

Computational Science-based Research on Dark Matter at KISTI

Kihyeon Cho^{1,2†}¹Korea Institute of Science and Technology Information, Daejeon 34141, Korea²University of Science and Technology, Daejeon 34113, Korea

The Standard Model of particle physics was established after discovery of the Higgs boson. However, little is known about dark matter, which has mass and constitutes approximately five times the number of standard model particles in space. The cross-section of dark matter is much smaller than that of the existing Standard Model, and the range of the predicted mass is wide, from a few eV to several PeV. Therefore, massive amounts of astronomical, accelerator, and simulation data are required to study dark matter, and efficient processing of these data is vital. Computational science, which can combine experiments, theory, and simulation, is thus necessary for dark matter research. A computational science and deep learning-based dark matter research platform is suggested for enhanced coverage and sharing of data. Such an approach can efficiently add to our existing knowledge on the mystery of dark matter.

Keywords: dark matter, computational science, nuclear physics, particle physics, astronomical physics

1. INTRODUCTION

Computational science is a new interdisciplinary field commonly used to explore problems that are difficult to approach using conventional theories or experimental simulations (Cho 2016b). As computers have evolved, computational science has become a more economical method of solving such problems, and its popularity has increased (Cho 2016b). Supercomputers and massive amounts of data help overcome human limitations (Lin & Yen 2009; Cho 2011; Cho et al. 2015; Cho 2016a) and computational science helps researchers and experts create theoretical models of complicated, expensive, dangerous, large-scale, and hyperfine phenomena (Cho 2016b). They help us to model the origin of the universe using simulations, and provide the only way for us to observe its evolution (Cho 2016b).

While the Standard Model of particle physics was completed with discovery of the Higgs boson (ATLAS collaboration 2012; CMS collaboration 2012), it is thought to be imperfect as a theory because gravitational force was excluded from the equation. Newton's equations of motion can be applied to the natural world, but yield results that deviate from common sense in the range of the speed of light, including a mass

increase and a length decrease. Likewise, the Standard Model has proven to be accurate for describing phenomena at the current energy scale, but is likely to be unsuitable at higher energy scales.

It is known that the universe today consists of 4 % Standard Model particles, 26 % dark matter, and 70 % dark energy (Cho 2016a). Specifically, dark matter research is increasingly reliant on computational science due to changes in the research environment and constant upgrades of data, facilities, equipment, and computing capacity.

Here, we introduce a methodology for dark matter research based on data, computational science, and deep learning. We suggest a dark matter research platform and discuss its role in modern science.

2. DARK MATTER DETECTION

To date, the known conditions required for dark matter are that it emits no type of light, escapes and does not collide with any matter, and is a cold model that was close to zero velocity in the early universe (Cho 2016b). Methods of detecting dark matter can be classified into direct detection, indirect

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 29 MAY 2017 Revised 5 JUN 2017 Accepted 7 JUN 2017

†Corresponding Author

Tel: +82-42-869-0722, E-mail: cho@kisti.re.kr

ORCID: <https://orcid.org/0000-0003-1705-7399>

detection, and accelerator detection. Fig. 1 shows the three methods of detecting dark matter and the interactions between the Standard Model and dark matter.

The first method is indirect detection. In cosmology, dark matter interacts and produces Standard Model particles. The Korea Astronomy and Space Science Institute (KASI) provides large-scale cosmology data for the indirect detection of dark matter.

The second method is direct detection. Since dark matter has mass, it interacts with detectors. The Institute for Basic Science (IBS) is currently performing underground direct detection and Axion experiments.

The third method is accelerator detection. Using the Standard Model particle collider with an accelerator, we can generate possible dark matter. The Korea Institute of Science and Technology Information (KISTI) provides accelerator data for researchers in Korea who use it to search for dark matter.

Combining research on dark matter using computational

science-based particle physics and that using cosmology increases the synergy effect between them. Regarding the infrastructure of each organization, the KASI uses the mass data center and supercomputer at the KISTI for proposed model research in cosmology. The KISTI also stores data observed at the KASI to compare with experimental data. To promote efficient national research, the development of calculation science, and the synergy effect of research, a dark matter research platform is established to allow resource and information sharing between researchers.

Fig. 2 shows the locations of astronomical and accelerator data retrieval. For astronomical data, Sloan digital sky survey (SDSS) dark energy spectroscopic instrument (DESI) data are produced in the USA and large synoptic survey telescope (LSST) are produced in Chile. For accelerator data, compact muon solenoid (CMS) data is produced at European Organization for Nuclear Research (CERN) in Europe, Belle data is produced at High Energy Accelerator Research Organization (KEK) in Japan, and Deep Underground Neutrino Experiment (DUNE) data

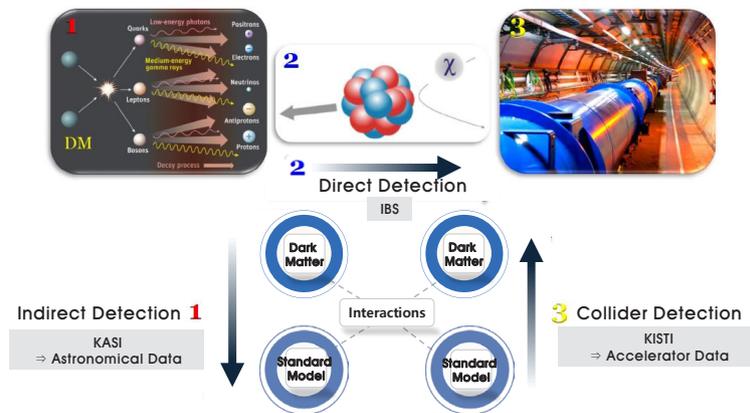


Fig. 1. Methods of dark matter detection: 1) Indirect detection, 2) Direct detection, and 3) Accelerator detection.

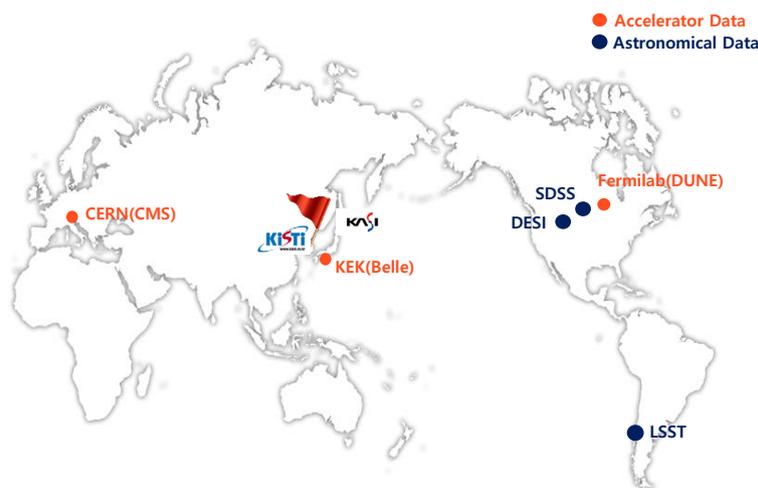


Fig. 2. Locations of astronomical and accelerator data generation.

Table 1. Limitations on sharing data of different types

Data	Astronomical Data	Astronomical data	Simulation Data
Input data	Collaboration only	Collaboration only	All researchers share
Output data	All researchers share	All researchers share	All researchers share

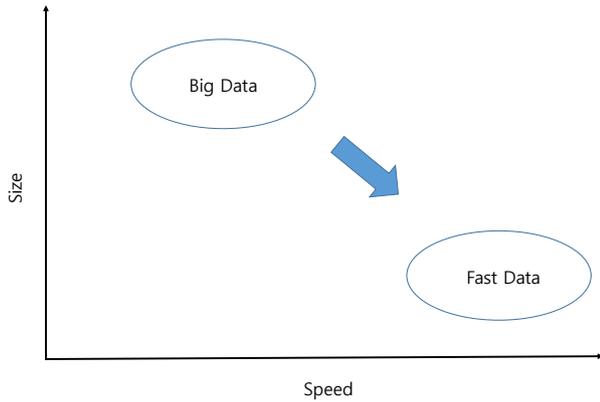


Fig. 3. The value of data over time.

is produced at Fermilab in USA. Table 1 shows the limitations of data sharing for astronomical data, accelerator data, and simulation data.

3. METHODOLOGY

3.1 Requirements of Dark Matter Research

In this section, we consider what we need for dark matter research. This is explained from the point of view of data, computational science, and deep learning. Regarding data, usage is moving from a theory-driven approach to a data-driven approach. However, the value of data over time drives automated low latency actions in response to events of interest. Big data employ a large historical dataset to make split second decisions on real time data, as shown in Fig. 3.

Data analysis has also changed from a theory-driven approach to a data-driven approach. The theory-driven approach, by which the Higgs boson was discovered (ATLAS collaboration 2012 ; CMS collaboration 2012), starts with the system and works towards the data. The data-driven approach starts with the data and works towards the system. Tables 2 and 3 show the planned future surveys of accelerator data (Song 2016). The size of LSST data in 2023–2030 will be 500 PB. The size of accelerator data will also increase enormously. Large hadron collider (CMS) data will be 100 PB/yr, which is 5–10 times its current size. Therefore, we need to prepare the

Table 2. Size of astronomical data (Song 2016)

Year	Observatory	Size of data (Total)
Current	SSDS	300 TB
2018 –	DESI (starting)	1 PB
2014–2020	SDSS4	500 TB
2017–2022	DESI	5 PB
2023–2030	LSST	500 PB

Table 3. Size of accelerator data

Year	Experiment	Size of Data
Current	Belle	~ 1 PB (total)
Current	LHC (CMS)	10–20 PB/yr
2018–2024	Belle II	100 PB (total)
2025–	LHC (CMS)	100 PB/yr

system for this increased data size.

Computational science represents the convergence of theory, experiments, and simulations (Cho 2016b). Here, we describe the dark matter research requirements from each point of view. Regarding theory, requirements include rapid updates and sharing of information, a specialist key point summary database, regular meetings for network feedback between experimentalists and theorists, rapid and easy comparisons of theoretical models, development of new numerical packages and computing power, and cross checks with astrophysics regarding theoretical models such as the N-body simulation. Regarding experiments, requirements include manpower for dark matter research using hardware and software, social networking at local and international institutions related to indirect, direct, and accelerator detection experiments, regular workshops for communication between the experimental and theoretical community, and big data computing resources for high-performance computing (HPC) and storage of experimental data. Regarding simulations, a large number of simulations with a theoretical background are required for dark matter detection, in which speed or memory issues are considered (Cho 2016b). Therefore, it is necessary to develop simulation tool kits (Cho 2016b).

This research cannot be performed by an individual or single research organization. The reason computational science is required in dark matter research is that the number of signal cases is small owing to the small cross-section (Cho et al. 2015). The Standard Model was established due to the discovery of the Higgs boson, which has the smallest scattering cross-section among standard models. However, the cross-section of dark matter is less than one thousandth that of the Higgs boson (Cho 2016b). Therefore, at least 1,000 times more experimental and simulation data are required than for the Higgs boson (Cho et al. 2015). In addition, the mass range and coupling constant range are wide (Baer et al. 2015). For dark matter simulation, a dark matter research

management system is required. By calculating the database and parameters according to the model, the experimental results of this system can restrain theoretical parameters and enhance precision of predicted values. With an automated tool, researchers can easily obtain quicker and more accurate results. The main goals are to develop a simulation toolkit involving a phenomenological analysis technique for large-scale simulation and testing of candidate dark matter models, and to conduct applied studies using the developed toolkit that are available to the wider research community. Software needs to be developed and applied to meet the demands of the evolving physics. This can be done by conducting a combined reconstruction of relevant components to simulate candidate dark matter models.

The third aspect of dark matter research is deep learning. Higgs boson detection has been applied to deep learning in High Energy Physics (Baldi et al. 2014). The DUNE experiment uses image processing. Fig. 4 shows the process of deep learning. From the theoretical model, we generate Monte Carlo events using the KISTI supercomputers and compare

the results with experimental data. The results are used as input data for the deep learning process.

3.2 Dark Matter Research Platform

As well as the requirements for dark matter research discussed above, astronomical data and accelerator data must be combined to make the processing and analysis more efficient. Using deep learning, we can process and analyze convergence data using a big data platform and research and development (R&D) for Intelligence Information to release output data. Fig. 5 shows the suggested dark matter research platform. The KISTI processes accelerator data, and the KASI processes astronomical data, using deep learning and put these data on the dark matter research big data platform. Simulation data is produced through the KISTI supercomputer using theory. We compare simulation data with astronomical data using deep learning. The dark matter research platform for big data consists of a deep learning and named data networking (NDN) system. Dark

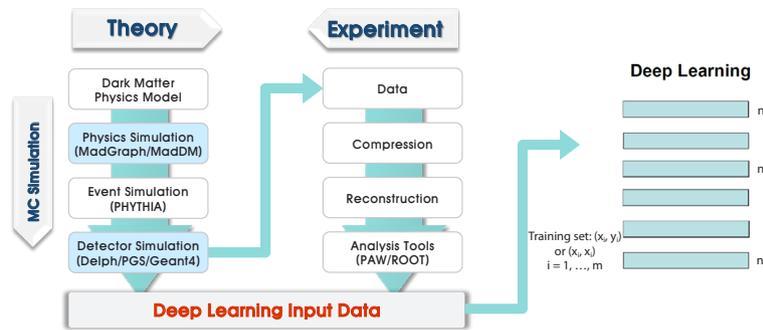


Fig. 4. Flow chart of the deep learning algorithm using Monte Carlo events.

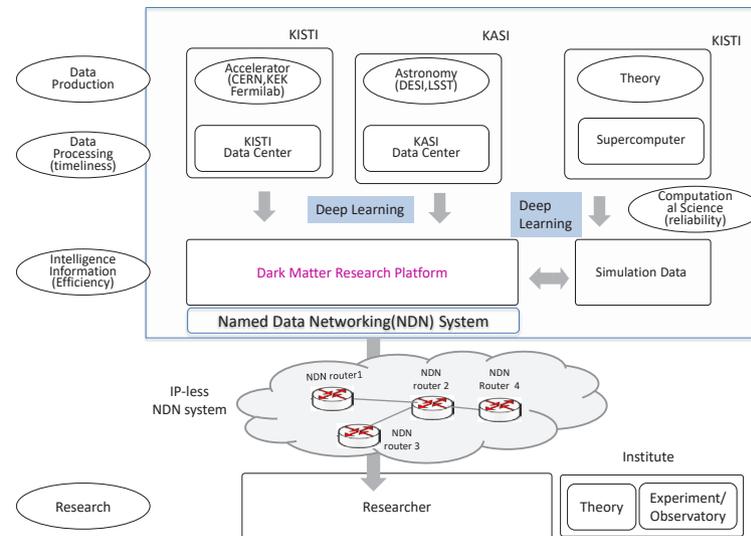


Fig. 5. The proposed dark matter research platform.

matter researchers can then use this system of combined theory-experiment-simulation.

4. RESULTS AND DISCUSSION

Here, we present some examples of computational science using astronomical data and accelerator data. Regarding astronomical data, some areas of space are dense, and others are not, because space is uneven. As the universe is still being created, doubts have risen as to whether the mass of standard model matter is sufficient for creating the galaxy. Recent simulation results show that dark matter plays a central role in creating the universe. Two massive cosmological simulations have been performed by a group of Korean scientists who studied the clustering of galaxies on cosmic time and length scales (Kim et al. 2012), using 8,240 CPU cores and 8.7 and 15 TB of memory, respectively, to simulate the gravitational evolutions of 216 and 374 billion particles, respectively. To compare with future galaxy surveys, mock luminous red galaxies (LRGs) in the pace light cone space were constructed (Kim et al. 2012) using the large-scale computational science method described earlier. In the simulation results assuming that dark matter exists, the density of matter evolved substantially with time and showed very similar density between the creation of the universe and today's large-scale galaxy structure (Cho 2016b).

The e^+e^- collider experiment is one example of computational science carried out using accelerator data. This aimed to discover new physical phenomena unexplained by the Standard Model. It tested related theories by comparing Monte Carlo simulation results based on phenomenological theories of the presence of dark matter and the corresponding accelerator data. Fig. 6 shows the dark matter production mechanism in the e^+e^- collider experiment (e.g., Belle II, International Linear Collider). It struggled to clarify the phenomenon of dark matter. Theories predicting unclarified regions were tested using simulation toolkits and Monte Carlo simulations were run using the KISTI supercomputer. The results of simulation outputs were then compared and analyzed to verify theories of dark matter and determine the characteristics of dark matter in search of new physical phenomena. Fig. 7 shows the Feynman diagrams of the

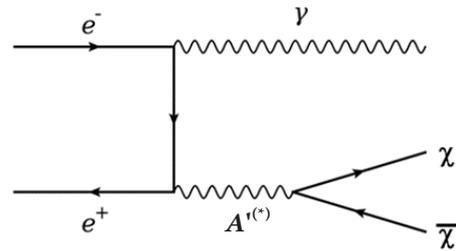


Fig. 6. Feynman diagrams of a dark matter event for the e^+e^- collider experiment ($e^+e^- \rightarrow \chi\bar{\chi}\gamma$) (Cho 2016b).

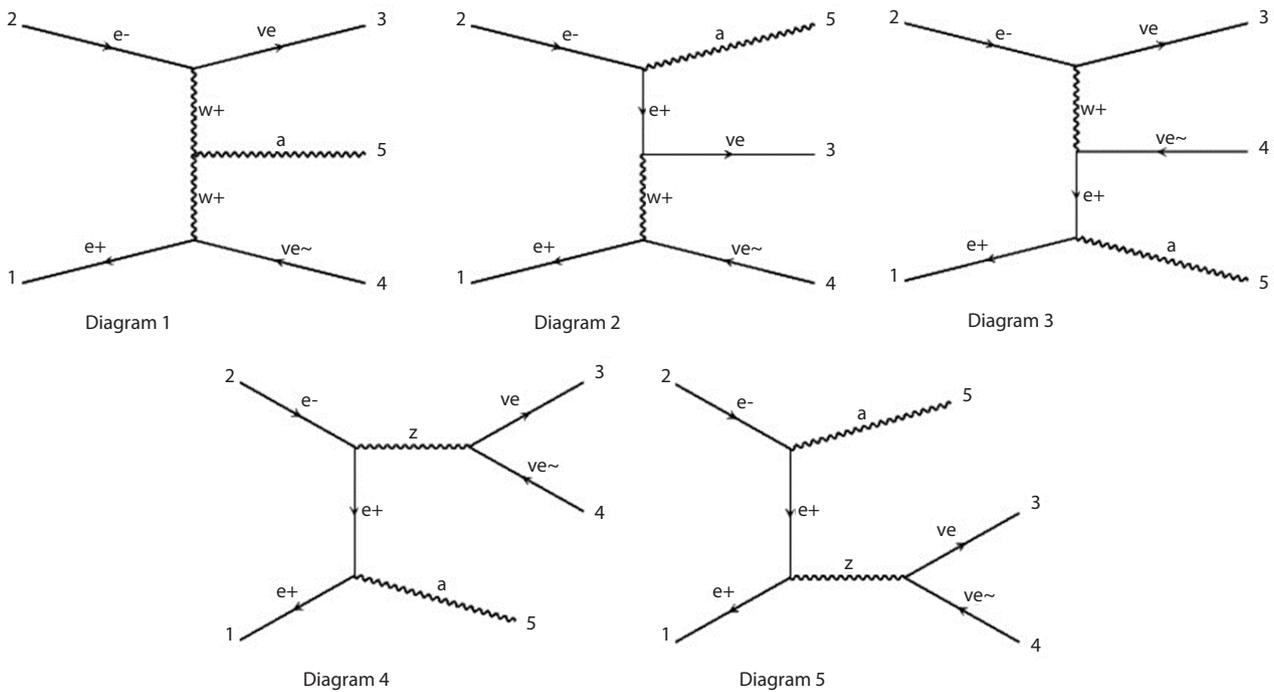


Fig. 7. Feynman diagrams of the standard model background at for the e^+e^- collider experiment ($e^+e^- \rightarrow \nu\bar{\nu}\gamma$).

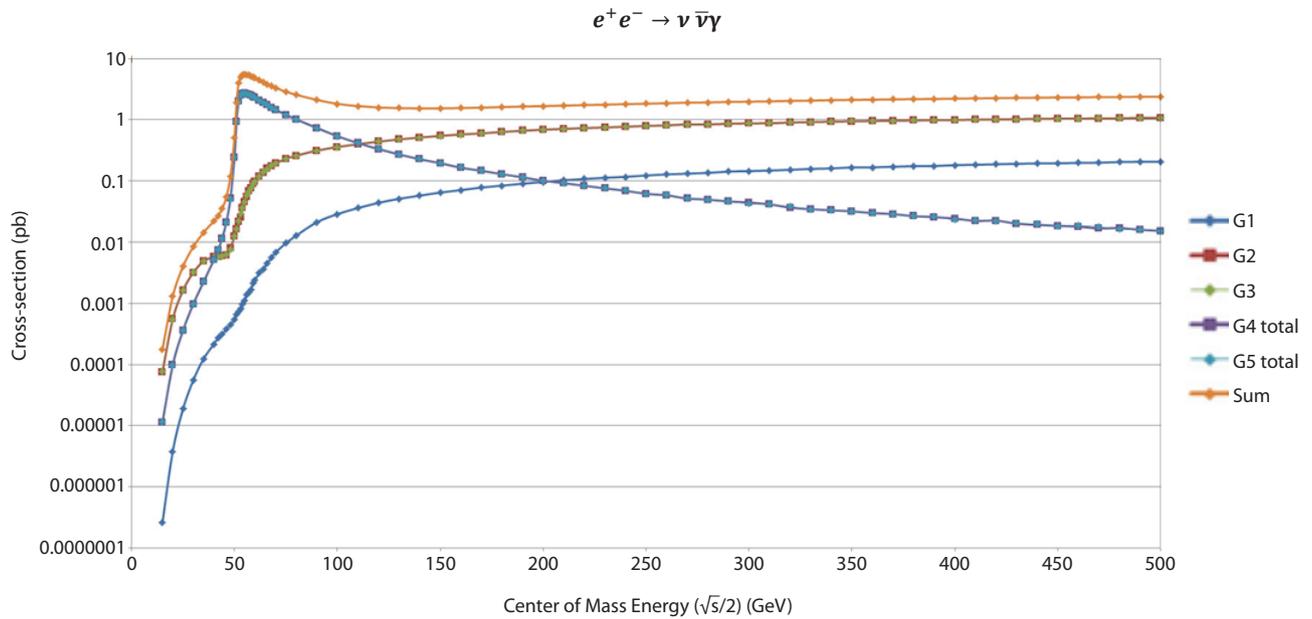


Fig. 8. Cross-section of the standard model background of for $e^+e^- \rightarrow \nu\bar{\nu}\gamma$ at in the e^+e^- collider experiment. The x- axis is denotes for the center of mass energy, in GeV, and the y- axis is represents for the cross-section, in pb. Each cross-section of each Feynman diagram described in Fig. 7 is also shown.

Standard Model background of a dark matter event in the e^+e^- collider experiment. The channel is $e^+e^- \rightarrow \nu\bar{\nu}\gamma$. We study the decay of neutrinos as the main background event of dark matter decay.

Fig. 8 shows the cross-section of the standard model background for $e^+e^- \rightarrow \nu\bar{\nu}\gamma$ simulated by MadGraph (Alwall et al. 2014). The computational science results using astronomical data and accelerator data can be used as inputs for the dark matter research platform as constraint conditions.

5. CONCLUSION

Computational science converging experiment-theory-simulation is required for research into dark matter detection, and relevant case studies are described in this paper. Computational science is essential for solving complicated problems in all fields beyond traditional science, including dark matter detection and core research on the evolution of the universe. Research on dark matter detection requires large data and computing resources. Moreover, software development required by the physics of dark matter detection are necessary. Since this cannot be performed by one particular research method, we suggest a dark matter research platform for synergizing all data from participating research institutes. Through it, we may prove the phenomenon of dark matter more efficiently by combining theoretical data, astronomical data, and accelerator data through computational science.

ACKNOWLEDGMENTS

This research was supported by the dark matter research cluster funded by the National Research Council of Science and Technology. The author thanks members of the dark matter research cluster.

REFERENCES

- Alwall J, Frederix R, Frixione S, Hirschi V, Maltoni F, et al., The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, *J. High Energy Phys.* 2014, 79 (2014). [https://doi.org/10.1007/JHEP07\(2014\)079](https://doi.org/10.1007/JHEP07(2014)079)
- ATLAS collaboration, Observation of a new particle in the search for the standard model Higgs boson with ATLAS detector at the LHC, *Phys. Lett. B* 716, 1-29 (2012). <https://doi.org/10.1016/j.physletb.2012.08.020>
- Baer H, Choi KY, Kim JE, Roszkowski L, Dark matter production in the early universe: beyond the thermal WIMP paradigm, *Phys. Rep.* 555, 1-60 (2015). <https://doi.org/10.1016/j.physrep.2014.10.002>
- Baldi P, Sadowski P, Witeson D, Searching for exotic particles in high-energy physics with deep learning, *Nature Commun.* 5, 4308 (2014). <https://doi.org/10.1038/ncomms5308>
- Cho K, Collider physics based on e-Science paradigm of experiment-computing-theory, *Comput. Phys. Commun.* 182, 1756-1759 (2011). <https://doi.org/10.1016/j.cpc.2010.12.019>

- Cho K, e-Science paradigm for astroparticle physics at KISTI, *J. Astron. Space Sci.* 33, 63-67 (2016a). <https://doi.org/10.5140/JASS.2016.33.1.63>
- Cho K, Computational science and the search for dark matter, *New Phys. Sae Mulli* 66, 950-956 (2016b). <https://doi.org/10.3938/NPSM.66.950>
- Cho K, Kim J, Kim J, Research and development of the evolving architecture for beyond the standard model, *J. Phys. Con. Ser.* 664, 072001 (2015). <https://doi.org/10.1088/1742-6596/664/7/072011>
- CMS collaboration, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC, *Phys. Lett. B* 716, 30-61 (2012). <https://doi.org/10.1016/j.physletb.2012.08.021>
- Kim J, Park C, Rossi G, Lee SM, The top-two biggest simulations ever performed for the study of the large-scale structures of the universe, *Phys. High Technol.* 21, 37-40 (2012). <https://doi.org/10.3938/PhiT.21.018>
- Lin SC, Yen E, e-Science for high energy physics in Taiwan and Asia, *J. Korean Phys. Soc.* 55, 2035-2039 (2009).
- Song Y, Study of astronomical big data and dark universe, in *Korea Supercomputing Conference 2016*, Seoul, Korea, 5-7 Oct 2016.