

A Distributed Learning Architecture for Scientific Imaging Problems

Athanasia Panousopoulou^{a,*}, Samuel Farrens^b, Konstantina Fotiadou^{a,c}, Arnaud Woiselle^d, Grigoris Tsagkatakis^{a,c}, Jean-Luc Starck^b, Panagiotis Tsakalides^{a,c}

^a*Institute of Computer Science, Foundation of Research and Technology - Hellas, Greece*

^b*Laboratoire AIM, UMR CEA-CNRS-Paris 7, Irfu, Service d'Astrophysique, CEA Saclay, France*

^c*Department of Computer Science, University of Crete, Greece*

^d*Safran Electronics & Defense, France*

Abstract

Current trends in scientific imaging are challenged by the emerging need of integrating sophisticated machine learning with Big Data analytics platforms. This work proposes an in-memory distributed learning architecture for enabling sophisticated learning and optimization techniques on scientific imaging problems, which are characterized by the combination of variant information from different origins. We apply the resulting, Spark-compliant, architecture on two emerging use cases from the scientific imaging domain, namely: (a) the space variant deconvolution of galaxy imaging surveys (astrophysics), (b) the super-resolution based on coupled dictionary training (remote sensing). We conduct evaluation studies considering relevant datasets, and the results report at least 60% improvement in time response against the conventional computing solutions. Ultimately, the offered discussion provides useful practical insights on the impact of key Spark tuning parameters on the speedup achieved, and the memory/disk footprint.

Keywords: Distributed Computing, Apache Spark, Distributed Learning, Dictionary Learning, Point Spread Function Deconvolution.

1. Introduction

The last decade has been earmarked by the significant technological advances on both expensive and cost-effective instrumentation, which pervasively collects, processes, and communicates massive streams of information. The resulting deluge of manifold data has provided new pathways for ground-breaking scientific discoveries in various fields, ranging from neuroscience and system

*Corresponding author

Email address: apanouso@ics.forth.gr (Athanasia Panousopoulou)

biology, to medicine and astrophysics, while challenging at the same time the computer engineering communities to enable the paradigm shift for accurate trends prediction over large-scale scientific datasets [1, 2].

The necessity of empowering distributed learning and inference over the petascales of scientific data is considered a game changer for distributed computing [3, 4] and respective management platforms, originally and vastly employed for retail and social networking services. Considering specifically the scientific imaging domain, the respective large-scale datasets (e.g., the Sloan Digital Sky Survey [5] and the upcoming Euclid space mission by the European Space Agency [6]) involve a rather small community of users, while confronting at the same time the challenge of sufficiently manipulating, and analyzing a significantly larger amount of information than the one considered in the social or Interner-based media [7, 8]. Interestingly, such datasets possess significant “V” properties of Big Data beyond their voluminous character; their production rate can readily reach the magnitudes of exabytes/day (velocity), while reflecting on different points of origin (variety), and often demanding robust algorithms for noisy, incomplete or inconsistent data (veracity). Ultimately, while *scientific imaging big data* can reflect on complex relationships (e.g., [9]), they are expensive to create, difficult to maintain, and laborious to infer contextual information.

Performing analytics over scientific imaging big data corresponds to a problem far more complex than upgrading hardware infrastructures to meet the latest technological trends in hardware acceleration, or exchanging conventional machine learning techniques with their sophisticated deep learning counterparts; The computational complexity will continue to increase along with the scale of the input problem. Instead, in parallel to the synthesis of new models for coping with the manifold characteristics of imaging signals, the challenge is how the current knowledge in optimization and machine learning research can be optimally exploited for building efficient imaging data processing in large-scale settings. Specifically, the role of computational approaches that adopt a “black-box” approach should be revisited from the perspective of current and emerging trends in big data analytics platforms, in order to optimally reuse sophisticated techniques that have been designed for small-scale problems over scientific imaging big data, and ultimately enable a rapid pace of innovation for scientific discovery.

Notably, big data technology has been employed for addressing key processing steps in voluminous imaging problems (e.g., SciSpark [10], Kira [11]). Even so, the migration from an academic implementation to a scalable solution over a distributed cluster of computers remains challenging;

the correct execution require careful control and mastery of low-level details of the distributed environment [12]. At the same time, although many common learning algorithms are supported for big data, there is still both the practical need and the research interest to expose less-widely or new machine learning algorithms to the scientific imaging domain [13]. In this work we address this gap by designing and developing an in-memory distributed learning architecture for applying sophisticated learning and optimization techniques on scientific imaging datasets. Specifically, we extend the work presented in [14], which introduced a scheme compliant to the Apache Spark computing framework [15] for solving the problem of removing distortion from noisy galaxy images towards three directions, namely: (a) we formulate the overall distributed learning framework for scientific imaging problems with dedicated emphasis on addressing both the volume and variety of the input data, (b) we explore how the resulting architecture can be utilized for the efficient parallelization of the learning problem at hand, considering two use cases from the scientific imaging domain associated with astrophysics and remote sensing, (c) we evaluate the proposed architecture using realistic datasets for each use case and provide useful technical insights on the impact of key experimental parameters on the speedup achieved, and the memory/disk footprint. The results highlight the benefits of the proposed architecture in terms of speedup and scalability for both use cases, while offering practical guidelines for enabling analytics over scientific imaging big data. While employing commodity hardware as the cornerstone of the distributed environment, we achieve more than 60% improvement in convergence rate terms with respect to conventional computing solutions, for both application scenarios.

The remainder of the paper is organized as follows: in Section 2, the recent trends on distributed computing tools for Sci-BD are outlined. The proposed distributed learning architecture is described in Section 3, while its instantiation for each use case along with the accompanying the evaluation studies are provided in Section 4. Conclusions are drawn in Section 5.

2. Background and Related Work

2.1. Big Data platforms for large-scale learning

Large-scale learning problems are typically empowered by distributed computing architectures, with the objective to disseminate the computational burden to a set of distributed resources. The resulting clusters of networked computing resources subsequently yield the underlying infrastructure, on top of which massive streams of raw data must be processed and transformed to meaningful

information. As such, processing models and accompanying programming abstractions are essential for implementing the application logic and facilitating the data analytics chain.

Three main categories of process models can be identified, namely: (a) generic (e.g., MapReduce [16], Dryad [3]) suitable for batch processing, (b) graph (e.g., GraphLab [17], GraphX [18]) accordingly expressing the relationship between data and computing tasks, and (c) streaming (e.g., S4 [19], Storm [20]) treating data as events and applying actor programming models. These programming models are instantiated on dedicated platforms for big data analytics. Depending on whether intermediate data are stored on the disk or in the memory, these platforms can support off-line, non-iterative, or on-line and iterative applications respectively. Hadoop and its variations [21, 22] is considered the mainstream MapReduce implementation and enables large-scale non-iterative computations (e.g., sorting) by the means of replicating data on disks of multiple hosts. Despite its popularity, the main drawback of Hadoop is its response time; all intermediate data are stored in the disk, thereby causing significant latency during processing [3].

An alternative to Hadoop, which allows in-memory analytics [23] over commodity hardware is the Spark framework [15]. Spark extends the MapReduce model by the use of an elastic persistence model, which provides the flexibility to persist these data records, either in memory, on disk, or both in memory and on disk. As such, Spark favors applications that need to read a batch of data multiple times, such as iterative processes met in machine learning and optimization algorithms.

Spark adopts a structured design philosophy, which allows, among others, its extension towards Spark MLLib, an open-source machine learning library [24, 25] suitable for both unstructured and semi-structured data. The key characteristic of Spark MLLib is that it can assemble into a single pipeline a sequence of algorithms, while supporting automatic parameter tuning for all algorithms within the defined pipeline. Despite implementing more than fifty-five common algorithms for model training, MLLib lacks of the essential data models for processing multi-dimensional arrays, such as the ones typically employed for expressing scientific and imaging data. To address this limitation, SparkArray [26] offers an extension of the Apache Spark framework that provides both a multi-dimensional array data model as well as a set of accompanying array operations on top of the Spark computing framework. Along similar lines, Misra et al. [27] identified that the primary bottleneck on matrix inversions over Spark and MLLib are the necessary matrix multiplications and, thus, proposed SPIN, a Spark-compliant distributed scheme for fast and scalable matrix inversion.

2.2. The positioning of Apache Spark in the distributed learning arena

Overall, the in-memory philosophy of Spark, along with the advances on offering a repository for machine learning techniques and useful matrix operations, have substantially empowered its positioning in the distributed learning arena. A key issue recently addressed by the computing community [28, 29, 30] relates to its performance against dedicated distributed learning platforms for scientific and imaging large-scale data. For instance, Petuum [28], is a representative counterpart of Spark, and a specialized distributed machine learning framework relying on the principles of parameter server and state synchronous parallelism [31] for enabling high-performance learning. Compared to Spark, Petuum takes advantage of the iterative-convergence properties of machine learning programs for improving both the convergence rate and per-iteration time for learning algorithms. This results to improved speed-up for large-scale machine learning models with many parameters, at the expense of a less convenient model for programming and code deployment than the one offered by Spark. Considering disk-based distributed platforms for array data analytics, ArrayBench [29] is a structured benchmark environment, over which a detailed analysis on the performance of Apache Spark against SciDB [32] is performed. The thorough analysis on different aspects of time response and scalability over voluminous workflows representing gene and biological networks highlight both the superiority of the persistence models of offered by Spark for data-intensive analytics over SciDB, as well as its the importance of suitably tuning the configuration parameters for achieving optimal performance with respect to the memory and disk usage. Ultimately, Caino-Lores et al. present in [30] the reproduction of an iterative scientific simulator for hydrological data, originally implemented based on the principles of message passing interface (MPI), into Apache Spark and study the performance against both private and public cloud infrastructures. The insights therein offered highlight both the benefits of Spark in terms of easing data parallelization and underlying task management against the MPI-based approach, while revealing at the same time the impact of the configuration parameters in the final performance with respect to memory and stability.

In parallel, current bibliography trends highlight the adoption of Spark for both key processing steps in large-scale imaging problems [10, 33, 11, 34, 35], as well as for parallelizing dedicated machine learning and optimization algorithms [36, 37, 38, 39]. Specifically, with regard to imaging data management over Spark, SciSpark [10, 33] pre-processes structured scientific data in network Common Format (netCDF) and Hierarchical Data Format (HDF). The result is a distributed com-

puting array structure suitable for supporting iterative scientific algorithms for multidimensional data, with applications on Earth Observation and climate data for weather event detection. Another representative framework is Kira [11, 34], which leverages on Apache Spark for speeding-up the source extraction process in astronomical imaging, while outperforming high performance computing approaches for near real-time data analysis over astronomical pipelines. Finally, Peloton et al. synthesize a native Spark connector to manage arbitrarily large astronomical datasets in Flexible Image Transport System (FITS) formatin [35]. The analysis therein provided indicates the capability of the proposed connector to automatically handle computation distribution and data decoding, thereby allowing the users to focus on the data analysis.

With regard to employing Spark for the parallelization of dedicated machine learning and optimization algorithms, authors in [36] and [37] provide new solutions for in-memory supervised learning, elaborating on exact k-nearest neighbors classification and Bayesian Network Classifiers [40], respectively. In both cases, the efficacy in terms of time to completion and scalability is highlighted over annotated datasets with pre-defined features. Shifting towards imaging optimization problems, the authors in [38] study the distributed asteroid detection as a complete framework over Spark and cloud computing that considers both the pre-processing of raw data, as well as parallelization of dedicated learning and optimization algorithms for asteroids detection. The resulting system provides the ability of both incrementally updating the new data from continuously observations, as well as the visual means for inspecting the position-linkage for the discovered asteroids. Finally, by employing a similar philosophy, Makkie et al. discuss in [39] how the sparse characteristics of a dictionary learning algorithm for functional network decomposition can be implemented over Spark, for satisfying the desired scalability and reproducibility requirements of neuroimaging big data analysis.

The discussion thus far highlights the potential of in-memory computing frameworks, in general, and Spark, in particular, to enable the paradigm shift for large-scale analytics over scientific imaging. Even so, in its vast majority the related bibliography neither explicitly addresses how the multivariate characteristics of imaging datasets can be profiled in the distributed frameworks, nor extensively reports the role of Spark tuning parameters in the performance of distributed architectures. With respect to the current state of art, our contributions are the following:

- We elaborate on the system perspective for empowering distributed learning over large-scale and multivariate imaging datasets, and provide the respective distributed architecture based

on Spark;

- We explore how the proposed architecture can be instantiated for two emerging imaging problems and respective state-of-art theoretical solutions, namely (a) the space variant deconvolution of noisy galaxy images (astrophysics), and (b) the super-resolution using dictionary learning (remote sensing);
- We evaluate the resulting implementations using commodity hardware and realistic datasets, highlighting the relationship between the size of the input problem and computational capacity of the distributed infrastructure;
- We study the role of Spark tuning parameters (i.e., number of partitions and persistence model) on the memory and time performance, with detailed discussion in terms of speedup, scalability, memory/disk usage and convergence behavior;
- Finally, we offer practical guidelines on the tuning of the distributed learning architecture for imaging problems, as derived from the evaluation procedure.

3. Distributed Learning for Scientific Imaging Problems

3.1. *Spark preliminaries*

Spark organizes the underlying infrastructure into a hierarchical cluster of computing elements, comprised of a master and a set of M workers (Fig. 1(a)). The master is responsible for implementing and launching the centralized services associated to the configuration and operation of the cluster, while the workers undertake the execution of the computing tasks across a large-scale dataset. The application program, which define the learning problem and the respective datasets, are submitted to the cluster through the *driver* program, which is essentially the central coordinator of the cluster for executing the specific computations. The driver is responsible for partitioning the dataset into smaller partitions which are disseminated to the workers, and sequentially sending tasks to the workers. The workers perform the learning task on their assigned data chunks and report back to the driver with the status of the task and the result of the computation.

The partitioning of the dataset relies on the Resilient Distributed Datasets (RDD) [41], which are defined as read-only collections of N data records that is created through deterministic operations (e.g., map, group with respect to a key), on either the initial dataset, or another RDD. The

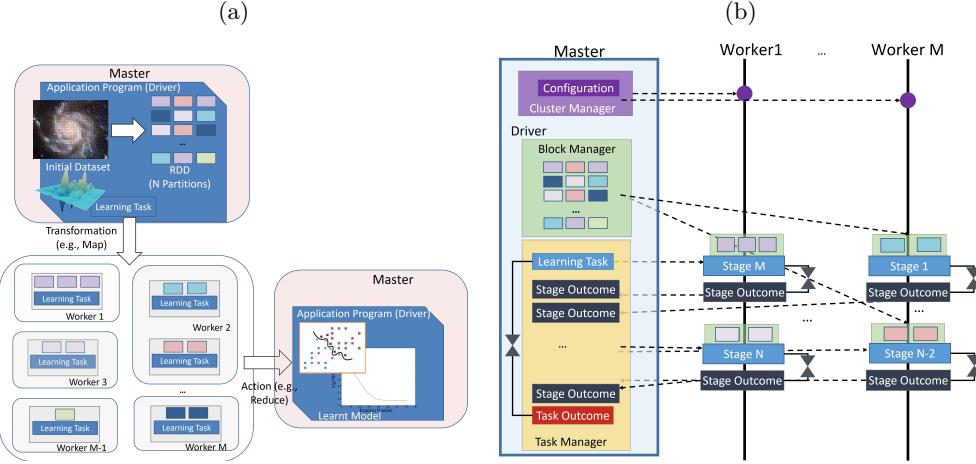


Figure 1: Overview of the Spark: (a) the master-slave architecture for deploying and executing a distributed learning task, (b) the interaction between the master and the workers during the execution of a learning task.

resulting data blocks are parceled into the M workers of the cluster. Each RDD is characterized by its lineage, which essentially conveys in a directed acyclic graph (DAG) enough information on how it was constructed from other RDD. As such, the lineage of an RDD can be utilized for reconstructing missing or lost partitions, without having to checkpoint any data. Through the driver, the application program can control how the initial data will be parallelized into one or multiple RDD, apply transformations on existing RDD. These transformations are lazy; RDD are only computed when they are used in *actions*, which are operations that return a value to the application program or export data to a storage system. As such, a set of pipelined transformations of an RDD will not be executed until an action is commanded.

The execution of an application program comprised of a set of sequential computing tasks is performed in three distinct phases (Fig. 1(b)) namely: (a) the configuration of the operational parameters; (b) the parallelization of the datasets and subsequent records into RDD; (c) the assignment and execution of the learning tasks. During the second phase, the Block Manager at the driver caters the workers with missing data blocks needed for computations throughout the lifetime of the application program. The third phases fires with the request of performing a learning task in the form of an action on the RDD. The Task Manager service calculates the DAG of the RDD lineage and accordingly assigns the execution of the learning task to the workers, in the form N stages. The result of the stage returns back to the driver program, and the Task Manager assigns

another stage of the same learning task, until all stages have been completed. This procedure is repeated for the remaining learning tasks. Notably, if the successive learning tasks are not independent from each other, e.g., part of an iterative optimization process, the lineage of the RDD will gradually increase, with a direct impact on the memory requirements.

3.2. Proposed architecture

The flexibility of Spark is partially grounded on the adopted data-parallel model, which is implemented as a pipeline of RDD through the appropriate combination of transformation and actions. Even so, a single transformation / action can handle neither multiple RDD at the same time, nor RDD nested within each other. This characteristic limits the applicability of the Spark framework on learning schemes over imaging data, which commonly rely on the combination of variant information from different origins for removing noisy artifacts and extracting essential information. Considering such problems, one can think of heterogeneous imagery that correspond to the same spatial information (e.g., patches of noisy and reference images) to be jointly processed for solving single/multi-objective optimization problems. As such, a substantial volume of bundled imaging data should become readily available in iterative processes for enabling large-scale imaging analytics.

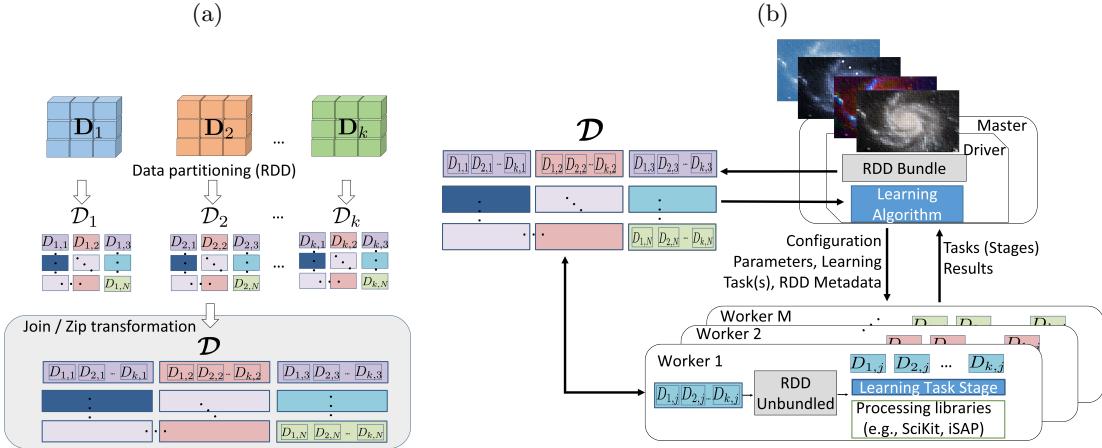


Figure 2: (a) The creation of the bundled RDD \mathcal{D} over Spark, (b) the architecture of the proposed scheme.

The herein proposed architecture considers the native RDD abstraction offered by Spark for performing distributed learning over bundled imaging datasets. Without loss of generality we consider

that a set of k imaging datasets and auxiliary structures (e.g., optimization variables, reference images), are modeled as multidimensional arrays $\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_k$. For each $\mathbf{D}_i, i \in \{1, 2, \dots, k\}$ the driver program creates the respective RDD \mathcal{D}_i , by essentially defining the partitioning of \mathbf{D}_i into N data blocks $D_{i,j}$, where $j \in \{1, 2, \dots, N\}$, i.e., $\mathcal{D}_i = [D_{i,1}; D_{i,2}; \dots, D_{i,N}]$. Applying a join- or zip-type of transformation across all individual \mathcal{D}_i results to the creation of the bundled RDD \mathcal{D} , which combines the parallelized set of the k input datasets (Fig. 2(a)):

$$\mathcal{D} = [\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_k] = \begin{bmatrix} D_{1,1} & D_{2,1} & \dots & D_{k,1} \\ D_{1,2} & D_{2,2} & \dots & D_{k,2} \\ \vdots & \vdots & \ddots & \vdots \\ D_{1,N} & D_{2,N} & \dots & D_{k,N} \end{bmatrix}$$

For the implementation of the proposed scheme over Spark both the learning problem and the k respective datasets, are submitted to the master of the cluster through the driver program. The initial imagery datasets can be either locally available on the master or stored in a Spark-compliant distributed file system (e.g., HDFS). The combination of the initial datasets and their subsequent parallelization into \mathcal{D} are undertaken by the RDD Bundle Component located at the side of the driver (Fig. 2(b)). Any transformation defining parts of the learning algorithm that should consider the combination of the k imaging datasets can in turn be utilized with the resulting bundled \mathcal{D} at the driver. When the respective action is fired, both the learning task and \mathcal{D} are parceled into the M workers of the cluster. Prior the execution of each learning task on the workers, each partition j of \mathcal{D} is separated into the original inputs $D_{1,j}, D_{2,j}, \dots, D_{k,j}$ for performing the task at hand. This is performed by the RDD Unbundle Component at the side of each worker. The calculation of the result of each task from the cluster is consistent to the architecture presented in Fig. 1(b); each learning task is split into N sequential computing stages, which are performed on different sets of k data blocks by different workers. The result of each stage returns back to the driver program, which assigns another stage of the same learning task, until all stages have been completed. In case the learning task is part of an iterative calculation, any updated part of the i -th partitioned blocks $D_{i,1}, D_{i,2}, \dots, D_{i,N}$ is bundled back into \mathcal{D} for the next iteration. Once all learning tasks are completed, the result is either returned to the driver program or stored in the distributed file system, using the build-in actions of Spark.

The proposed scheme essentially provides a Spark-native approach for applying pipelined RDD transformations over a set of different imagining datasets that should be jointly processed. Much like

any other Spark-based architecture, the driver program defines the application-specific tasks, along with the input datasets for parallelization and bundling, and their partitioning scheme, by defining for instance the value of N . In addition to the application-defined tasks, dedicated processing libraries (e.g., Astropy [42], iSAP [43], SciKit-Learn [44]) can be deployed at each worker for solving the learning problem at hand. Notably, the use of the RDD Bundle and RDD Unbundle components helps retaining the core principles of the original learning algorithm intact, thereby facilitating the re-usability of existing and novel approaches in machine learning, originally designed and evaluated for small-scale scenarios. Even so, as detailed in Section 4, the problem and data characteristics have a key role in instantiating the proposed scheme for designing and developing in-memory distributed sophisticated learning techniques for imaging big data.

4. Use Cases

The herein proposed distributed learning platform has been employed for addressing the large-scale challenges of two application scenarios in imaging and respective datasets, namely: (a) the space variant deconvolution of galaxy survey images, and (b) the joint dictionary training for image super-resolution over low- and high-resolution data, entailing video streaming in either different bands (hyperspectral) or standardized RGB coloring. The respective software libraries and datasets are publicly available at [45].

4.1. Astrophysics: Space variant deconvolution of galaxy survey images

Even in ideal conditions, astronomical images contain aberrations introduced by the Point Spread Function (PSF) of the telescope. Therefore, obtaining accurate and unbiased properties of extended sources, such as galaxies, from these images is predicated on the ability to remove or compensate for the effects of the PSF.

The PSF describes the response of an imaging system to point sources. The PSF of an astronomical instrument can either be modelled, assuming complete knowledge of the imaging system, or measured from distant stars in the field, which can be approximated as point sources. However, removing the PSF, a process referred to as deconvolution, is a non-trivial problem given the presence of random noise in the observed images that prevents the use of analytical solutions. This problem is even more complicated for upcoming space surveys, such as the one corresponding to the Euclid mission, for which the PSF will additionally vary across the field of view [46].

Currently, very few efficient methods exist for dealing with a space variant PSF. Nevertheless, an elegant solution is the Object-Oriented Deconvolution [47], which relies on the assumption that galaxies belong to a low-rank dimensional space, given the similarities between various images. The low-rank approximation can then be used as a method of regularization for an inverse problem of the form:

$$\mathbf{Y} = \mathcal{H}(\mathbf{X}) + \mathbf{N}, \quad (1)$$

where $\mathbf{Y} = [\mathbf{y}^0, \mathbf{y}^1, \dots, \mathbf{y}^n]$ is a stack of observed noisy galaxy images, $\mathbf{X} = [\mathbf{x}^0, \mathbf{x}^1, \dots, \mathbf{x}^n]$ is a stack of the true galaxy images, $\mathbf{N} = [\mathbf{n}^0, \mathbf{n}^1, \dots, \mathbf{n}^n]$ is the noise corresponding to each image and $\mathcal{H}(\mathbf{X}) = [\mathbf{H}^0\mathbf{x}^0, \mathbf{H}^1\mathbf{x}^1, \dots, \mathbf{H}^n\mathbf{x}^n]$ is an operator that represents the convolution of each galaxy image with the corresponding PSF for its position.

In order to solve a problem of this type one typically attempts to minimize some convex function such as the least squares minimization problem:

$$\operatorname{argmin}_{\mathbf{X}} \frac{1}{2} \|\mathbf{Y} - \mathcal{H}(\mathbf{X})\|_2^2, \quad (2)$$

which aims to find the solution $\hat{\mathbf{X}}$ that gives the lowest possible residual $(\mathbf{Y} - \mathcal{H}(\hat{\mathbf{X}}))$.

This problem is ill-posed as even the tiniest amount of noise will have a large impact on the result of the operation. Therefore, to obtain a stable and unique solution to Eq. (1), it is necessary to regularize the problem by adding additional prior knowledge of the true images. Farrens et al. proposed in [48] two alternatives for addressing the regularization issues, namely: (a) a sparsity approximation, (b) a low-rank approximation. Briefly, the sparsity approximation imposes that the desired solution should be sparse when transformed by a given dictionary, which in the case of galaxy images, this dictionary corresponds to an isotropic wavelet transformation [49]. In this case, the solution to Eq. (2) becomes:

$$C(\mathbf{X}) = \operatorname{argmin}_{\mathbf{X}} \frac{1}{2} \|\mathbf{Y} - \mathcal{H}(\mathbf{X})\|_2^2 + \|\mathbf{W}^{(k)} \odot \Phi(\mathbf{X})\|_1 \quad \text{s.t. } \mathbf{X} \geq 0 \quad (3)$$

where: $\|\bullet\|_2^2$ denotes the Frobenius norm; the Φ operator realizes the isotropic undecimated wavelet transform without the coarse scale; $\mathbf{W}^{(k)}$ is a weighting matrix related to the standard deviation of the noise in the input images; \odot denotes the Hadamard (entry-wise) product; k is a reweighting index, necessary to compensate for the bias introduced by using the $l1$ -norm.

The low-rank approximation, on the other hand, simply assumes that with a sufficiently large sample of galaxy images some properties of these objects should be similar and therefore a matrix containing all of the image pixels will not be full rank [48]. In this case, the solution to Eq. (2) becomes:

$$C(\mathbf{X}) = \underset{\mathbf{X}}{\operatorname{argmin}} \quad \frac{1}{2} \|\mathbf{Y} - \mathcal{H}(\mathbf{X})\|_2^2 + \lambda \|\mathbf{X}\|_* \quad \text{s.t.} \quad \mathbf{X} \geq 0 \quad (4)$$

where λ is a regularization control parameter and $\|\bullet\|_*$ denotes the nuclear norm.

As explained in [48], both of these techniques can be used to improve the output of the optimisation problem and hence the quality of the image deconvolution, which is considered a key factor for the success of the Euclid mission.

4.1.1. Parallelization using the distributed learning architecture

A primal-dual splitting technique [50] can be adopted for solving both Eq. (3) and Eq. (4), taking into account the inherent sparse or low-rank properties, respectively. The sequential approach for implementing it reflects on an iterative optimization process. During each iteration step all requested input, i.e., noisy data (\mathbf{Y}), PSF data, primal (\mathbf{X}_p) and dual (\mathbf{X}_d) optimization variables, and weighting matrix (\mathbf{W}) -for the case of the sparsity-prior regularization- are jointly fed to the solver in order to calculate the value $C(\mathbf{X}_p)$ of the cost function described by either Eq. (3) or Eq. (4), as illustrated in Fig. 3.

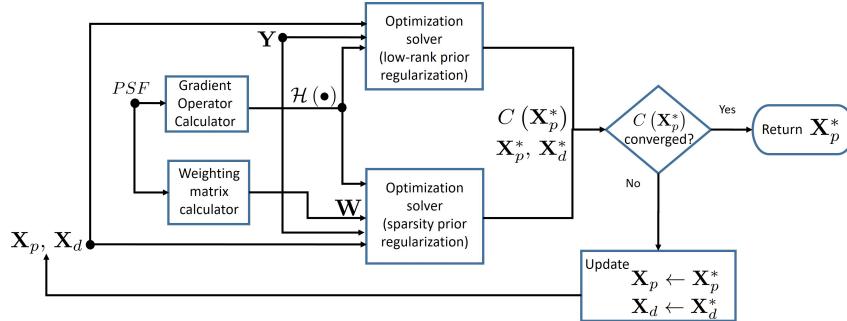


Figure 3: The flow diagram for the sequential implementation of the space variant deconvolution of noisy galaxy images.

This approach would be inefficient when the size of the noisy input data and respective space-variant PSF increases. In order to consider the necessary interaction between the different inputs for

solving this optimization problem, we herein propose Algorithm 1 for distributing the optimization phase according to the proposed learning architecture.

Algorithm 1: The PSF algorithm parallelization (steps are performed on the driver, while steps are performed over the cluster).

Data: The data \mathbf{Y} , the respective PSF, the maximum number of iterations i_{\max} , and the cost tolerance ϵ .

Typically: $i_{\max} = 300$, $\epsilon = 10^{-4}$

Result: The estimated images \mathbf{X}_p^* that minimize Eq. (3) or Eq. (4).

```

1 Initialize  $\mathbf{X}_p$ ,  $\mathbf{X}_d$  and extract the operator  $\mathcal{H}(\bullet)$ .
2 Define the RDD for  $\mathbf{Y}$ , PSF,  $\mathbf{X}_p$ ,  $\mathbf{X}_d$  as  $\mathcal{D}_{\mathbf{Y}}$ ,  $\mathcal{D}_{PSF}$ ,  $\mathcal{D}_{\mathbf{X}_p}$ ,  $\mathcal{D}_{\mathbf{X}_d}$ , respectively, with  $N$  partitions per RDD.
3 if Sparsity regularization: then
    | Apply the weighting matrix calculator to  $\mathcal{D}_{PSF}$ :  $\mathcal{D}_{\mathbf{W}} = \mathcal{D}_{PSF}.map(\lambda x : \mathbf{W}^{(k)}(x))$ .
4    | Create the RDD bundle:  $\mathcal{D} = \mathcal{D}_{\mathbf{Y}}.zip(\mathcal{D}_{PSF}).zip(\mathcal{D}_{\mathbf{W}}).zip(\mathcal{D}_{\mathbf{X}_p}).zip(\mathcal{D}_{\mathbf{X}_d})$ .
5 else if Low-rank regularization: then
    | Create the RDD bundle:  $\mathcal{D} = \mathcal{D}_{\mathbf{Y}}.zip(\mathcal{D}_{PSF}).zip(\mathcal{D}_{\mathbf{X}_p}).zip(\mathcal{D}_{\mathbf{X}_d})$ .
end
6 for  $i : 1 : i_{\max}$  do
7    | Update  $\mathcal{D}_{\mathbf{X}_p}$ ,  $\mathcal{D}_{\mathbf{X}_d}$  in  $\mathcal{D}$  using [50]:  $\mathcal{D} = \mathcal{D}.map(\lambda x : \text{Update } x)$ .
8    | Update  $C(\mathbf{X}_p^*)$  (Eq. (3) or Eq. (4), depending on sparsity or low-rank regularization):
9    |  $C(\mathbf{X}_p^*) = \mathcal{D}.map(\lambda x : \text{Calculate } C(x)).reduce(\lambda x, y : x + y)$ .
10   | if  $C(\mathbf{X}_p^*) \leq \epsilon$  then
11      |   | break
12   | end
13 end
14 Save  $\mathcal{D}$  to disk or distributed file system and return  $\mathbf{X}_p^*$  to driver.

```

Algorithm 1 entails the parallelization of \mathbf{Y} , PSF data, \mathbf{X}_p , \mathbf{X}_d into $\mathcal{D}_{\mathbf{Y}}$, \mathcal{D}_{PSF} , $\mathcal{D}_{\mathbf{X}_p}$, $\mathcal{D}_{\mathbf{X}_d}$ respectively on the side of the driver program. When the sparsity-prior regularization solution is adopted, the inherited dependency of the weighting matrix \mathbf{W} on the PSF data, leads to the transformation of \mathcal{D}_{PSF} to the corresponding weighting data blocks $\mathcal{D}_{\mathbf{W}}$. All requested input is in turn compressed into \mathcal{D} , which essentially contains tuples of the form $\langle \mathbf{Y}, PSF, \mathbf{W}, \mathbf{X}_p, \mathbf{X}_d \rangle$ (i.e., $k = 5$, sparsity solution) or $\langle \mathbf{Y}, PSF, \mathbf{X}_p, \mathbf{X}_d \rangle$ (i.e., $k = 4$, low-rank solution). The resulting RDD \mathcal{D} is used to calculate the updated value of the optimization variable \mathbf{X}_p^* based on the sparsity prior (Eq. (3)) or low-rank (Eq. (4)) regularization on each worker. The firing of a *reduce* action, triggers

the value of the cost function $C(\mathbf{X}_p^*)$. This process that relies on the interaction between the driver and the workers (map to $\mathbf{D} \rightarrow$ reduce to $C(\mathbf{X}_p^*)$) is repeated until either the value of $C(\mathbf{X}_p^*)$ converges to ϵ , or until the maximum number of iterations is reached. The resulting stack of stack of the true galaxy images \mathbf{X}_p^* is directly saved on the disk of the driver program.

4.1.2. Evaluation Studies

The evaluation studies emphasize on the performance of the PSF algorithm parallelization implemented over the herein proposed learning architecture. The data used for this work consists of two samples of 10,000 and 20,000 simulated 41×41 pixel postage stamps, respectively. Each stamp contains a single galaxy image obtained from Great3 challenge [51]. These images were convolved with a Euclid-like spatially varying and anisotropic PSF and various levels of Gaussian noise were added to produce a simplified approximation of Euclid observations. In total there are 600 unique PSFs [52], which are down-sampled by a factor of 6 to avoid aliasing issues when convolving with the galaxy images [53].

The distributed learning architecture features Spark 2.1.0 [54], deployed over a cluster of $M = 5$ workers. The driver allocates 8GB RAM, while 4 out of 5 workers allocate 2.8GB RAM and 4 CPU cores. The fifth worker allocates 2.8GB RAM and 8 CPU cores, thereby yielding in total 24 CPU cores and 14GB RAM. With regard to the RDD persistence model, for all experiments conducted we considered the default storage level (memory-only), and as such, RDDs are persistent in memory. In case this exceeds the memory availability, some partitions will not be cached and will be recomputed on the fly each time they're needed.

The evaluation procedure emphasizes on the time performance with respect to the sequential implementation¹ in terms of speedup, execution time per optimization loop, and scalability; the memory usage; the convergence behavior of the cost function. The key experimental parameters are the solution approach (sparsity or low-rank), and the data size (10,000 or 20,000 stack of galaxy images), with respect to the number of partitions N partitions per RDD. We herein consider $N = \{2x, 3x, 4x, 6x\}$, where x corresponds to the total number of cores available for parallel calculations, i.e. $x = 24$. Notably, considering other parallel computing frameworks for comparison (e.g., Hadoop) is beyond the scope of this work, as they are either unsuitable for the essential iterative computations

¹<https://github.com/sfarrens/psf>

typically met in learning imaging problems, or focus on the extraction of astronomical imaging data.

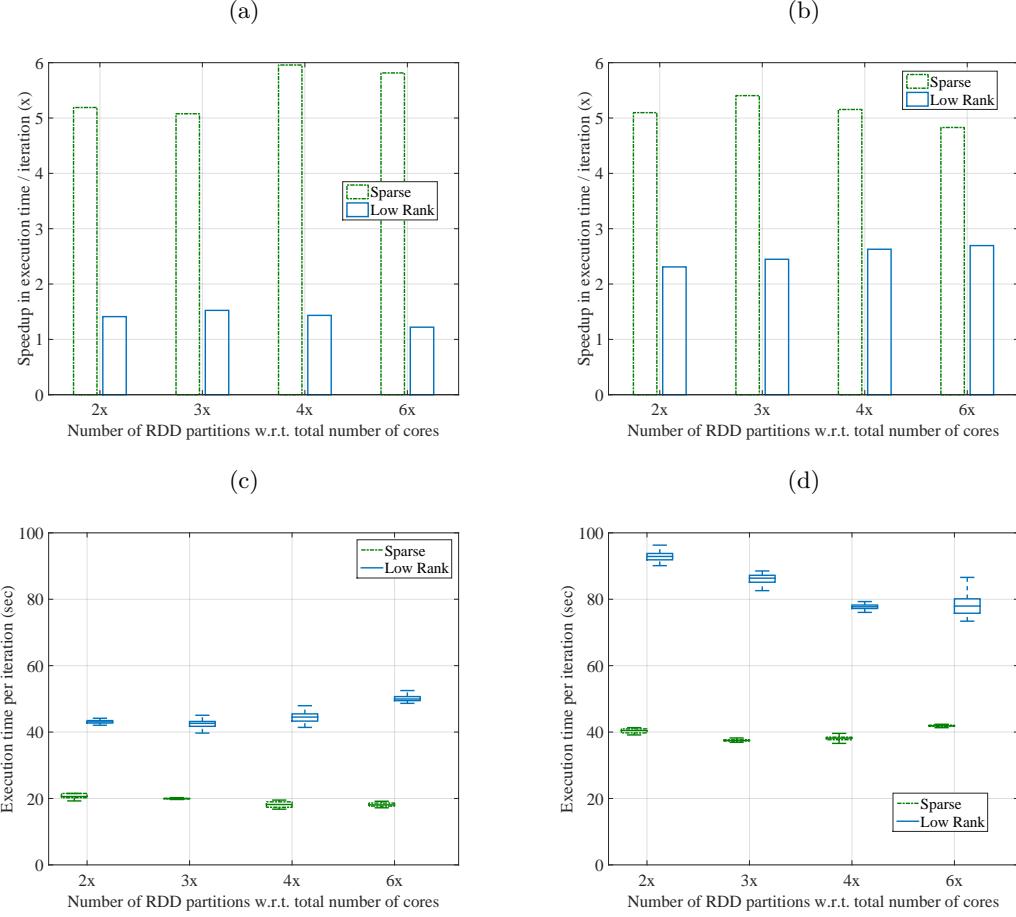


Figure 4: Time performance for the parallelization of the PSF deconvolution based on the proposed architecture: (a) speedup for the stack of 10,000 images, (b) speedup for the stack of 20,000 images, (c) time execution per optimization loop for the stack of 10,000 images, (d) time execution per optimization loop for the stack of 20,000 images.

Time performance. Figure 4.1.2 presents the time performance achieved for both sparse- and low rank-based solutions in terms of speedup (Fig. 4.1.2(a) for the stack of 10,000 images, and Fig. 4.1.2(b) for the stack of 20,000 images), and execution time per iteration of the optimization process for the calculation of $C(\mathbf{X}_p^*)$ (Fig. 4.1.2(c) for the stack of 10,000 images, and Fig. 4.1.2(d) for the stack of 20,000 images). The initial observation to make is that the herein proposed technique

offers increased speed up. Greater improvement is achieved for the case of the sparsity-based solution ($\text{speedup} \geq 5x$) than for the low-rank based solution ($\text{speedup} \in [1.2, 2.5]$) for all cases of N examined and for both sizes of datasets considered (Fig. 4.1.2(a) and (b)). This is due to the nature of the problem and the adopted solution; the sparsity-based solution has inherent parallelizable properties since the algorithm can work independently on different portions of the dataset. By contrast, the low-rank based solution performs SVD over the entire stack of images, and as such the parallelized data need to be reassembled at the driver program for the calculation of $\mathbf{X}_p^* - \mathbf{X}_d^*$. Nevertheless, as the number of noisy images increases from 10,000 to 20,000, the speedup of the low-rank solution improves to more than $2x$ for all cases of N considered. This is consistent to nature of the Spark framework, and highlights that the scale of an input imaging dataset is relevant to the capacity of the cluster, and the demands of the solving approach; the overhead introduced by Spark for the distributed calculations may hinder the overall performance, when the dataset is relatively small compared to the computational capacity of the cluster.

With respect to the time execution per optimization loop (Fig. 4.1.2(c) and Fig. 4.1.2(d)) we observe that the distributed architecture yields a stable performance, since for all combinations of solution approach, data size, and level of N , the time execution has limited statistical dispersion between the 25-th and 75-th percentiles. The impact of the parameter N on the time performance is more evident for the low-rank solution, exhibiting a contradictory pattern; for the case of 10,000 images (Fig. 4.1.2(c)) as the number N of partitions per RDD increases ($2x \rightarrow 6x$), the median execution time per iteration is increased by approximately 8 secs ($\sim 42\text{sec} \rightarrow 50\text{sec}$). This is consistent to the speedup results, and suggest that the increase of N implies more data chunks of smaller size need to be exchanged among the workers, thereby introducing unnecessary shuffling overhead due to the SVD computations. By contrast, when 20,000 images are considered (Fig. 4.1.2(c)) the increase of N between $2x \rightarrow 6x$ results to a drop of the median execution time per iteration by approximately 10 secs ($\sim 98\text{sec} \rightarrow 88\text{sec}$). In this case, the partitioning into more data chunks implies less burden on memory per task, which substantially compensates any shuffling overhead.

Ultimately, scalability aspects on deploying Algorithm 1 over the cluster infrastructure are illustrated in Fig 5 for the stack of 20,000 images and $N = 4x$, for both sparse and low-rank approaches. As expected, as the total number of cores increases with respect to the cores considered for the evaluation of the sequential approach (i.e., 4 cores), the distributed learning approach offers substantial improvements in terms of time performance. Specifically, increasing the number of

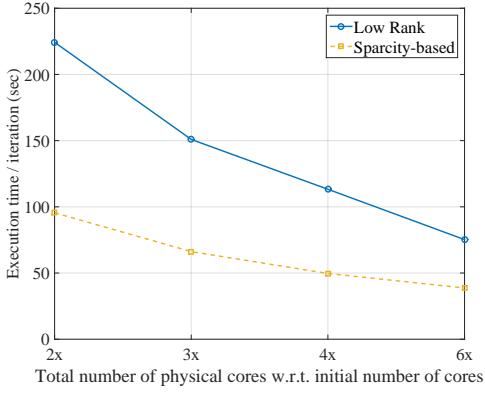


Figure 5: Scalability of execution time per iteration for the stack of 20,000 images. available cores in the cluster from $2x$ to $6x$ results into a 50% and 65% improvement for the sparse and the low-rank approach, respectively.

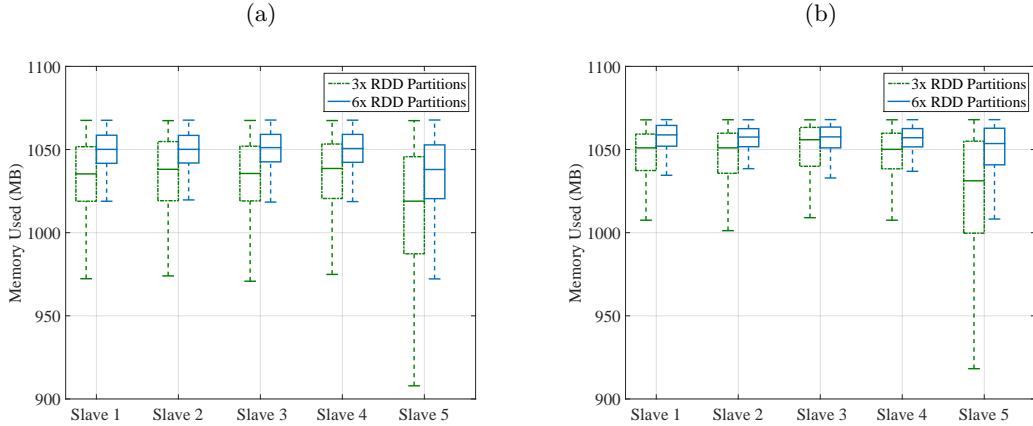


Figure 6: Memory usage per slave when the stack of 20,000 images is considered for (a) sparsity based solution , (b) low-rank based solution.

Memory Usage. Figure 6 presents the memory usage per worker throughout the experiment duration considering the stack of 20,000 images and $N = \{3x, 6x\}$ for the sparsity (Fig. 6(a)), or the low rank-based solution (Fig. 6(b)). For both solution approaches the use of memory remains consistently at $\sim 1.068\text{GB}$, which is the maximum amount of memory allocated per worker for the

computations of the assigned tasks². This is aligned to the iterative nature of the optimization problem at hand; during each optimization step the lineage of \mathcal{D} is updated with the new values of $\mathcal{D}_{\mathbf{x}_p}, \mathcal{D}_{\mathbf{x}_d}$. As a result, the length of the respective DAG and the respective number of computations undertaken by each worker increase over time. Even so, due to the adopted persistence model, the memory usage does not exceed the maximum allowable limit, and intermediate variables not fitting in memory will be recomputed each time they're needed.

With regard to the deviations on memory usage across the different slaves we observe that four out of five slaves, allocating 4 cores on the cluster, have a similar memory usage in terms of statistical dispersion between the 25-th and 75-th percentiles, which does not exceed the value of 35MB for both approaches of optimization. Interestingly, Slave 5, which allocates 8 cores on the cluster, exhibits a greater dispersion, reaching up to ~60MB for the case of sparse-based optimization and $N = 3x$. This deviated behavior is associated with the fact that the master assigns on this slave more jobs per tasks than the remaining ones. As such, with this configuration, in order for Slave 5 to free memory space for handling more jobs, both the persistence model performs more on-the-fly computations, as well as the garbage collector cleans more frequently any unnecessary variable, thereby resulting into greater deviations (e.g. drops) on the memory usage during the execution of the experiment.

Finally, interesting remarks are derived with regard to the impact of the number N of partitions per RDD on the memory usage. For both solution approaches the deviation on the memory usage is greater for $N = 3x$ than the one recorder for $N = 6x$. Indicatively, considering the sparse solution on Slave 5 the dispersion of the memory usage equals to 58.4MB for $N = 3x$, opposed to 32.3MB (Fig. 6(a)), which is observed for the case of $N = 6x$. This is due to the fact that a smaller number of partitions $N = 3x$ results in fewer data blocks with relatively large size and increased demands in memory. This in turn stimulates the persistence model to perform more on-the-fly calculations. On the other hand, as the number of partitions increases $N = 6x$, the size of the data blocks decreases, and subsequently, the persistence model become more relaxed. However, for the case of the sparse approach, more stages are needed to complete an action, which results into a slight increase in the execution time per loop (Fig. (b)). Even so, the benefit in this case is a more reliable behavior on

²According to the unified memory management of Spark 2.1.0, the remaining amount of memory is reserved for cached blocks immune to being evicted by execution.

each worker in terms of memory usage, while additionally highlighting the trade off between the RAM availability and the size of the input problem; When the memory per worker is limited with respect to size of the input data it is considered preferable to increase the number of partitions of the RDD. It may increase the time response, due to the increased number of tasks, however it will yield a more reliable solution within the lifetime of the program execution.

Convergence Behavior. Figure 7 illustrates the convergence behavior of the value of $C(\mathbf{X}_p^*)$ versus the time elapsed when $i_{\max} = 150$, and either the sequential or PSF parallelization (Algorithm 1, $N = 3x$) sparsity-based approach is adopted for 20,000 images. When the proposed architecture is adopted the cost function starts converging within the first hour of experiment, opposed to the sequential approach which can only complete a few iterations within the same time period (Fig. 7(a)). Overall, the distributed learning approach is 75% faster than the conventional one; the completion time of the standalone approach equals to ~ 8 hours, opposed to the parallelized version, which does not exceed 2 hours (Fig. 7(b)). These results highlight the fact that despite the memory overhead for storing intermediate results on each worker, the herein proposed solution is extremely beneficial in terms of time response for enabling large-scale PSF deconvolution over noisy galaxy images.

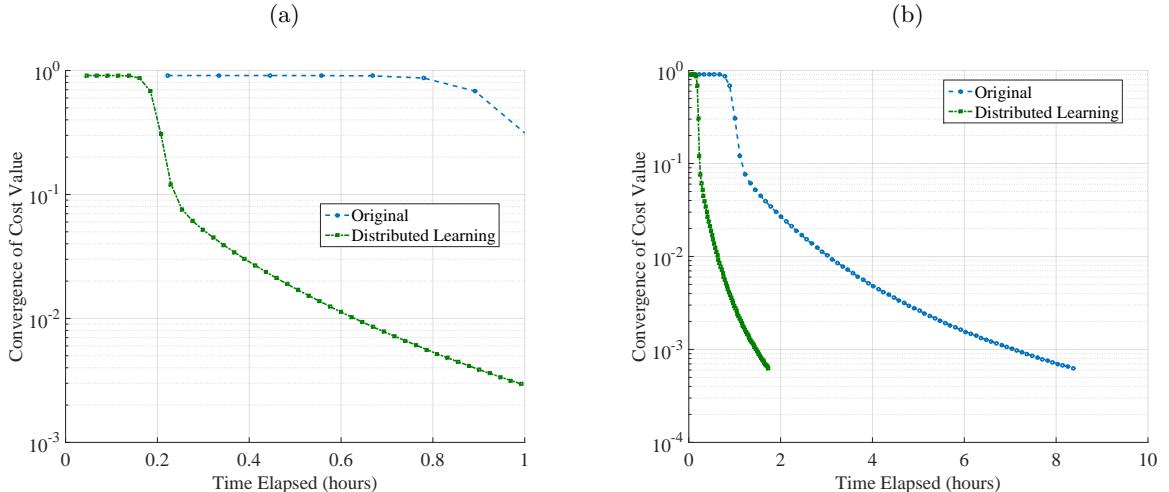


Figure 7: The convergence behavior $C(\mathbf{X}_p^*)$ for the sparsity-based solution, with respect to time considering the case of 20,000 images (a) 1st hour of experiment, (b) total experiment.

4.2. Remote Sensing: Super-resolution using Sparse Coupled Dictionary Training

Current state of art in imaging systems can provide us with images from different spectral bands, giving information on the reflectivity, color, temperature or other physical properties depending on the wavelength. These images have an initial resolution ranging from a few hundred pixels (e.g., thermal infrared) to a few thousands (e.g., visible light). Even so, the highest of these resolutions is not always sufficient to distinguish small objects in a large field of view, such as in the case of an airborne carrier which looks at a distance of a few kilometers.

While single image super-resolution techniques have been developed to enhance numerically the resolution of the sensors, their main drawback is that they do not take into account structural information of the data, such as sharp borders of objects. As such, learning jointly the structure of low and high resolution images in joint dictionary learning schemes can overcome this problem and result in sharper and more visually pleasing results.

A solution to this challenge is centred around the fusion of low and high resolution training examples, in an innovative scheme [55, 56] which introduces a Coupled Dictionary Learning process. Briefly, the Sparse Coupled Dictionary Learning (SCDL) algorithm formulates the super-resolution problem within the highly efficient Alternating Direction Method of Multipliers (ADMM) optimization framework [57]. As described in [55, 56], this approach synthesizes a high resolution hypercube from its low resolution acquired version, by exploiting the Sparse Representations theory [58]. Traditional approaches consider a set of K low and high resolution image pairs and assume that these images are generated by the same statistical process under different resolution. As such, they share the same sparse coding, with respect to their corresponding low $\mathbf{X}_l \in \mathbb{R}^{P \times A}$, and high $\mathbf{X}_h \in \mathbb{R}^{M \times A}$ resolution dictionaries, where A is the number of atoms in each dictionary.

The coupled dictionary learning technique relies on the ADMM formulation, to yield an unconstrained version of the dictionary learning problem, which can be efficiently solved via alternating minimization. Formally, we consider the observation signals, $\mathbf{S}_l = \{\mathbf{s}_l\}_{i=1}^M$, and $\mathbf{S}_h = \{\mathbf{s}_h\}_{i=1}^P$. The main task of coupled dictionary learning is to recover both the dictionaries \mathbf{X}_h and \mathbf{X}_l with their corresponding sparse codes \mathbf{W}_h and \mathbf{W}_l , by the means of the following ℓ_1 -minimization problem:

$$\begin{aligned} & \min_{\mathbf{X}_h, \mathbf{W}_h, \mathbf{X}_l, \mathbf{W}_l} \|\mathbf{S}_h - \mathbf{X}_h \mathbf{W}_h\|_F^2 + \|\mathbf{S}_l - \mathbf{D}_l \mathbf{W}_l\|_F^2 + \lambda_l \|\mathbf{Q}\|_1 + \lambda_h \|\mathbf{P}\|_1 \\ & \text{s.t. } \mathbf{P} - \mathbf{W}_h = 0, \mathbf{Q} - \mathbf{W}_l = 0, \mathbf{W}_h - \mathbf{W}_l = 0, \|\mathbf{D}_h(:, i)\|_2 \leq 1, \|\mathbf{D}_l(:, i)\|_2 \leq 1, \end{aligned} \quad (5)$$

where λ_h , λ_l are the sparsity balance terms. The ADMM scheme considers the separate structure of each variable in Eq.(5), relying on the minimization of its augmented Lagrangian function:

$$\begin{aligned} L(\mathbf{X}_h, \mathbf{X}_l, \mathbf{W}_h, \mathbf{W}_l, \mathbf{P}, \mathbf{Q}, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3) = & \|\mathbf{X}_h \mathbf{W}_h - \mathbf{S}_h\|_F^2 + \|\mathbf{X}_l \mathbf{W}_l - \mathbf{S}_l\|_F^2 + \lambda_h \|\mathbf{P}\|_1 \\ & + \lambda_\ell \|\mathbf{Q}\|_1 + \langle Y_1, \mathbf{P} - \mathbf{W}_h \rangle + \langle Y_2, \mathbf{Q} - \mathbf{W}_l \rangle + \langle Y_3, \mathbf{W}_h - \mathbf{W}_l \rangle + \frac{c_1}{2} \|\mathbf{P} - \mathbf{W}_h\|_F^2 \\ & + \frac{c_2}{2} \|\mathbf{Q} - \mathbf{W}_l\|_F^2 + \frac{c_3}{2} \|\mathbf{W}_h - \mathbf{W}_l\|_F^2, \end{aligned} \quad (6)$$

where \mathbf{Y}_1 , \mathbf{Y}_2 , and \mathbf{Y}_3 stand for the Lagrange multiplier matrices, while $c_1 > 0$, $c_2 > 0$, and $c_3 > 0$ are the respective step size parameters. In each optimization step, the updated dictionary matrices are the following (Algorithm 1 in [56]):

$$\mathbf{X}_h = \mathbf{X}_h + \frac{\mathbf{S}_h \times \mathbf{W}_h}{\phi_h + \delta} \quad (7)$$

$$\mathbf{X}_l = \mathbf{X}_l + \frac{\mathbf{S}_l \times \mathbf{W}_l}{\phi_l + \delta}, \quad (8)$$

where $\phi_h = \mathbf{W}_h \times \mathbf{W}_h^T$, $\phi_l = \mathbf{W}_l \times \mathbf{W}_l^T$, and δ is the regularization factor.

As explained in [55, 56] Sparse Representations and Coupled Dictionary Learning are powerful tools for reconstructing spectral profiles from their corresponding low-resolution, and noisy versions, with tolerance in extreme noise scenarios.

4.2.1. Parallelization using the distributed learning architecture

The general algorithmic strategy of ADMM scheme for calculating \mathbf{X}_l and \mathbf{X}_h seeks a stationary point by solving for one of the variables, while keeping the others fixed. The sequential approach for solving Eq. 7 reflects on an iterative process. As elaborated in [55, 56], during each iteration step all variables involved (i.e., \mathbf{S}_h , \mathbf{S}_l , \mathbf{W}_h , \mathbf{W}_l , \mathbf{P} , \mathbf{Q} , $\mathbf{Y}_{1,2,3}$) need to be jointly processed in matrix operations for updating the values of \mathbf{X}_h and \mathbf{X}_l (Fig 8).

Both the flow diagram presented in Fig. 8 as well as the calculation steps explained in [56] highlight the importance of the intermediate sparse coding and Laplacian multiplier matrices for calculating the dictionaries. Notably these intermediate matrices have size $K \times A$ and as such, the sequential approach could readily become inefficient as the number of data samples K and the combination of P , M and A increases.

The parallelization of this scheme is described in Algorithm 2 for distributing the essential matrix calculations, according to the proposed distributed learning architecture. In compliance to

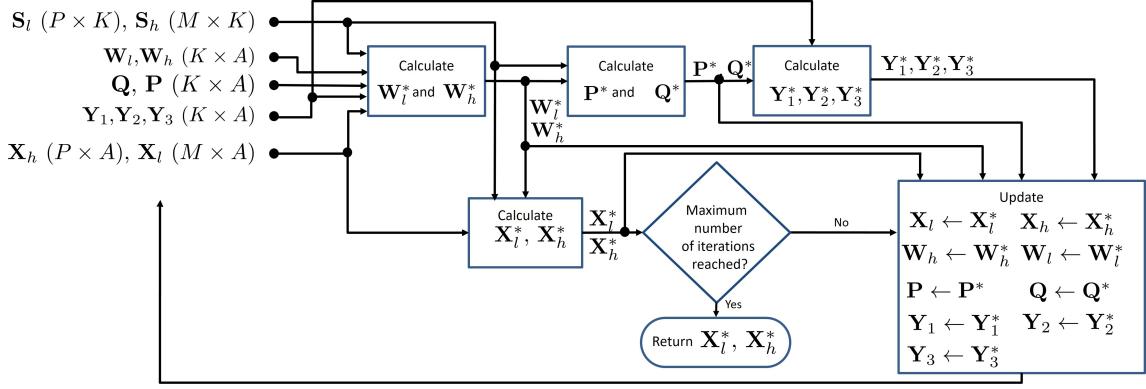


Figure 8: The flow diagram for the sequential implementation of the sparse coupled dictionary learning algorithm.

parallel efforts (e.g., [39]) we assume that dictionaries fit into the memory of a single computer. The first step considers the parallelization of the input imaging 3D cubes $\mathbf{S}_h, \mathbf{S}_l$ into $\mathcal{D}_{\mathbf{S}_h}, \mathcal{D}_{\mathbf{S}_l}$, on the side of the driver program and their compression into the \mathcal{D} , which essentially contains tuples of the form $\langle \mathbf{S}_h, \mathbf{S}_l \rangle$ (i.e., $k=2$). The initial values of the dictionaries $\mathbf{X}_h, \mathbf{X}_l$ contain random samples of $\mathcal{D}_{\mathbf{S}_h}, \mathcal{D}_{\mathbf{S}_l}$ collected over the cluster at the driver program. The RDD bundle is \mathcal{D} is transformed into a SCSDL object \mathcal{D}_{SCSDL} which enriches \mathcal{D} with the parallelized version of all auxiliary matrices i.e. $\mathcal{D}_{\mathbf{W}_h}, \mathcal{D}_{\mathbf{X}_l}, \mathcal{D}_{\mathbf{P}}, \mathcal{D}_{\mathbf{Q}}, \mathcal{D}_{\mathbf{Y}_1}, \mathcal{D}_{\mathbf{Y}_2}, \mathcal{D}_{\mathbf{Y}_3}$. During each iteration, the driver broadcasts the current dictionaries $\mathbf{X}_h, \mathbf{X}_l$ to the cluster, which are in turn employed for updating the auxiliary matrices contained in \mathcal{D}_{SCSDL} in a distributed fashion, according to Algorithm 1 in [56]. By exploiting the properties of outer products calculations, the firing of the respective combination of map-reduce action over \mathcal{D}_{SCSDL} (Step 9 in Algorithm 2) triggers the calculation of auxiliary structures $\mathbf{S}_h \times \mathbf{W}_h$, $\mathbf{S}_h \times \mathbf{W}_h$, and ϕ_h, ϕ_l , which are essential for updating the dictionary matrices $\mathbf{X}_h, \mathbf{X}_l$ in Eq.(7), Eq.(8), respectively. This process is repeated until the maximum number of iterations is reached, while the final dictionaries $\mathbf{X}_h^*, \mathbf{X}_l^*$ are directly saved on the memory of the driver program.

4.2.2. Evaluation Studies

Similarly to the case of the space variant deconvolution, the objective the benchmark studies is to evaluate the performance of the SCSDL algorithm parallelization implemented over the herein proposed learning architecture. Two types of datasets are herein considered, namely: (a) hyperspectral (HS) data, comprised of video streaming frames on different spectral bands, captured by

Algorithm 2: The SCDL algorithm parallelization (steps are performed on the driver, while steps are performed over the cluster).

Data: The high \mathbf{S}_h ($\in \mathbb{R}^{P \times K}$) and the low \mathbf{S}_l ($\in \mathbb{R}^{M \times K}$) 3D cubes, the number of dictionary atoms A , and the maximum number of iterations i_{\max} . Typically: $i_{\max} = 100$.

Result: The high \mathbf{X}_h^* ($\in \mathbb{R}^{P \times A}$) and low \mathbf{X}_l^* ($\in \mathbb{R}^{M \times A}$) resolution dictionary matrices.

- 1 Define the RDD for \mathbf{S}_l and \mathbf{S}_h as $\mathcal{D}_{\mathbf{X}_h}, \mathcal{D}_{\mathbf{X}_l}$ respectively, with N partitions per RDD.
- 2 Initialize \mathbf{X}_h and \mathbf{X}_l , considering $P \times A$ and $M \times K$ random samples from $\mathcal{D}_{\mathbf{X}_h}$ and $\mathcal{D}_{\mathbf{X}_l}$ respectively.
- 3 Return $\mathbf{X}_h, \mathbf{X}_l$ to driver.
- 4 Create the RDD bundle: $\mathcal{D} = \mathcal{D}_{\mathbf{X}_h}.\text{zip}(\mathcal{D}_{\mathbf{X}_l})$.
- 5 Map \mathcal{D} into $\mathcal{D}_{\mathbf{Z}}$ for $\mathbf{Z} = \{\mathbf{W}_h, \mathbf{W}_l, \mathbf{P}, \mathbf{Q}, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3\}$, by creating a RDD SCDL object \mathcal{D}_{SCDL} , containing N blocked parallel partitions for all intermediate matrices \mathbf{Z} , i.e., $\mathcal{D}_{SCDL} \triangleq [\mathcal{D}_{\mathbf{X}_h}, \mathcal{D}_{\mathbf{X}_l}, \mathcal{D}_{\mathbf{W}_h}, \mathcal{D}_{\mathbf{W}_l}, \mathcal{D}_{\mathbf{P}}, \mathcal{D}_{\mathbf{Q}}, \mathcal{D}_{\mathbf{Y}_1}, \mathcal{D}_{\mathbf{Y}_2}, \mathcal{D}_{\mathbf{Y}_3}]$; $\mathcal{D}_{SDCL} = \mathcal{D}.\text{map}(\lambda x : \text{Start SCDL}(x, A))$.
- 6 **for** $i : 1 : i_{\max}$ **do**
- 7 Broadcast \mathbf{X}_h and \mathbf{X}_l , and respective auxiliary transposed and inversed variables [56].
- 8 Update $\mathcal{D}_{\mathbf{Z}}$, for $\mathbf{Z} = \{\mathbf{W}_h, \mathbf{W}_l, \mathbf{P}, \mathbf{Q}, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3\}$ in \mathcal{D}_{SCDL} using Alg. 1, [56]:

$$\mathcal{D}_{CSDL} = \mathcal{D}_{SCDL}.\text{map}(\lambda x : \text{Update } x)$$
- 9 Calculate the auxiliary variables $\mathbf{X}_h \times \mathbf{W}_h, \mathbf{X}_l \times \mathbf{W}_l, \phi_h = \mathbf{W}_h \times \mathbf{W}_h^T, \phi_l = \mathbf{W}_l \times \mathbf{W}_l^T$ (Alg. 1, [56]):

$$[\mathbf{X}_h \times \mathbf{W}_h, \mathbf{X}_l \times \mathbf{W}_l, \phi_h, \phi_l] = \mathcal{D}_{SCDL}.\text{map}(\lambda x : \text{Calc. Outer Product}(x)).\text{reduce}(\lambda x, y : \text{Calc. Outer Product}(x, y))$$
- 10 Update the dictionary matrices $\mathbf{X}_h, \mathbf{X}_l$ according to Eq. (7)-(8).
- end
- 11 Save final dictionary matrices $\mathbf{X}_h^* \leftarrow \mathbf{X}_h, \mathbf{X}_l^* \leftarrow \mathbf{X}_l$.

snapshot mosaic sensors that feature IMEC Spectrally Resolvable Detector Array [59] ($P = 5 \times 5$, $M = 3 \times 3$, $K \simeq 40,000$), and (b) grayscale (GS) data ($P = 17 \times 17$, $M = 9 \times 9$, $K \simeq 40,000$), taken from an airborne carrier, looking at the ground at a few kilometers using a high definition color video sensor at 25 images per second.

The distributed learning architecture retains the same characteristics of the private cluster described in Section 4.1.2, i.e. Apache Spark 2.1.0. 2017 release over 5 slaves. In order to better accommodate the computational needs of this use case, we configure one of the slaves to generate 2 Spark workers, each allocating 2.8GB RAM and 4 CPU cores. As a result, the cluster yields in total 24 CPU cores and 16.8GB RAM. With regard to the data persistence model, we consider both the memory only, as well as the memory-and-disk model, according to which RDD data not fitting in memory are instead stored on disk and read from there when they are needed.

Similarly to the astrophysics use case (Section 4.1), the evaluation procedure emphasizes on the speedup with respect to the sequential implementation³, the memory and disk usage, and the convergence rate of the normalized root mean square error in both low and high resolution. The key experimental parameters are the type of input data (HS or GS) and respective parameters (i.e., M , P), and the dictionary size A with respect to the number of partitions N partitions per RDD. We herein consider $N = \{2x, 3x, 4x, 6x\}$, where x corresponds to the total number of cores available for parallel calculations, i.e. $x = 24$.

Speedup. Figure 9 presents the time performance achieved for both HS as well as GS imaging data in terms of speedup (Fig. 9(a), Fig. 9(b)) and execution time per iteration (Fig. 9(c), Fig. 9(d)) for the calculation of \mathbf{X}_h^* , and \mathbf{X}_l^* . Increased speed up is offered for both for the HS as well as the GS data, as the number of dictionary atoms (D) increases for all cases of N . Specifically, for the HS data (Fig. 9(a)) the speedup increases $\sim 2.5x \rightarrow 5x$ when D increases 1024 to 2056. The impact of the number of partitions N per RDD on the time performance is more evident for lower values of D ; for example for $D = 512$, the optimal speedup is achieved for $N = 2x$, and decreases as the number of RDD blocks increases. This is consistent to the PSF use case (10,000 images), and highlights that when the size of the tasks fits in the memory of each worker, it is more beneficial to retain a small value of N , in order to avoid the unnecessary Spark network and computational overheads. Interestingly, as the size D of the dictionaries increases speedup becomes less variant on the number of partitions; for $D = 2056$ the speedup achieved is $\sim 5x$ for all cases of N examined. Similar remarks can be derived for the GS data (Fig. 9(b)), wherein we additionally observe that due to increased size of the problem ($P = 17 \times 17$, $M = 9 \times 9$) the speedup values offered are smaller than the ones provided for the HS data. Nevertheless, when $D = 2056$, the proposed scheme yields speedup that remains higher than $2x$.

With respect to the time execution per optimization loop (Fig. 9(c) and Fig. 9(d)) the distributed architecture yields a stable performance, since for all combinations of data type, dictionary size, and value of N the time execution has limited statistical dispersion between the 25-th and 75-th percentiles. As expected, an increase to the size D of the dictionaries leads to an increase to the execution time, since the size of the problem significantly changes. The impact of the parameter N on the time performance is more evident for the case of the GS data and $D = 2056$, indicating that

³<https://github.com/spl-icsforth/SparseCoupledDictionaryLearning>

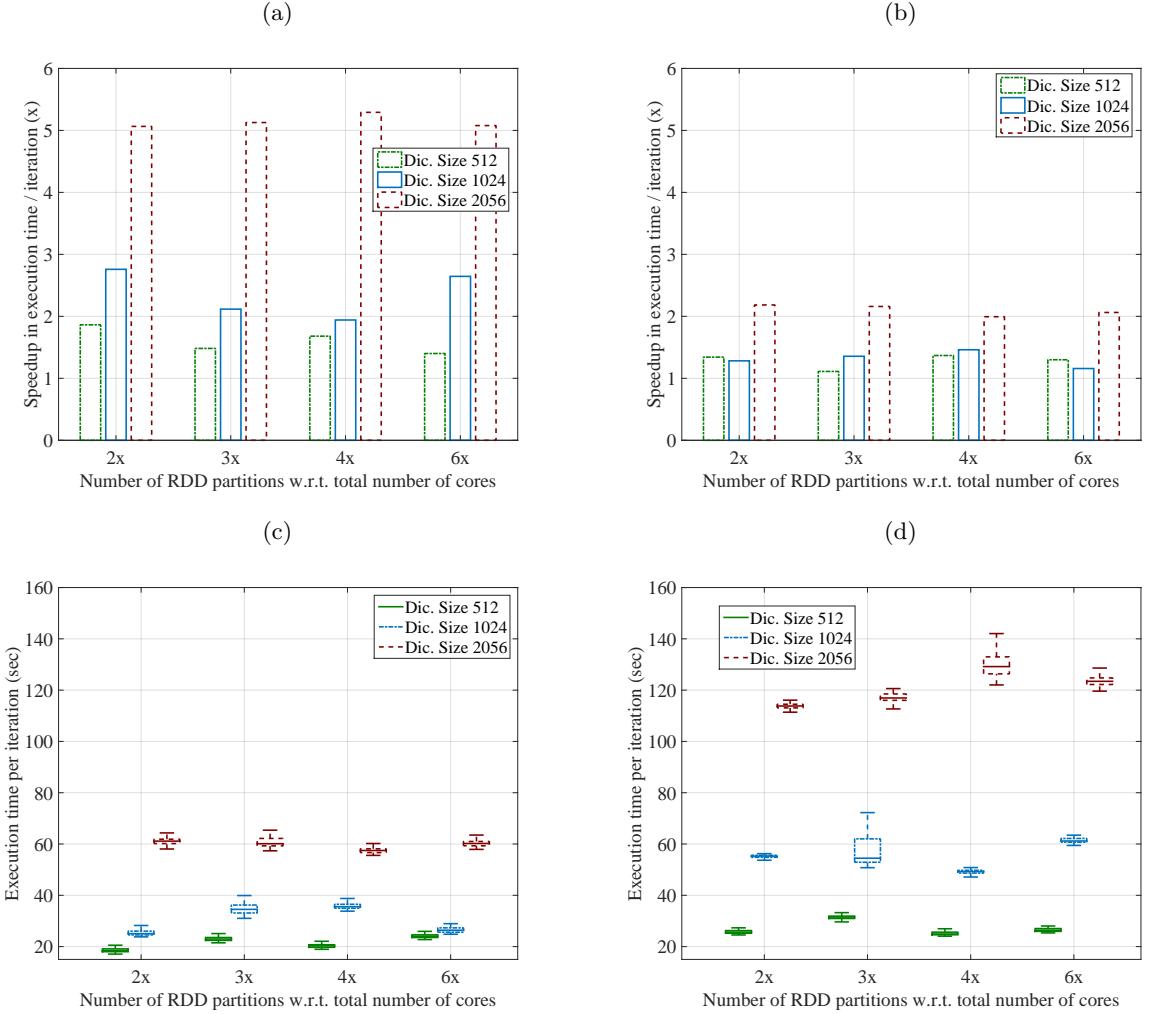


Figure 9: Time performance for the parallelization of the SCDL-based super resolution based on the proposed architecture: (a) speedup for the HS imaging data, (b) speedup for the GS imaging data, (c) time execution per optimization loop for the HS data, (d) time execution per optimization loop for the GS data.

for this problem, it is more beneficial to retain a smaller number of RDD partitions. Indicatively, when $N = 2x$ the median of execution time remains below 120secs, opposed to case of $N = 4x$ wherein the execution time per iteration reaches 130secs. This is consistent to the speedup results, and suggest that the increase of N implies more data chunks of smaller size need to be exchanged among the workers, thereby introducing unnecessary shuffling overhead. By contrast, when the value of N is retained lower, the partitioning into less data chunks compensates such phenomena.

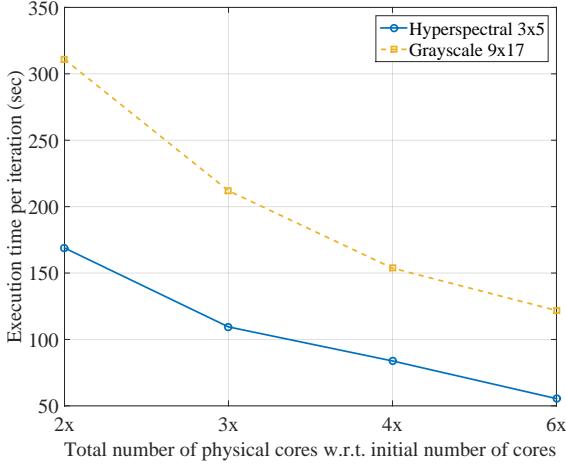


Figure 10: Scalability of execution time per iteration for the HS and GS imaging data.

Scalability aspects on deploying Algorithm 2 over the cluster infrastructure are illustrated in Fig 10 for the $D = 2056$ and $N = 4x$, for both HS and GS data. As expected, as the total number of cores increases with respect to the cores considered for the evaluation of the sequential approach (i.e., 4 cores), the distributed learning approach offers substantial improvements in terms of time performance. Specifically, increasing the number of available cores in the cluster from $2x$ (i.e., 8 cores) to $6x$ (i.e., 24 cores) results into a 61.3% ($310\text{sec} \rightarrow 120\text{sec}$) and 72.2% ($180\text{sec} \rightarrow 50\text{sec}$) improvement for the GS and HS data, respectively.

Memory & Disk Usage. Figure 11 and Figure 12 present the memory usage per worker throughout the experiment duration for the HS data and GS data respectively, for $N = \{3x, 6x\}$.

Considering the HS data (Fig. 11) the use of memory remains consistently at the maximum amount of memory allocated per worker for the computations of the assigned tasks. Similarly to the astrophysics use case (Section 4.1, Fig. 6), this result is aligned to the adopted persistence model (memory only) and the iterative nature of the optimization problem at hand, which entails subsequent increase of the respective number of computations over time. Nevertheless, opposed to the PSF use case, we observe similar memory usage in terms of statistical dispersion between the 25-th and 75-th percentiles, across all slaves involved. This is due to the current configuration of the cluster, which considers homogeneous resources allocation for all workers involved (i.e., 6 workers, each allocating 2.8GB RAM and 4 CPU cores). Finally, with regard to the impact of the number N of partitions per RDD on the memory usage, the deviation on the memory usage is greater for

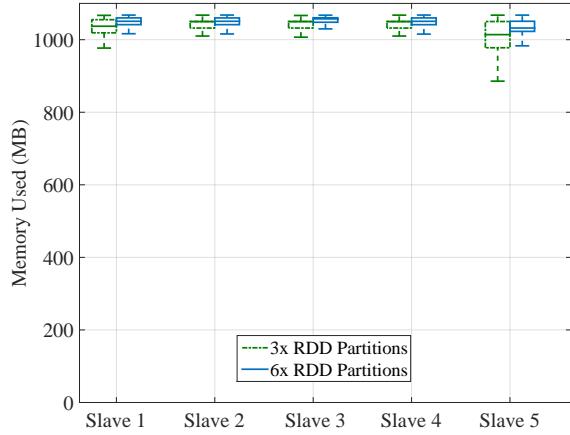


Figure 11: Memory usage per slave for the HS data

$N = 3x$ than the one recorder for $N = 6x$. Indicatively, considering Slave 5 the dispersion of the memory usage equals to $\sim 60\text{MB}$ for $N = 3x$, opposed to $\sim 30\text{MB}$ for $N = 6x$. Similarly to the PSF use case, the smaller number of partitions ($N = 3x$) results in fewer data blocks of greater size, thereby stimulating the persistence model to perform more on-the-fly calculations. By contrast, when $N = 6x$, more data blocks of smaller size are generated, and subsequently, the persistence model become more relaxed.

With regard to the case of the GS data (Fig. 12) interesting remarks can be derived depending on whether the memory-only or the memory-and-disk model is applied. Figures 12(a) and 12(b) present the memory and disk usage respectively for the memory-only model, while Fig. 12(c)-(d) illustrate the memory and disk usage for the memory-and-disk model. Similarly to the case of HS data and the astrophysics use case, the adoption of the memory-only model results in employing the maximum amount of memory per worker for the computations of the assigned tasks, while the smaller the value of N , the more stable the behaviour of each worker. Nevertheless, opposed to the case of HS data and the astrophysics use case, the increased size of the problem ($P = 17 \times 17$, $M = 9 \times 9$) result into using disk space for storing intermediate results not fitting into the memory (Fig. 12(b)). When the memory-and-disk model is applied we observe that the use of memory decreases to less than 500MB on all slaves, while the value of the number of partitions N has no essential impact on the statistical variation of the memory usage. Specifically, the dispersion between the 25-th and 75-th

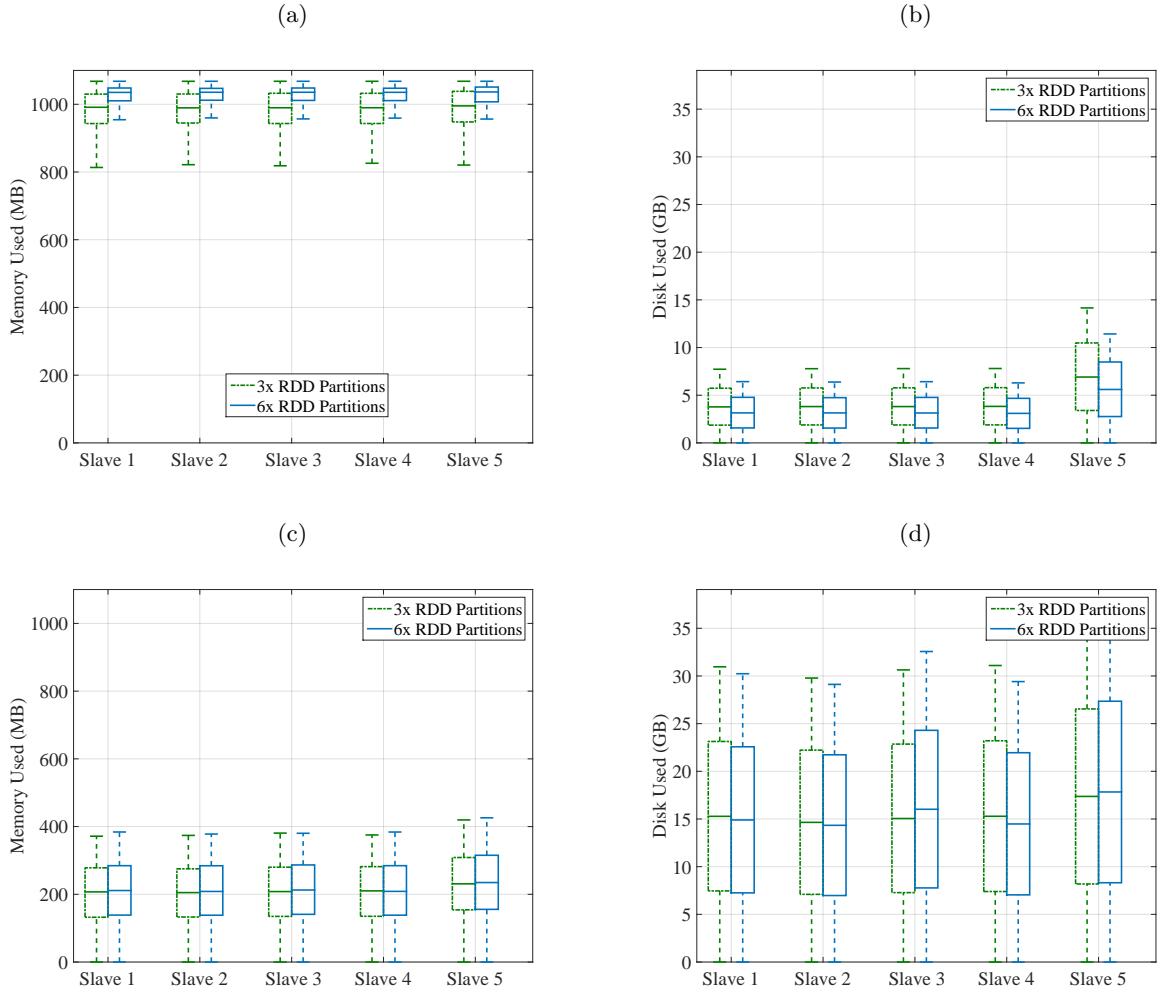


Figure 12: Memory and disk usage per slave for the GS data: (a) memory usage when the memory-only model is applied, (b) disk usage when the memory-only model is applied, (c) memory usage when the memory-and-disk model is applied, (d) disk usage when the memory-and-disk model is applied.

percentile remains at ~ 160 MB for all slaves and different values of N . The decreased use of memory is accompanied by a substantial increase in the disk usage, as illustrated in Fig. 12(c)-12(d); the median disk volume increases from ~ 5 GB to ~ 15 GB when the cluster configuration transits from the memory-only to the memory-and-disk model.

The increased disk usage is consistent to the persistence model and in order to investigate how the disk overhead affects the time performance, we present in Fig. 13 the time execution

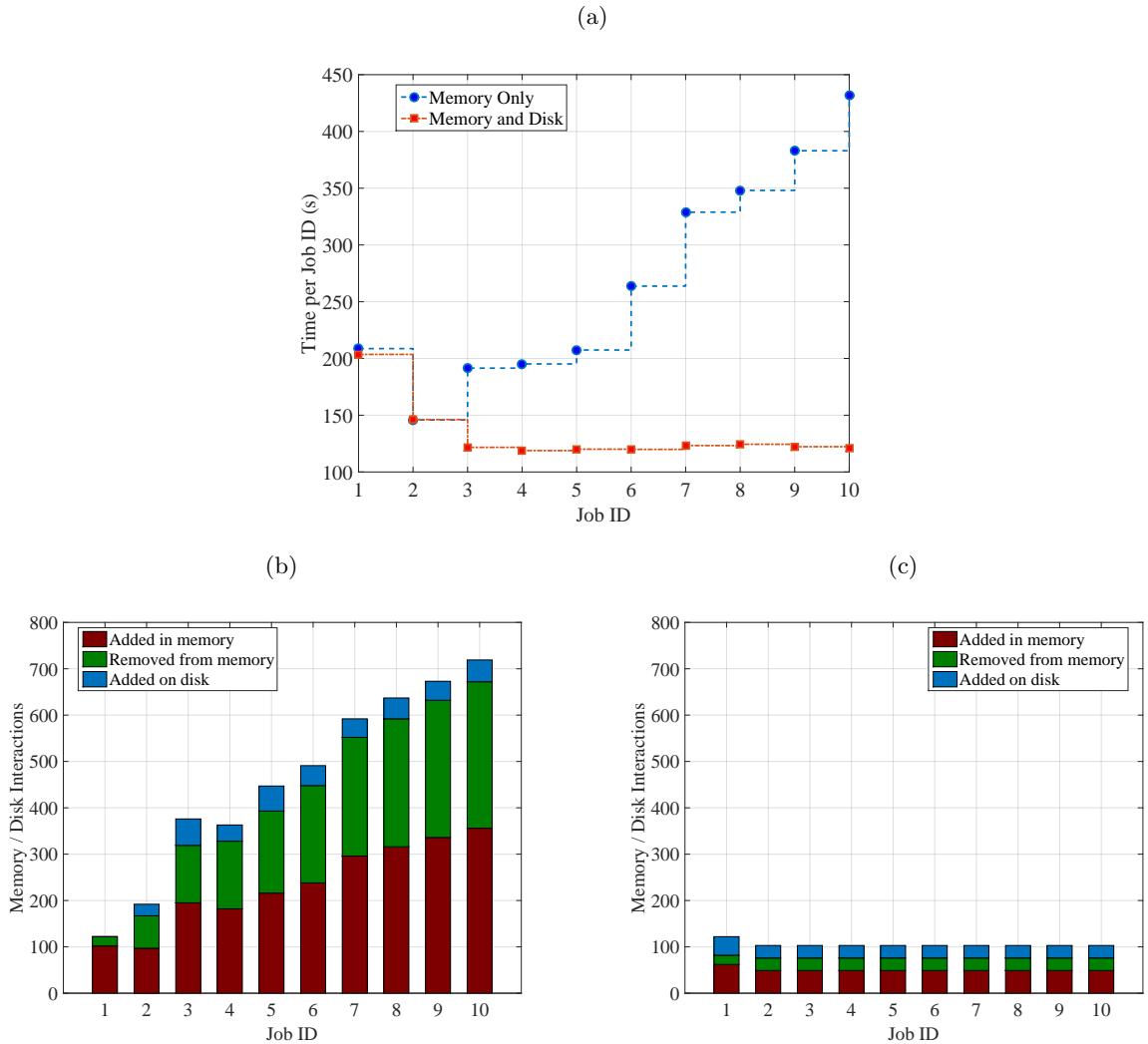


Figure 13: Time execution and memory-disk interactions for the GS data: (a) time execution on Slave 1 for 10 subsequent Spark jobs, (b) memory and disk interactions when the memory-only model is applied, (c) memory and disk interactions when the memory-and-disk model is applied.

(Fig. 13(a)), along with the memory and disk interactions when either the memory-only (Fig. 13(b)) or the memory-and-disk (Fig. 13(c)) model is applied, over ten subsequent Spark tasks. The results indicate that *the use of memory-only model is not advisable for this scenario*; the time for the execution of the subsequent iterations for calculating \mathbf{X}_h^* , and \mathbf{X}_l^* increases over time. Specifically, the time needed for completing 10 subsequent Spark jobs increases 200 sec → 430 sec when the memory-only model is applied. By contrast, when the memory-and-disk model is applied, the

execution of the same 10 subsequent jobs reduces 200 sec→120 sec. This behaviour relates to the needed memory interactions (add, remove, add on disk) that take place depending on the persistence model. As shown in Fig. 13(b), the memory-only model, which entails the on-demand recalculation of intermediate results, imposes an increasing trend of adding-removing results from memory, in order to free resources for the execution of subsequent iterations. On the other hand, when the memory-and-disk model is applied (Fig. 13(c)) the number of add-remove memory interactions remains consistent throughout the execution of the Spark jobs, since intermediate results are directly stored on the disk instead of recalculating them when they are needed.

Convergence Rate. Ultimately, Fig. 14 illustrates the convergence behavior of reconstruction error between the calculated dictionaries (\mathbf{X}_h^* , \mathbf{X}_l^*) and the augmented Lagrangian function versus the time elapsed, when either the sequential or the distributed SCDL approach (Algorithm 2, $N = 3x$) is adopted for the GS data, considering the memory-and-disk persistence model. For both high (Fig. 14(a)) and low (Fig. 14(c)) resolution dictionaries, the reconstruction error starts converging within the first 20 minutes of experiment when the proposed architecture is adopted. This is in sharp contrast to the sequential approach which can only complete two iterations within the same time period. Overall, the distributed learning approach is 65.7% faster than the conventional one; the completion time of the standalone approach equals to ∼3.5 hours, opposed to the parallelized version, which does not exceed 0.8 hours (Fig. 14(b), Fig. 14(d)). These results highlight that the herein proposed solution is extremely beneficial in terms of time response for enabling large-scale joint dictionary training schemes.

4.3. Discussion and Practical Guidelines

The evaluation studies indicate the effectiveness of the proposed architecture on jointly processing multiple imaging datasets. Even so, each application scenario requires a different approach in order to be treated as a distributed learning problem. Profiling both the problem and the data characteristics in order to compensate the performance bottlenecks is considered a critical aspect for providing an alternative, distributed solution that outperforms conventional approaches. Comprehensively, the key findings derived by this study can be summarized as follows:

- While being a generic purpose distributed computing framework, the core module of Spark herein adopted yields extremely promising results for solving image processing problems at scale. As highlighted in Fig. 5 and Fig. 10 for the astrophysics and remote sensing use cases

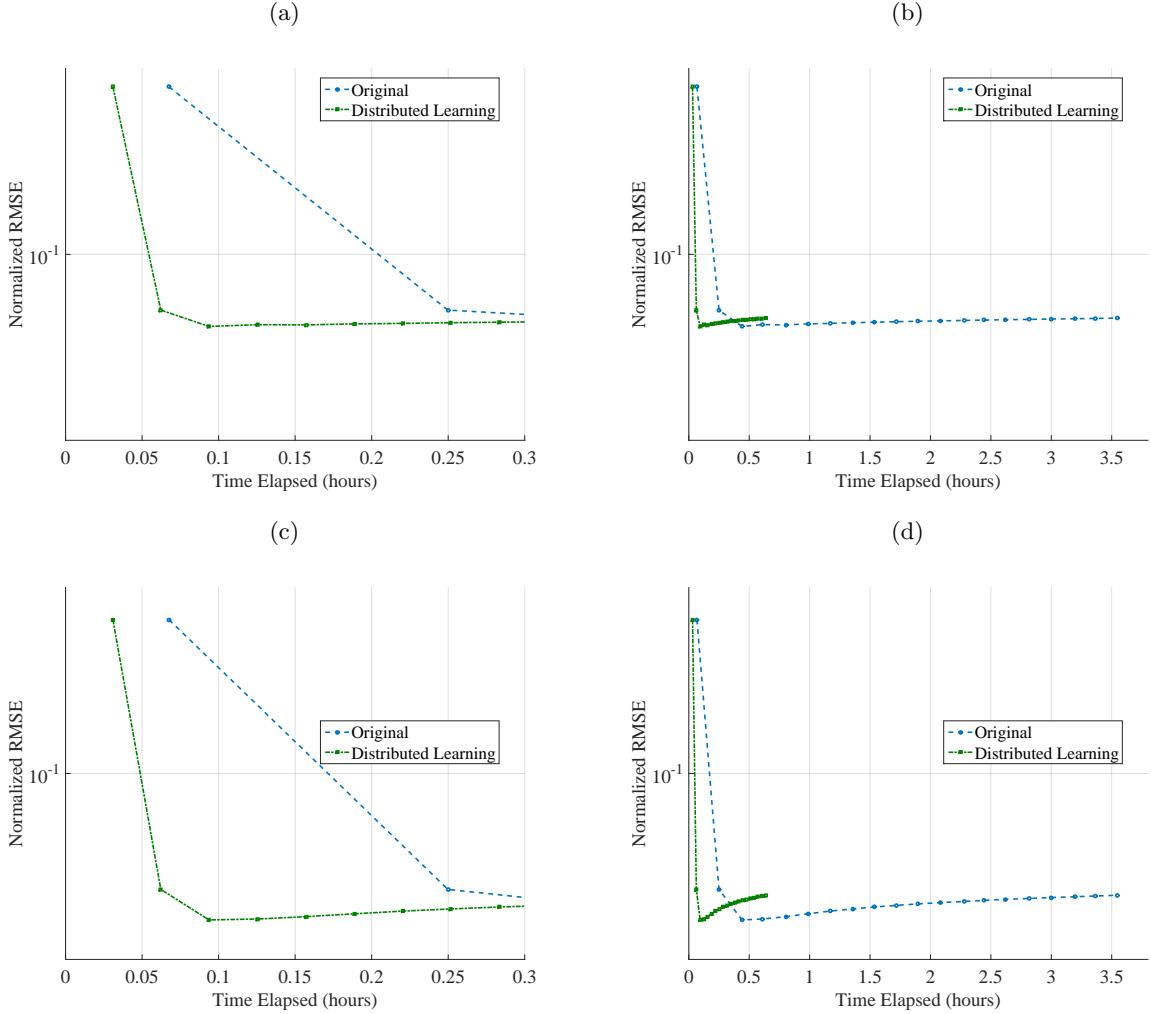


Figure 14: The reconstruction error between calculated dictionaries and augmented Lagrangian function for the GS data and memory-and-disk persistence model: (a) high resolution during the first 20 minutes of the experiment, (b) high resolution, total duration of the experiment, (c) low resolution during the 20 minutes of the experiment, (d) low resolution total duration of the experiment.

respectively, our architecture offers a scalable solution, providing more than 50% improvement in time response, when the number of available cores becomes 6 times greater than the original approach. At the same time, the convergence behavior of the optimization process becomes substantially faster compared to the sequential counterpart (75% for the astrophysics use case, 65.7% for the remote sensing use case).

- With regard to the cluster behavior, the scale of an input imaging dataset is relevant to the capacity of the cluster, and the demands of the solving approach; the overhead introduced by Spark for the distributed calculations may hinder the overall performance, when the dataset is relatively small compared to the computational capacity of the cluster. In such cases it is preferable to retain a small number of partitions N , in order to avoid introducing unnecessary shuffling overhead.
- On a similar note, the benefit of having an increased number of partitions is related to the memory usage per worker. Results on both use cases (Fig. 6 and Fig 11-12) point out the trade off between the RAM availability and the size of the input problem; When the memory per worker is limited with respect to size of the input data it is considered preferable to increase the number of partitions of the RDD. It may increase the time response, due to the increased number of tasks, however it will yield a more reliable solution within the lifetime of the program execution.
- The cluster configuration in terms of homogeneity of computational resources can also affect the memory behavior. Providing workers with similar RAM size and number of computational cores results in smaller deviation on the memory usage across all slaves involved. This is evident when comparing the memory behavior results (Fig. 6 versus Fig. 11-12) when either the astrophysics (five slaves→five workers, wherein one worker provides double number of CPU cores than the remaining four) or the remote sensing (five slaves→six workers, all workers provide equivalent number of CPU cores) problem is considered.
- As highlighted in the super-resolution use case, the impact of the persistence model is crucial when the memory per worker is limited and the use of disk space is unavoidable. In such cases, the memory-and-disk model, responsible for storing intermediate results on disk, is preferable; it eliminates the necessity of adding-removing results from memory, and thereby improve the time performance of subsequent calculations (e.g. Fig. 13).

5. Conclusions

In this work we propose an Spark-compliant architecture for performing distributed learning over bundled scientific imaging data. We present the respective algorithm parallelization for two high-impact use cases namely: (a) the computational astrophysics domain (space-variant deconvolution

of noisy galaxy images), (b) the remote sensing domain (super-resolution using sparse coupled dictionary training). The respective software libraries and datasets are publicly available at [45]. The evaluation studies highlight the practical benefits of changing the implementation exemplar and moving towards distributed computational approaches; while employing commodity hardware the results for both application scenarios indicate an improvement greater than 60% in time response terms against the conventional computing solutions. In addition, the trade off between memory availability and computational capacity has been revealed, while the offered insights draft the roadmap for further improvements on distributed scientific imaging frameworks.

Acknowledgment

This work was funded by the DEDALE project (no. 665044) within the H2020 Framework Program of the European Commission.

References

References

- [1] D. Chen, Y. Hu, C. Cai, K. Zeng, X. Li, Brain big data processing with massively parallel computing technology: challenges and opportunities, *Software: Practice and Experience* (2016) n/a–n/aSpe.2418.
- [2] V. Marx, Biology: The big challenges of big data, *Nature* 498 (7453) (2013) 255–260.
- [3] C. P. Chen, C.-Y. Zhang, Data-intensive applications, challenges, techniques and technologies: A survey on big data, *Information Sciences* 275 (2014) 314–347.
- [4] H. Hu, Y. Wen, T. S. Chua, X. Li, Toward scalable systems for big data analytics: A technology tutorial, *IEEE Access* 2 (2014) 652–687. doi:10.1109/ACCESS.2014.2332453.
- [5] The Sloan Digital Sky Survey, <https://classic.sdss.org/>.
- [6] The Euclid Space Mission, <http://sci.esa.int/euclid/>.
- [7] P. Huijse, P. A. Estevez, P. Protopapas, J. C. Principe, P. Zegers, Computational intelligence challenges and applications on large-scale astronomical time series databases, *IEEE Computational Intelligence Magazine* 9 (3) (2014) 27–39.

- [8] H.-D. Guo, L. Zhang, L.-W. Zhu, Earth observation big data for climate change research, *Advances in Climate Change Research* 6 (2) (2015) 108 – 117, special issue on advances in Future Earth research. doi:<https://doi.org/10.1016/j.accre.2015.09.007>.
- [9] X. Song, R. Shibasaki, N. J. Yuan, X. Xie, T. Li, R. Adachi, Deepmob: Learning deep knowledge of human emergency behavior and mobility from big and heterogeneous data, *ACM Trans. Inf. Syst.* 35 (4) (2017) 41:1–41:19. doi:[10.1145/3057280](https://doi.org/10.1145/3057280).
- [10] B. Wilson, R. Palamuttam, K. Whitehall, C. Mattmann, A. Goodman, M. Boustani, S. Shah, P. Zimdars, P. Ramirez, Scispark: highly interactive in-memory science data analytics, in: *Big Data (Big Data)*, 2016 IEEE International Conference on, IEEE, 2016, pp. 2964–2973.
- [11] Z. Zhang, K. Barbary, F. A. Nothaft, E. R. Sparks, O. Zahn, M. J. Franklin, D. A. Patterson, S. Perlmutter, Kira: Processing astronomy imagery using big data technology, *IEEE Transactions on Big Data PP* (99) (2016) 1–1.
- [12] E. P. Xing, Q. Ho, P. Xie, D. Wei, Strategies and principles of distributed machine learning on big data, *Engineering* 2 (2) (2016) 179 – 195. doi:<https://doi.org/10.1016/J.ENG.2016.02.008>.
- [13] L. Zhou, S. Pan, J. Wang, A. V. Vasilakos, Machine learning on big data: Opportunities and challenges, *Neurocomputing* 237 (2017) 350 – 361. doi:<https://doi.org/10.1016/j.neucom.2017.01.026>.
- [14] A. Panousopoulou, S. Farrens, Y. Mastorakis, J.-L. Starck, P. Tsakalides, A distributed learning architecture for big imaging problems in astrophysics, in: *Signal Processing Conference (EUSIPCO)*, 2017 25th European, IEEE, 2017, pp. 1440–1444.
- [15] M. Zaharia, R. S. Xin, P. Wendell, T. Das, M. Armbrust, A. Dave, X. Meng, J. Rosen, S. Venkataraman, M. J. Franklin, A. Ghodsi, J. Gonzalez, S. Shenker, I. Stoica, Apache spark: A unified engine for big data processing, *Commun. ACM* 59 (11) (2016) 56–65. doi:[10.1145/2934664](https://doi.org/10.1145/2934664).
- [16] J. Dean, S. Ghemawat, Mapreduce: simplified data processing on large clusters, *Communications of the ACM* 51 (1) (2008) 107–113.

- [17] Y. Low, D. Bickson, J. Gonzalez, C. Guestrin, A. Kyrola, J. M. Hellerstein, Distributed graphlab: a framework for machine learning and data mining in the cloud, *Proceedings of the VLDB Endowment* 5 (8) (2012) 716–727.
- [18] R. S. Xin, D. Crankshaw, A. Dave, J. E. Gonzalez, M. J. Franklin, I. Stoica, Graphx: Unifying data-parallel and graph-parallel analytics, arXiv preprint arXiv:1402.2394.
- [19] L. Neumeyer, B. Robbins, A. Nair, A. Kesari, S4: Distributed stream computing platform, in: Data Mining Workshops (ICDMW), 2010 IEEE International Conference on, IEEE, 2010, pp. 170–177.
- [20] A. Toshniwal, S. Taneja, A. Shukla, K. Ramasamy, J. M. Patel, S. Kulkarni, J. Jackson, K. Gade, M. Fu, J. Donham, et al., Storm@ twitter, in: Proceedings of the 2014 ACM SIGMOD international conference on Management of data, ACM, 2014, pp. 147–156.
- [21] L. Wang, J. Tao, R. Ranjan, H. Marten, A. Streit, J. Chen, D. Chen, G-hadoop: Mapreduce across distributed data centers for data-intensive computing, *Future Generation Computer Systems* 29 (3) (2013) 739–750.
- [22] R. Ahuja, Hadoop framework for handling big data needs, in: *Handbook of Research on Big Data Storage and Visualization Techniques*, IGI Global, 2018, pp. 101–122.
- [23] H. Zhang, G. Chen, B. C. Ooi, K. L. Tan, M. Zhang, In-memory big data management and processing: A survey, *IEEE Transactions on Knowledge and Data Engineering* 27 (7) (2015) 1920–1948.
- [24] X. Meng, J. Bradley, B. Yavuz, E. Sparks, S. Venkataraman, D. Liu, J. Freeman, D. Tsai, M. Amde, S. Owen, et al., Mllib: Machine learning in apache spark, *The Journal of Machine Learning Research* 17 (1) (2016) 1235–1241.
- [25] M. Assefi, E. Behravesh, G. Liu, A. P. Tafti, Big data machine learning using apache spark mllib, in: *Big Data*, 2017 IEEE International Conference on, IEEE, 2017, pp. 3492–3498.
- [26] W. Wang, T. Liu, D. Tang, H. Liu, W. Li, R. Lee, Sparkarray: An array-based scientific data management system built on apache spark, in: *Networking, Architecture and Storage (NAS)*, 2016 IEEE International Conference on, IEEE, 2016, pp. 1–10.

- [27] C. Misra, S. Haldar, S. Bhattacharya, S. K. Ghosh, Spin: A fast and scalable matrix inversion method in apache spark, in: Proceedings of the 19th International Conference on Distributed Computing and Networking, ICDCN '18, ACM, New York, NY, USA, 2018, pp. 16:1–16:10. doi:10.1145/3154273.3154300.
- [28] E. P. Xing, Q. Ho, W. Dai, J. K. Kim, J. Wei, S. Lee, X. Zheng, P. Xie, A. Kumar, Y. Yu, Petuum: A new platform for distributed machine learning on big data, *IEEE Transactions on Big Data* 1 (2) (2015) 49–67.
- [29] X. Zhang, U. Khanal, X. Zhao, S. Ficklin, Making sense of performance in in-memory computing frameworks for scientific data analysis: A case study of the spark system, *Journal of Parallel and Distributed Computing* doi:<https://doi.org/10.1016/j.jpdc.2017.10.016>.
- [30] S. Cano-Lores, A. Lapin, J. Carretero, P. Kropf, Applying big data paradigms to a large scale scientific workflow: Lessons learned and future directions, *Future Generation Computer Systems*. doi:<https://doi.org/10.1016/j.future.2018.04.014>.
- [31] Q. Ho, J. Cipar, H. Cui, S. Lee, J. K. Kim, P. B. Gibbons, G. A. Gibson, G. Ganger, E. P. Xing, More effective distributed ml via a stale synchronous parallel parameter server, in: *Advances in neural information processing systems*, 2013, pp. 1223–1231.
- [32] M. Stonebraker, P. Brown, D. Zhang, J. Becla, Scidb: A database management system for applications with complex analytics, *Computing in Science & Engineering* 15 (3) (2013) 54–62.
- [33] R. Palamuttam, R. M. Mogrovejo, C. Mattmann, B. Wilson, K. Whitehall, R. Verma, L. McGibney, P. Ramirez, Scispark: Applying in-memory distributed computing to weather event detection and tracking, in: *Big Data (Big Data), 2015 IEEE International Conference on*, IEEE, 2015, pp. 2020–2026.
- [34] Z. Zhang, K. Barbary, F. A. Nothaft, E. Sparks, O. Zahn, M. J. Franklin, D. A. Patterson, S. Perlmutter, Scientific computing meets big data technology: An astronomy use case, in: *2015 IEEE International Conference on Big Data (Big Data)*, 2015, pp. 918–927.
- [35] J. Peloton, C. Arnault, S. Plaszczynski, Analyzing astronomical data with apache spark, arXiv preprint arXiv:1804.07501.

- [36] J. Maillo, S. Ramrez, I. Triguero, F. Herrera, knn-is: An iterative spark-based design of the k-nearest neighbors classifier for big data, *Knowledge-Based Systems* 117 (2017) 3 – 15, volume, Variety and Velocity in Data Science. doi:<https://doi.org/10.1016/j.knosys.2016.06.012>.
- [37] J. Arias, J. A. Gamez, J. M. Puerta, Learning distributed discrete bayesian network classifiers under mapreduce with apache spark, *Knowledge-Based Systems* 117 (2017) 16 – 26, volume, Variety and Velocity in Data Science. doi:<https://doi.org/10.1016/j.knosys.2016.06.013>.
- [38] C.-S. Huang, M.-F. Tsai, P.-H. Huang, L.-D. Su, K.-S. Lee, Distributed asteroid discovery system for large astronomical data, *Journal of Network and Computer Applications* 93 (2017) 27 – 37. doi:<https://doi.org/10.1016/j.jnca.2017.03.013>.
- [39] M. Makkie, X. Li, S. Quinn, B. Lin, J. Ye, G. Mon, T. Liu, A distributed computing platform for fmri big data analytics, *IEEE Transactions on Big Data* doi:[10.1109/TB DATA.2018.2811508](https://doi.org/10.1109/TB DATA.2018.2811508).
- [40] J. Pearl, Probabilistic reasoning in intelligent systems: networks of plausible inference, Elsevier, 2014.
- [41] M. Zaharia, M. Chowdhury, T. Das, A. Dave, J. Ma, M. McCauley, M. J. Franklin, S. Shenker, I. Stoica, Resilient distributed datasets: A fault-tolerant abstraction for in-memory cluster computing, in: Proceedings of the 9th USENIX Conference on Networked Systems Design and Implementation, NSDI'12, USENIX Association, Berkeley, CA, USA, 2012, pp. 2–2.
- [42] T. P. Robitaille, E. J. Tollerud, P. Greenfield, M. Droettboom, E. Bray, T. Aldcroft, M. Davis, A. Ginsburg, A. M. Price-Whelan, W. E. Kerzendorf, et al., Astropy: A community python package for astronomy, *Astronomy & Astrophysics* 558 (2013) A33.
- [43] O. Fourt, J.-L. Starck, F. Sureau, J. Bobin, Y. Moudden, P. Abrial, J. Schmitt, isap: Interactive sparse astronomical data analysis packages, *Astrophysics Source Code Library* 1 (2013) 03029.
- [44] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research* 12 (2011) 2825–2830.

- [45] The distributed learning toolbox for scientific imaging big data, <https://github.com/dedalefet/Distributed-Learning-Toolbox> (2018).
- [46] R. Laureijs, J. Amiaux, S. Arduini, J. . Augu  res, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet, et al., Euclid Definition Study Report, ArXiv e-prints [arXiv:1110.3193](https://arxiv.org/abs/1110.3193).
- [47] J.-L. Starck, A. Bijaoui, I. Valtchanov, F. Murtagh, A combined approach for object detection and deconvolution, *Astronomy and Astrophysics Supplement Series* 147 (1) (2000) 139–149.
- [48] S. Farrens, F. M. Ngol  Mboula, J.-L. Starck, Space variant deconvolution of galaxy survey images, *Astronomy & Astrophysics* 601 (2017) A66. [arXiv:1703.02305](https://arxiv.org/abs/1703.02305), doi:10.1051/0004-6361/201629709.
- [49] J.-L. Starck, F. Murtagh, J. Fadili, *Sparse Image and Signal Processing: Wavelets and Related Geometric Multiscale Analysis*, Cambridge University Press, 2015.
- [50] L. Condat, A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms, *Journal of Optimization Theory and Applications* 158 (2) (2013) 460–479.
- [51] R. Mandelbaum, B. Rowe, J. Bosch, C. Chang, F. Courbin, M. Gill, M. Jarvis, A. Kannawadi, et al., The Third Gravitational Lensing Accuracy Testing (GREAT3) Challenge Handbook, *ApJS* 212 (2014) 5. [arXiv:1308.4982](https://arxiv.org/abs/1308.4982).
- [52] T. Kuntzer, M. Tewes, F. Courbin, Stellar classification from single-band imaging using machine learning, *A&A* 591 (2016) A54. [arXiv:1605.03201](https://arxiv.org/abs/1605.03201), doi:10.1051/0004-6361/201628660.
- [53] M. Cropper, H. Hoekstra, T. Kitching, R. Massey, J. Amiaux, L. Miller, Y. Mellier, J. Rhodes, B. Rowe, S. Pires, C. Saxton, R. Scaramella, Defining a weak lensing experiment in space, *MNRAS* 431 (2013) 3103–3126. [arXiv:1210.7691](https://arxiv.org/abs/1210.7691), doi:10.1093/mnras/stt384.
- [54] Apache Spark 2.1.0 (2017).
URL <https://spark.apache.org/docs/2.1.0/index.html>

- [55] K. Fotiadou, G. Tsagkatakis, P. Tsakalides, Multi-source image enhancement via coupled dictionary learning, in: Proc. Signal Processing with Adaptive Sparse Structured Representations (SPARS), 2017.
- [56] K. Fotiadou, G. Tsagkatakis, B. Moraes, F. Abdalla, P. Tsakalides, Denoising galaxy spectra with coupled dictionary learning, in: European Signal Processing Conference (EUSIPCO), 2017.
- [57] M. Hong, Z.-Q. Luo, M. Razaviyayn, Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems, SIAM Journal on Optimization 26 (1) (2016) 337–364.
- [58] M. Elad, Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing, 1st Edition, Springer Publishing Company, Incorporated, 2010.
- [59] A. Lambrechts, P. Gonzalez, B. Geelen, P. Soussan, K. Tack, M. Jayapala, A cmos-compatible, integrated approach to hyper-and multispectral imaging, in: 2014 IEEE International Electron Devices Meeting, IEEE, 2014, pp. 10–5.