

Pier Luigi Dragotti, Ph.D.
Editor-in-Chief
IEEE Transactions on Signal Processing
IEEE Signal Processing Society
445 Hoes Lane
Piscataway, NJ 08854, USA
January 25, 2020

Dear Dr. Dragotti,

I am writing to submit our manuscript entitled, “High Dimensional Data Sharing: Multi-Task Learning with Theoretical Guarantee” for consideration as an IEEE Transactions on Signal Processing Regular Paper. We introduced an estimator for recovering signals from a system of coupled superposition models known as the linear “Data Sharing” model. Our results establish the sufficient geometric condition for the recovery of signals, and we provide the first complete sample complexity, estimation error bound, and computational complexity analysis for such an estimator.

From the statistical machine learning perspective, data sharing is a form of *multi-task learning* where the underlying parameter of each task is a summation of two parameters, one of which is task-specific and the other one is shared among all tasks. From the signal processing point of view, data sharing is a generalization of *the demixing* problem. The goal of demixing is to identify two structured signals, given only the sum of the two signals and prior information about their structures. Data sharing expands this framework by considering a coupled system of mixed signals with the premise of leveraging all of the samples to recover the shared signal more efficiently.

In the high dimensional linear regression setting, the Data Shared LASSO (DSL) has been suggested by Gross and Tibshirani recently and has found many applications while lacking proper theoretical analysis. DSL idea is to have a shared regression parameter across different tasks which gets supplemented by individual (sparse) parameters per task. Our estimator generalizes DSL by letting the parameters to have any structure promoted by arbitrary convex function.

Given the recent interest in leveraging data from heterogeneous sources in learning and signal recovery, we believe that the findings presented in our paper will appeal to the statistical signal processing community and specifically those who are interesting in the recovery of structured low dimensional signals.

This manuscript expands on the prior published work such as:

1. Gross SM, Tibshirani R. Data Shared Lasso: A Novel Tool to Discover Uplift. *Comput Stat Data Anal.* 2016;101: 226–235.
2. McCoy, M. B., & Tropp, J. A. (2014). Sharp recovery bounds for convex demixing, with applications. *Foundations of Computational Mathematics*, 14(3), 503-567.
3. Oymak, Samet, Benjamin Recht, and Mahdi Soltanolkotabi. Sharp time–data tradeoffs for linear inverse problems. *IEEE Transactions on Information Theory* 64.6 (2017): 4129-4158.

Each of the authors confirms that this manuscript has not been previously published and is not currently under consideration by any other journal. Additionally, all of the authors have approved the contents of this paper and have agreed to the IEEE Transactions on Signal Processing’s submission policies. Moreover, to the best of our knowledge, the named authors have no conflict of interest, financial or otherwise.

Sincerely,

Amir Asiaee
Corresponding Author
Postdoctoral Fellow
The Ohio State University
Mathematical Biosciences Institute
1735 Neil Avenue Columbus, OH 43210
<https://www.amirasiaee.com/>
asiaeetaheri.1@osu.edu