

# Mixed Precision Iterative Refinement for Linear Inverse Problems

*Lucas Onisk and James G. Nagy*

## Abstract

We are interested in linear discrete inverse problems which involve the reconstruction of objects or signals from noisy observed data. The available linear system is given by

$$Ax + e = b, \quad (1)$$

where  $A \in \mathbb{R}^{m \times n}$ ,  $m \geq n$  is a matrix whose singular values decay without significant gap and cluster at the origin (i.e., the matrix is ill-conditioned). Discrete inverse problems can arise through the discretization of Fredholm integral equations of the first kind; see [5, 3], but can also arise in massive data streaming problems such as the training of the random feature model in machine learning [8] or limited angle imaging problems including, for example, those from medical imaging [2]. To derive a meaningful solution from the available problem, regularization is needed.

In Tikhonov regularization, the least-squares problem associated with (1) is replaced by the penalized least-squares problem

$$\min_{x \in \mathbb{R}^n} \left\{ \|Ax - b\|_2^2 + \alpha^2 \|Lx\|_2^2 \right\} \quad (2)$$

where  $\alpha > 0$  is a regularization parameter that balances the sensitivity of the solution vector to the error in  $b$ , as well as the closeness to the desired solution of the unavailable error-free problem. When the regularization matrix  $L \in \mathbb{R}^{s \times n}$  is chosen so that the null spaces of  $A$  and  $L$  trivially intersect then the solution of (2) may be written in closed form.

Iterative refinement (IR) has long been utilized as an iterative strategy to improve the accuracy of numerical solutions to linear systems of equations. Recent works by Higham and collaborators have considered the use of IR in conjunction with mixed precision computing in light of recent advancements in hardware capabilities; see [6, 1]. Our interests of studying IR applied to the Tikhonov problem were motivated by the work [7] which considered the solution of symmetric positive definite linear systems and least-squares problems in mixed precision which showed regularization to be a key requirement when computing low precision factorizations.

The  $k^{th}$  iterate of IR applied to the Tikhonov problem in standard form,  $(A^T A + \alpha^2 I)x^{(k)} = A^T b$ , where  $L = I$  may be written recursively as

$$x^{(k)} = x^{(k-1)} + (A^T A + \alpha^2 I)^{-1} A^T r^{(k-1)} - \alpha^2 (A^T A + \alpha^2 I)^{-1} x^{(k-1)} \quad (3)$$

where  $r^{(k-1)} = b - Ax^{(k-1)}$  denotes the  $(k-1)^{th}$  residual. Riley in [9] and Golub in [4] note that the IR procedure in (3) is equivalent to iterated Tikhonov regularization in exact arithmetic whose  $k^{th}$  iterate is given by

$$x^{(k)} = x^{(k-1)} + (A^T A + \alpha^2 I)^{-1} A^T r^{(k-1)}$$

which may be interpreted as a preconditioned Landweber method, or, from a mathematical optimization point of view - a preconditioned gradient descent method with fixed step size.

To better understand the application of mixed precision IR applied to the Tikhonov problem we derive a methodology to formulate the iterates as filtered solutions by writing them as a recursive

relationship between the iterates of preconditioned Landweber with a Tikhonov-type preconditioner and previous iterates. A filtered solution is of the form

$$x_{filt} = \sum_j \phi_j \frac{u_j^T b}{\sigma_j} v_j$$

where vectors  $u_j$  and  $v_j$  correspond to left and right singular vectors of  $A$ , respectively. The  $\sigma_j$  correspond to the singular values of  $A$ . An intelligent selection of the filter factors  $\phi_j$  can remove deleterious components of the approximate solution to the least-squares problem stemming from (1). By considering a filtered solution, we are able to study the effect that each level of precision utilized in IR has on (i) the quality of the approximate solution and (ii) the number of iterations the algorithm requires to terminate according to some termination criterion. We demonstrate in our numerical results that mixed precision IR on the Tikhonov problem gives comparable or superior accuracy against results computed in double precision as well as another benchmark which supports its use in modern applications that natively support mixed precision floating-point arithmetic.

## References

- [1] E. Carson, N. J. Higham, and S. Pranesh, *Three-precision GMRES-based iterative refinement for least squares problems*, SIAM Journal on Scientific Computing, 42 (2020), pp. A4063–A4083.
- [2] J. Chung and L. Nguyen, *Motion estimation and correction in photoacoustic tomographic reconstruction*, SIAM Journal of Imaging Sciences, 10 (2017), pp. 216–242.
- [3] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.
- [4] G. Golub, *Numerical methods for solving linear least squares problems*, Numerische Mathematik, 7 (1965), pp. 206–216.
- [5] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*, SIAM, Philadelphia, 1998.
- [6] N. Higham and T. Mary, *Mixed precision algorithms in numerical linear algebra*, Acta Numerica, 31 (2022), pp. 347–414.
- [7] N. J. Higham and S. Pranesh, *Exploiting lower precision arithmetic in solving symmetric positive definite linear systems and least squares problems*, SIAM Journal on Scientific Computing, 43 (2021), pp. A258–A277.
- [8] A. Rahim and B. Recht, *Random features for large-scale kernel machines*, Advances in Neural Information Processing Systems, 20 (2007), p. 1177–1184.
- [9] J. D. Riley, *Solving systems of linear equations with a positive definite, symmetric, but possibly ill-conditioned matrix*, Mathematical Tables and other Aids to Computation, (1955), pp. 96–101.