Registration No.	Name	Affiliation	Title of presentation	Abstract
1	Yoshiyuki Kabashima	Tokyo Institute of Technology, Japan	Online compressed sensing	An online scheme to carry out compressed sensing is developed on the basis of the Bayesian inference framework and mean field theory. **Co-authors: Paulo Rossi (San Paulo University) and Jun-ichi Inoue (Hokkaido University).**
2	Y-h. Taguchi	Chuo University	Principal component analysis based unsupervised feature extraction applied to budding yeast temporally periodic gene expression	Principal component analysis (PCA) based unsupervised feature extraction (FE) was successfully applied to a wide ranged problems ranging from circulating microRNA biomarker identification for diseases, proteome analysis, genes associated with aberrant promoter methylation over three autoimmune diseases, identification of disease causing genes of PTSD mediated heart failure, integrated analysis of genotype/DNA methylation of cancer, identification of ALS causing genes, identification of genes associated with differential gene expression as well as aberrant promoter methylation during transgenerational epigenetics and comparison between control/treated cancer cell lines. PCA based unsupervised FE was also suggested to be equivalent to FE based upon time consuming variational Bayes frameworks and its FE is extremely stable (i.e., relatively independent of sample selections), while no other methods that were compared with PCA based unsupervised FE could be competitive with it. In this poster, we summarize usefulness, superiorities towards conventional supervised methodologies, and its application to translational bioinformatics (e.g., findings of desase biomarker and in silico drug discoveries).
3	Shaogao Lyu	Southwestern University of Finance and Economics, China	Prediction performance and structure selection consistency for high dimensional partially varying coefficient models in reproducing kernel Hilbert space	Partially varying coefficient models (PVCM) provide a useful class of tools for modelling complex data by incorporating a combination of constant and time varying covariate effects. One natural question is that how to decide which covariates correspond to constant coefficients and which correspond to time-dependent coefficient functions. The structure selection problem is fundamentally important, since tackling this problem enhances model interpretation and avoids overfitting, as well as keeps the model flexibility. To address this issue, this paper proposes a new approach to estimation and structure selection for PVCM. Within a highdimensional framework, we derive convergence rates for the prediction risk of the proposed method when each unknown time-dependent coefficient lies in a reproducing kernel Hilbert space. Our upper bounds in $\|\cdot\|_2$ and $\|\cdot\|_n$ norms are established under two different kinds of settings, and are shown to be the optimality of our method under their individual settings. Under certain regularity conditions, we also show that the proposed estimator is able to identify the underlying structure correctly with high probability.
4	Kengo Nakamura	Tohoku University	Extraction of heavy metals characteristics of the 2011 Tohoku tsunami deposits using multiple classification analysis	Tsunami deposits were accumulated due to the impact of the Tohoku-oki earthquake at Tohoku coast areas in Japan. This paper reports an application of principal component analysis (PCA) and Cluster analysis (CA) to heavy metals content amount and dissolution concentration from tsunami deposits. Results of the heavy metals content and dissolution concentrations from tsunami deposits suggested low environmental risk, geometric means could be calculated to be 16mg/kg and 0.003ml/L for Pb, 1.8mg/kg and 0.004ml/L for As, and 0.17mg/kg and 0.0001ml/L for Cd. The CA using excluding outliers from PCA can extract characteristic of heavy metals relationships at content amount and dissolution concentration. The first clusters of content amount are Ni, Fe, Cd, Cu, Al, Cr, Zn and Mn; the second cluster is Pb, Sb, As and Mo. First cluster of dissolution concentration is Ni, Fe, Al and Cr; the second cluster is Mo, Sb, As, Cu, Zn Pb and Mn. Therefore, heavy metals with higher relationship can be simulated by elution trends, and heavy metals with lower relationship do not show same elution trends. Obviously, Ni and Fe content amounts are assumed for controlling Cd content amount. Pb and As of content amount are considered to reveal a kind of dissolution mechanism using 1N HCl in tsunami deposits.
5	Ryoko Nakata	JAMSTEC, Japan	Geodetic data inversion for spatial distribution of slip under smoothness, discontinuity, and sparsity constraints	In geodetic data inversion, insufficient observational data and smoothness constraints for model parameters make it difficult to clearly resolve small-scale heterogeneous structures with discontinuous boundaries. We therefore developed a novel regularization scheme for the inversion problem that uses discontinuity, sparsity, and smoothness constraints. The proposed method was applied to synthetic displacements calculated by a ring-shaped afterslip. The afterslip was obtained from reasonable numerical simulation of earthquake generation cycle with a rate- and state- dependent friction law and 3-D plate interface geometry. Obtained afterslip distribution that minimizes the evaluation function showed a ring-shaped distribution with sharply varying boundary. The same inversion test was conducted with smoothly varying circular slip distribution. Our method well reproduced both discontinuous and continuous slip distributions.
6	Takamitsu Araki	National Institute of Advanced Industrial Science and Technology, Japan	Simultaneous estimation of spatio-temporal distribution and duration of fault slip by switching model	Slow slip event (SSE), that is, aseismic transient slip of faults occurs over days, is an important crustal event in subduction zone, and Seismologists have needed to estimate accurately duration and spatio-temporal distribution of the SSE using only crustal measurement data. Estimation methods of the spatio-temporal distribution and the duration have been developed separately, and those methods use inappropriate models that increase the estimation errors. We propose the estimation method that estimates the spatio-temporal distribution and the duration of the SSE simultaneously by using the switching model. For the observation data measuring one SSE, the switching model represents three forms in three periods. In the first and third periods, the fault is fixed, and in the second period, the fault is slipping slowly. The time points at which the model changes as well as the parameters of the switching model are estimated by the maximum likelihood method using the Expectation-Maximization algorithm. Our method is compared to the conventional methods through their application to synthetic data.

7	Hiromichi Suetani	Oita University, Japan	Manifold learning approach for collective chaos	It has been known that a certain class of high-dimensional chaotic systems exhibits lower-dimensional motion at a macroscopic level whereas it also keeps high-dimensional chaos at a microscopic level, called "collective chaos". In this study, we propose an approach based on manifold learning for modeling collective chaos.
8	Koh Takeuchi, Yoshinobu Kawahara, and Tomoharu Iwata	NTT and Osaka University	Higher Order Fused Regularization for Supervised Learning with Grouped Parameters	We often encounter situations in supervised learning where there exist possibly overlapping groups that consist of more than two parameters. For example, we might work on parameters that correspond to words expressing the same meaning, music pieces in the same genre, and books released in the same year. Based on such auxiliary information, we could suppose that parameters in a group have similar roles in a problem and similar values. In this paper, we propose the Higher Order Fused (HOF) regularization that can incorporate smoothness among parameters with group structures as prior knowledge in supervised learning. We define the HOF penalty as the Lov\'{a}sz extension of a submodular higher-order potential function, which encourages parameters in a group to take similar estimated values when used as a regularizer. Moreover, we develop an efficient network flow algorithm for calculating the proximity operator for the regularized problem. We investigate the empirical performance of the proposed algorithm by using synthetic and real-world data.
9	Chongke Bi	RIKEN	Interactive In-situ Visualization for Large-Scale Simulations on the K Computer	A large-scale simulation usually needs several days, or even several months to be executed on the supercomputer. It is desired by researchers to observe the visualization results of such kind of simulations in real time for analysis and adjustment of initial simulation parameters. However, currently most of such kinds of simulation datasets are analyzed by post-visualization. It is usually necessary to execute these large-scale simulations several times for selecting a set of good initial simulation parameters and visualization parameters. This leads to a time consuming process for researchers. Interactive in-situ visualization can help us to minimize this issue by enabling the visualization of the simulation results immediately after several time steps are finished. By analyzing these real time visualization results, researchers can adjust the simulation parameters and restart the simulation at any time. As a result, it becomes possible to avoid the repetitive execution of the entire simulation just for the parameter selection process. In this paper, a compression based interactive in-situ visualization method is proposed. This is achieved by using proper orthogonal decomposition (POD), which can extract data features for getting a good compression ratio for in-situ visualization. We also proposed a data sharing approach between the simulation code and the compression code to obtain the best performance. Furthermore, the visualization framework, called Heterogeneously Integrated Visualization Environment (HIVE) is employed in our system for the rendering and the interactive visualization.
10	Tomoki Tokuda	OIST	Multiple co-clustering based on nonparametric mixture models with heterogeneous marginal distributions	We propose a novel method for multiple clustering that assumes a co-clustering structure (partitions both in rows and columns of data matrix) in each view, which is applicable to high dimensional data. This method can be considered as a sparse model for clustering by means of partitioning features that contribute to several sample clustering solutions. The method is based on a nonparametric Bayesian approach in which the number of views and the number of feature-/subject-clusters are inferred in a data-driven manner. We simultaneously model different types of distribution family such as Gaussian, Poisson and multinomial in each cluster block. This leads our method to be applicable to datasets consisting of both numerical and categorical variables which is common in biomedical data analysis. The clustering solutions are obtained based on variational inference with mean field approximation. We apply our method to a real dataset with no true cluster structure available, from which useful implication is drawn about possible clustering structures of the data.
11	Thong Pham	Osaka University	Efficient joint estimation of rich-gets-richer and fit-gets-richer effects in complex networks	The famous "rich-gets-richer" and "fit-gets-richers" phenomena have long been conjectured to be the main driving forces behind the evolution of various kinds of systems founded in diverging fields such as biology, ecology, sociology, economics, etc. The field of complex network analysis is of course no exception. "Rich-gets-richer" and "fit-gets-richers" phenomena, now under the new names of "preferential attachment" and "fitness", have been hypothesized to be the underlying mechanisms of the evolution of some important types of complex networks. As in any scientific conjectures, the hypothesized presences of preferential attachment and fitness mechanisms call for the real need of validating such hypotheses in real-world networks. In our case, this amounts to actually estimate the preferential attachment (PA) function and node fitnesses from observed data. In the first part of our research, we provide a statistical method to jointly estimate PA function and fitnesses from a time-series of network snapshots. In the second part, we consider the case when we could only observe a single snapshot. Even in this case, which is frequently encountered in real-world situations, we show that not all hope is lost. By introducing the key concept of timeline, one can actually recover the PA function and node fitnesses to some extent.
12	Kourosh Meshgi	Kyoto University	Imitation Game: How a Imitate any Visual Tracker using a few Videos	Every year many trackers are introduced to address many visual tracking challenges such as illumination variances and partial occlusion of the objects. The biggest obstacle to construct the holy-grail of tracking from current trackers is the complexity of integrating mechanisms and contradictory objectives pursued by each of them. In this study, we propose a framework, MIMIC, to imitate the behaviors of an arbitrary black-box tracker, with a relatively simple non-linear tracker benefiting from a pool of various features. This study is based on a premise that a linear combination of sufficiently expressive features along with a flexible but still simple non-linear observation model can roughly approximate the behaviors of many popular trackers. To suppress over-fitting in this high-dimensional learning task, we employed the "dropout" technique. By emphasizing on optimization over model ensemble, this algorithm investigates different combinations of the features and prevents over-fitting by approximating the training dataset by means of a virtual ensemble of the models employing variety of feature subsets. The proper combination of the L1 regularization and the dropout algorithm is expected to work well in the simultaneous task of feature selection and feature weighting, by avoiding over-fitting even when the number of available videos is limited.

13	Takumi Ueda	Graduate School of Pharmaceutical Sciences, the University of Tokyo	Development of an NMR spectra reconstruction method to elucidate the CCR1- and CCR5- binding modes of MIP-1a	Conventional methods for the reconstruction of multidimensional NMR spectra from undersampled time-domain data were not effective for the accurate determination of the signal intensity ratios of the crowded two-dimensional spectra of proteins. Here, we developed an NMR spectra reconstruction method, “Conservation of experimental data in ANAlysis of FOuRier” (Co-ANAFOR), to reconstruct the crowded spectra from the undersampled time-domain data. The number of sampling points required for the transferred cross-saturation (TCS) experiments between membrane proteins, photosystem I and cytochrome b6f, and their ligand, plastocyanin, with Co-ANAFOR was half of that needed for linear prediction, and the peak height reduction ratios of the spectra reconstructed from truncated time-domain data by Co-ANAFOR were more accurate than those reconstructed from non-uniformly sampled data by compressed sensing (Ueda et al., Journal of Biomolecular NMR, 2015). Application of Co-ANAFOR to the TCS experiments of chemokine receptors, CCR1 and CCR5, and their ligand MIP-1a revealed that the residues on the N-loop and b-sheets of MIP-1a are close to both CCR1 and CCR5, and those in the C-terminal helix region are close to CCR5, although the NMR measurement time for CCR1 and CCR5 samples was limited to 24 hours, due to their low stability (Yoshiura et al., Journal of Biomolecular NMR, 2015).
14	Tomoya Sakai	Nagasaki University	Sparse modeling of lung sounds and their separation by convex optimization	We present and demonstrate a convex optimization technique for separating lung sounds acquired through an electronic stethoscope. Our technique exploits sparse and low-rank properties of the lung sounds in the Fourier, wavelet, and time-frequency domains. Co-authors: Senya Kiyasu, Sueharu Miyahara, and Yasushi Obase (Nagasaki University)
15	Masaaki Nagahara, Katsuyuki Kunida^A	Kyoto University, The University of Tokyo^A	Dynamical Sparse Modeling	The core idea of sparse modeling is to elicit a few essential parameters from a very large data set, often referred to as Big Data, with a fast and efficient algorithm. In this study, we consider data from dynamical systems, and propose a sparse modeling method based on dynamical models. This can be described as sparse optimization subject to differential equations, which can be solved via sparse optimal control theory that the presenter has recently proposed. We call this new sparse modeling "dynamical sparse modeling." The optimization can be efficiently solved by recent optimization algorithms such as alternating direction method of multipliers (ADMM), after discretization. The poster presentation shows the mathematical formulation of dynamical sparse modeling, efficient algorithms, and numerical simulation results.
16	Motoki Shiga	Gifu University	Nonnegative Matrix Factorization for Spectral Imaging Data Analysis	Advances in scanning transmission electron microscopy (STEM) techniques have enabled us to automatically obtain electron energy-loss (EELS)/energy-dispersive X-ray (EDX) spectral datasets from a specified region of interest (ROI) at an arbitrary step width, called spectral imaging (SI). Instead of manually identifying the potential constituent chemical components from the ROI and determining the chemical state of each spectral component from the SI data stored in a huge three-dimensional matrix, it is more effective and efficient to use a statistical approach for the automatic resolution and extraction of the underlying chemical components. Among many different statistical approaches, we adopt a non-negative matrix factorization (NMF) technique, mainly because of the natural assumption of non-negative values in the spectra and cardinalities of chemical components, which are always positive in actual data. We propose a new NMF model with two penalty terms: i) an automatic relevance determination (ARD) prior, which optimizes the number of components, and ii) a soft orthogonal constraint, which clearly resolves each spectrum component. Numerical experiments using real STEM-EDX/EELS SI datasets demonstrate that the ARD prior successfully identifies the correct number of physically meaningful components. The soft orthogonal constraint is also shown to be effective, particularly for STEM-EELS SI data, where neither the spatial nor spectral entries in the matrices are sparse.
17	Jun-ya Gotoh, Akiko Takeda, and Katsuya Tono	Chuo University, Japan;The University of Tokyo, Japan;The University of Tokyo, Japan	Proximal DC algorithm for Sparse Optimization	We propose a non-convex sparsity-inducing regularizer and an iterative algorithm we call proximal DC (Difference of Convex functions) algorithm, which repeatedly linearizes the non-convex terms at a subgradient. Our algorithm can be regarded as a generalized iterative shrinkage thresholding algorithm (ISTA) and converges linearly to a critical point from any initial point. Through experiments, we demonstrate that the proposed algorithm finds a better solution with desired sparsity in almost the same order of computational time compared to ISTA.
18	Takuma Kasai and Takanori Kigawa	RIKEN Quantitative Biology Center	NMR spectrum analysis with error detection mechanism	Nuclear magnetic resonance (NMR) is a useful method for protein analysis since properties of NMR signals; intensities, chemical shifts, and line widths, reflect their structural and dynamic properties. However, low signal-to-noise ratio often leads to wrong interpretation of the NMR spectra. Here we report that methods from information science, such as error detection, can be incorporated for reliable analyses of the NMR spectra. We previously developed a distinctive stable-isotope labeling strategy, SiCode (Stable isotope encoding), to determine amino-acid type of amide signals with a small number of labeled protein samples. In the SiCode strategy, information of amino-acid type is encoded to stable-isotope labeling ratios of the samples and decoded from signal intensities of the observed NMR spectra. We introduced an additional labeled sample to the SiCode corresponding to a check digit to certainly improve the noise tolerance through error detection in amino-acid typing especially under the condition of low signal-to-noise ratio.

19	Kazunori Iwamitsu^A^, Shingo Aihara^A^, Masato Okada^B,C^ and Ichiro Akai^A^	^A^Kumamoto University, Japan, ^B^The University of Tokyo, Japan, ^C^RIKEN, Japan.	Bayesian estimation of excitonic absorption spectra with the Metropolis algorithm	In our study, we analyzed the excitonic absorption spectra of Cu2O thin films sandwiched by MgO plates with the Metropolis algorithm of a Bayesian estimation. Consequently, we clarify probability distributions of all the spectral parameters only using one absorption spectrum.
20	Katsuaki Koike	Kyoto University	Spatial modeling of metal contents in a kuroko-type deposit with application to estimating	Demand for metal resources has largely increased because of constructing sustainable society and innovative technologies. Based on that background, this study is aimed to (i) develop a method for highly precise spatial modeling of metal contents in a metal deposit and (ii) clarify a physical law that formed a deposit. Matsumine mine, the largest kuroko deposit in northern Japan, is selected as a case study site. The metal content data of Cu, Zn, and Pb at 1457 points were used for the spatial analyses over a region of 500 m× 1000 m×300 m. High metal-content zones by geostatistics were revealed to be overlapped with the patchy kuroko zones and extend horizontally as connecting the zones. Assuming that the transport of ore solutions and the deposition of metals are approximated by an advective-diffusion spread phenomenon, the advective velocities and the diffusion coefficients were calculated from the metal content model. The result suggested the main paths of ore solutions in the distributions of silicious ore and rhyolite. Our next step is to improve the metal content estimation by combining a physical law of the deposit formation and kriging.
21	Hiromichi Nagao	The University of Tokyo	Seismic wave field imaging in the Tokyo metropolitan area based on lasso	Rapid prediction of damage due to large earthquakes on constructions through a numerical simulation would provide important information for making a decision relating to rescue and rehabilitation activities. Such a simulation requires ground motion input to each construction, which usually distributes much denser than seismometers. We propose a new method based on lasso for the purpose of data-driven imaging of seismic wave field in an urban area from seismograms obtained by a dense array. The coefficient matrix in a Taylor's expansion model is estimated by group lasso rather than the ordinary lasso in order to avoid the dependency of subjective coordinate settings in time and space. We have applied the proposed method to synthetic seismograms obtained from an analytic solution of seismic wave field assuming that an earthquake of M7 class occurs in a horizontally-layered underground. The resulting image reproduced better than those obtained by the ordinary least square, ridge regression or lasso. We have also applied to the actual MeSO-net data when the 2011 Earthquake off the Pacific coast of Tohoku occurred, and confirmed that group lasso reproduced the actual seismic wave field more accurately than the other methods.
22	Wataru Kurebayashi	Aomori University	Efficient estimation of Koopman operators from time series data	The mode decomposition of time series data based on the Koopman operator has been a hot topic in applied mathematics and fluid mechanics. We propose an efficient kernel-based method for estimating Koopman operators from time series data.
23	Makoto Uemura	Hiroshima University, Japan	Period analysis of variable stars using the group LASSO	The period analysis is a fundamental tool to study variable stars in which we estimate one or a few periodic signals in the time-series data of stellar brightness (so-called the "light curve"). In the case of ground-based astronomical observations, the sampling pattern of the light curve is so irregular that aliases occasionally disturb the detection of real signals. Here, we report the capability and application of our period analysis methods using the LASSO and group-LASSO for variable stars. They are definitely useful for the reconstruction of power spectra because we can assume them sparse when we know the objects are periodic variables. We demonstrate that, for periods longer than 1 day, the astronomical sampling pattern achieves a reconstruction quality similar to that by random sampling. For periods shorter than 1 day, however, the reconstruction quality is significantly degraded. Multi-longitude observations can improve it even if the data size is small for each observatory. We also introduce applications to real data, in which sharp signals in the LASSO periodogram emphasize a small, but significant period change in periodic humps in dwarf novae.
24	Motonobu Kanagawa	SOKENDAI (ISM), Japan	Convergence Rates of Quasi Monte Carlo Integration under Misspecified Assumptions	Quasi Monte Carlo (QMC) methods deterministically generate samples, thereby achieving numerical integration with convergence rates faster than Monte Carlo integration. These samples are generated based on smoothness assumptions on integrands. We discuss situations where such assumptions are violated. We show that QMC can be consistent even for such situations, and can achieve certain convergence rates.

25	Peng Hong	The University of Tokyo, Japan	Physical understanding of reflectance spectra of solar system small bodies based on data-driven approach	Large amounts of visible and near-infrared reflectance spectra data of solar system small bodies have been obtained by ground- and space-based telescopic observations and close-up observations by spacecraft in the past few decades. Small bodies, including asteroids, comets, dwarf planets and trans-Neptunian objects have been classified into several types, based on the shape of reflectance spectra. The physical meaning of reflectance spectra, however, is still poorly constrained because of the lack of forward models. In addition, different types of meteorite samples can show similar shape of reflectance spectra, suggesting the application limit of classification scheme based solely on reflectance spectra. In this study, we show the relationship between spectral type and geometric albedo mainly obtained by ground- and space-based observations, in order to separate similar reflectance spectra inherent in small bodies. Furthermore, we compare the observational results obtained by close-up observations with spectra-albedo relationship to extract physical meaning of reflectance spectra, such as the highest temperature that a small body has experienced.
26	Koujin Takeda, Toshiyuki Tanaka	Ibaraki University, Kyoto University	Application of partial parallel interference cancellation to sparse signal recovery	We propose a scheme for how to construct sparse signal recovery algorithms in compressed sensing under general measurement matrix, by applying the idea of partial parallel interference cancellation in CDMA demodulation to iterative soft thresholding. As a result, we arrive at a novel algorithm, which is equivalent to approximate message passing (AMP) under an appropriate assumption on the measurement matrix. This means that our novel algorithm is one of generalizations of AMP. In addition, via numerical experiment we verify that the proposed algorithm shows better convergence than the original AMP when the measurement matrix does not satisfy the assumption of AMP.
27	Shinsuke Uda	Kyushu University, Japan	Estimation method of sparse partial correlation matrix from omics data set with missing values	Omics data analysis is important to understand whole picture of living system. However, analytical methods of omics data do not fully developed owing to high-dimension-low-sample-size setting and missing values. In this study, we developed an estimation method of sparse partial correlation matrix from omics data set with missing values, taken together with data base of biological knowledge. The partial correlations between molecular species is estimated by sparse regression. In addition, prior knowledge of network structure stored in database such as the KEGG is used by tuning the hyper parameter of L1 norm regularization for each pathway. The missing values in omics data set are estimated simultaneously with the partial correlations by low rank approximation of data matrix.
28	Hajime Tamaki	Graduate School of Life Science, Hokkaido University, Japan	Application of compressed sensing to solid-state NMR measurements of heptahelical membrane protein	In nuclear magnetic resonance (NMR) spectroscopy, data sampling reduction by combination of compressed sensing (CS) and non-uniform sampling (NUS) is an attractive approach to obtain high-resolution spectra and gain sensitivity. This approach successfully introduced to solution NMR measurement of many types of multi-dimensional spectra. However, the effectivity of the approach is still unclear in solid-state NMR due to its broadened line width and low sensitivity, especially in the case of large size protein measurement. We applied NUS to a 13C-13C chemical shift correlation spectrum of a heptahelical membrane protein recorded by magic-angle sample spinning solid-state NMR and reconstructed by CS using L1-norm as a regularization term. Our results show that the spectral features include linearity of the signal intensity were reproduced by 30 % of a fully linearly sampled data set. Combination of NUS and CS is useful even the case of the heptahelical membrane protein measurements in solid-state.
29	Yasuko Sugase-Miyamoto1, Narihisa Matsumoto1, Kenji Kawano2, Masato Okada3	1AIST, Japan; 2Kyoto University, Japan; 3Univ Tokyo, Japan	Global category of faces and upright versus inverted category of faces were sparsely represented by neurons in monkey temporal visual cortex.	Face-responsive neurons in monkey temporal visual cortex differentiate both global category of faces (human vs. monkey faces) and their rotation in the picture-plane (upright vs. inverted presentations). To examine how neurons contribute to global categorization (GL), upright vs. inverted categorization of human faces (HUI), and upright vs. inverted categorization of monkey faces (MUI), we analyzed activities of 119 face-responsive neurons in area TE of two rhesus monkeys, performing a fixation task. The test stimuli were colored pictures of monkey and human faces, and their inverted pictures. The population activity vectors consisting of mean spike counts across trials to each stimulus were computed in 115-165 ms time window after stimulus onset. L1-regularized logistic regression revealed that the minimal number of neurons to categorize the GL, HUI, and MUI was 2, 2, and 10, respectively. The two neurons for the GL were different from the two for the HUI and from the 10 for the MUI. But the two neurons for the HUI were identical with two of the 10 neurons for the MUI. The results suggest that different members of the neuronal population contribute to categorize GL and HUI or MUI, and that HUI and MUI have contributing members in common.
30	Dai Yonebayashi, Peter Davis, Tatsuto Murayama	University of Toyama and Telecognix Corporation	Noisy Data Aggregation with Polar Codes	The quality of collective estimation of a source state involves a difficult tradeoff between sensing quality which sometimes requires a sufficient allocation of our total bandwidth for each of individual sensors, and aggregation quality which requires the use of as many devices as possible under the resource constraint. Therefore it makes sense to consider a strategy for optimal aggregation for an ensemble of independent noisy observation with constrained system capacity. In this poster presentation, we examine such a tradeoff by applying the ``polar codes,'' widely known as the state-of-the-art in the field of lossy source coding, and show some numerical results.

31	Chako Takahashi	Yamagata University	Mean-field approximation for Gaussian-Bernoulli restricted Boltzmann machine	A machine learning technique referred to as deep learning has been widely used in many research fields. Restricted Boltzmann machine (RBM), that is the Boltzmann machine defined on a bipartite graph, is one of the most fundamental components of it. Gaussian-Bernoulli restricted Boltzmann machine (GBRBM) is the extended model of RBM that allows us to treat continuous data. As with RBM, GBRBM consists of two different layers: a visible layer and a hidden layer. The visible layer consists of real-valued visible variables and the hidden layer consists of binary hidden variables. Many effective learning algorithms, such as the contrastive divergence, have been proposed. However, inference algorithms have not been much sophisticated. In this study, we consider inference algorithms in GBRBM on the basis of the naïve mean-field approximation. We derive two types of mean-field approximations: the one is the mean-field approximation for whole variables in GBRBM and the other is that for the marginal distribution of GBRBM. We compare the two different mean-field approximations qualitatively and quantitatively, and show that the latter one is better in both points of view.
32	Satoru Tokuda, Kenji Nagata, Masato Okada	The University of Tokyo, Japan	A theory of phase transitions and crossovers in statistical estimation: Toward a data-driven approach for physical science	Extracting hidden structures behind observed data is what natural sciences is all about, though the data are always limited in accuracy and amount. Any structures are not identified if an amount of data is not enough large. Some structures are identified if the amount is enough large. The qualitative change in statistical estimation like this cannot be explained by the asymptotic theory of statistical estimation. In this study, we develop a theory of the qualitative change in statistical estimation based on the fact that statistical estimation is mathematically equivalent to statistical physics. We introduce a novel function, which is called Bayesian specific heat, and explain the qualitative change as a crossover in statistical physics. We show that Bayesian specific heat asymptotically corresponds to real log canonical threshold, and explains phase transitions, which occur in infinite systems. We apply a finite size scaling to Bayesian specific heat and show that the scaling function explains crossovers. We take a regression problem, which is called Bayesian spectral deconvolution, as a case study and show the effectiveness of Bayesian specific heat. We discuss a necessary data amount and noise level from a viewpoint of our theory.
33	Hikaru Takenaka^A^, Kenji Nagata^A^, Takashi Mizokawa^B^, and Masato Okada^A^	^A^The University of Tokyo, Japan, ^B^Waseda University, Japan	Auto-Extraction of Effective Spin Hamiltonians from Electronic Structural Calculations by Bayesian Inference	We describe a novel method for extracting effective classical spin Hamiltonians automatically from mean-field type electronic structural calculations by means of Bayesian inference. The method is applied to a NiS_2 triangular lattice in NiGa_2S_4 with a spin disordered ground state. Starting from unrestricted Hartree-Fock calculations for the spin configurations of 16 Ni sites, we estimated that not only the strongest superexchange interaction between the third nearest neighbor sites but also those between the nearest and the second nearest neighbor sites should be taken into account to extract effective classical spin Hamiltonians for NiGa_2S_4. We also show results obtained from the above calculations with the Boltzmann factor. This method may enable mean-field type electronic structural calculations and magnetic experiments to be compared seamlessly because the Boltzmann factor takes into account the finite temperature effect. As a result, we estimated that the superexchange interaction between the nearest neighbor sites is ferromagnetic, which is consistent with magnetic experiment results. This supports the theory that the competition between the antiferromagnetic third neighbor interaction and the ferromagnetic nearest neighbor interaction may lead to a novel magnetic order, called the quantum spin liquid.
34	Yasuhiro Matsunaga	RIKEN AICS, Japan	Sequential data assimilation for single-molecule FRET photon-counting data	Data assimilation is a statistical method designed to improve the quality of numerical simulations in combination with real observations. We have developed a sequential data assimilation method that incorporates one-dimensional time-series data of single-molecule FRET (smFRET) photon-counting into conformational ensembles of biomolecules derived from "replicated" molecular dynamics (MD) simulations. A particle filter using a large number of replicated MD simulations with a likelihood function for smFRET photon-counting data is employed to screen the conformational ensembles that match the experimental data. In the poster, we show the details of our method as well as preliminary results obtained by an application to the smFRET data of protein folding dynamics.
35	Masashi Sato(1), Okito Yamashita(2,3), Masa-aki Sato(2), Yoichi Miyawaki(4)	1. Graduate School of Informatics and Engineering, The University of Electro-Communications, Japan, 2. ATR Neural Information Analysis Laboratories, Japan, 3. Brain Functional Imaging Technologies Group, CiNet, Japan	Information spreading of magnetoencephalography source localization and its effect on neural decoding.	Magnetoencephalography (MEG) is widely used as a tool to investigate dynamics of human neural information processing. For this purpose, source localization methods are used to estimate spatial patterns of the cortical current corresponding to recorded MEG signals. However, it remains unclear whether spatial patterns of the cortical current are estimated accurately. Here, we simulated MEG signals for artificial experimental conditions and applied neural decoding analysis to examine how accurately spatial patterns of the cortical current are estimated by several MEG source localization methods. All methods successfully localized cortical current with large amplitude in target areas where the current sources were assumed. However, the trained neural decoder predicted the experimental conditions from spatial patterns of the estimated cortical current in non-target areas as well as in the target areas. These results show information represented in the cortical activity patterns in the specific areas spread to other areas through MEG source localization. To avoid such false positiveness, we propose a method to evaluate the most informative cortical area for neural decoding. Acknowledgements: This study was partially supported by JSPS KAKENHI Grant Number 26120514.
36	Hirotaka Sakamoto^A^, Yoshinori Nakanishi-Ohno^A,B^, Masato Okada^A,C^	The University of Tokyo^A^, Research Fellow of JSPS^B^, RIKEN BSI^C^, Japan	Bayesian hyper-parameter estimation on Markov random field influenced by using image data of low quality	Recently, a lot of image data are acquired in various areas of natural science. Since these image data are typical of high-dimensional data, it is important to estimate the values of latent variables such as diffusion coefficients. We study a Bayesian framework to estimate the values of diffusion coefficients, represented by hyper-parameters in Markov random field models, from image data. This framework allows us to evaluate the reliability of the estimates as well. Our motivation is to examine how the estimated values and their reliability are influenced by methods used in data analysis and acquisition. In this study, we focus on the pre-processing of image averaging in the case of the measurement of low S/N ,and the use of down-sampled data. First, we show that the method of averaging has no effect on image restoration in terms of root mean square error but decreases the reliability of hyper-parameter estimation. Next, we show that down-sampled data bias the estimates of hyper-parameters. We discuss the relation between this bias and the dimension of image data using the concept of renormalization group developed in statistical mechanics.

37	Yuri Kokura, Yukinojo Kitakami, Takashi Ohnishi, Hideaki Haneishi	Chiba University, Japan	Enhanced Angiography Under Natural Respiration Using Robust Principal Component Analysis	In digital subtraction angiography (DSA) targeting thoracoabdominal organs, patients have to hold their breath in order to achieve the subtraction successfully. However, steady breath hold is sometimes difficult especially for elder patients. In this paper, we propose a blood vessel enhancement method with only consecutive digital angiographic images under the natural breathing. The method consists of two-steps. In the first step, we use a robust principal component analysis to separate the consecutive images into a low-rank (L) component and a sparse (S) component. The S component mainly corresponds to the pattern of the contrast agent. At this step, deformation between sequential images due to respiration is corrected using L component. In the second step, we generate a widely-spread blood vessel pattern from the sequence of motion-corrected S-component images. We decomposed many sets of angiographic images to L + S components and could observe satisfactory enhanced angiography images.
38	Yukinojo Kitakami, Takashi Ohnishi, Hideaki Haneishi	Chiba University,Japan	4D-MRI Reconstruction Using Low-Rank and Sparse Matrix Decomposition	We have previously developed a method for obtaining 4D-MRI to visualize and quantify the three-dimensional motion of the thoracoabdominal organ due to respiration. However, a data collection time of approximately 30 min is needed. Now, we propose to reduce the number of encoded samples in the k-space for time shortening and to apply a sparse model-based reconstruction algorithm for preserving image quality. The applied technique is a low-rank plus sparse matrix decomposition (L+S). We performed a simulation experiment where the encoded data were reduced to one-third of the full sampling. We compared the ideal image reconstructed with the full sampling data, the image reconstructed by the conventional method with the k-space data filling the missing space with zero, and the image reconstructed by the L+S technique. We confirmed that L+S technique can reduce the artifacts and noise and provide image quality similar to that of the ideal image.
39	Masamichi J. Miyama and Koji Hukushima	University of Tokyo, Japan	Sparse modeling approach for STM data analysis by using LARS-LASSO	We propse the new framework of the real-space data analysis of the STM (Scanning Tunneling Microscope) topography image, which enables us to determine the positions of atoms to the degree where we can extract the locally strain of the atom's configuration from the noisy datasets. In our analysis, we utilize the relevant vector machine (RVM) as a data model, a 2D Gaussian function as a base function that corresponds to a position of a single atom peak, and LASSO as the prior information. In our framework, it is necessary to choose the suitable coefficient of LASSO term for obtaining the desirable solution of the RVM in order to determine the positions of atoms. One solution for this procedure is a cross-validation technique, however it takes high cost when a model containes the large number of variables as in our case. In our poster presentation, we present the application of LARS for our problems and discuss the result of the optimization of our inference tequnique.
40	Satoshi Takabe	University of Tokyo, Japan	Statistical-mechanical analysis of Boolean compressed sensing	Detecting a small number of defects among items is a traditional but essential issue in statistical inference. Given the number of items N and defects K, $O(N)$ tests are naively required for detection. If some of items are tested at one by pooling them, however, the number of tests reduces to $O(K \log N)$. This technique is called the group testing and is applied to the DNA cloning library screening, for example. Recently, focusing on the sparsity of defects, Boolean compressed sensing (BCS) is proposed to improve detection performance of the group testing. In this presentation, we report its implementation based on the Max-Sum algorithm and analysis of its typical behavior averaged over random pools. Our analysis successfully estimates its detection ratio even in the case of a finite number of items. It also shows that our method based on BCS asymptotically needs less number of tests than other greedy methods; a half of the combinatorial matching pursuit and a quarter of the combinatorial basis pursuit.
41	Toshiya Takami	Kyushu University, Japan	Parallelization of Irregular Sparse Computations with One-sided Communications	Sparse linear algebra is one of the basic operations for managing large-scale data sets in physics and chemistry. We often encounter non-symmetric sparse matrices in actual applications. In this poster, we concentrated on parallel computations for those irregular sparse operations by the use of one-sided communications. Over many years, send/receive-type communications in Message Passing Interface (MPI) has been widely used for parallel implementations in scientific computations. Recently, MPI3 provides a proper implementation for one-sided communications, where we can use put/get-type operations without disturbing corresponding processes. We will analyze how non-symmetric sparse operations are effectively parallelized by the use of those one-sided operations.
42	Kazuki Kuramochi	The University of Tokyo /NAOJ, Japan	Imaging the silhouette of the Galactic supermassive black hole with the sparse modeling	At the heart of almost every major galaxy lurks a supermassive black hole, which is the most extreme object predicted by Einstein's General Theory of Relativity. The super-massive black hole Sgr A* residing the center of our Milkey way galaxy exhibits the largest apparent angular size in the Universe. Sgr A* provides an excellent opportunity to image the immediate vicinity of the event horizon with an Earth-size radio interferometer Event Horizon Telescope (EHT). Two important issues have remained for imaging the silhouette of Sgr A*. First, its size is comparable to the diffraction limit of EHT, requiring a super-resolution imaging. This issue can be resolved with LASSO, a technique of the sparse modeling, providing a good fidelity image even in a super-resolution regime (Honma et al. 2014). Another issue is the interstellar scattering effects caused by the tenuous interstellar plasma in our line of sight that scatters radio waves from Sgr A*. It blurs an interferometric image and reduces its effective resolution. We present realistic simulated observations of Sgr A* with EHT including the interstellar effects. We found that LASSO can overcome also the effect of interstellar scattering, making it promising to image the giant black hole in our galaxy.

43	Takeo Hoshi	Tottori University, Japan	Data problems that appeared in large-scale quantum material simulations on the K supercomputer	Unsettled data problems appeared in large-scale quantum material simulations mainly of organic device materials. The simulations were carried out on the K supercomputer by our novel simulator ELSES(http://www.elses.jp/) with original massively parallel numerical algorithms for upto 100-nm-scale systems. The present paper focuses on industrial organic materials and related ones. Since they are characterized by large-scale disordered pi-electron networks, the data analysis methods like sparse modeling are crucial so as to obtain material design principle for industrial products. Our preliminary results indicate proto-typical data problems on (i) sp2-sp3 nano-composite carbon solids and (ii) quantum (hole wavepacket) dynamics in organic device materials of polymers (1D-like system) and thin films (2D-like system). The first topic gives a static (time-independent) data problem for spatial domain analysis. The second one gives dynamical (time-dependent) data problems for electronic device performance (mobility). These topics require the collaboration between computational science and data science. As a future outlook, the softwares and the data should be integrated on supercomputers as a cloud-type service for academic and industrial researchers. The author thanks to Koji Hukushima (U. Tokyo) for ongoing discussions on the topics.
44	Akio Taniguchi	University of Tokyo	FMLO: a New HD^3 Spectroscopic Method for Radio Astronomy with Removing Correlated Noises	We develop a new observing method for (sub)millimeter spectroscopy in radio astronomy using high-dimensional time series data. We focus on the behavior of noises that they are fluctuated slowly in time (< 1 Hz) and correlated in spectrum (correlated noise). FMLO, our new method, obtains highly-sampled (10 Hz) time series spectra where astronomical signals are frequency-modulated (FM), which is realized by changing the frequency of local oscillator (LO) in heterodyne receiver. Correlated noises are therefore removed in time series spectra by principal component analysis (PCA). FMLO never requires off-point (referential) observations and thus dramatically improves efficiency of single-dish spectroscopy, especially faint & broad molecular line observations toward cosmological distant galaxies. In this poster we will present the principle and the data reduction method of FMLO. For the faster and more objective data reduction, we introduce an iterative algorithm and probabilistic PCA (PPCA), which is used for machine learning field. PPCA can correctly determine the number of principal component used to reconstruct the true correlated noise and thus model the astronomical signals. We will also show some observational results of FMLO implemented on the Nobeyama 45m telescope in Nagano, including optimization of FM patterns and FMLO-mapping, which will be more efficient on-the-fly observations.
45	Masahiro Haze^1^, Yoshinori Nakanishi-Ohno^2^, Yasuo Yoshida^2^, Koji Hukushima^3^, Masato Okada^2^ and Yukio Hasegawa^1^	ISSP, Univ. of Tokyo^1^, Grad. Sch. Front. Sci., Univ. Tokyo^2^, Grad. Sch. Arts. Sci., Univ. Tokyo^3^	Compressed sensing for efficient measurements of quasiparticle interference using scanning tunneling microscopy / spectroscopy	Scanning tunneling microscopy and spectroscopy (STM/S) enables us to observe quasi-particle interference (QPI) patterns, which are formed by scattering and interference of quasi-particles at surfaces. By taking the Fourier transformation of QPI patterns at various energies, we obtain the energy dispersion of the quasi-particles, which is related to the electronic band structure. This technique is widely used in a community of condensed matter physics to obtain the electronic structure of materials such as unconventional superconductors and topological insulators, However, QPI measurements are time-consuming because tunneling spectroscopy has to be performed over wide area in high spatial and energy resolutions to obtain information in sufficient momentum space (q-space). In order to improve the time-efficiency, we have used compressed sensing (CS); a signal-processing technique for the recovery of signals from a small number of data points. Our application of CS is based on the sparseness of Fourier-transformed QPI in q space. We applied a CS technique called least absolute shrinkage and selection operator (LASSO) to analyze an STM image of the Ag(111) surface whose Fourier-transformed QPI shows a circular pattern. Through numerical simulations, we revealed that LASSO performs well in reconstructing the circular pattern from reduced number of data points.
46	Masaaki Imaizumi	University of Tokyo	Nonparametric multivariate regression with tensor-product RKHS	We consider a nonparametric multivariable regression problem, and estimate a function in tensor-product reproducing kernel Hilbert space (RKHS). The multivariate regression function is represented by tensor-product of elements in several RKHSs, and the representer theorem enables us to solve the optimization problem with the regression function. We introduce elastic-net type regularization for the each element of the RKHSs, and investigate the convergence of the estimator. We show that our method is suitable for the variable selection problem, and it is also applicable for complex data analysis.
47	Kazuyuki Hara	Nihon University	Dropout acts as an ensemble learning	Dropout is used in the deep net act as a regularizer. However, the dropout updates the hidden units asynchronous way and this property may produce the diversity of the hidden units activities. We propose novel function of dropout to act as an ensemble learning.
48	Longyin Xu	Keio University, Japan	TimeTubes: Preliminary Design of Visualization Tool for Time-dependent, Multivariate Blazar Datasets	Blazars are active galactic nuclei whose relativistic jets ejected from the central black hole are pointing towards the earth. It is still a challenge to analyze time-dependent, multivariate datasets for the blazars with conventional visualization methods such as scatter plot matrices. This poster presents TimeTubes, a new visualization scheme that allows users to analyze dynamic changes and causality among representative time-varying dimensions. In our initial attempt to visualize blazar datasets provided by Hiroshima Astrophysical Science Center, we selected six dimensions from the originals, including polarization parameters (Q/I and U/I) and their corresponding errors, intensity and color. Our basic idea is to align ellipse-transformed and colorized snapshots of these dimensions in parallel, to form a tube in 3D space. The 2D coordinates of the center of each ellipse and its major and minor axes are given by (Q/I, U/I) and the error of Q/I and U/I, respectively. The resulting tube is then colorized by the observed intensities and colors of the blazar. We designed a designated interface with 10 feature functions to control the view of TimeTubes. In the presentation, the usability of TimeTubes along with feedback from experts from the center will also be discussed. **Co-authors: Masanori NAKAYAMA: Keio University, Japan, Hsiang-Yun WU: Keio University, Japan, Makoto UEMURA: Hiroshima University, Japan, Issei FUJISHIRO: Keio University, Japan.**

49	Yusuke Niibe	Keio University, Japan	Making many-to-many parallel coordinate plots scalable by asymmetric biclustering	Many-to-many parallel coordinate plots enable users to grasp at a time correlations between arbitrary pairs of dimensions in a given multi-dimensional dataset.However, the more pairs of dimensions we visualize, the smaller area for displaying each of the pairs becomes. This problem makes it difficult for the users to discover specific highly-correlated pairs in the datasets. This poster presents our attempt to make the many-to-many parallel coordinate plots scalable by biclustering the underlying dataset in terms of data samples and dimensions asymmetrically with different methods. The effectiveness of this approach is empirically proven through applications of the current prototype system to several practical datasets. **Co-authors: Hsiang-Yun Wu, Keio University, Japan, Kazuho Watanabe, Toyohashi University of Technology, Japan, Shigeo Takahashi, The University of Aizu, Japan, Issei Fujishiro, Keio University, Japan**
50	Noriko Katsumata	RIKEN BSI, Japan	Comparison of visual response properties of inferior temporal cortex (IT) in Awake and Anesthetized Macaque	Visual object representation in higher visual cortex, inferior temporal cortex (IT), has long been studied either with awake or anesthetized monkeys. Neural responses to objects in awake and anesthetized conditions are generally considered to have identical property. However, there are no studies directly comparing responses under these two conditions with the same neuronal populations. In addition, time courses of responses were not well studied in anesthetized animals while it has been reported that object responses in early and late time phases were different in awake animals. Here, we chronically implanted mul ti-electrode-array to IT and compared object responses in awake and anesthetized conditions. We found that object response tuning curves calculated from mean firing rates were significantly correlated in 82% of the sites (r=0.6±0.13). However, time-to-time analysis revealed that correlation was not consistent across entire time period. Responses with mean firing rates in anesthetized condition response was correlated particularly in early phase of responses in the awake condition. We will also discuss time courses of spatial pattern of responses in awake condition in relation to previously reported spatial patterns of responses with anesthetized animals.
51	Takenori Oida, Kazuhiro Tamiwa	Kyoto University, Japan	Sampling strategy of compressed sensing for improving the contrast of T1-weighted images in ultra-low field MRI	Ultra-low field MRI (ULF-MRI) is one of the modalities in bio-medical images, which has an advantage to obtain high contrast T1-wighted images. However, since it is difficult to utilize wide bandwidth, long readout time is required to obtain enough frequency resolution. On the other hand, compressed sensing MRI has been proposed by Lustig et. al. for rapid scan of high field MRI. In this study, we propose a sampling strategy of compressed sensing for improving the contrast of T1-weighted images in ULF-MRI. To evaluate our strategy, the reconstruction simulations with the fast composite splitting algorithms (FCSA) were carried out. The results of reconstruction simulations show that our sampling strategy is effective to reduce readout time. In ULF-MRI, since the readout time occupy the large part of the repetition time (TR), TR can be reduced by short readout time. Therefore, T1 contrast in ULF-MRI is able to be improved by our sampling strategy.
52	Toshiyuki TANAKA, Shashi PRABH, Yiyan LIU	Kyoto University, Shiv Nadar University	Mean-field hard-core models for throughput equalization in CSMA-based wireless networks	Hard-core models for wireless ad-hoc networks represent interference between users in terms of hard-core interactions. In mean-field hard-core models one ignores geometric aspects of networks such as locations of wireless nodes, which is justifiable in modeling scattering-rich environments. We study the problem of throughput equalization in mean-field hard-core models. A node with a large degree will suffer from more frequent interference from nodes neighboring in the conflict graph, so that it has to try transmission more frequently in order to equalize throughput with nodes with smaller degrees. We have obtained via a statistical-mechanics approach a simple scheme to determine frequencies of transmission attempts which achieves throughput equalization in conflict graphs in the large-system limit.
53	Chihiro Nakajima	Tohoku University	Compressed sensing and semi-supervised learning in material science	Two applications of informatics for material science, an analysis of rock-textures by image segmentation based on semi-supervised learning and 3D-reconstruction of atomic arrangement of a gold nano-cluster, are introduced.
54	Koji Fujimoto	New York University	Dynamic digital phantom for optimization of GRASP DCE-MRI reconstruction	A realistic digital phantom for optimization of Golden-Angle Radial Sparse Parallel (GRASP) Dynamic Contrast Enhanced (DCE) MRI of the upper abdomen was created using a set of existing dynamic MR images as a template. Anatomical Region of Interest (ROI)s were manually drawn for the aorta, liver, kidney, muscles, and the averaged signal change over time was measured. By using this phantom and simulated coil sensitivity maps, k-space data simulating golden-angle radial acquisition was generated. Images were reconstructed by enforcing temporal sparsity using different values for the regularization parameter (λ). Signal changes for each ROI were compared to determine the optimal λ value. Results obtained from the digital phantom were similar to the results using experimental GRASP data.