Math 76 Final Project:

Wade Williams, Grace Faulkner, Ryan Sorkin, Isabelle Lewitt,
with help from Jonathan Lindbloom

August 2022

1 Introduction

The goal of our project was to compare the performance of the selected numerical algorithms for solving three inverse problems (i) an MRI-like tomographic reconstruction problem, (ii) a computed tomography problem, and (iii) an image de-blurring problem. Specifically, we compare the performance of the Split-Bregman algorithm, the alternating direction method of multipliers (ADMM), linearized ADMM (LADMM), and the primal-dual hybrid gradient method (PDHG). For both problems, we apply anisotropic total variation (TV) regularization (without boundary conditions) to turn the ill-posed inverse problem into a well-posed one. For the Split-Bregman iteration, we use the implementation provided by Rick Archibald (in MATLAB, Jonathan translated it to Python for us). For the remaining optimizers, rather than coding these algorithms from scratch we use Los Alamos National Laboratory's (LANL) Scientific Computational Imaging Code (SCICO) [1] that Jonathan helped us learn how to use.

For the MRI-like problem, we assume that the Fourier data has been generated according to

$$\hat{f} = PFx + \varepsilon, \tag{1}$$

where $x \in \mathbb{R}^{m \times n}$ is the unobserved true signal, F is a two-dimensional discrete Fourier transform (DFT), P is an under-sampling matrix that discards components of the DFT Fx, and ε represents i.i.d. noise. To recover the unknown signal x using TV regularization, we seek the solution to the convex optimization problem

$$x^\star = \text{argmin}_x \; \|PFx - \hat{f}\|_2^2 + \lambda \|\text{TV}(x)\|_1 \tag{2}$$

which can be thought of as a blend between the data and the regularization. An example diagram for the MRI problem is given in Figure 1.

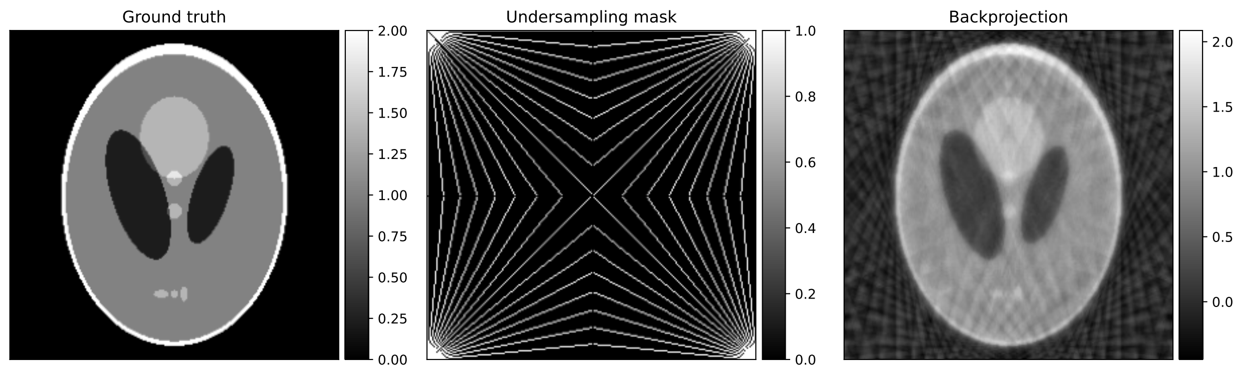

Figure 1: The problem setup for the MRI-like tomographic reconstruction problem. (a) the ground truth image. (b) The binary undersampling mask P. (c) A simple backprojection $\hat{x}$ obtained by computing $\hat{x} = F^H P^T \hat{f}$.

$$x^\star = \mathrm{argmin}_x \ \frac{1}{\sigma_y^2}\|Ax - \hat{f}\|_\Lambda^2 + \lambda\|\mathrm{TV}(x)\|_1 \tag{3}$$

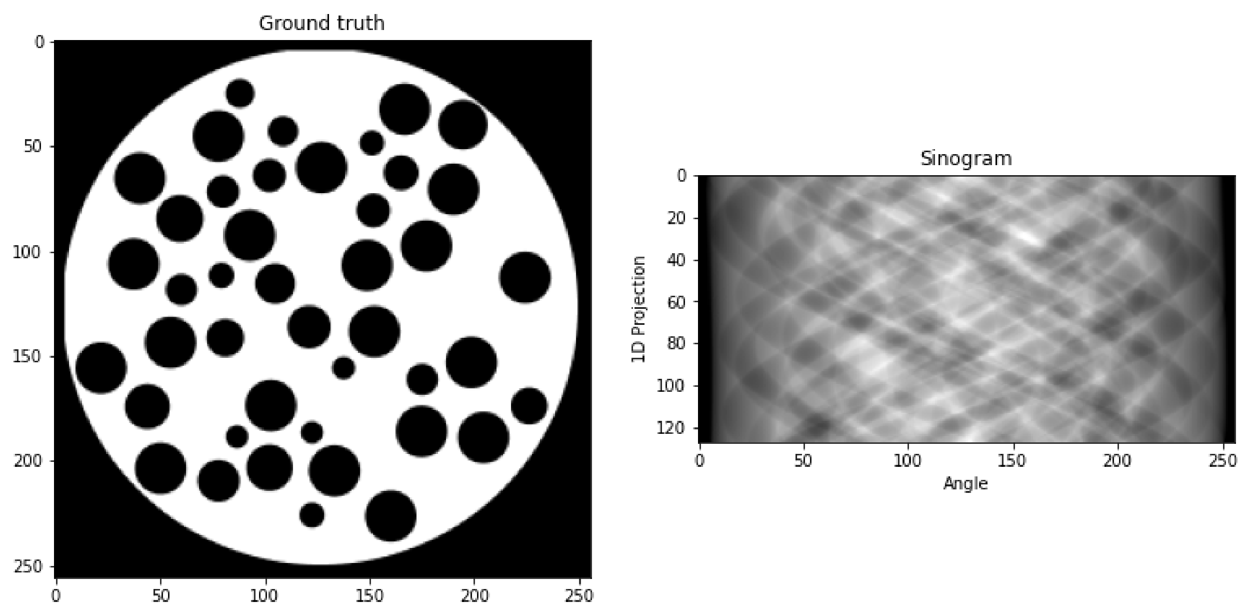

Figure 2: Sinogram Image Data

For the image de-blurring problem, we took the large 675×1200 pixel `vader` test image x and applied a circulant box blur (periodic boundary conditions) of radius 20, which is shown in Figure 3. This corresponds to the observation

$$\hat{y} = Ax \tag{4}$$

where A is the circulant blurring matrix and $\hat{y}$ is the observation, and our TV-deblurring problem is to solve the convex optimization problem

$$x^{\star} = \operatorname{argmin}_x \ \|Ax - \hat{y}\|_2^2 + \lambda \|\mathrm{TV}(x)\|_1. \tag{5}$$

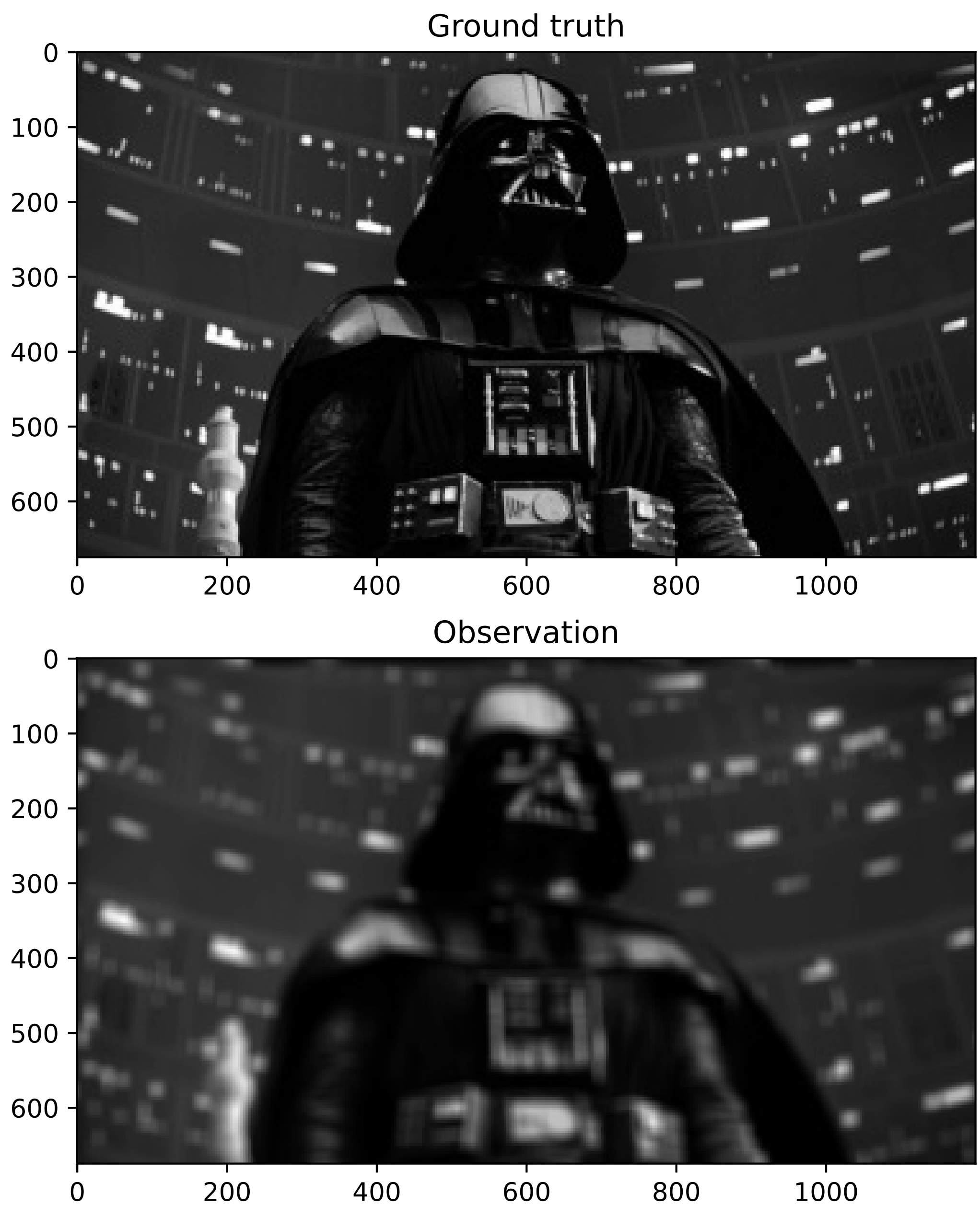

Figure 3: Darth Vader test image along with circulant box blur (radius 20) of test image.

The benefit of using a circulant blur is that the optimization algorithms can avoid using the conjugate gradient (CG) method for one of the sub-problems and instead use discrete Fourier transforms, resulting in much quicker iterations.

2 Algorithms

2.1 Split Bregman

2.2 ADMM

The Alternating Direction Method of Multipliers, or ADMM, is the algorithm intended to blend the decomposability of dual ascent with the convergence properties of the method of multipliers.The ADMM algorithm solves problems in the following form;

$$\text{argmin}_{(x,z)} f(x) + g(z), \tag{6}$$

$$Ax + Bz = c \tag{7}$$

2.2.1 Assumptions and Convergence

In order for the ADMM to work, we rely on the notion that both f(x) and g(z) are convex, and more specifically on the two following assumptions.

Assumption 1 : The extended real valued functions

$$\text{f} : R^n \rightarrow R \cup (+\infty) \;\; and \;\; g : R^m \rightarrow R \cup (+\infty) \tag{8}$$

This assumption suggests that there exists x and z, not necessarily unique, that minimise the augmented Lagrangian. Furthermore, the assumption allows for f and g to be non-differentiable and to assume the value of $+\infty$.

Assumption 2 : The unaugmented Lagrangian, L_0 has a saddle point. There exists, (x^*, y^*, z^*), not necessarily unique, for which;

$$\text{L}_0(x^*, z^*, y) \leq L_0(x^*, z^*, y^*) \leq L_0(x, z, y^*) \tag{9}$$

Thus by Assumption 1, it follows that $L_0(x^*, z^*, y^*)$ is finite for any saddle point, (x^*, z^*, y^*). Therefore, (x^*, z^*) is a solution to the aforementioned (6) and (7), and $f(x^*) < \infty$ and $g(z^*) < \infty$.

Together Assumptions 1 and 2, demonstrate that the ADMM satisfies the following:

- Residual Convergence; this shows that the iterates approach feasibility.

$$r^k \rightarrow 0 \text{ as } k \rightarrow \infty$$

- Objective Convergence;this shows that the objective function of the iterates approaches the optimal value.

$$f(x^k) + g(z^k) \rightarrow p^* \text{ as } k \rightarrow \infty$$

- Dual Variable Convergence; this shows that y^* is a dual optimal point.

$$y^k \rightarrow y^* \text{ as } k \rightarrow \infty$$

2.2.2 Stopping Criterion and Optimality

The optimality solutions in the ADMM problem posed in (6) and (7), are defined by three conditions, as outlined below, split into primal and dual feasibility. The final condition holds for $(x^{k+1}, y^{k+1}, z^{k+1})$, and the residuals of the first two conditions are the primal and dual residuals respectively, r^{k+1} and s^{k+1} , and both of these residuals converge to 0 as the ADMM proceeds. Here the delta, δ , acts as the subdifferential operator.

$$
\begin{array}{ccc}
Ax^* + Bz^* - c = 0 & \rightarrow & PrimalFeasibility \\
0 \in \delta\mho(\curvearrowleft^*) + A^T y^* \text{ and } 0 \in \delta\eth(F^*) + B^T y^* & \rightarrow & DualFeasibility
\end{array}
$$

The residuals of the optimality conditions above, can be related to a bound on the objective suboptimality of the current point. The convergence proof shows that when the residuals r^k and s^k are small, the objective suboptimality must also be small. As a result of this we can guess that $||x^k - x^*|| \leq d$ and thus,

$$
f(x^k) + g(z^k) - p^* \leq -(y^k)^T r^k + d||s^k|| \leq ||y^k||||r^k|| + d||s^k||
$$

By using the middle or right hand terms as an approximate bound on the objective suboptimatility, we can determine that a reasonable termination criterion is when the primal and dual residuals are small.

2.3 LADMM

Though the ADMM algorithm as expanded upon above is often the most effective method used, sometimes the proximal mapping of the function $f(x)$ cannot be explicitly computed, thus making ADMM inefficient. In such scenarios, we instead linearise $f(x)$ therefore using the Linear Alternating Direction Method of Multiples (LADMM). The LADMM algorithm solves the problems presented in (6) and (7) by generating a sequence $(x^{k+1}, y^{k+1}, z^{k+1})$. The algorithm is as follows;

Initialize an iteration counter $k \leftarrow 0$ and a bounded starting point (x^0, λ^0, z^0) .
Repeat.
Update x^{k+1} according to its closed form solution: $x^{k+1} = \operatorname{argmin}_x \left\{ \|PFx - \hat{f}\|_2^2 + \frac{\varrho}{2}\|z^k - \mathrm{TV}(x)\|_2^2 \right\}$
.
$\lambda^{k+1} \leftarrow \lambda^k - B(((B^T B)^{-1}(B^T A))x^{k+1} - z^k)$
Update z^{k+1} according to the generated sequence: $z^{k+1} = \operatorname{argmin}_z \left\{ \lambda\|z\|_1 + \frac{\varrho}{2}\|z - \mathrm{TV}(x^{k+1})\|_2^2 \right\}$.
if some stopping criterion is satisfied; then
Break;
else
$k \leftarrow k + 1$;
end if
until exceed the maximum number of outer loop

2.4 PDHG

PDHG is a first order method, meaning that is only requires the functional and gradient evaluations. It is also a primal-dual method, meaning that each iteration updates both a primal and a dual variable. As a result of this the PDHG method is often able to avoid many of the difficulties that arise when only using a primal or dual method. Such a difficulty can occur for example in TV minimisation when the gradient descent is applied to the primal functional runs into issues where the gradient of the solution is zero because the functional is not differentiable at that point. The PDHG method begins with a saddle point formulation of the problem at hand and then proceeds by alternating proximal steps that alternately minimise and maximise a penalised form of the saddle point. The PDHG algorithm for TV deblurring is as follows;

$$\mathrm{p}^{k+1} = \Pi_x(p^k + \tau_k \lambda D u^k) \tag{10}$$

$$\mathrm{u}^{k+1} = ((1-\Theta_k) + \Theta_k K^T K)^{-1}((1-\Theta_k)u^k + \Theta_k(K^T f - 1/\lambda D^T p^{k+1})) \tag{11}$$

3 Numerical Results

3.1 MRI Problem

For the MRI problem, we looked at all four of the aforementioned optimization algorithms. For all algorithms, we used the same number of iterations and value for the regularization parameter $\lambda = 1000$, which was selected through experimentation with the codes. A caveat here is that while we were able to get the Split-Bregman code to work, we were unable to determine where the λ parameter was being controlled in the code. Thus, in terms of our analysis we present (i) Split-Bregman and (ii) ADMM/LADMM/PDHG separately, as we are unsure if Split-Bregman is truly solving the same problem with the same value for λ.

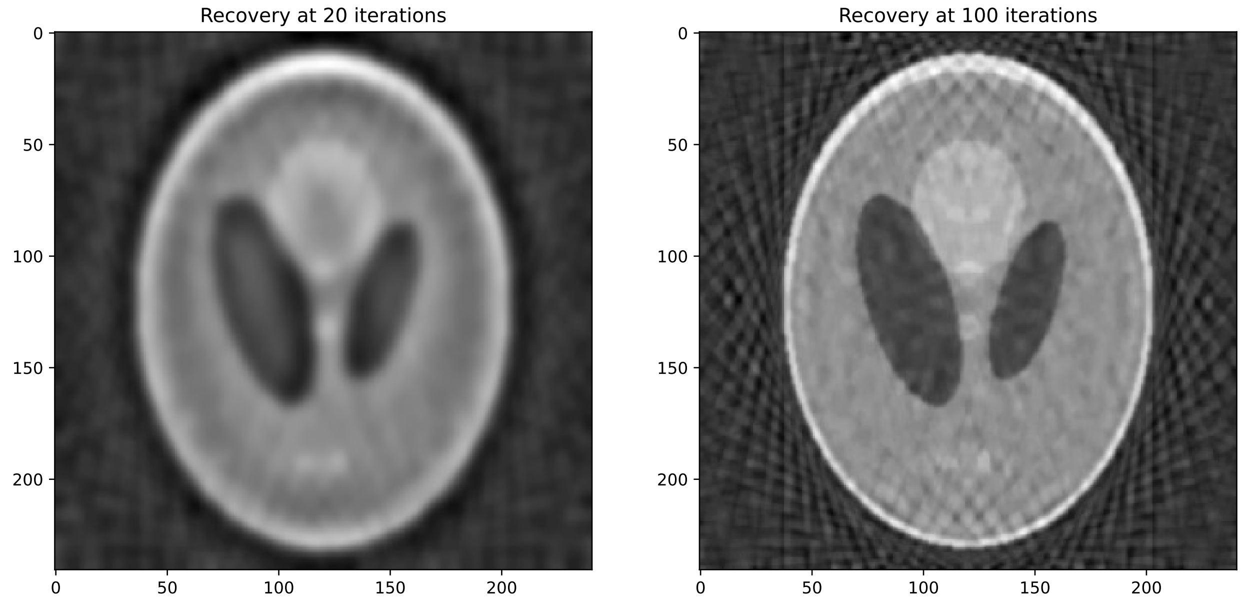

Figure 4: Split-Bregman iteration (fast algorithm, but iterations take too long since we didn't use GPU code for this.

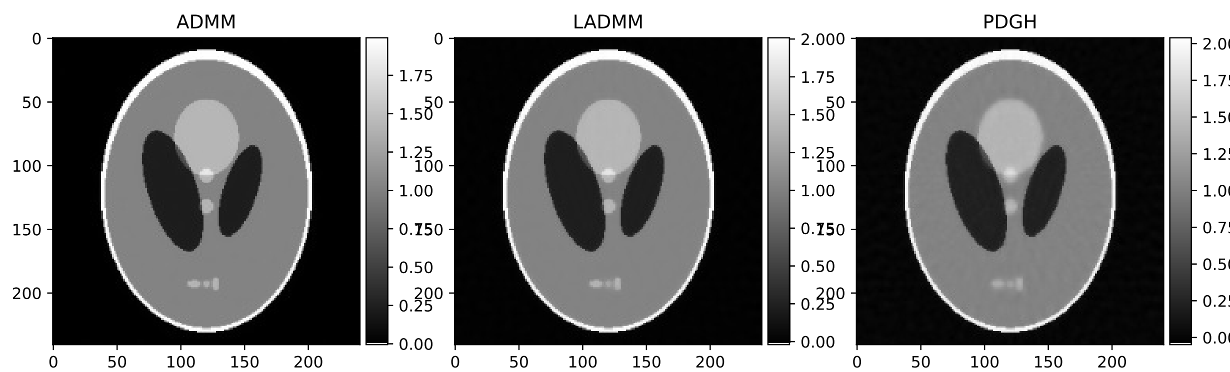

Figure 5: ADMM, LADMM, PDHG iterations. All achieve final reconstructions of similar accuracy when not constraining the number of iterations or optimizing the λ parameter.

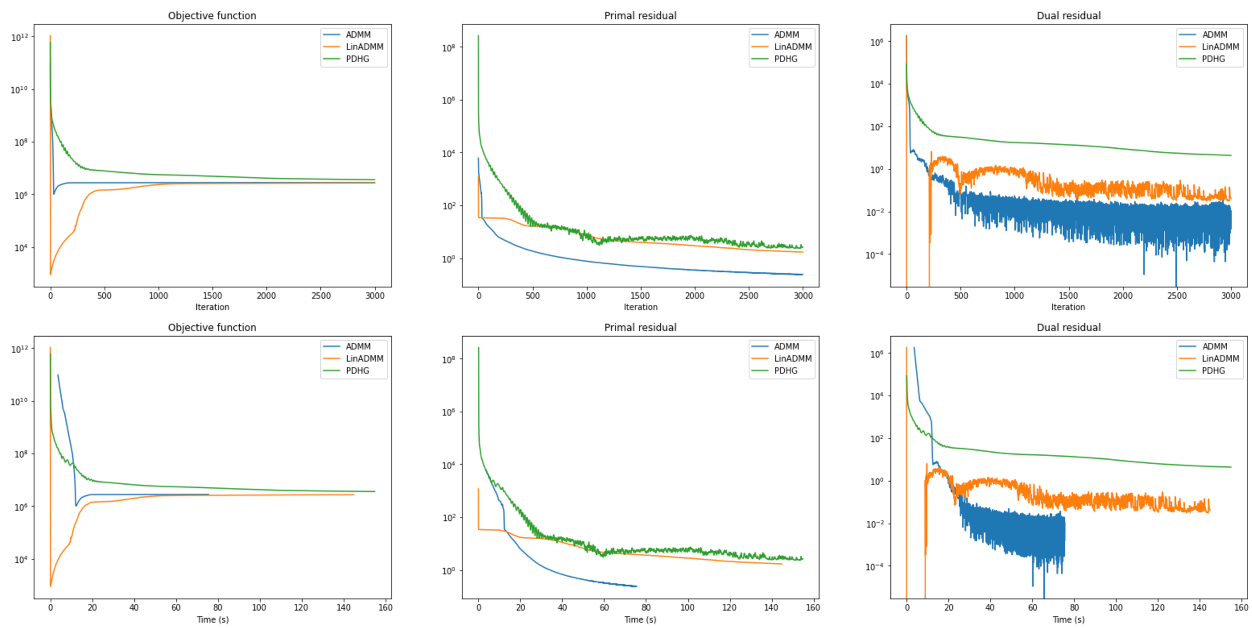

Figure 6: MRI solver objective comparison

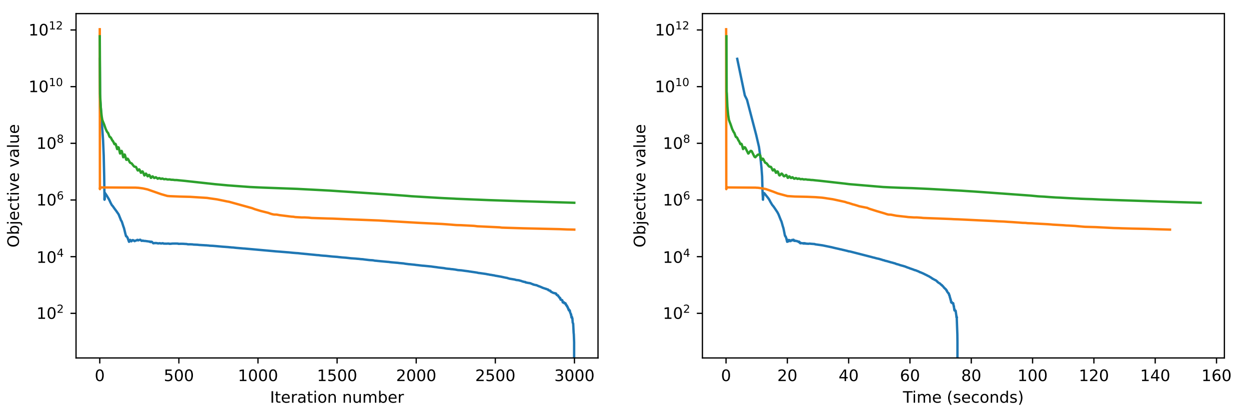

Figure 7: MRI solver objective comparison relative to ADMM

While figure 5 shows little visual difference between the ADMM, the LADMM, and the PDHG algorithms, and all appear to achieve relatively similar reconstructions when the number of iterations is not restricted. When the number of iterations is restricted, there is a distinct difference between the three methods. As demonstrated in figure 7, the objective value is higher when using PDHG and LADMM than ADMM for nearly all iterations. Furthermore, figure 9 shows that with both time (minus the first 10 seconds) and iteration, the objective function, primal residual and dual residual are more effective with ADMM, followed by the Linear ADMM and then PDHG. This therefore suggests that the ADMM is the most effective method to solve these inverse problems, followed by the LADMM and then the PDHG.

3.2 CT Problem

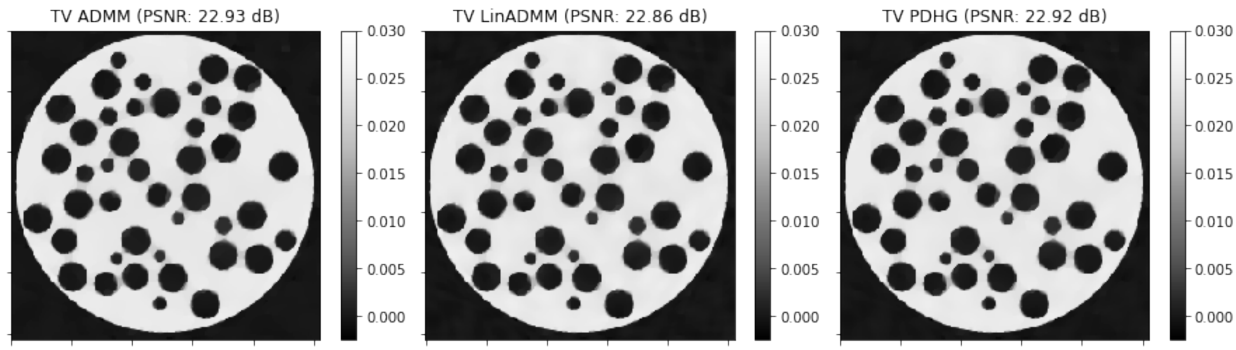

Figure 8: ADMM, LADMM, PDHG iterations. Again, all sinogram reconstructions achieve similar accuracy.

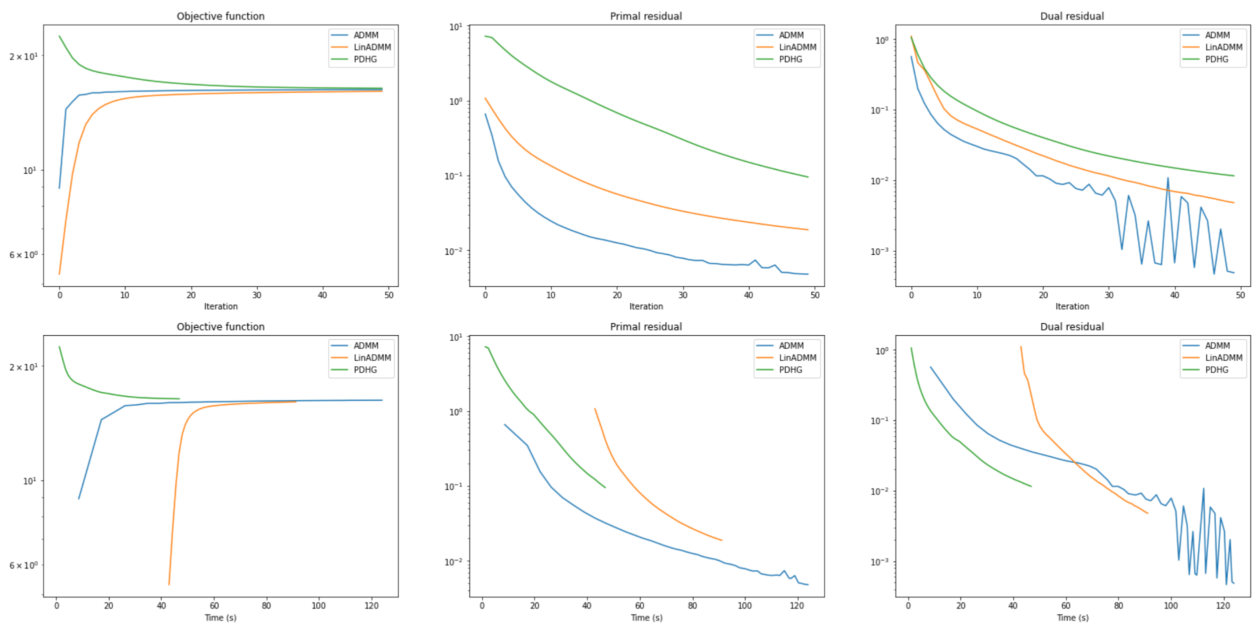

Figure 9: CT solver objective comparisons

For the CT problem, we see again that all three algorithms reach reconstructions of similar accuracy when not restricting the number of iterations or run time. In this problem, however, ADMM, LADMM, and PDHG each reach nearly the exact same objective function in very few iterations. For few iterations, ADMM still performs the best as it did in the MRI problem, but PDHG actually performs better than ADMM with respect to time. This isn't entirely unexpected as ADMM uses the conjugate gradient twice (once in each subproblem), making it more costly with very few iterations, but generally more efficient. In this case, PDHG is much faster than both ADMM and LADMM with respect to time.

3.3 De-blurring Problem

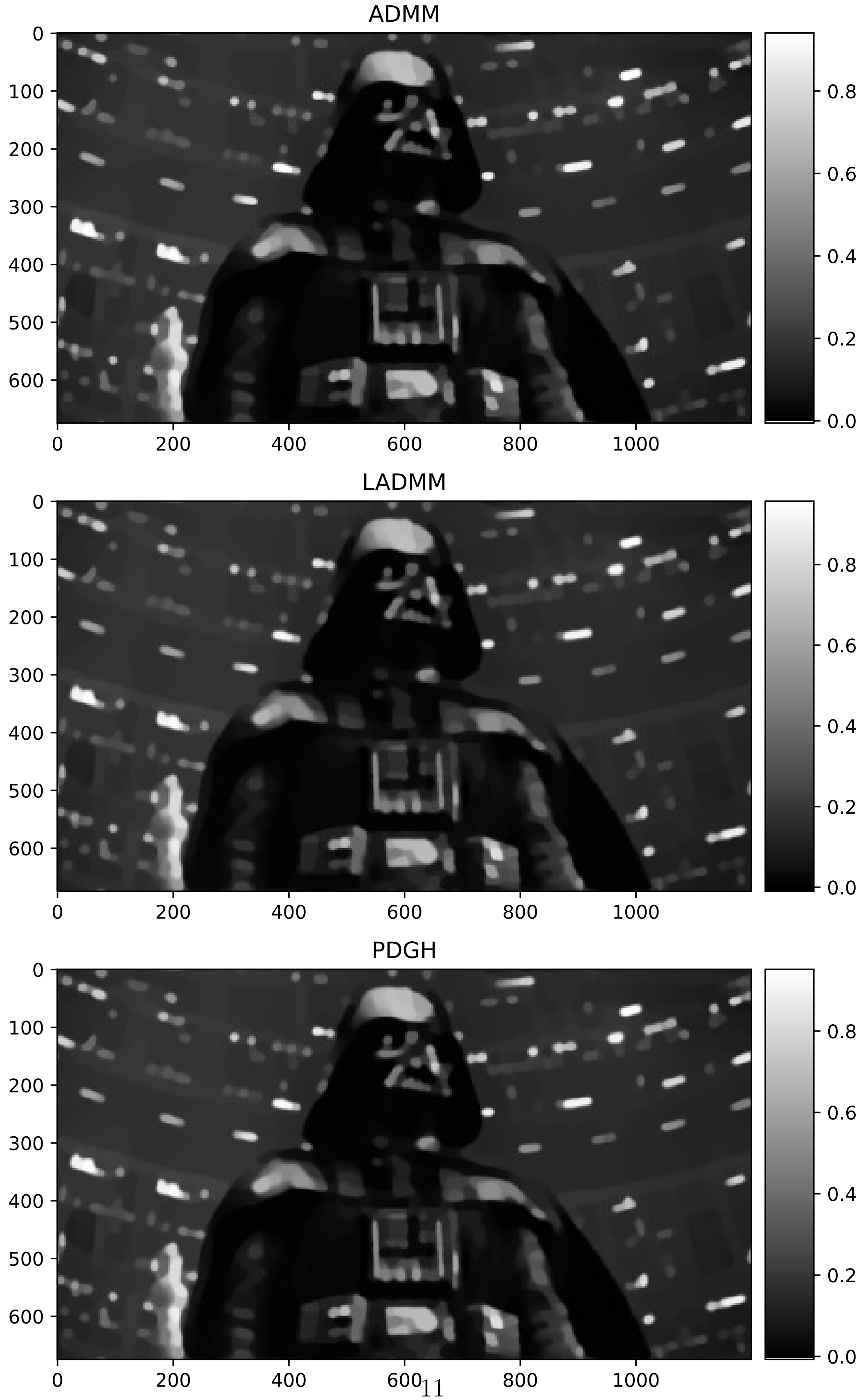

Figure 10: Deblurring reconstructions

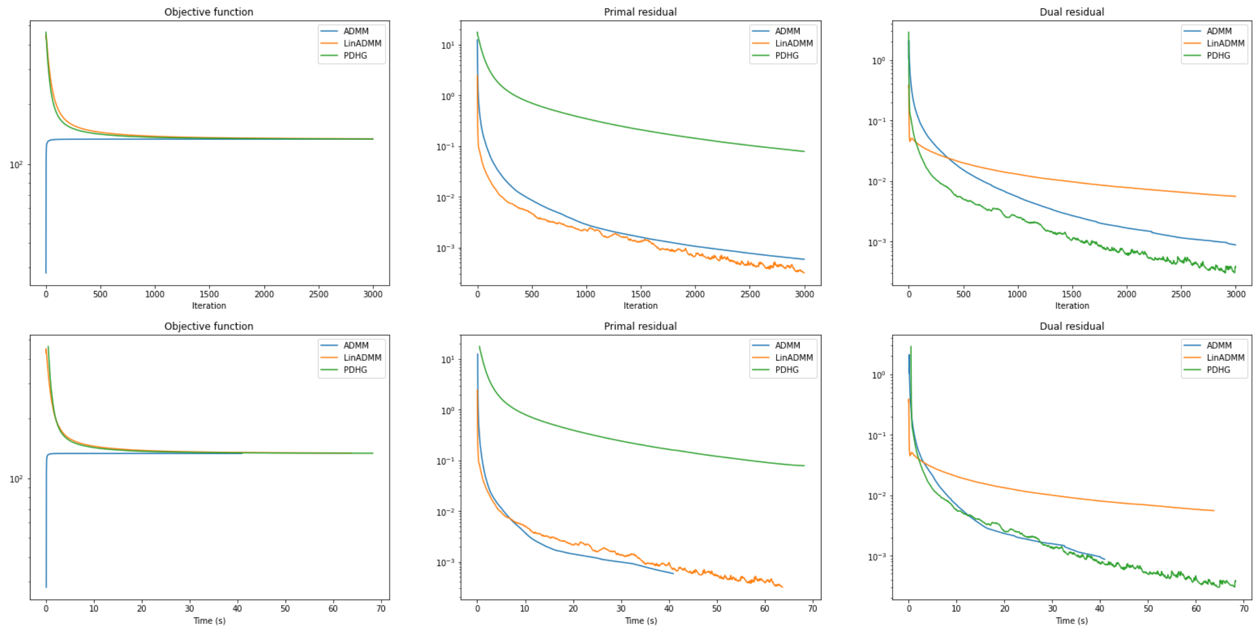

Figure 11: Deblurring solver objective comparisons

For our de-blurring problem, we see that ADMM is superior to LADMM and PDHG with respect to both iterations and time, but each algorithms does eventually reach a very similar solution to the others. The theme of the performance of these three algorithms is consistent across problems. In each case, ADMM converges with the fewest iterations, but takes much longer to run for a small number of iterations. Its superior rate of convergence with respect to the number of iterations eventually overcomes its initial longer run time as the number of iterations increases.

4 Conclusion

MRI and CT reconstruction and image de-blurring are important applications for mathematicians, and the efficiency and accuracy at which these problems and similar ones can be performed is a critical area of research. In this project, we compared the performance of the Split-Bregman algorithm, the alternating direction method of multipliers (ADMM), linearized ADMM (LADMM), and the primal-dual hybrid gradient method (PDHG) algorithms for solving (i) an MRI-like tomographic reconstruction problem, (ii) a computed tomography problem, and (iii) an image de-blurring problem. We have shown that the performance of these algorithms varies by the nature of the problem and its constraints. We have also investigated the accuracy of each, and while we do not draw clear cut conclusions from the data, it is evident that the computational cost and accuracy significantly varies across problems and methods.

References

[1] Thilo Balke et al. *Scientific Computational Imaging COde (SCICO)*. Software library available from https://github.com/lanl/scico. 2022.