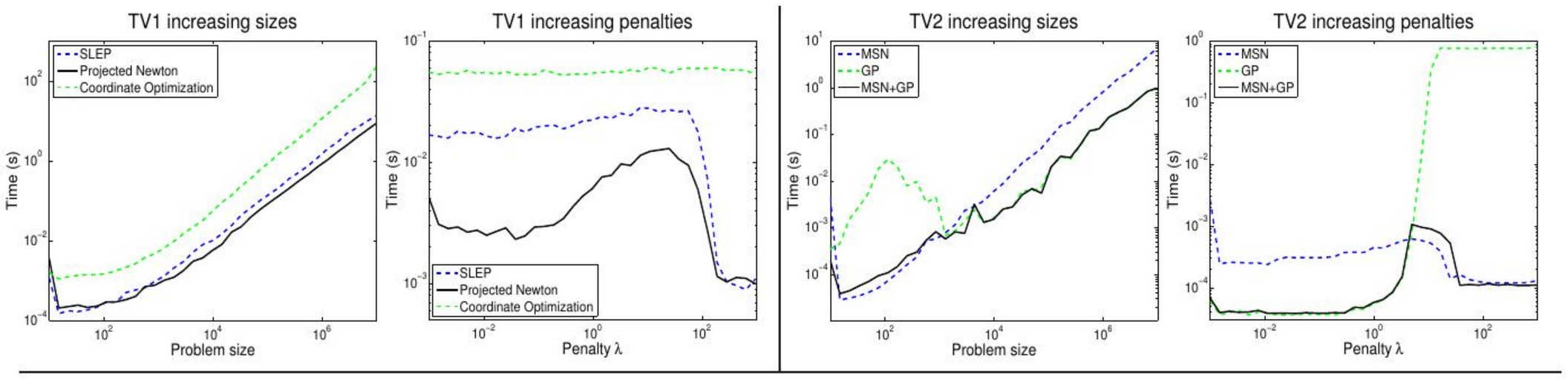

Figure 1. Running times (in secs). Left panel: PN, SLEP and Coordinate Optimization solvers for Tv_1^{1D}-proximity with increasing input sizes and penalties. Right panel: MSN, GP, and hybrid MSN-GP solvers for Tv_2^{1D}-proximity with increasing input sizes and penalties.

Table 1. Running times (in milliseconds) for PN, SLEP and Coordinate Optimization solvers for TV-L1 problems with increasing input sizes (in log-scale); n denotes problem size.

$\log_{10} n$	SLEP	PN	COORD.
1.00	**1.19**	3.6561	1.63
1.53	**0.17**	0.2476	1.37
2.00	0.30	**0.23**	1.52
2.58	0.59	**0.41**	2.69
3.11	1.33	**1.04**	6.74
3.64	5.25	**3.10**	22.20
4.17	15.10	**8.22**	92.41
4.70	67.60	**39.35**	359.50
5.23	221.58	**137.81**	1550.27
5.75	759.62	**464.32**	5678.25
6.28	2874.83	**1655.25**	23285.00
6.81	9457.11	**5659.42**	93366.00

plicitly takes advantage of this set by updating only the not indexed by I. On the other hand, for large λ, PN's strategy becomes similar to that of SLEP, hence the similar performance. Finally, as Coordinate Optimization computes the full regularization path, its runtime is invariant to λ.

4.2. Results for TV-L2

To compare the running times of MSN and GP, we again use duality gap of 10^{-5} as the stopping criterion. Further, as MSN might generate infeasible solutions during the optimization, we also apply a boundary proximity criterion for MSN with tolerance 10^{-6}. Looking at the results it can be seen that the performance of MSN and GP differs noticeably in the two experimental scenarios. While Figure 1 (first plot; right panel) might indicate that GP converges faster than MSN for large inputs, it does so depending on the size of λ relative to $\|\boldsymbol{y}\|_2$. Indeed, Figure 1 (last plot) shows that although for small values of λ, GP runs faster than MSN, as λ increases, GP's performance worsens dramatically, so much that for moderately large λ it is unable to find an acceptable solution even after 10000 iterations (an upper limit imposed in our implementation). Conversely, MSN finds a solution satisfying the stopping criterion under every situation, thus showing a more robust behavior. Therefore, we propose a hybrid approach that combines the strengths of MSN and GP. This hybrid is guided using the following (empirically determined) rule of thumb: if $\lambda < \|\boldsymbol{y}\|_2$ use GP, otherwise use MSN. Further, as a safeguard, if GP is invoked but fails to find a solution within 50 iterations, the hybrid should switch to MSN. This combination guarantees rapid convergence in practice. Results for this hybrid approach are included in the plots in Figure 1, and we see that it successfully mimics the behavior of the better algorithm amongst MSN and GP.

5. Numerical Results: Applications

To highlight the potential benefits of our algorithms we show below three main applications: (i) fused-lasso; (ii) image denoising; and (iii) image deblurring. However, we note here that the exact application itself is not as much a focus as the fact that our solvers apply to a variety of applications while leading to noticeable empirical speedups.

5.1. Results for 1D Fused-Lasso

Our first application is to fused-lasso for which we plug in our algorithms as subroutines into the generic TRIP solver of Kim et al. (2010). We then apply TRIP to solve the following variants of fused-lasso:

1. **Fused-lasso (FL):** Here $L(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2$, and $R(\boldsymbol{x}) = \lambda_1\|\boldsymbol{x}\|_1 + \lambda_2\|\boldsymbol{D}\boldsymbol{x}\|_1$; this is the original fused-lasso problem introduced in (Tibshirani et al., 2005), and used in several applications, such as in bioinformatics (Tibshirani & Wang, 2008; Rapaport & Vert, 2008; Friedman et al., 2007).
2. **ℓ_2-variable fusion (VF):** Same as FL but with $\lambda_2\|\boldsymbol{D}\boldsymbol{x}\|_2$ instead. This FL variant seems to be new.
3. **Logistic-fused lasso (LFL):** A logistic loss $L(\boldsymbol{x}, c) = \sum_i \log\left(1 + e^{-\boldsymbol{y}_i(\boldsymbol{a}_i^T\boldsymbol{x}+c)}\right)$ can be introduced in the FL formulation to obtain a more appropriate model for