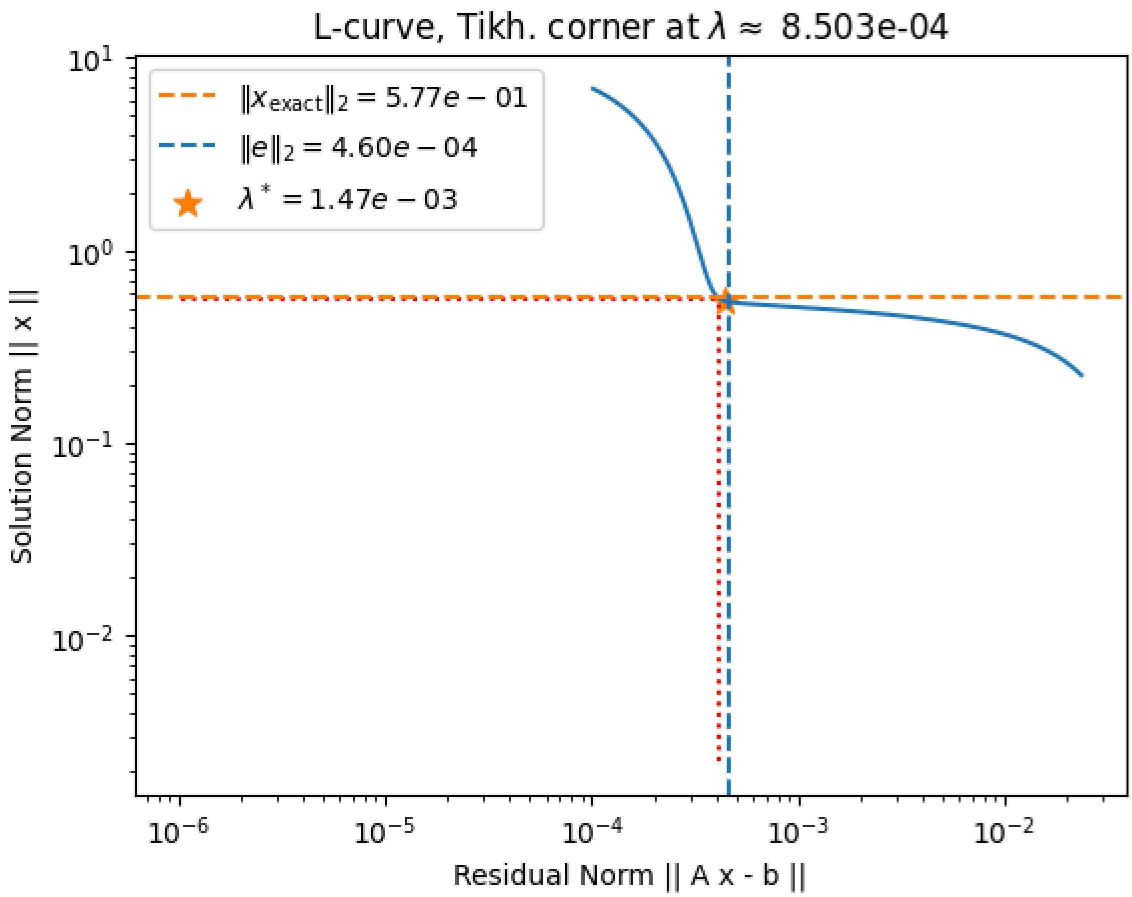


Response: More noise shift the elbow to larger λ. The vertical $\|e\|_2$ line marks the residual floor and the horizontal $\|x_{\text{exact}}\|_2$ line marks a resonable solution scale. The elbow near their intersection is a justifiable choice for λ.

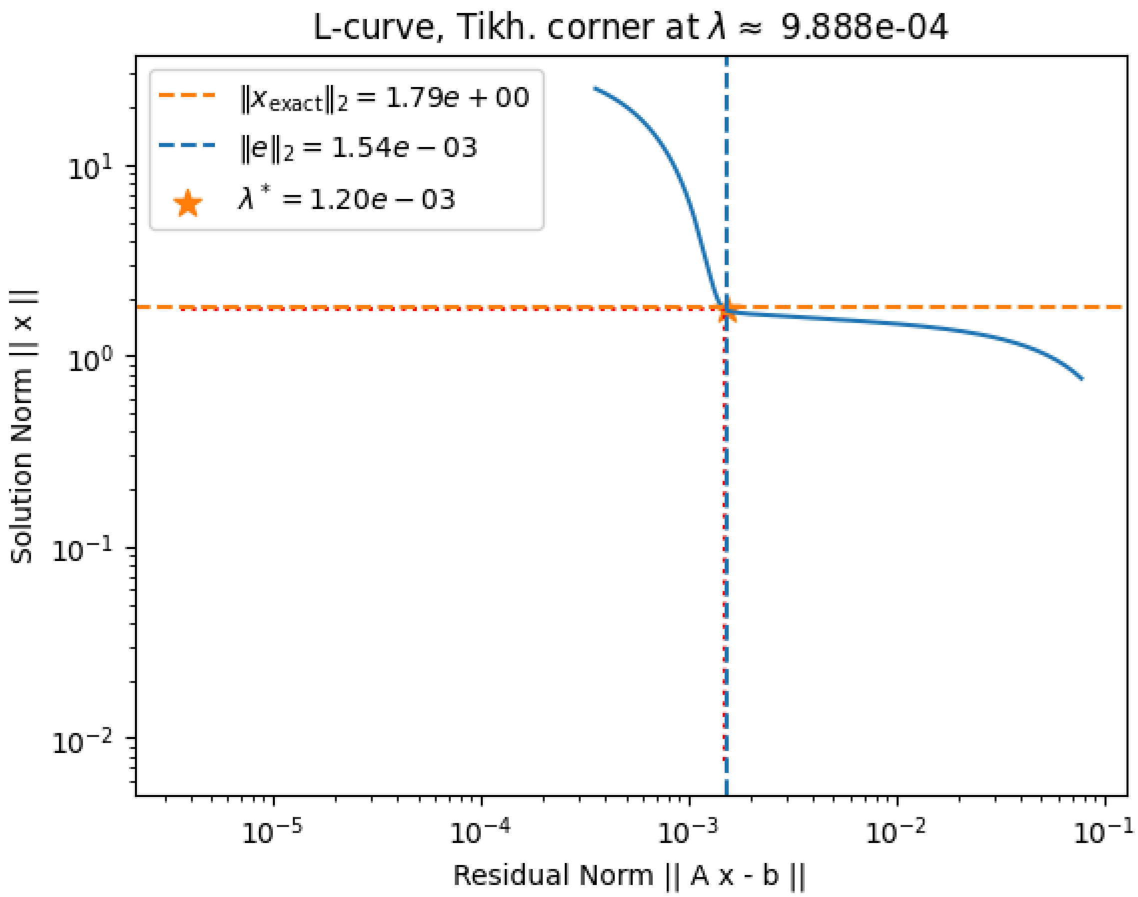
Part C

Find (by trial and error) a Tikhonov regularization parameter λ^* that approximately minimizes the error $\|x_{\text{exact}} - x_\lambda\|_2$ between the exact solution x_{exact} and the regularized solution x_λ. Add the point $(\|Ax_{\lambda^*} - b\|_2, \|x_{\lambda^*}\|_2)$ to the L-curve (it must lie on the L-curve corresponding to b). Is it near the corner? (Note: here b denotes the noisy RHS data vector, *not* the noiseless RHS vector)

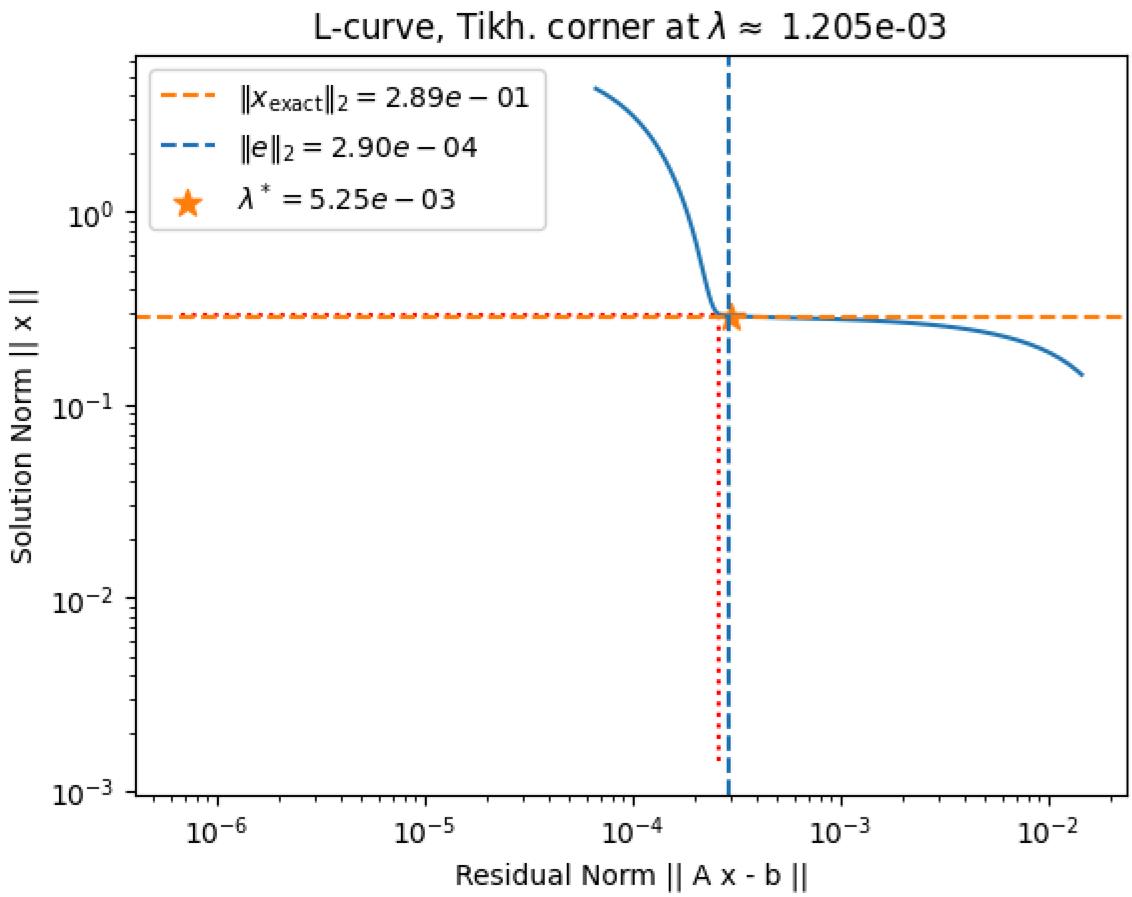
```
lambda* ≈ 1.445e-03
min error ||x_exact - x_{λ*}||_2 ≈ 1.597e-01

[RHS 2] lambda* ≈ 1.202e-03
[RHS 2] min error ||x_exact - x_λ*||_2 ≈ 3.852e-01

[RHS 3] lambda* ≈ 5.248e-03
[RHS 3] min error ||x_exact - x_λ*||_2 ≈ 1.499e-02
```



L-curve, Tikh. corner at $\lambda \approx$ 8.503e-04
$\|x_{\text{exact}}\|_2 = 5.77e-01$
$\|e\|_2 = 4.60e-04$
$\lambda^* = 1.47e-03$
Solution Norm || x ||
Residual Norm || A x - b ||
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-6}
10^{-5}
10^{-4}
10^{-3}
10^{-2}



L-curve, Tikh. corner at $\lambda \approx$ 9.888e-04
$\|x_{\text{exact}}\|_2 = 1.79e+00$
$\|e\|_2 = 1.54e-03$
$\lambda^* = 1.20e-03$
Solution Norm || x ||
Residual Norm || A x - b ||
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}



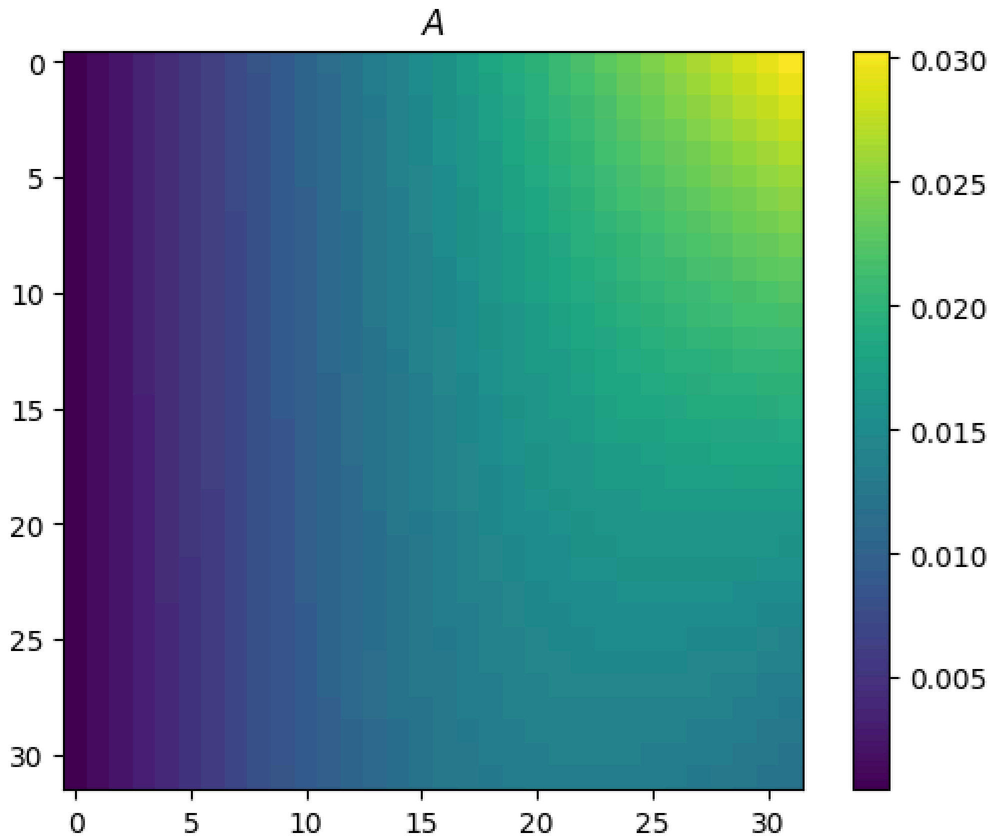
Response: Across the three RHS, the L-curve corner lies close to, but not always exactly at, the true-error λ^*.

Problem 4.7 (Limitations of TSVD and Tikhonov Methods)

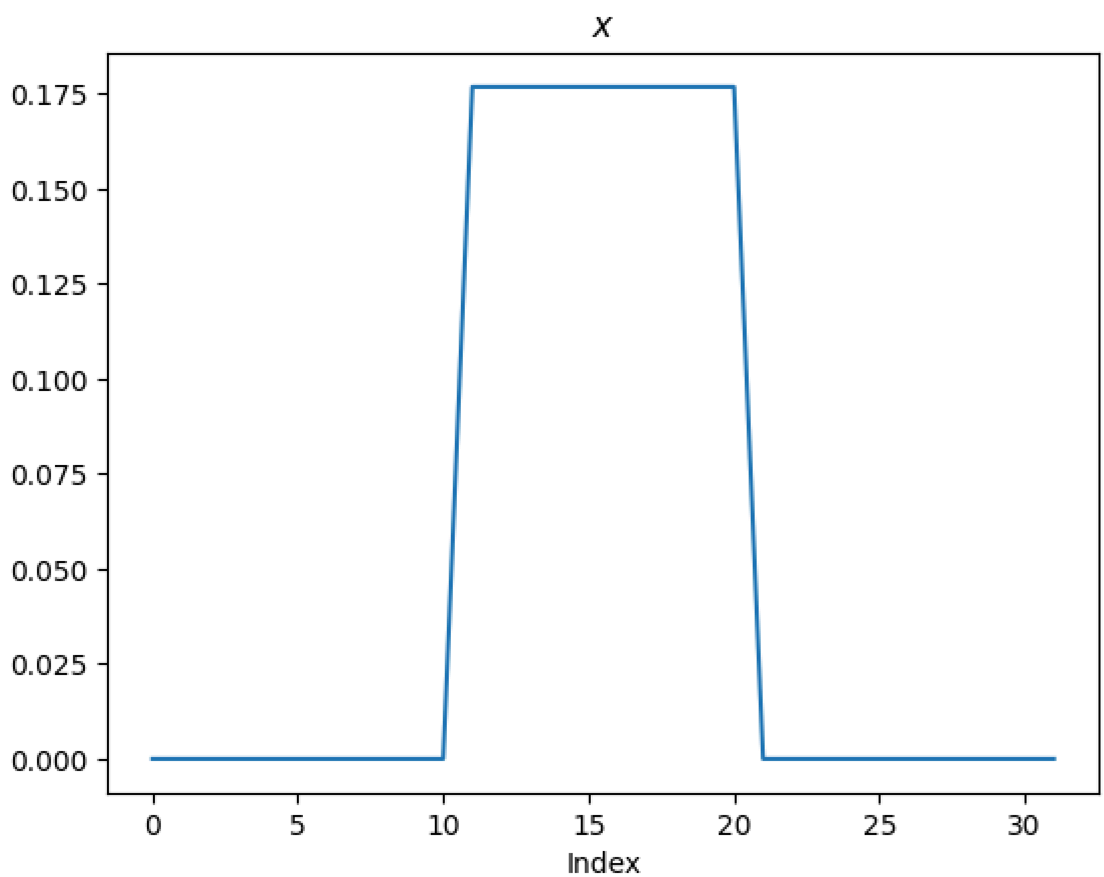
This exercise illustrates one of the limitations of TSVD and Tikhonov solutions, namely, that they are not so well suited for computing regularized solutions when the exact solution is discontinuous. We use the model problem `wing`, whose solution has has two discontinuities. Since we are mainly interested in the approximation properties of the TSVD and Tikhonov solutions, we do not add any noise in this exercise.

Part A

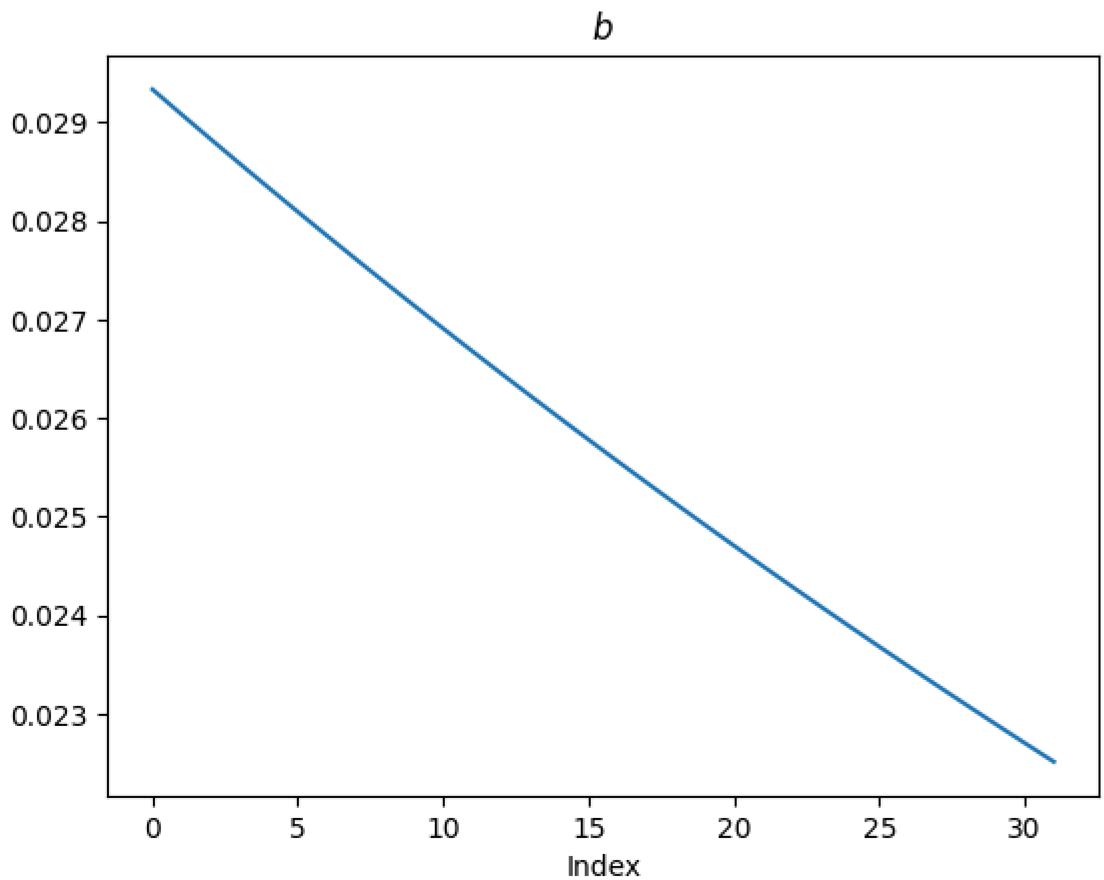
Generate the model problem using `wing`, plot the exact solution, and notice its form. Compute TSVD and Tikhonov solutions for various regularization parameters. Monitor the solutions and try to find the "best" value of k and λ. Notice how difficult it is to reconstruct the discontinuities.



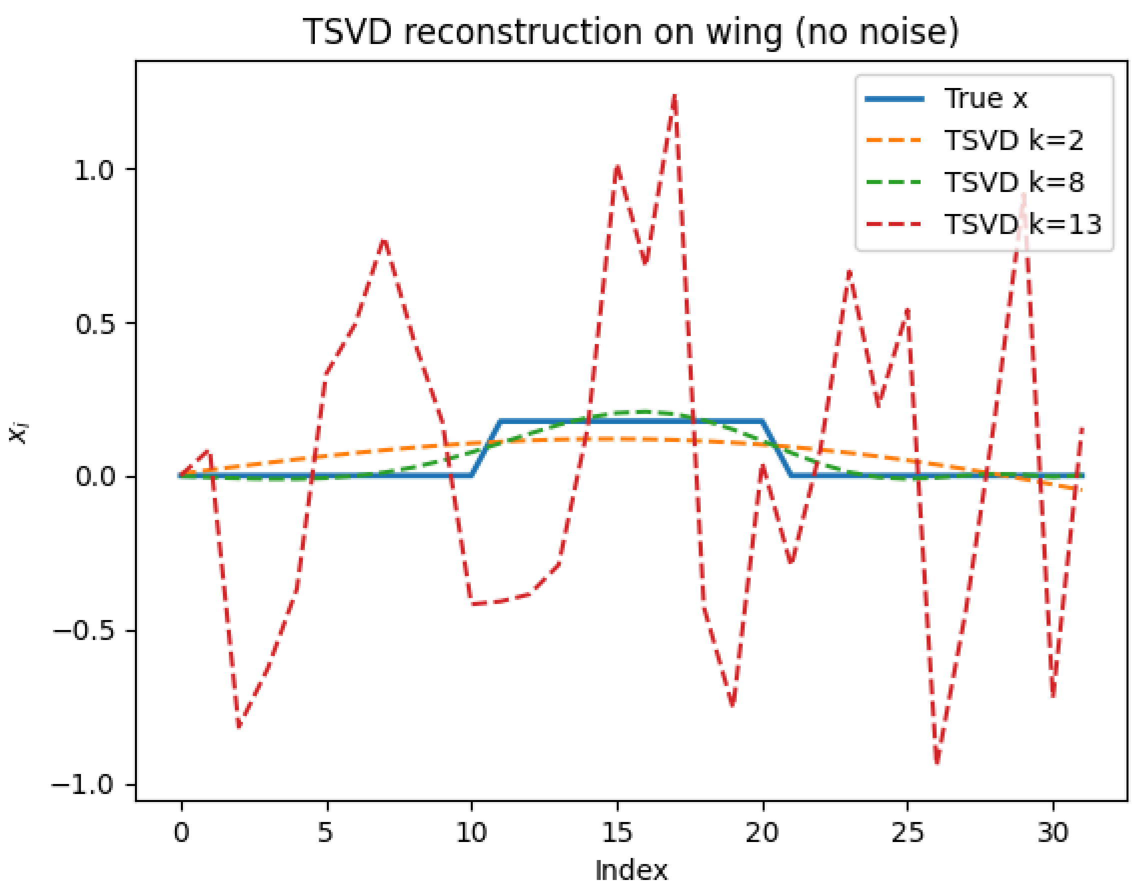
A
0.030
0.025
0.020
0.015
0.010
0.005

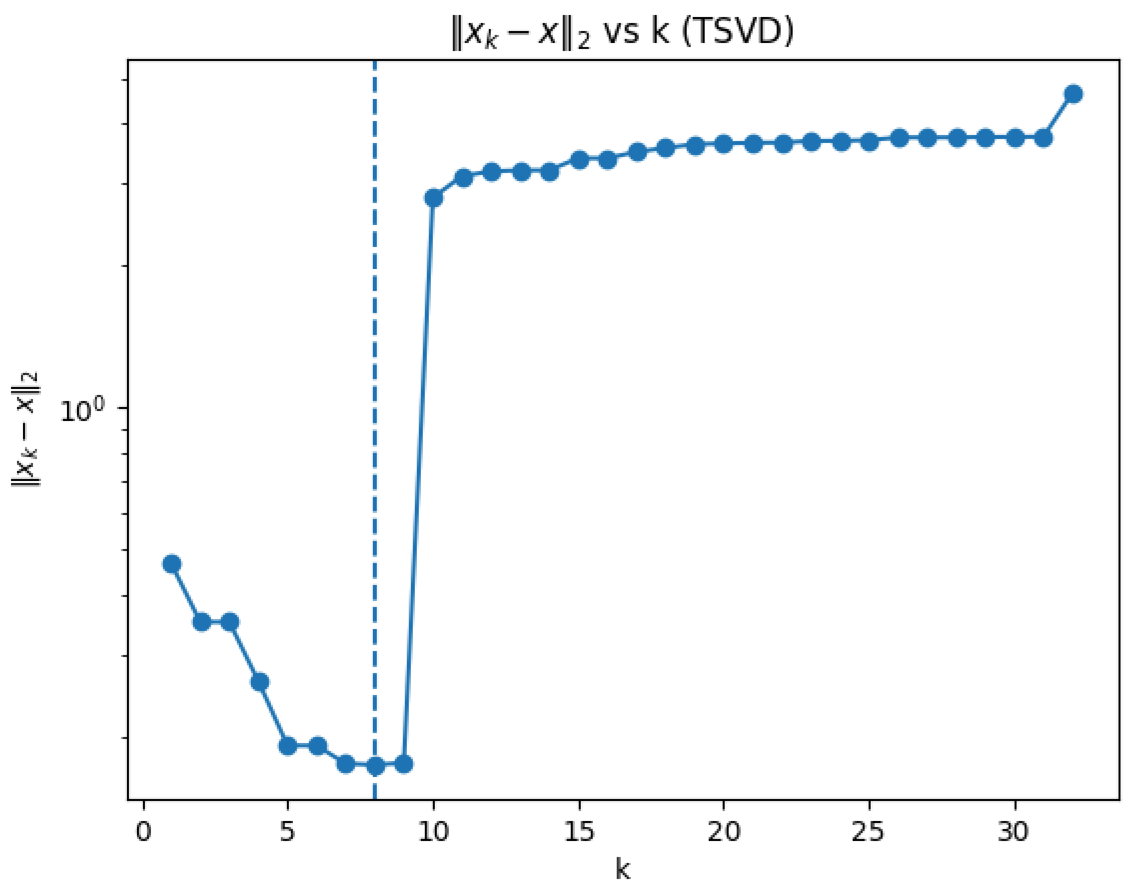


x
Index

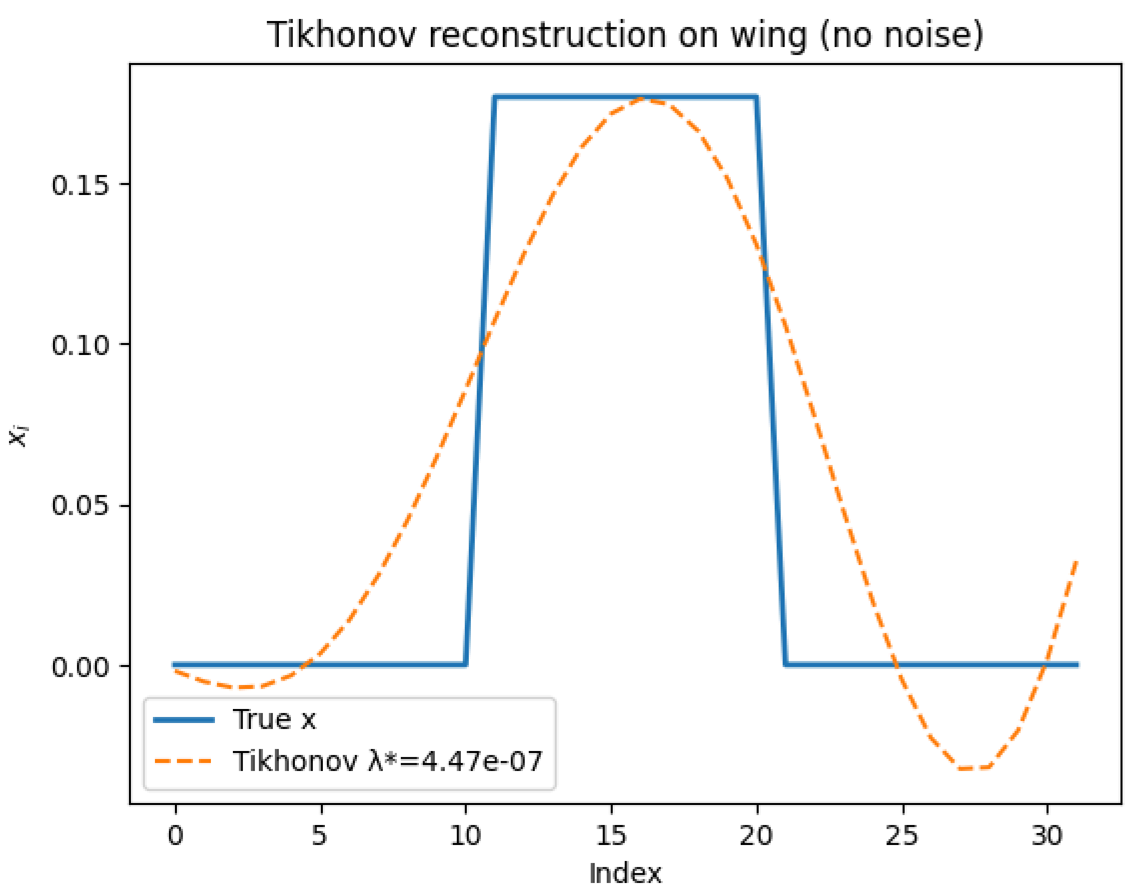


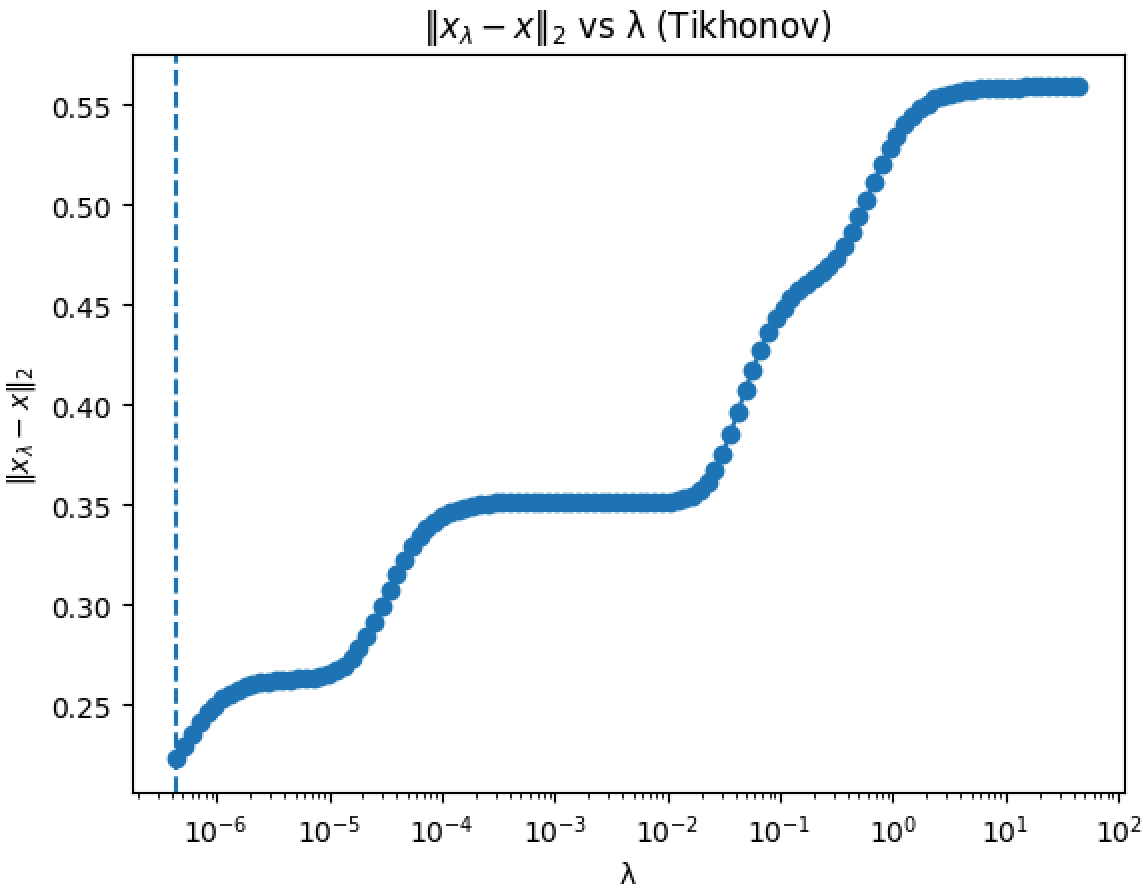

```
Best k (by true-error): k* = 8 with ||x_k* - x||_2 = 1.752e-01
```






```
Best lambda (by true-error): λ* = 4.470e-07 with ||x_λ* - x||_2 = 2.232e-01
```





Response: On the discontinuous wing problem (no noise), both TSVD and Tikhonov struggle at the jump points. TSVD needs a moderate truncation to avoid oversmoothing (small k) or oscillatory ringing (large k), while Tikhonov's quadratic penalty rounds the edges even at its optimum. In our run, the true-error optima were

$$k^* = 8, \qquad \|x_{k^*} - x\|_2 \approx 1.75 \times 10^{-1}, \qquad \lambda^* \approx 4.47 \times 10^{-7}, \qquad \|x_{\lambda^*} - x\|_2 \approx 2.23 \times 10^{-1},$$

showing TSVD preserves edges slightly better but neither method perfectly recovers the jumps.

let $b = Ax_{exact} + \eta e$, $E[e] = 0$, $Cov(e) = I_m$ $\longrightarrow$ Noise Model assumptions

Tikhonov Solution

$$x_\lambda = \min_x \{ \|Ax - b\|_2^2 + \lambda^2 \|x\|_2^2 \}$$

convert to lsqr, then normal eq $(A^T A x = A^T b)$

$$\min_x \left\| \begin{pmatrix} A \\ \lambda I \end{pmatrix} x - \begin{pmatrix} b \\ 0 \end{pmatrix} \right\|_2 = \begin{pmatrix} A \\ \lambda I \end{pmatrix}^T \begin{pmatrix} A \\ \lambda I \end{pmatrix} x = \begin{pmatrix} A \\ \lambda I \end{pmatrix}^T \begin{pmatrix} b \\ 0 \end{pmatrix}$$

$$(A^T A + \lambda^2 I) x = A^T b$$

$$x_\lambda = (A^T A + \lambda^2 I)^{-1} A^T b$$

Insert SVD of A into normal equations and use $I = VV^T$

$A = U \Sigma V^T$, where U and V are orthonormal and Σ is diag of singular values $\sigma_1 \dots \sigma_n$.

$$A^T A = (U \Sigma V^T)^T (U \Sigma V^T) \qquad U^T U = I$$

$$= V \Sigma^T U^T U \Sigma V^T$$

$$= V (\Sigma^T \Sigma) V^T$$

$$= V \Sigma^2 V^T$$

$$x_\lambda = (V \Sigma^2 V^T + \lambda^2 V V^T)^{-1} V \Sigma U^T b$$

$$= V (\Sigma^2 + \lambda^2 I)^{-1} V^T V \Sigma U^T b$$

$$= V (\Sigma^2 + \lambda^2 I)^{-1} \Sigma U^T b$$

Since Σ = diagonal of $\sigma_1 \dots \sigma_n$,

$$(\Sigma^2 + \lambda^2 I)^{-1} \Sigma = \operatorname{diag}\left(\frac{\sigma_i}{\sigma_i^2 + \lambda^2} \right)_{i=1}^{n} = D$$

let $g := U^T b \in \mathbb{R}^n$,

$$g_i = (U^T b)_i = u_i^T b$$

Diagonal Scale

$$y := Dg \quad , \text{where} \quad y_i = d_i g_i = \frac{\sigma_i}{\sigma_i^2 + \lambda^2} (u_i^T b)$$

$$Vy = \sum_{i=1}^{n} y_i v_i = \sum_{i=1}^{n} \frac{\sigma_i}{\sigma_i^2 + \lambda^2} (u_i^T b) v_i$$

Define filter factor:

$$\varphi_i^{[\lambda]} = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}, \qquad \text{Hence } \frac{\sigma_i}{\sigma_i^2+\lambda^2} = \frac{\varphi_i^{[\lambda]}}{\sigma_i}$$

$$x_\lambda = \sum_{i=1}^{n} \varphi_i^{[\lambda]} \frac{u_i^T b}{\sigma_i} v_i$$

$$A^T u_i = (U\Sigma V^T)^T u_i = V\Sigma U^T u_i = V\Sigma e_i = \sigma_i v_i$$

$$(A^T u_i)^T = u_i^T A = (\sigma_i v_i)^T = \sigma_i v_i^T$$

Using $b = Ax_{exact} + \eta e$ and $u_i^T A = \sigma_i v_i^T$

$$u_i^T b = \sigma_i (v_i^T x_{exact}) + \eta u_i^T e$$

$$x_\lambda = \underbrace{\sum_{i=1}^{n} \varphi_i^{[\lambda]} (v_i^T x_{exact}) v_i}_{\text{Deterministic Signal}} + \underbrace{\eta \sum_{i=1}^{n} \varphi_i^{[\lambda]} \frac{u_i^T b}{\sigma_i} v_i}_{\text{Random Noise}}$$

Because $x_{exact} = \sum_{i=1}^{n} (v_i^T x_{exact}) v_i$ and $E[e] = 0$

Expectation $\rightarrow$ $$E[x_\lambda] = \sum_{i=1}^{n} \varphi_i^{[\lambda]} (v_i^T x_{exact}) v_i$$

Mean Bias $\rightarrow$ $$\boxed{E(x_\lambda) = x_{exact} - \sum_{i=1}^{n} (1 - \varphi_i^{[\lambda]})(v_i^T x_{exact}) v_i}$$

The looking at the random noise segment,

$$x_\lambda - E[x_\lambda] = \sum_{i=1}^{n} \varphi_i^{[\lambda]} \frac{u_i^T b}{\sigma_i} v_i$$

Let $z := U^T e \in \mathbb{R}^n$, so $z_i = u_i^T e$, and since U orthonormal $\rightarrow$ $\mathrm{cov}(e) = \eta^2 I$

$E[z] = 0$, $\quad \mathrm{cov}(z) = \eta^2 I \quad \rightarrow \quad E[z_i z_j] = \eta^2 \delta_{ij}$

$$\begin{aligned}\mathrm{cov}(x_\lambda) &= E[(x_\lambda - E[x_\lambda])(x_\lambda - E[x_\lambda])^T] \\ &= \sum_{i,j=1}^{n} \frac{\varphi_i^{[\lambda]} \varphi_j^{[\lambda]}}{\sigma_i \sigma_j} E[z_i z_j] v_i v_j^T \\ &= \eta^2 \sum_{i=1}^{n} \frac{(\varphi_i^{[\lambda]})^2}{\sigma_i^2} v_i v_i^T\end{aligned}$$

$$\boxed{\mathrm{cov}(x_\lambda) = \eta^2 \sum_{i=1}^{n} (\varphi_i^{[\lambda]})^2 \sigma_i^{-2} v_i v_i^T}$$