Journal of Statistical Mechanics:
Theory and Experiment

PAPER: DISORDERED SYSTEMS, CLASSICAL AND QUANTUM • OPEN ACCESS

Sparse approximation based on a random overcomplete basis

To cite this article: Yoshinori Nakanishi-Ohno *et al J. Stat. Mech.* (2016) 063302

OPEN ACCESS

Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

PAPER: Disordered systems, classical and quantum

Sparse approximation based on a random overcomplete basis

Yoshinori Nakanishi-Ohno[1,2], Tomoyuki Obuchi[3], Masato Okada[1] and Yoshiyuki Kabashima[3]

[1] Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba 277-8561, Japan
[2] Research Fellow of Japan Society for the Promotion of Science, Chiyoda, Tokyo 102-0083, Japan
[3] Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, Yokohama, Kanagawa 226-8502, Japan
E-mail: nakanishi@mns.k.u-tokyo.ac.jp, obuchi@sp.dis.titech.ac.jp, okada@k.u-tokyo.ac.jp and kaba@dis.titech.ac.jp

Received 28 July 2015
Accepted for publication 30 January 2016
Published 10 June 2016

Online at stacks.iop.org/JSTAT/2016/063302
doi:10.1088/1742-5468/2016/06/063302

Abstract. We discuss a strategy of sparse approximation that is based on the use of an overcomplete basis, and evaluate its performance when a random matrix is used as this basis. A small combination of basis vectors is chosen from a given overcomplete basis, according to a given compression rate, such that they compactly represent the target data with as small a distortion as possible. As a selection method, we study the ℓ_0- and ℓ_1-based methods, which employ the exhaustive search and ℓ_1-norm regularization techniques, respectively. The performance is assessed in terms of the trade-off relation between the distortion and the compression rate. First, we evaluate the performance analytically in the case that the methods are carried out ideally, using methods of statistical mechanics. The analytical result is then confirmed by performing numerical experiments on finite size systems, and extrapolating the results to the infinite-size limit. Our result clarifies the fact that the ℓ_0-based method greatly outperforms the ℓ_1-based one. An interesting outcome of our analysis is that any small value of distortion is achievable for any fixed compression rate r

J. Stat. Mech. (2016) 063302

in the large-size limit of the overcomplete basis, for both the ℓ_0- and ℓ_1-based methods. The difference between these two methods is manifested in the size of the overcomplete basis that is required in order to achieve the desired value for the distortion. As the desired distortion decreases, the required size grows in a polynomial and an exponential manners for the ℓ_0- and ℓ_1-based methods, respectively. Second, we examine the practical performances of two well-known algorithms, orthogonal matching pursuit and approximate message passing, when they are used to execute the ℓ_0- and ℓ_1-based methods, respectively. Our examination shows that orthogonal matching pursuit achieves a much better performance than the exact execution of the ℓ_1-based method, as well as approximate message passing. However, regarding the ℓ_0-based method, there is still room to design more effective greedy algorithms than orthogonal matching pursuit. Finally, we evaluate the performances of the algorithms when they are applied to image data compression.

Keywords: analysis of algorithms, heuristics, source and channel coding, statistical inference

Contents

J. Stat. Mech. (2016) 063302

1. Introduction

Information processing based on the sparseness of various data is an active area of research. This sparseness means that data are typically expressed by a small combination of non-zero components when a proper basis is used. The significance of sparseness for information processing had already begun to be noted when principal component analysis was invented, in 1901 [1]. Low-rank approximation of a matrix is known to be a useful method of collaborative filtering for recommendation systems [2–4]. In neuroscience, the sparse-coding hypothesis has gradually been accepted as a method of elucidating visual and auditory systems [5–10]. Recent interest in information processing with sparse data has been triggered by compressed sensing, since it was demonstrated that ℓ_1-norm minimization can give exact solutions in a reasonable time, under appropriate conditions [11–14].

In this study, we discuss sparse data processing from a different viewpoint, namely that of sparse approximation. This refers to a technique of approximately representing target data by a small number of non-zero elements, the purpose of which is to achieve a better trade-off relation between the distortion caused by the approximation and the compression rate [15–24]. We adopt a strategy of sparse approximation that utilizes an overcomplete basis (OCB), which is also termed a 'frame' in the field of signal processing. OCBs contain more basis vectors than the dimension of target data. This means that a better and smaller set of basis vectors may be chosen to compactly express the data. Therefore, in terms of the trade-off relation, the OCB-based strategy is expected to outperform naive strategies such as random projection.

For selecting basis vectors from an overcomplete basis, we discuss the ℓ_0- and ℓ_1-based methods, which employ the exhaustive search and ℓ_1-norm regularization techniques, respectively. Our adoption of these methods is motivated by their application in compressed sensing [25, 26]. Focusing on the trade-off relation, we evaluate the performance of sparse approximation from two different viewpoints. First, we theoretically analyze the ideal performance that is achieved when the ℓ_0- and ℓ_1-based methods are performed exactly, by using methods of statistical mechanics. We regard the distortion

J. Stat. Mech. (2016) 063302

and the compression rate as the thermal averages of physical quantities derived from partition functions. In the large-system limit, these are assessed by the replica method and the saddle-point method [27, 28]. In order to validate the results of our analysis, we extrapolate physical quantities in the limit, from finite-size results obtained using the exchange Monte Carlo method [29, 30] and quadratic programming. Second, we investigate the practical performance of the OCB-based strategy. We examine the performances of two well-known algorithms, orthogonal matching pursuit [31, 32] and approximate message passing [33], when they are employed to approximately execute the ℓ_0- and ℓ_1-based methods, respectively. We also apply the approximate algorithms to a task of image data compression and evaluate their performances, as a practical example.

The rest of this paper is organized as follows. In section 2, we set up the problem of sparse approximation that we will focus on, and explain the ℓ_0- and ℓ_1-based methods and related work. In section 3, we analyze the ideal performances of these methods, in terms of the trade-off relation. In section 4, we discuss the practical performance of the OCB-based strategy, and its application to image data. In section 5, we conclude this paper.

2. Problem setting

2.1. Sparse approximation using a random overcomplete basis

Given a data vector $\boldsymbol{y} \in \mathbb{R}^M$ and a compression rate r, the purpose of sparse approximation is to obtain a sparse representation $\boldsymbol{x} \in \mathbb{R}^N$ using a basis matrix $\boldsymbol{A} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N) \in \mathbb{R}^{M \times N}$, while keeping the distortion ϵ caused by the approximation as small as possible. The compression rate r is defined as the ratio of the number of non-zero components of $\boldsymbol{x}$ to the dimension of the data vector. That is,

$$r = \frac{\|\boldsymbol{x}\|_0}{M}, \tag{1}$$

where $\|\cdot\|_0$ denotes the so-called ℓ_0-norm of a vector. The ℓ_0-norm represents the number of non-zero elements of a vector, defined as $\|\boldsymbol{v}\|_0 = \sum_i |v_i|_0$, where $|v_i|_0$ is equal to 0 ($v_i = 0$) or 1 ($v_i \neq 0$). We measure the distortion using the mean squared error, as

$$\epsilon = \frac{1}{2M}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2, \tag{2}$$

where $\|\cdot\|_2$ is the ℓ_2-norm of a vector, defined as $\|\boldsymbol{v}\|_2 = \sqrt{\sum_i v_i^2}$. The distortion given by (2) indicates how similar the approxitame expression $\boldsymbol{A}\boldsymbol{x}$ is to the original data $\boldsymbol{y}$. Note that this is different from the reconstruction error, which is often used to measure the proximity between an original sparse signal $\boldsymbol{x}^0$ and an estimated sparse representation $\hat{\boldsymbol{x}}$ in research on compressed sensing. For our purpose of an analytical evaluation of ϵ, we consider the case where the elements of the data vector $\boldsymbol{y}$ are independently and identically distributed (i.i.d.) random variables from the normal distribution, whose mean and variance are 0 and σ_y^2, respectively, and together are denoted by $\mathcal{N}(0, \sigma_y^2)$.

J. Stat. Mech. (2016) 063302

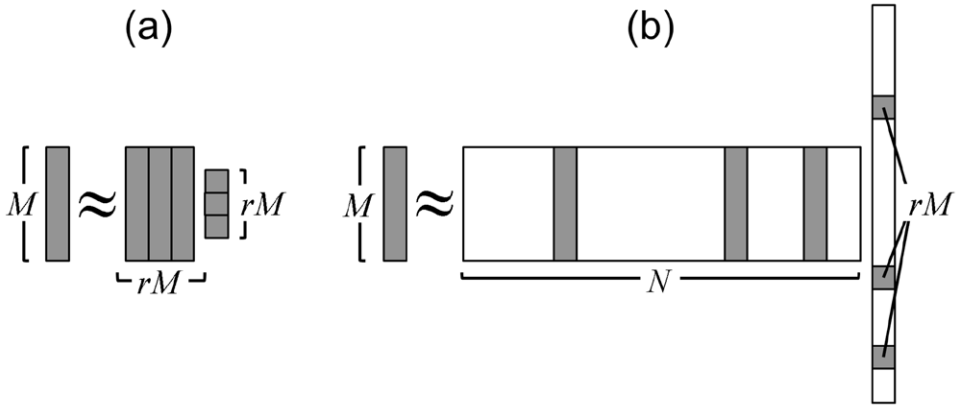

Figure 1. Schematic diagrams of sparse approximation. (a) Naive method. (b) OCB-based strategy.

The elements of the basis matrix $\boldsymbol{A}$ are also i.i.d. random variables from $\mathcal{N}(0, M^{-1})$. Then, the matrix $\boldsymbol{A}$ is almost surely of rank $\min(M, N)$, and the distortion becomes a random variable.

If $N = rM$, the minimization of (2) is nothing but the method of least squares (LS), and the corresponding compressed vector is easily obtained as

$$\hat{\boldsymbol{x}} = \boldsymbol{A}^{+}\boldsymbol{y}, \tag{3}$$

where $\boldsymbol{A}^{+}$ is the pseudoinverse of $\boldsymbol{A}$, given by

$$\boldsymbol{A}^{+} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}. \tag{4}$$

Let us call this the naive method, which is illustrated in figure 1(a). In the large-size limit $M \to \infty$, the corresponding distortion converges to

$$\epsilon_{\mathrm{naive}} = \frac{1-r}{2}\sigma_y^2, \tag{5}$$

with probability one. In general, in the limit $M \to \infty$ certain random variables, such as ϵ, have the so-called self-averaging property, and will almost surely converge to their average values. This enables us to present a clear discussion, and hereafter we focus on this limit.

On the other hand, for $N > rM$ we have a lot of options in choosing a combination of rM basis vectors from the matrix, as illustrated in figure 1(b). If the chosen combination is more suitable for representing the data vector than one that is chosen randomly, then the distortion becomes smaller than $\epsilon_{\mathrm{naive}}$. This is the idea behind the OCB-based strategy. However, this strategy presents the problem of how to choose the combination of basis vectors. As representative solutions for this problem, we focus on the ℓ_0- and ℓ_1-based methods below.

2.2. Methods

2.2.1. ℓ_0-based method. The basic idea of the ℓ_0-based method is to minimize the distortion by choosing the best combination of rM basis (column) vectors from a given

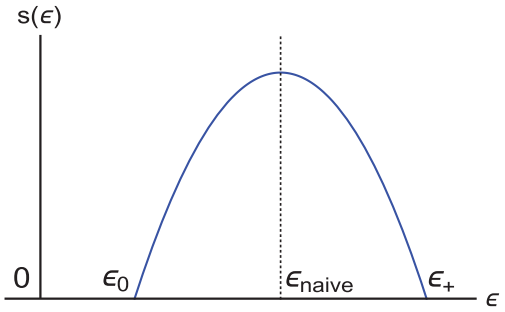

Figure 2. A schematic shape of the entropy function. The smaller zero point of the entropy, ϵ_0, corresponds to the minimum of the distortion connected to the best combination of the basis vectors. The point giving the largest entropy value accords to $\epsilon_{\rm naive}$ because the typical combinations of rM basis vectors chosen from a typical sample of $M \times N$ random matrices constitute typical samples of $M \times rM$ random matrices.

OCB. More generally, we would like to define the distortion as a function of the chosen combination of basis vectors, and to control it in a simple manner. This motivates us to introduce a binary vector $\boldsymbol{c} \in \{1, 0\}^N$, to store information on whether each basis vector is chosen ($c_i = 1$) or not ($c_i = 0$). We also introduce a distortion, labelled by $\boldsymbol{c}$, with

$$\epsilon(\boldsymbol{c}|\boldsymbol{y}, \boldsymbol{A}) = \min_{\boldsymbol{x}} \left\{ \frac{1}{2M} \|\boldsymbol{y} - \boldsymbol{A}(\boldsymbol{c} \circ \boldsymbol{x})\|_2^2 \right\}, \tag{6}$$

where $\circ$ is the Hadamard product of two vectors, defined as $(\boldsymbol{v} \circ \boldsymbol{w})_i = v_i w_i$. In addition, we define an entropy function $s(\epsilon|\boldsymbol{y}, \boldsymbol{A})$ to represent the number of configurations $\boldsymbol{c}$ that give a value of ϵ for the distortion, as follows:

$$s(\epsilon|\boldsymbol{y}, \boldsymbol{A}) = \frac{1}{M} \ln(\#\{\boldsymbol{c} \mid \|\boldsymbol{c}\|_0 = rM \wedge \epsilon(\boldsymbol{c}|\boldsymbol{y}, \boldsymbol{A}) = \epsilon\}), \tag{7}$$

where # denotes the number of elements of the following set.

This entropy function is expected to be analytic and convex upward with respect to ϵ, and cannot be negative, by definition. A typical shape of the entropy is depicted in figure 2. There are two zero points in the entropy function and the smaller and larger ones are denoted by ϵ_0 and ϵ_+, respectively. The smaller zero point ϵ_0 of the entropy function, $s(\epsilon_0) = 0$, gives the minimum value of the distortion

$$\epsilon_0(\boldsymbol{y}, \boldsymbol{A}) = \min_{\boldsymbol{c}} \epsilon(\boldsymbol{c}|\boldsymbol{y}, \boldsymbol{A}) \quad \text{subj. to } \|\boldsymbol{c}\|_0 = rM. \tag{8}$$

Hence, our original motivation for introducing the ℓ_0-based method, to find the minimum distortion led by the best combination of basis vectors, can be achieved through the evaluation of the entropy function. In addition, the evaluation of the entropy function is easier than the direct evaluation of ϵ_0, and moreover the entropy function provides more information about the space of the variables $\boldsymbol{c}$, which can be useful for practical applications such as designing algorithms. Thus, the entropy function $s(\epsilon)$ is the primary object of our analysis in the ℓ_0-based method. A similar analysis has been proposed for examining the weight space structure of multilayer perceptrons [34].

J. Stat. Mech. (2016) 063302

2.2.2. ℓ_1-based method. The ℓ_0-based method is the most closely matched to the original idea of the OCB-based strategy. However, its algorithmic realization of searching combinations of basis vectors is computationally inefficient, because it requires an exponentially growing computational cost as the system size N increases. In practical situations, instead of the ℓ_0-based method, a method based on ℓ_1-norm regularization can be employed. This motivates us to examine the following ℓ_1-based method.

Our ℓ_1-based method arises from the following minimization problem:

$$\hat{\boldsymbol{\xi}} = \arg\min_{\boldsymbol{\xi}} \left\{ \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{\xi}\|_2^2 + \lambda\|\boldsymbol{\xi}\|_1 \right\}, \tag{9}$$

where $\|\cdot\|_1$ is the ℓ_1-norm of a vector, defined as $\|\boldsymbol{v}\|_1 = \sum_i |v_i|$, with the absolute value denoted by $|\cdot|$. The solution of this minimization problem, $\hat{\boldsymbol{\xi}}$, provides useful information for finding the sparse vector we desire. This minimization problem is equivalent to the least absolute shrinkage and selection operator, also known as LASSO [35]. The main benefit of this approach represented by (9) is the computational ease of performing the minimization. As the objective function of (9) is convex, its minimization can be exactly carried out with a computational time in $O(N^3)$, using versatile algorithms of quadratic programming. Furthermore, the ℓ_1-norm term in (9) results in a sparsifying effect in $\hat{\boldsymbol{\xi}}$, and its coefficient λ is adjusted according to the compression rate. Namely, λ is chosen so that $\|\hat{\boldsymbol{\xi}}\|_0 = rM$.

Our aim in the analysis in the ℓ_1-case is to evaluate the distortion resulting from $\hat{\boldsymbol{\xi}}$. The expression of the distortion is given by

$$\epsilon_1 = \frac{1}{2M}\|\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{\xi}}\|_2^2. \tag{10}$$

An inconvenience presented by this distortion is that it is not minimized on the set of basis vectors chosen by $\hat{\boldsymbol{\xi}}$, owing to the presence of the ℓ_1-norm term. In order to remove this extra distortion, we determine again the values of the non-zero components by purely minimizing the distortion, after the support estimation of the sparse vector by the ℓ_1-norm regularization. This procedure is described as follows:

$$\epsilon_1^{\mathrm{LS}} = \min_{\boldsymbol{x}} \left\{ \frac{1}{2M}\|\boldsymbol{y} - \boldsymbol{A}(|\hat{\boldsymbol{\xi}}|_0 \circ \boldsymbol{x})\|_2^2 \right\}, \tag{11}$$

where $|\cdot|_0$ of a vector is defined by $(|\boldsymbol{v}|_0)_i = |v_i|_0$. This can be carried out by the method of LS for the sub-matrix of $\boldsymbol{A}$ that is composed of columns corresponding to $|\hat{\xi}_i|_0 = 1$. These two quantities, ϵ_1 and ϵ_1^{LS}, are the objects of our analysis in the ℓ_1 case.

2.3. Related work

Sparse approximation, which is also known as N-term approximation [17, 18, 36], has been studied widely in the fields of signal processing, statistics and information theory. To our best knowledge, the problem of approximating $\boldsymbol{y}$ by a linear combination of a small number of column vectors of $\boldsymbol{A}$ was first addressed by [15], in which finding the optimal support (combination of the columns) was shown to be NP-hard. Since then,

much effort has been made for analyzing properties of approximate algorithms. The approximation performance of a greedy algorithm termed orthogonal matching pursuit (OMP) [31, 32] has been examined actively by using various bounding techniques [21–23, 37, 38]. A sufficient condition that guarantees OMP to successfully find the optimal support was also provided in [23]. Reference [24] examined the condition for correctly recovering the true support for the ℓ_1-based method and matching pursuit algorithms in the case where $\boldsymbol{y}$ is generated as $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}^0 + \boldsymbol{n}$ from a true sparse vector $\boldsymbol{x}^0 \in \mathbb{R}^N$ and a noise vector $\boldsymbol{n} \in \mathbb{R}^M$.

All of the above studies offer mathematically rigorous sufficient conditions for guaranteeing a certain desired performance. However, empirically, such sufficient conditions are often overly cautious for explaining results of experiments, and therefore, another approach is necessary for accurately clarifying the abilities and limitations of sparse approximation. The aim of this paper is to accomplish this task in the large system limit utilizing methods of statistical mechanics while the mathematical guarantee of the obtained results is suspended.

3. Analysis of ideal performance

3.1. Analytical treatment in the limit $M \to \infty$

We investigate the limit $M \to \infty$, as stated above. For this purpose, we employ some statistical mechanical tools, which provide useful assistance investigating this limit. According to the terminology of statistical mechanics, we call the limit $M \to \infty$ the thermodynamic limit, and the average over $\boldsymbol{y}$ and $\boldsymbol{A}$ the configurational average, which is denoted by $[\cdot]_{\boldsymbol{y},\boldsymbol{A}}$. In taking the limit $M \to \infty$, the aspect ratio of the basis matrix, $\alpha = M/N$, is fixed.

3.1.1. ℓ_0-based method. A versatile technique of statistical mechanics is to introduce a generating function Z of an energy function $\mathcal{H}$, called a partition function. This defines a canonical distribution p. In the ℓ_0 case, we define the energy function, partition function, and canonical distribution respectively as follows:

$$\mathcal{H}_0(\boldsymbol{c}; \beta \,|\, \boldsymbol{y}, \boldsymbol{A}) = -\frac{1}{\beta} \ln \int \mathrm{d}_{\boldsymbol{c}}\boldsymbol{x} \mathrm{e}^{-\frac{\beta}{2}\|\boldsymbol{y}-\boldsymbol{A}(\boldsymbol{c}\circ\boldsymbol{x})\|_2^2}, \tag{12}$$

$$Z_0(\mu, \beta \,|\, \boldsymbol{y}, \boldsymbol{A}) = \sum_{\boldsymbol{c}} \delta(Mr - \|\boldsymbol{c}\|_0) \mathrm{e}^{-\mu\mathcal{H}_0(\boldsymbol{c};\beta|\boldsymbol{y},\boldsymbol{A})} \equiv \operatorname*{Tr}_{\boldsymbol{c}} \mathrm{e}^{-\mu\mathcal{H}_0(\boldsymbol{c};\beta|\boldsymbol{y},\boldsymbol{A})}, \tag{13}$$

$$p_0(\boldsymbol{c}; \mu, \beta \,|\, \boldsymbol{y}, \boldsymbol{A}) = \frac{1}{Z_0(\mu, \beta \,|\, \boldsymbol{y}, \boldsymbol{A})} \delta(Mr - \|\boldsymbol{c}\|_0) \mathrm{e}^{-\mu\mathcal{H}_0(\boldsymbol{c},\beta|\boldsymbol{y},\boldsymbol{A})}, \tag{14}$$

where $\int \mathrm{d}_{\boldsymbol{c}} x_i$ is equal to $\int \mathrm{d}x_i$ ($c_i = 1$) or 1 ($c_i = 0$).

The energy function is related to the distortion of a given basis-vector choice $\boldsymbol{c}$ as follows:

J. Stat. Mech. (2016) 063302

$$\frac{1}{M}\lim_{\beta\to\infty}\mathcal{H}_0(\boldsymbol{c};\beta|\boldsymbol{y},\boldsymbol{A})=\epsilon(\boldsymbol{c}\,|\boldsymbol{y},\boldsymbol{A}). \tag{15}$$

In addition, (7) means that the number of $\boldsymbol{c}$'s that provide $\epsilon(\boldsymbol{c}|\boldsymbol{y},\boldsymbol{A})=\epsilon$ is given as $\exp(Ms(\epsilon|\boldsymbol{y},\boldsymbol{A}))$. In the limit of $\beta\to\infty$, these provide us with another expression of (13) as

$$\lim_{\beta\to\infty}Z_0(\mu,\beta|\boldsymbol{y},\boldsymbol{A})=\int \mathrm{d}\epsilon\exp(M(s(\epsilon|\boldsymbol{y},\boldsymbol{A})-\mu\epsilon)), \tag{16}$$

which, in conjunction with employment of the saddle point evaluation for the integration with respect to ϵ, leads to a formula

$$\lim_{\beta\to\infty}\frac{1}{M}\ln Z_0(\mu,\beta|\boldsymbol{y},\boldsymbol{A})=\max_{\epsilon_0\leqslant\epsilon\leqslant\epsilon_+}\{s(\epsilon|\boldsymbol{y},\boldsymbol{A})-\mu\epsilon\}\equiv\phi_0(\mu|\boldsymbol{y},\boldsymbol{A}), \tag{17}$$

where $\phi_0(\mu|\boldsymbol{y},\boldsymbol{A})$ plays the role of the cumulant generating function of ϵ. Note that the maximization problem of (17) that originates from the saddle point assessment of (16) must be solved on the well-defined region of s, which requires appropriate bounds the minimum value of distortion ϵ_0 and the maximum value of distortion ϵ_+. Overall, we can calculate the object of our analysis, $s(\epsilon)$, through the inverse Legendre transformation, once we have obtained ϕ_0. Therefore, we turn our attention to the calculation of ϕ_0.

The cumulant-generating function has the self-averaging property, as does the entropy, and we assess the configurational average, given by

$$\phi_0(\mu)=[\phi_0(\mu|\boldsymbol{y},\boldsymbol{A})]_{\boldsymbol{y},\boldsymbol{A}}. \tag{18}$$

We employ the replica method in order to calculate this average, and a detailed analysis is provided in appendix A. Under the replica symmetric (RS) ansatz, which may not be appropriate, $\phi_0(\mu)$ is evaluated as

$$\begin{aligned}\phi_0(\mu)=\operatorname*{extr}_{\hat{\Theta}_0}\Bigg\{&\frac{1}{2}\ln\frac{1+\chi}{1+\chi+\mu(Q-q)}-\frac{1}{2}\frac{\mu(q+\sigma_y^2)}{1+\chi+\mu(Q-q)}\\&+\frac{1}{2}\left(\hat{r}r+\hat{Q}Q-\frac{\hat{\chi}}{\mu}\chi+\hat{q}q\right)+\frac{1}{\alpha}\int \mathrm{D}z\ln(1+Y)\Bigg\},\end{aligned} \tag{19}$$

where $\operatorname{extr}_{\Theta}\{\cdot\}$ denotes the operation of extremization with respect to Θ, $\hat{\Theta}_0=\{Q,\chi,q,\hat{r},\hat{Q},\hat{\chi},\hat{q}\}$, and $\int \mathrm{D}z=\int\frac{\mathrm{d}z}{\sqrt{2\pi}}\mathrm{e}^{-\frac{z^2}{2}}$, and we set

$$Y\equiv\sqrt{\frac{\hat{\chi}+\hat{Q}}{\hat{Q}+\hat{q}}}\,\mathrm{e}^{-\frac{1}{2}\hat{r}+\frac{1}{2}\frac{\hat{q}}{\hat{Q}+\hat{q}}z^2}. \tag{20}$$

By applying the extremization condition, we obtain the following equations of state (EOSs):

$$\hat{\chi}=\mu^2\left\{\frac{\Delta}{(1+\chi)(1+\chi+\mu\Delta)}+\frac{\sigma_y^2+q}{(1+\chi+\mu\Delta)^2}\right\}, \tag{21a}$$

J. Stat. Mech. (2016) 063302

$$\hat{Q} = \mu\left\{\frac{1}{1+\chi+\mu\Delta} - \frac{\mu(\sigma_y^2+q)}{(1+\chi+\mu\Delta)^2}\right\}, \tag{21b}$$

$$\hat{q} = \mu^2\frac{\sigma_y^2+q}{(1+\chi+\mu\Delta)^2}, \tag{21c}$$

$$r = \frac{1}{\alpha}\int \mathrm{D}z\frac{Y}{1+Y}, \tag{21d}$$

$$\chi = \frac{\mu r}{\hat{\chi}+\hat{Q}}, \tag{21e}$$

$$Q = r\frac{\hat{\chi}-\hat{q}}{(\hat{\chi}+\hat{Q})(\hat{Q}+\hat{q})} + \frac{1}{\alpha}\frac{\hat{q}}{(\hat{Q}+\hat{q})^2}\int \mathrm{D}z\, z^2\frac{Y}{1+Y}, \tag{21f}$$

$$q = \frac{1}{\alpha}\frac{\hat{q}}{(\hat{Q}+\hat{q})^2}\int \mathrm{D}z\, z^2\left(\frac{Y}{1+Y}\right)^2. \tag{21g}$$

where we write $\Delta = Q - q$. From the EOSs, we obtain some simple and general relations, which we summarize here for later convenience:

$$\hat{\chi}+\hat{Q} = \frac{\mu}{1+\chi}, \tag{22a}$$

$$\hat{Q}+\hat{q} = \frac{\mu}{1+\chi+\mu\Delta}, \tag{22b}$$

$$\hat{\chi}-\hat{q} = \frac{\mu^2\Delta}{(1+\chi)(1+\chi+\mu\Delta)}, \tag{22c}$$

$$\chi = \frac{r}{1-r}. \tag{22d}$$

The relation involving the entropy, (17), enables us to employ a convenient parametric form of $\epsilon(\mu)$ and $s(\mu) = s(\epsilon(\mu))$, and (21) and (22) allow us to simplify $\epsilon(\mu)$, as

$$\epsilon(\mu) = -\frac{\partial\phi_0(\mu)}{\partial\mu} = \frac{\hat{\chi}}{2\mu^2}, \tag{23}$$

$$s(\mu) = \phi_0(\mu) + \mu\epsilon(\mu). \tag{24}$$

The explicit form of $s(\mu)$ is not enlightening, and therefore we omit it. As the value of μ is increased from $\mu = 0$, the point of (ϵ, s) moves along the entropy curve from the maximum point ($\mu = 0$) in the direction of decreasing the distortion ($\mu > 0$) as shown in

J. Stat. Mech. (2016) 063302

figure 2. When the entropy curve crosses the zero-entropy line at $\mu = \mu_0$, the minimum distortion is given by

$$\epsilon_0 = \epsilon(\mu_0). \tag{25}$$

Here, we make a technical remark on the derivation of (19). In contrast to the usual prescription of the replica method, we require two different replica numbers for the present analysis, because we have two different integration variables, $\boldsymbol{x}$ and $\boldsymbol{c}$, in the calculation of ϕ_0. Using (15) and (18), and introducing a variable $\nu = \mu/\beta$, we can rewrite $\phi_0(\mu)$ as

$$\begin{aligned}\phi_0(\mu) &= \lim_{\nu\to 0}\frac{1}{M}\left[\ln \operatorname*{Tr}_{\boldsymbol{c}}\left(\int \mathrm{d}_{\boldsymbol{c}}\boldsymbol{x}\, \mathrm{e}^{-\frac{1}{2}\frac{\mu}{\nu}\|\boldsymbol{y}-\boldsymbol{A}(\boldsymbol{c}\circ\boldsymbol{x})\|_2^2}\right)^{\nu}\right]_{\boldsymbol{y},\boldsymbol{A}} \\ &= \lim_{n\to 0}\lim_{\nu\to 0}\frac{1}{Mn}\ln\left[\left\{\operatorname*{Tr}_{\boldsymbol{c}}\left(\int \mathrm{d}_{\boldsymbol{c}}\boldsymbol{x}\, \mathrm{e}^{-\frac{1}{2}\frac{\mu}{\nu}\|\boldsymbol{y}-\boldsymbol{A}(\boldsymbol{c}\circ\boldsymbol{x})\|_2^2}\right)^{\nu}\right\}^{n}\right]_{\boldsymbol{y},\boldsymbol{A}}.\end{aligned} \tag{26}$$

In the last line, we use the replica identity $[\ln X]_{\boldsymbol{y},\boldsymbol{A}} = \lim_{n\to 0}(1/n)\ln[X^n]_{\boldsymbol{y},\boldsymbol{A}}$. We identify n and ν as the two replica numbers, and assume that they are natural numbers, which enables us to expand the powers and to calculate the configurational average. The remaining calculations follow the usual procedure of the replica method, and we assume the RS ansatz in the order parameters.

Our present framework in calculating ϕ_0 is actually similar to the one-step replica-symmetry-breaking (1RSB) ansatz. In this identification, ν is identified as the 1RSB breaking parameter (usually written as m), and each configuration of $\boldsymbol{c}$ corresponds to a pure state in the 1RSB free-energy landscape; the entropy can be regarded as complexity. The analytical results obtained on the basis of RS assumption will be justified later, in a comparison with numerical calculations.

3.1.2. ℓ_1-based method.

Derivation of ϵ_1. Similarly to the case of the ℓ_0-based method, the energy function, partition function, and canonical distribution of the ℓ_1 case are defined respectively as

$$\mathcal{H}_1(\boldsymbol{\xi}|\boldsymbol{y},\boldsymbol{A}) = \frac{1}{2}\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{\xi}\|_2^2 + \lambda\|\boldsymbol{\xi}\|_1, \tag{27}$$

$$Z_1(\mu,\kappa|\boldsymbol{y},\boldsymbol{A}) = \int \mathrm{d}\boldsymbol{\xi}\ \mathrm{e}^{-\mu(\mathcal{H}_1(\boldsymbol{\xi}|\boldsymbol{y},\boldsymbol{A})+\kappa\|\boldsymbol{\xi}\|_0)}, \tag{28}$$

$$p_1(\boldsymbol{\xi};\mu,\kappa|\boldsymbol{y},\boldsymbol{A}) = \frac{1}{Z_1(\mu,\kappa|\boldsymbol{y},\boldsymbol{A})}\mathrm{e}^{-\mu(\mathcal{H}_1(\boldsymbol{\xi}|\boldsymbol{y},\boldsymbol{A})+\kappa\|\boldsymbol{\xi}\|_0)}. \tag{29}$$

The parameter κ is introduced for technical convenience in evaluating the compression rate r, and the limit of $\kappa\to 0$ is taken in the end. The energy function $\mathcal{H}_1$ is exactly the object for the minimization in (9). We also introduce the averaged free-energy density, given by

$$f_1(\mu,\kappa) = -\frac{1}{M\mu}[\ln Z_1(\mu,\kappa|\boldsymbol{y},\boldsymbol{A})]_{\boldsymbol{y},\boldsymbol{A}}, \tag{30}$$

J. Stat. Mech. (2016) 063302

which plays the role of the cumulant-generating function that is given by ϕ_0 in the ℓ_0 case. In the limit $\mu \to \infty$, the minimizer of the energy function becomes dominant in p_1, and we focus on this limit. Any quantity of interest can be calculated from f_1. For example, the compression rate r and the distortion ϵ_1 are calculated as

$$r = \lim_{\mu\to\infty}\lim_{\kappa\to 0}\frac{\partial}{\partial\kappa}f_1(\mu,\kappa), \tag{31}$$

$$\epsilon_1 = \lim_{\mu\to\infty}\left(1+\mu\frac{\partial}{\partial\mu}-\lambda\frac{\partial}{\partial\lambda}\right)f_1(\mu,0). \tag{32}$$

An analytically compact form of f_1 is assessed by using the replica method in the limit $M\to\infty$, through the replica identity, as

$$f_1(\mu,\kappa) = -\lim_{n\to 0}\frac{1}{M\mu n}\ln[Z_1^n(\mu,\kappa|\boldsymbol{y},\boldsymbol{A})]_{\boldsymbol{y},\boldsymbol{A}}\,. \tag{33}$$

As in the ℓ_0 case, we assume the replica-symmetric solution. The details of the necessary calculations are presented in appendix B. The result is given by

$$f_1(\mu\to\infty,\kappa) = \operatorname*{extr}_{\hat{\Theta}_1}\left\{\frac{1}{2}\frac{P+\sigma_y^2}{1+\chi_p}-\frac{1}{2}(\hat{P}P-\hat{\chi}_p\chi_p)-\frac{\hat{\chi}_p}{2\alpha\hat{P}}\left((1+2\theta_+\theta_-)\mathrm{erfc}(\theta_+)-\theta_-\frac{2}{\sqrt{\pi}}\mathrm{e}^{-\theta_+^2}\right)\right\}, \tag{34}$$

where $\hat{\Theta}_1 = \{P,\chi_p,\hat{P},\hat{\chi}_p\}$, $\theta_\pm = \frac{\lambda\pm\sqrt{2\kappa\hat{P}}}{\sqrt{2\hat{\chi}_p}}$, and $\mathrm{erfc}(\cdot)$ is the complementary error function, defined as $\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty \mathrm{d}t\mathrm{e}^{-t^2}$. The extremization condition gives the following EOSs for the present case:

$$\hat{\chi}_p = \frac{P+\sigma_y^2}{(1+\chi_p)^2}, \tag{35a}$$

$$\hat{P} = \frac{1}{1+\chi_p}, \tag{35b}$$

$$\chi_p = \frac{1}{\alpha\hat{P}}\left(\mathrm{erfc}(\theta_+)+\sqrt{\frac{\kappa\hat{P}}{\hat{\chi}_p}}\frac{2}{\sqrt{\pi}}\mathrm{e}^{-\theta_+^2}\right), \tag{35c}$$

$$P = \frac{\hat{\chi}_p}{\alpha\hat{P}^2}\left((1+2\theta_+\theta_-)\mathrm{erfc}(\theta_+)-\theta_-\frac{2}{\sqrt{\pi}}\mathrm{e}^{-\theta_+^2}\right)+\frac{\kappa}{\alpha\hat{P}}\mathrm{erfc}(\theta_+). \tag{35d}$$

By using (31) and (32), we obtain

$$r = \frac{1}{\alpha}\mathrm{erfc}(\theta) \tag{36}$$

J. Stat. Mech. (2016) 063302

$$\epsilon_1 = \frac{1}{2}\frac{P+\sigma_y^2}{1+\chi_p} - \frac{1}{2}(\hat{P}P - \hat{\chi}_p\chi_p) - \frac{\hat{\chi}_p}{2\alpha\hat{P}}\left((1-2\theta^2)\mathrm{erfc}(\theta) + \theta\frac{2}{\sqrt{\pi}}\mathrm{e}^{-\theta^2}\right), \tag{37}$$

where $\theta = \frac{\lambda}{\sqrt{2\hat{\chi}_p}}$. In addition, a simple formula

$$\epsilon_1 = \frac{1}{2}\hat{\chi}_p. \tag{38}$$

is derived from the EOSs of (35) in the limit of $\kappa \to 0$, and a useful relation

$$\chi_p = \frac{r}{1-r}, \tag{39}$$

which is similar to (22*d*), is offered by (35) and (36).

Derivation of ϵ_1^{LS}. We also evaluate ϵ_1^{LS}, as defined in (11). The computations are rather technical, and there we defer the details to appendix B. Here, we present an outline of the analysis, and the result.

Again, we use the energy function defined in the ℓ_0 case, but here the argument is $|\boldsymbol{\xi}|_0$, determined by $p_1(\boldsymbol{\xi})$. Thus, we obtain

$$\mathcal{H}_0(|\boldsymbol{\xi}|_0;\beta|\boldsymbol{y},\boldsymbol{A}) = -\frac{1}{\beta}\ln\int \mathrm{d}_{|\boldsymbol{\xi}|_0}\boldsymbol{x}\,\mathrm{e}^{-\frac{\beta}{2}\|\boldsymbol{y}-\boldsymbol{A}(|\boldsymbol{\xi}|_0\circ\boldsymbol{x})\|_2^2}. \tag{40}$$

Since the vector $\boldsymbol{\xi}$ is drawn from p_1, we calculate the average value of $(1/M)\mathcal{H}_0(|\boldsymbol{\xi}|_0)$ over p_1, in addition to the configurational average. Taking the limits of $\mu \to \infty$ and then $\beta \to \infty$ afterward, we obtain the desired distortion ϵ_1^{LS} as follows:

$$\epsilon_1^{\mathrm{LS}} = \lim_{\beta\to\infty}\lim_{\mu\to\infty}\frac{1}{M}\left[\int \mathrm{d}\boldsymbol{\xi}\; p_1(\boldsymbol{\xi};\mu,0|\boldsymbol{y},\boldsymbol{A})\mathcal{H}_0(|\boldsymbol{\xi}|_0;\beta|\boldsymbol{y},\boldsymbol{A})\right]_{\boldsymbol{y},\boldsymbol{A}}. \tag{41}$$

By utilizing the replica method again, we can calculate this. We defer the details of the calculations to appendix B, and here write down the resultant formula:

$$\begin{aligned}\epsilon_1^{\mathrm{LS}} = \underset{\hat{\Theta}_1^{\mathrm{LS}}}{\mathrm{extr}}\Bigg\{&\frac{1}{2}\frac{P+\sigma_y^2}{1+\chi_q}\left(\frac{\chi_c}{1+\chi_p}\right)^2 - \frac{C+\sigma_y^2}{1+\chi_q}\frac{\chi_c}{1+\chi_p} + \frac{1}{2}\frac{Q+\sigma_y^2}{1+\chi_q}\\ &- (\hat{C}C - \hat{\chi}_c\chi_c) - \frac{1}{2}(\hat{Q}Q - \hat{\chi}_q\chi_q)\\ &- \frac{\hat{\chi}_p}{2\alpha\hat{Q}}\left(\left(\frac{\hat{\chi}_q}{\hat{\chi}_p} - 2\frac{\hat{\chi}_c}{\hat{\chi}_p}\frac{\hat{C}}{\hat{P}} + (1+2\theta^2)\frac{\hat{C}^2}{\hat{P}^2}\right)\mathrm{erfc}(\theta) + \theta\left(\frac{\hat{\chi}_c^2}{\hat{\chi}_p^2} - \frac{\hat{C}^2}{\hat{P}^2}\right)\frac{2}{\sqrt{\pi}}\mathrm{e}^{-\theta^2}\right)\Bigg\},\end{aligned} \tag{42}$$

where $\hat{\Theta}_1^{\mathrm{LS}} = \{C, \chi_c, Q, \chi_q, \hat{C}, \hat{\chi}_c, \hat{Q}, \hat{\chi}_q\}$, and $\theta = \frac{\lambda}{\sqrt{2\hat{\chi}_p}}$. One point to remark on is that we should not take the extremization condition with respect to $\hat{\Theta}_1 = \{P, \chi_p, \hat{P}, \hat{\chi}_p\}$ in this expression. Instead, we should substitute the extremizer of (34) into it. Applying the extremization condition with respect to $\hat{\Theta}^{\mathrm{LS}}$ gives

J. Stat. Mech. (2016) 063302

$$\hat{\chi}_q = \frac{P+\sigma_y^2}{(1+\chi_q)^2}\left(\frac{\chi_c}{1+\chi_p}\right)^2 - 2\frac{C+\sigma_y^2}{(1+\chi_q)^2}\frac{\chi_c}{1+\chi_p} + \frac{Q+\sigma_y^2}{(1+\chi_q)^2}, \tag{43a}$$

$$\hat{Q} = \frac{1}{1+\chi_q}, \tag{43b}$$

$$\hat{\chi}_c = -\frac{P+\sigma_y^2}{1+\chi_q}\frac{\chi_c}{(1+\chi_p)^2} + \frac{C+\sigma_y^2}{1+\chi_q}\frac{1}{1+\chi_p} \tag{43c}$$

$$\hat{C} = -\frac{1}{1+\chi_q}\frac{\chi_c}{1+\chi_p}, \tag{43d}$$

$$\chi_q = \frac{1}{\alpha\hat{Q}}\mathrm{erfc}(\theta), \tag{43e}$$

$$Q = \frac{\hat{\chi}_p}{\alpha\hat{Q}^2}\left(\left(\frac{\hat{\chi}_q}{\hat{\chi}_p} - 2\frac{\hat{\chi}_c}{\hat{\chi}_p}\frac{\hat{C}}{\hat{P}} + (1+2\theta^2)\frac{\hat{C}^2}{\hat{P}^2}\right)\mathrm{erfc}(\theta) + \theta\left(\frac{\hat{\chi}_c^2}{\hat{\chi}_p^2} - \frac{\hat{C}^2}{\hat{P}^2}\right)\frac{2}{\sqrt{\pi}}\mathrm{e}^{-\theta^2}\right), \tag{43f}$$

$$\chi_c = \frac{1}{\alpha\hat{Q}}\left(-\frac{\hat{C}}{\hat{P}}\mathrm{erfc}(\theta) + \theta\frac{\hat{\chi}_c}{\hat{\chi}_p}\frac{2}{\sqrt{\pi}}\mathrm{e}^{-\theta^2}\right), \tag{43g}$$

$$C = -\frac{\hat{\chi}_p}{\alpha\hat{Q}}\left(\left(-\frac{\hat{\chi}_c}{\hat{\chi}_p}\frac{1}{\hat{P}} + (1+2\theta^2)\frac{\hat{C}}{\hat{P}^2}\right)\mathrm{erfc}(\theta) - \theta\frac{\hat{C}}{\hat{P}^2}\frac{2}{\sqrt{\pi}}\mathrm{e}^{-\theta^2}\right). \tag{43h}$$

From the EOSs, we can obtain the following simple relations:

$$\chi_q = \frac{r}{1-r} = \chi_p, \tag{44a}$$

$$\hat{Q} = \hat{P} = \frac{1}{1+\chi_p}, \tag{44b}$$

$$\epsilon_1^{\mathrm{LS}} = \frac{1}{2}\hat{\chi}_q. \tag{44c}$$

We now make some comments regarding the derivation of (42). In order to calculate the configurational average, we are required to deal with two different factors, Z_1 in $p_1 = (1/Z_1)\mathrm{e}^{-\mu\mathcal{H}_1}$, and the logarithm in $\mathcal{H}_0$. Correspondingly, as in the ℓ_0 case, we introduce replicas of two different kinds: n replicas to handle $1/Z_1$, and ν replicas to handle the logarithm. Using them, we can rewrite (41) as

$$\epsilon_1^{\mathrm{LS}} = \lim_{\beta\to\infty}\lim_{\mu\to\infty}\lim_{n\to 0}\lim_{\nu\to 0} -\frac{1}{M\beta\nu}\ln\left[Z_1^{n-1}(\mu,0|\boldsymbol{y},\boldsymbol{A})\int \mathrm{d}\boldsymbol{\xi}\,\mathrm{e}^{-\mu\mathcal{H}_1(\boldsymbol{\xi}|\boldsymbol{y},\boldsymbol{A})}\left(\int \mathrm{d}_{|\boldsymbol{\xi}|_0}\boldsymbol{x}\,\mathrm{e}^{-\frac{\beta}{2}\|\boldsymbol{y}-\boldsymbol{A}(|\boldsymbol{\xi}|_0\circ\boldsymbol{x})\|_2^2}\right)^{\nu}\right]_{\boldsymbol{y},\boldsymbol{A}}. \tag{45}$$

It is now possible to calculate the configurational average by assuming n and ν are natural numbers, and we can follow the usual prescription of the replica method. However, there remains a technical point concerning the limits $n\to 0$ and $\nu\to 0$ in the present formulation. The region around $n=\nu=0$ has an unusual property. The extremization condition with respect to the order parameters yields several different solutions. Among these solutions, by employing a versatile tool of spin-glass theory to analyze a probabilistic model conditioned by another probabilistic model, called the Franz–Parisi potential, we should choose the one analytically connected to $\hat{\Theta}_1$ in (34) in the limit $\nu\to 0$. This is achieved by the remark given below (42) [39].

3.2. Numerical validation using simulations on finite *M*

3.2.1. ℓ_0-based method. We examine the analytical results, using numerical simulations of finite-size systems. When M is sufficiently small, we can obtain the cumulant-generating function ϕ_0 by exhaustively searching all possible combinations of basis vectors. In cases where M is less small, we use the exchange Monte Carlo (MC) method to sample basis vector combinations obeying the canonical distribution at various temperature points [29, 30], and then estimate the cumulant-generating function ϕ_0 using the multi-histogram method [40].

In all simulations, we set $\alpha=0.5$ and $\sigma_y^2=1$. We treat two values of r equal to 0.2 and 0.4. In the case of $r=0.2$ (0.4), we calculate cumulant-generating function values at 15 temperature points, which are distributed according to the geometric progression in the range between 1 and 10 (between 1 and 35) in the value of μ. We conduct the exhaustive search for $M\leqslant 25$ (15), and use the exchange MC method for larger M. The configurational average is calculated by taking the median over 1000 different samples of $(\boldsymbol{y},\boldsymbol{A})$. The error bars are estimated by the Bootstrap method.

The procedure for our MC method will now be explained. At every temperature point, we randomly choose the initial vector $\boldsymbol{c}$ among those satisfying $\|\boldsymbol{c}\|_0=rM$. For $r=0.2$, the number of MC steps required for thermalization and sufficient sampling is $2,3,4,7,10\times 10^4$ for $M=30$, 35, 40, 45, 50, respectively, while for $r=0.4$ it is $2,4,8,15,30\times 10^4$ for $M=20$, 25, 30, 35, 40, respectively. The first half of the MC steps are discarded for thermalization. One MC step consists of two parts. First, updating once at every temperature point, and then exchanging once between every pair of neighboring temperature points. In each update of $\boldsymbol{c}$, we randomly choose one index i such that $c_i=1$ and another j such that $c_j=0$ to flip into the opposite state. That is, we set $c_i=0$ and $c_j=1$, and accept or reject this trial according to the Metropolis criterion based on the energy values calculated from $\mathcal{H}_0$ (12). The Metropolis criterion is also used in the exchange of $\boldsymbol{c}$s of different temperature points.

The results of the numerical simulations are presented in figure 3. Figure 3(a) shows the results of the cumulant-generating function value at $\mu=1$. On the vertical axis, the circles represent extrapolated values from finite-size results. The extrapolation lines are

given by the linear regression using an asymptotic form $\phi_0 \approx a + bM^{-1} + cM^{-1}\ln M^{-1}$. The regression is conducted by employing the method of least squares, as follows:

$$\min_{a,b,c} \frac{1}{2} \sum_{M} \left(a + b\frac{1}{M} + c\frac{1}{M}\ln\frac{1}{M} - \phi_0(M) \right)^2. \tag{46}$$

The asymptotic form is based on the Stirling's formula and is exact at $\mu = 0$, which motivates us to use the form even for $\mu \neq 0$. The cumulant-generating function and entropy density in the limit $M \to \infty$ are presented in figures 3(b) and (c), respectively. The lines represent the analytical results. The circles represent the extrapolated values from the numerical results. The analytical solutions are seen to be consistent with the numerical ones. Hence, the numerical results clearly validate the analytical results in the ℓ_0-based method.

3.2.2. ℓ_1-based method. Similarly to the case of the ℓ_0-based method, we examine the analytical results of the ℓ_1-based method by performing numerical simulations on finite-size systems. We carry out the ℓ_1-norm regularization using quadratic programming, and evaluate the distortion before the method of LS, ϵ_1; the distortion after the method of LS, ϵ_1^{LS}; and the compression rate r.

The values of α and σ_y^2 are fixed as $\alpha = 0.5$ and $\sigma_y^2 = 1$ for all simulations. We treat two values of λ equal to 1 and 2. We calculate (9) and (11) using quadratic programming and the method of LS for $M = 50, 100, \ldots, 250$.

The results of the numerical simulations are shown in figure 4. Figures 4(a)–(c) plot the numerically evaluated distortion before the method of LS, distortion after the method of LS, and the compression rate, respectively, against the system size M. On the vertical axes, the circles and crosses represent extrapolated and analytical values in the $M \to \infty$ limit, respectively. The extrapolation lines are given by the linear regression using the asymptotic forms $\epsilon_1 \approx a + bM^{-1}$, $\epsilon_1^{\mathrm{LS}} \approx c + dM^{-1}$, and $r \approx e + fM^{-1}$. We see that the analytical solutions are very close to the extrapolated values. This correlation clearly demonstrates the reliability of the analytical results.

3.3. Comparison in the trade-off relation

We compare the ideal performance in the $M \to \infty$ limit for different methods in terms of the trade-off relation between the distortion and the compression rate. Figure 5(a) shows the trade-off relations in the case of $\alpha = 0.5$. We see that both of the OCB-based methods achieve a better trade-off relation than the naive one. In the OCB-based strategy, the ℓ_0-based method significantly outperforms the ℓ_1-based one, even if the method of LS is operated after carrying out support estimation by the ℓ_1-norm regularization. We attribute the inferiority of the ℓ_1-based method to the regularization term. Indeed, as shown in figure 5(b), the regularization term is necessary to decrease the rate, but it distorts the original purpose of minimizing the distortion, as clearly seen from (27).

For a further comparison of the OCB-based methods, figure 6 shows the trade-off relations where different values of α control the degree of overcompleteness. Figures 6(a) and (b) present the results of the ℓ_0- and ℓ_1-based methods, respectively. In the ℓ_1-based method, the method of LS has been operated after the support estimation. Both methods

J. Stat. Mech. (2016) 063302

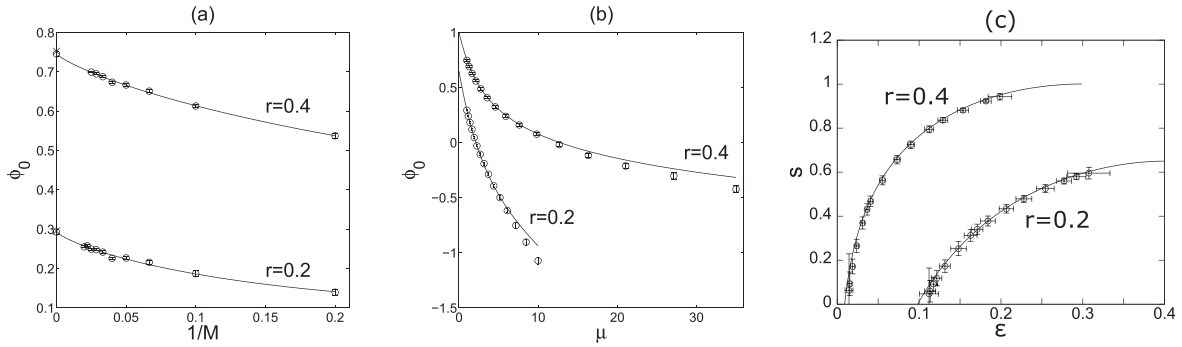

Figure 3. Cumulant-generating function ϕ_0 and entropy density s of the ℓ_0-based method with $\sigma_y^2 = 1$, $\alpha = 0.5$, and $r = 0.2$, 0.4. (a) Plots of numerically evaluated ϕ_0 at $\mu = 1$. The lines are given by the linear regression. On the vertical axis, the circles and crosses represent the extrapolated and analytical values in the $M \to \infty$ limit, respectively. The lengths of the error bars are comparable to the sizes of symbols. (b) Plots of ϕ_0 in the $M \to \infty$ limit. The lines and circles represent the analytical and extrapolated values, respectively. The lengths of the error bars are comparable to the sizes of symbols. (c) Plots of s against ϵ in the $M \to \infty$ limit. The lines and circles represent the analytical and extrapolated values, respectively. These are calculated from the values of ϕ_0 in (b).

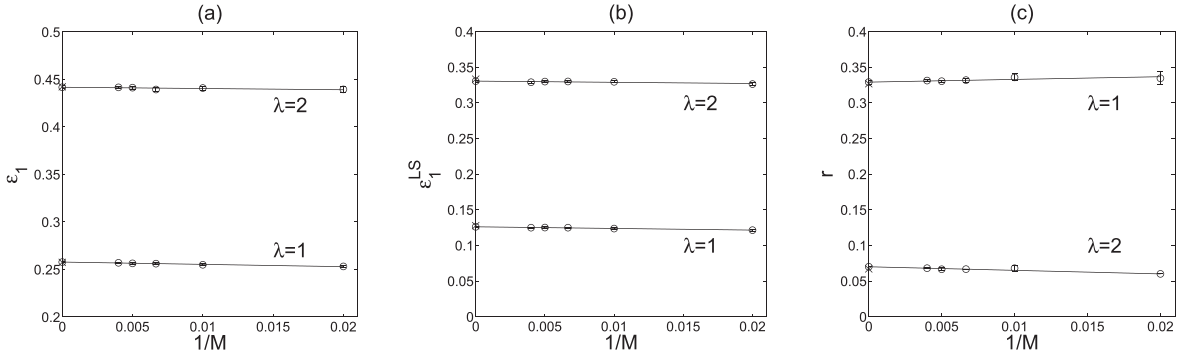

Figure 4. Plots of numerically evaluated values of the ℓ_1-based method with $\sigma_y^2 = 1$, $\alpha = 0.5$, and $\lambda = 1, 2$. The extrapolation lines are given by the linear regression. On the vertical axes, the circles and crosses represent the extrapolated and analytical values in the $M \to \infty$ limit, respectively. The lengths of the error bars are comparable to the sizes of symbols. (a) Distortion before the method of LS, ϵ_1. (b) Distortion after the method of LS, ϵ_1^{LS}. (c) Compression rate, r.

achieve a better trade-off relation as the degree of overcompleteness increases, or α decreases. Another interesting observation is the superiority of the ℓ_0-based method compared to the ℓ_1-based one, regardless of the degree of overcompleteness.

3.3.1. In the large limit of the degree of overcompleteness, $\alpha \to 0$. From figure 6 we see that the distortion becomes smaller as α decreases, both for the ℓ_0- and ℓ_1-based methods. An interesting question is whether the distortion vanishes or not in the limit $\alpha \to 0$, or more quantitatively, how ϵ is scaled by α in the small limit.

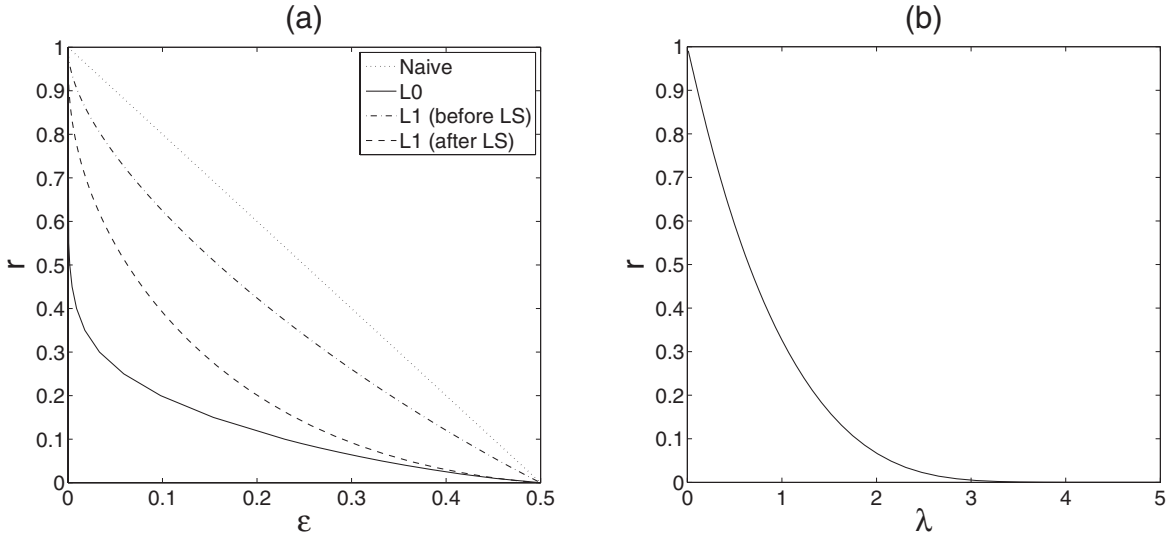

Figure 5. Results of the analysis in the $M \to \infty$ limit with $\sigma_y^2 = 1$ and $\alpha = 0.5$. (a) Trade-off relations of the naive, ℓ_0-based, and ℓ_1-based methods (before and after the method of LS). (b) Relation between the rate and the regularization coefficient in the ℓ_1-based method.

Deferring the detailed calculations to appendix A.2 and B.2, here we summarize our analytical results on the behavior of ϵ in the limit $\alpha \to 0$

$$\epsilon_0 \propto \alpha^{\frac{2r}{1-r}} \to 0, \tag{47}$$

$$\epsilon_1 \to \frac{1}{2}(1-r)^2\sigma_y^2 = O(1), \tag{48}$$

$$\epsilon_1^{\mathrm{LS}} \propto |\ln \alpha|^{-1} \to 0. \tag{49}$$

The asymptotic behaviors of ϵ_0 and ϵ_1^{LS} are examined using numerical solutions of the corresponding EOSs, (21) and (43a–43h), in figure 7. Our analytic formulas show an excellent agreement with the numerical results.

We stress the consequence of (47)–(49). First, they give a firm indication that it is reasonable to apply the method of LS after the ℓ_1-norm regularization, which is heuristically employed in related problems such as compressed sensing in practical situations. The difference in (48) and (49) indicates that the method of LS actually diminishes the distortion, and even eliminates it in the ideal limit $\alpha \to 0$, which never happens with only the use of ℓ_1-norm regularization. Second, (47) provides a general bound for the computational cost of searching the appropriate basis vectors. From (47), given a target value of the distortion $\hat{\epsilon}$ and some data on the length M, the required size $N_{\mathrm{req}}(\hat{\epsilon}, M)$ of the basis matrix to achieve this distortion value is scaled as

$$N_{\mathrm{req}}(\hat{\epsilon}, M) \propto M\hat{\epsilon}^{-\frac{1-r}{2r}}. \ (\ell_0) \tag{50}$$

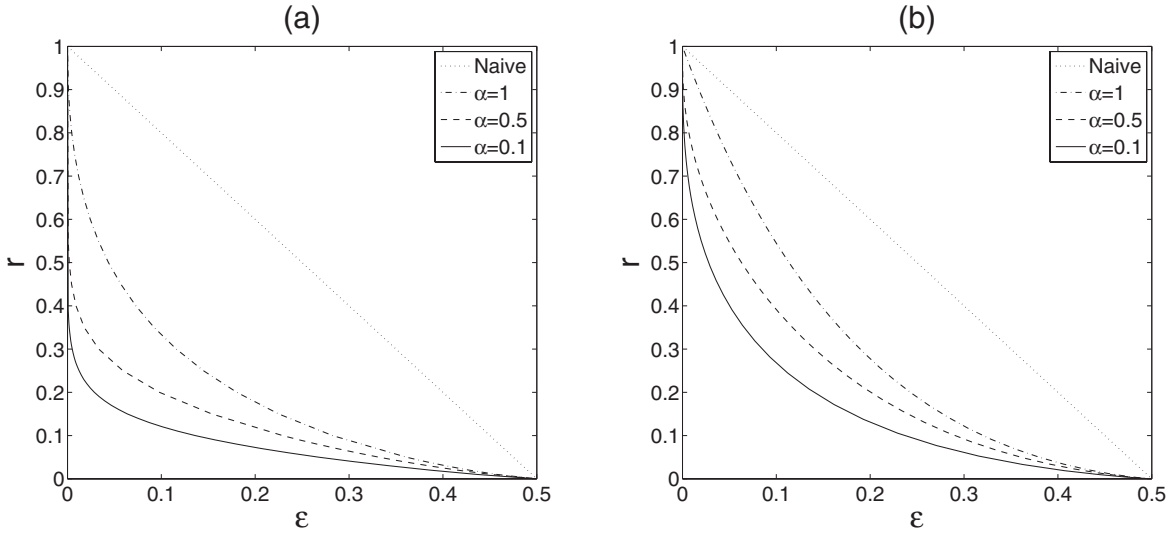

Figure 6. Trade-off relations at various values of α in the case of $\sigma_y^2 = 1$. (a) Results of the ℓ_0-based method. (b) Results of the ℓ_1-based method (after the method of LS). For both the methods, against a fixed r, the distortion ϵ becomes smaller as α decreases.

This grows in a polynomial manner as the target distortion value $\hat{\epsilon}$ decreases, and the exponent of the polynomial negatively grows as the compression rate r decreases. This quantitative information will provide a theoretical basis in designing algorithms. Finally, (49) manifests the limit of the ℓ_1-based method. The size $N_{\rm req}$ required to achieve the target distortion $\hat{\epsilon}$ in this case is scaled as

$$N_{\rm req}(\hat{\epsilon}, M) \propto M \mathrm{e}^{\frac{1}{\hat{\epsilon}}}, \ (\ell_1 + \mathrm{LS}). \tag{51}$$

This grows exponentially as $\hat{\epsilon}$ decreases, which is considered to be reasonable. If it were a polynomial, versatile algorithms exactly solving the ℓ_1-norm regularization could be applied to solve the problem with a computational cost of a polynomial order of the system size and the precision, which is believed not to be possible. However, (51) can still be useful, because it provides a quantitative comparison between the data size M and the acceptable distortion $\hat{\epsilon}$ in an unified manner.

4. Examination of practical performance

4.1. Algorithms and their performances

A lot of computational time is required to conduct the exhaustive search used in the ℓ_0-based method. However, it is considered that certain greedy algorithms might work well for practical applications. Orthogonal matching pursuit (OMP, figure 8) is a greedy algorithm that may be suitable for the present purpose [31, 32]. OMP only requires a computational time of order $O(M^4)$ for the current purpose. We compare the performance of OMP with the ideal performances of both the ℓ_0- and ℓ_1-based methods.

J. Stat. Mech. (2016) 063302

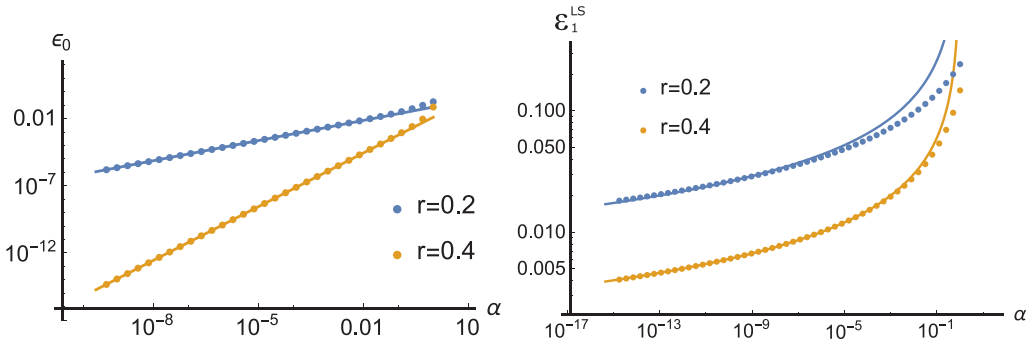

Figure 7. Plots of ϵ against α in the small α limit, derived by solving (21) and (43*a*–43*h*) numerically. The left panel represents ϵ_0, and the right panel represents ϵ_1^{LS}. The lines are the fits based on our analytical formulas, (47) and (49), and these show excellent agreement with the points obtained by the numerical evaluations.

In addition to OMP, we also examine approximate message passing (AMP, figure 9), as a representative algorithm carrying out the ℓ_1-norm regularization. From the viewpoint of quadratic programming, ℓ_1-norm regularization is solved exactly using versatile algorithms, which require a computational time of order $O(M^3)$. In contrast, AMP only requires a computational time of order $O(M^2)$ per update. Despite the low computational cost, AMP is known to be able to recover the results of those versatile algorithms, in certain reasonable situations [33]. The present case, where the basis matrix $\boldsymbol{A}$ and the data vector $\boldsymbol{y}$ are generated from i.i.d. normal distributions, is expected to be one such situation. Hence, we can fairly compare the result of AMP with the ideal performance of the ℓ_1-based method, and therefore with that of OMP.

We evaluate the performances of OMP and AMP when they are employed for sparse approximation with the OCB-based strategy. We examine the case with $\sigma_y^2 = 1$ and $\alpha = 0.5$. Figure 10 presents the results of the performance evaluations of OMP and AMP. Figure 10(a) shows the results for finite-size systems, namely $M = 50, 100, \ldots, 250$, and the extrapolation by the linear regression using an asymptotic form of $\epsilon \approx a + bM^{-1}$. The compression rate is set to $r = 0.5$ when evaluating OMP, and the regularization coefficient λ is set to 0.65 when evaluating AMP, so that $r \approx 0.5$. We evaluate the performance of AMP based on the distortion after the method of LS. In figure 10(b), we compare the extrapolated performances of OMP and AMP at various rates with the achievable trade-off relation analyzed in section 3. The AMP result compares well with the ideal performance of the ℓ_1-based method, while that for OMP does not reach the ideal result of the ℓ_0-based method. However, a notable finding is that OMP considerably outperforms the ℓ_1-based results. This motivates the exploration of better algorithms for the ℓ_0-based method, in the context of sparse approximation. Such exploration is currently under way.

4.2. Application to image data

We investigate the performance of sparse approximation, when it is applied to a task of image data compression. We compress image data composed of 256×256 pixels. The experimental procedure of compression is as follows. First, image data are normalized

Input: a data vector $\boldsymbol{y}$, a basis matrix $\boldsymbol{A}$, a rate r.

Initialization: $\boldsymbol{x}^{(0)} = \boldsymbol{0}$, $U = \{1, 2, \ldots, N\}$, $S^{(0)} = \emptyset$.

Iteration: repeat from $n = 1$ until $n = rM$:

$$\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}^{(n-1)},$$

$$j = \underset{k \in U \setminus S^{(n-1)}}{\arg\max} \{|\boldsymbol{a}_k^{\mathrm{T}} \boldsymbol{r}|\},$$

$$S^{(n)} = S^{(n-1)} \cup \{j\},$$

$$\boldsymbol{x}^{(n)} = \underset{\boldsymbol{x}}{\arg\min}\{||\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}||_2\} \quad \text{subj. to } \mathrm{supp}(\boldsymbol{x}) \subset S^{(n)}.$$

Output: a sparse vector $\hat{\boldsymbol{x}} = \boldsymbol{x}^{(rM)}$.

Figure 8. The procedure of OMP. $\varnothing$ is the empty set. $\mathrm{supp}(\cdot)$ is the support set.

Input: a data vector $\boldsymbol{y}$, a basis matrix $\boldsymbol{A}$, a regularization coefficient λ, a tuning parameter δ.

Initialization: $\boldsymbol{x}^{(0)} = \boldsymbol{0}$, $\chi^{(0)} = 0$, $\boldsymbol{r}^{(0)} = \boldsymbol{y}$.

Iteration: repeat until convergence at $n = \hat{n}$:

$$\hat{Q} = \frac{1}{1+\chi^{(n-1)}},$$

$$\boldsymbol{r}^{(n)} = (1 - \hat{Q})\boldsymbol{r}^{(n-1)} + \hat{Q}(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}^{(n-1)}),$$

$$\boldsymbol{h} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{r}^{(n)} + \hat{Q}\boldsymbol{x}^{(n-1)},$$

$$\chi^{(n)} = (1 - \delta)\chi^{(n-1)} + \delta \frac{1}{\hat{Q}} \frac{1}{M} \sum_i \Theta(|h_i| - \lambda),$$

$$x_i^{(n)} = (1 - \delta)x_i^{(n-1)} + \delta \frac{1}{\hat{Q}} \mathrm{sign}(h_i)(|h_i| - \lambda)\Theta(|h_i| - \lambda) \quad \text{for } i = 1, \ldots, N.$$

Output: a sparse vector $\hat{\boldsymbol{x}} = \boldsymbol{x}^{(\hat{n})}$.

Figure 9. The procedure of AMP. $\mathrm{sign}(\cdot)$ is the sign function. $\Theta(\cdot)$ is the Heaviside step function.

so as to set the mean and variance to 0 and 1, respectively. Next, 256×256 pixels are randomly permuted, in order to obtain 1024 column vectors, whose dimension is 64. Following these operations, the data can be regarded as random numbers with a mean and variance of 0 and 1, which approximates the properties of the data to the situation which we have already studied theoretically and numerically. Finally, setting $r = 0.5$, we compress each of the column vectors into a sparse vector by using a 64×128 random matrix, namely $\alpha = 0.5$. We examine the performances of OMP and AMP. When applying AMP, we set the regularization coefficient to 0.65, so that $r \approx 0.5$, and the method of LS is operated after the support estimation by the ℓ_1-norm regularization. The results of experiments are presented in figure 11. Although OMP requires a computational time that is several times larger than that of AMP, OMP outperforms AMP in terms of appearance and peak signal-to-noise ratio (PSNR), defined as

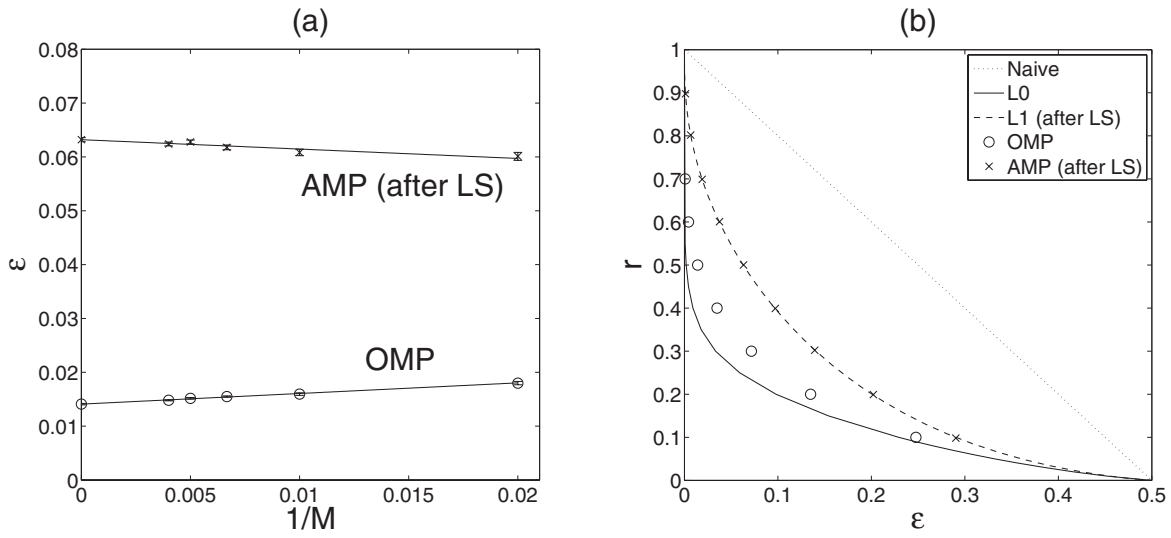

Figure 10. Performances of OMP and AMP in the case with $\sigma_y^2 = 1$ and $\alpha = 0.5$. The performance of AMP is evaluated after the method of LS. (a) Plots of the numerically evaluated distortions in the case of $r = 0.5$. The extrapolation lines are given by the linear regression. On the vertical axis, the symbols represent extrapolated values in the $M \to \infty$ limit. The lengths of the error bars are comparable to the sizes of symbols. In OMP r is set to $r = 0.5$, and in AMP λ is set to 0.65, so that $r \approx 0.5$. (b) Trade-off relations in the $M \to \infty$ limit. The circles and crosses represent extrapolated values of OMP and AMP, respectively.

$$\mathrm{PSNR} = 10 \log_{10} \frac{255^2}{\frac{1}{N}\sum_{ij}(\hat{I}_{ij} - I_{ij})^2}, \tag{52}$$

where $\boldsymbol{I} = \{I_{ij}\}$ and $\hat{\boldsymbol{I}} = \{\hat{I}_{ij}\}$ represent an original image and a sparse-approximated image, respectively, and N is the number of image pixels.

If the scope of application is limited to image data compression, more convenient bases, such as a discrete wavelet transformation, will achieve much better results in the performance and computational time [41, 42]. However, in general contexts it is not easy to find a proper basis for sparse approximation in advance. A solution to this problem is to use blind compressed sensing and related techniques such as dictionary learning [43–45], but the computational costs are rather high. Our OCB-based strategy may overcome this difficulty, because it avoids the learning of the dictionary by preparing many candidates for basis vectors and choosing a suitable combination. Our theoretical analysis and numerical experiments positively support this possibility.

5. Conclusion

In the present paper, sparse-data processing has been discussed from the viewpoint of sparse approximation. We have focused on a strategy of sparse approximation that is based on a random OCB, and have explored the abilities and limitations of the ℓ_0- and

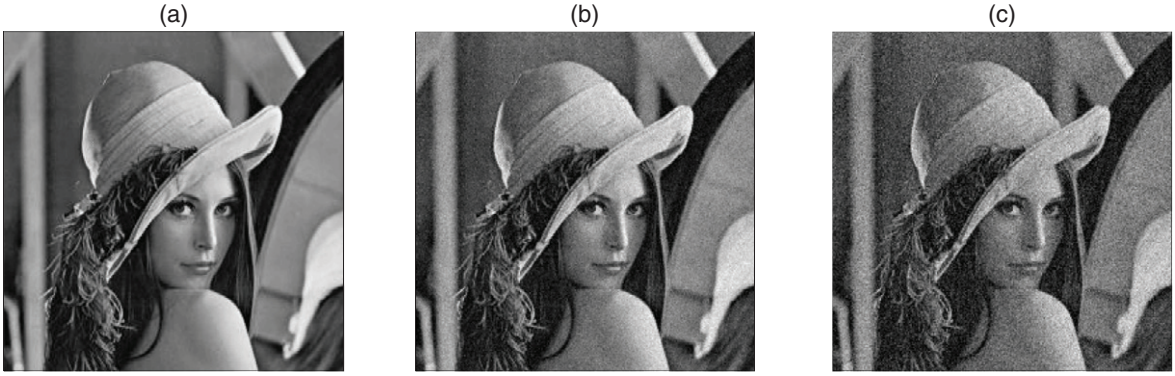

Figure 11. Application of sparse approximation with the OCB-based strategy to image data compression. The degree of overcompleteness is $\alpha = 0.5$. (a) Original image data. (b) Sparsely-approximated image data obtained using OMP. The compression rate is $r = 0.5$. PSNR is 28.2. The time required is approximately 55 sec. (c) Sparsely-approximated image data obtained using AMP. The regularization coefficient is $\lambda = 0.65$, so that $r \approx 0.5$. The AMP-based sparse representation is given after the method of LS. PSNR is 22.9. The time required is approximately 4.5 sec. Copyright Playboy Enterprises, Inc.

ℓ_1-based methods. We have analyzed the ideal performances of these methods in the large-system limit in a statistical-mechanical manner, which has been validated by numerical simulations on finite-size systems and their extrapolation to the infinite-size limit. Our results have indicated that the ℓ_0-based method outperforms the naive and ℓ_1-based methods in terms of the trade-off relation between the distortion and the compression rate. A notable result is that any small distortion is achievable for any finite fixed value of the compression rate, by increasing the degree of overcompleteness, for both the ℓ_0- and ℓ_1-based methods. This result allows us to determine both the theoretical limit of the OCB-based strategy and the limit for practical algorithms based on the ℓ_1 regularization. In addition, it provides a firm basis for the use of the method of LS after the ℓ_1 regularization, which is frequently applied in related problems such as compressed sensing in practical situations.

In addition to the ideal performance analyzed in section 3, we also investigated the practical performance of our strategy in section 4. We evaluated the performances of OMP and AMP as algorithms to approximately perform the ℓ_0- and ℓ_1-based methods, respectively. Our evaluation showed that OMP surpasses both AMP and the exact execution of the ℓ_1-based method, in terms of the trade-off relation. This suggests that greedy algorithms are more suitable for sparse approximation using our strategy than convex relaxation algorithms, although there is still room to design more effective greedy algorithms than OMP. We are currently undertaking further research in this direction.

We considered the application of our method to image data compression, as a practical example, and evaluated its performance when OMP and AMP are utilized. OMP outperforms AMP in appearance and PSNR, although OMP requires a computational time that is several times larger. In order to efficiently decrease the computational time of our strategy, it is important to find a proper basis. This suggests the use of some prior knowledge in constructing the overcomplete basis. Some further possibilities, such as combining our methods with dictionary learning, are still open, and would be interesting to address in future work.

J. Stat. Mech. (2016) 063302

Acknowledgments

This work was supported by JSPS KAKENHI Grant Numbers 13J04920 (YN-O), 26870185 (TO), 25120009 (MO), and 25120013 (YK).

Appendix A. Calculations for the ℓ_0-based method

A.1. Derivation of ϕ_0

Based on (26), we define

$$\psi_0(n,\nu,\mu) = \frac{1}{M}\ln\left[\left\{\mathop{\mathrm{Tr}}_{\boldsymbol{c}}\left(\int \mathrm{d}_{\boldsymbol{c}}\boldsymbol{x}\,\mathrm{e}^{-\frac{1}{2}\frac{\mu}{\nu}\|\boldsymbol{y}-\boldsymbol{A}(\boldsymbol{c}\circ\boldsymbol{x})\|_2^2}\right)^{\nu}\right\}^{n}\right]_{\boldsymbol{y},\boldsymbol{A}}. \tag{A.1}$$

The cumulant-generating function ϕ_0 is recovered from ψ_0, as $\phi_0(\mu) = \lim_{n,\nu\to 0}(1/n)\psi_0(n,\nu,\mu)$. When (n,ν) are positive integers, we obtain

$$\psi_0(n,\nu,\mu) = \frac{1}{M}\ln \mathop{\mathrm{Tr}}_{\{\boldsymbol{c}^a\}}\mathop{\mathrm{Tr}}_{\{\boldsymbol{x}^{a\alpha}\}|\{\boldsymbol{c}^a\}}\left[\mathrm{e}^{-\frac{\mu}{2\nu}\sum_{j=1}^{M}\sum_{a=1}^{n}\sum_{\alpha=1}^{\nu}\left(y_j-\sum_i A_{ji}c_i^a x_i^{a\alpha}\right)^2}\right]_{\boldsymbol{y},\boldsymbol{A}}, \tag{A.2}$$

where $\mathrm{Tr}_{\{\boldsymbol{c}^a\}} = \prod_{a=1}^{n}\sum_{\boldsymbol{c}^a}\delta\left(Mr-\sum_i c_i^a\right)$, and $\mathrm{Tr}_{\{\boldsymbol{x}^{a\alpha}\}|\{\boldsymbol{c}^a\}} = \prod_{a=1}^{n}\prod_{\alpha=1}^{\nu}\int \mathrm{d}_{\boldsymbol{c}^a}\boldsymbol{x}^{a\alpha}$. Let us introduce the variables $s_j^{a\alpha} = \sum_i A_{ji}c_i^a x_i^{a\alpha}$ and $Q_{(a\alpha)(b\beta)} = \frac{1}{M}\sum_i (c_i^a x_i^{a\alpha})(c_i^b x_i^{b\beta})$. According to the central limit theorem, we regard the variables $\{s_j^{a\alpha}\}$ as random variables that follow a zero-mean multivariate normal distribution, with covariances $[s_j^{a\alpha}s_k^{b\beta}]_{\boldsymbol{A}} = \delta_{jk}Q_{(a\alpha)(b\beta)}$. Using these variables, we obtain

$$\psi_0(n,\nu,\mu) = \frac{1}{M}\ln \mathop{\mathrm{Tr}}_{\{\boldsymbol{c}^a\}}\mathop{\mathrm{Tr}}_{\{\boldsymbol{x}^{a\alpha}\}|\{\boldsymbol{c}^a\}}\mathop{\mathrm{Tr}}_{\boldsymbol{Q}}\left(\left[\mathrm{e}^{-\frac{\mu}{2\nu}\sum_a\sum_\alpha(y-s^{a\alpha})^2}\right]_{y,\{s^{a\alpha}\}|\boldsymbol{Q}}\right)^{M}, \tag{A.3}$$

where $\mathrm{Tr}_{\boldsymbol{Q}} = \prod_{(a\alpha),(b\beta)}\int \mathrm{d}Q_{(a\alpha)(b\beta)}\delta\left(MQ_{(a\alpha)(b\beta)} - \sum_i (c_i^a x_i^{a\alpha})(c_i^b x_i^{b\beta})\right)$, and the brackets $[\cdot]_{y,\{s^{a\alpha}\}|\boldsymbol{Q}}$ denote the average over y and $s^{a\alpha}$, which is conditioned by the variance $Q_{(a\alpha)(b\beta)}$ as explained above.

After introducing the Fourier representation of the delta function, $\delta(\cdot)\propto\int \mathrm{d}\tilde{X}\mathrm{e}^{\frac{\tilde{X}}{2}(\cdot)}$, the saddle-point method is employed to obtain

$$\begin{aligned}\psi_0(n,\nu,\mu) = \mathop{\mathrm{extr}}_{\Theta_0}\Bigg\{&\ln\left[\mathrm{e}^{-\frac{\mu}{2\nu}\sum_a\sum_\alpha(y-s^{a\alpha})^2}\right]_{y,\{s^{a\alpha}\}|\boldsymbol{Q}} + \sum_a\frac{\tilde{r}_a}{2}r \\ &+ \sum_{(a\alpha),(b\beta)}\frac{\tilde{Q}_{(a\alpha)(b\beta)}}{2}Q_{(a\alpha)(b\beta)} + \frac{1}{\alpha}\ln\sum_{\{c^a\}}\mathop{\mathrm{Tr}}_{\{x^{a\alpha}\}|\{c^a\}}\mathrm{e}^{-\sum_a\frac{\tilde{r}_a}{2}c^a-\sum_{(a\alpha),(b\beta)}\frac{\tilde{Q}_{(a\alpha)(b\beta)}}{2}(c^a x^{a\alpha})(c^b x^{b\beta})}\Bigg\},\end{aligned} \tag{A.4}$$

where $\Theta_0 = \{\boldsymbol{Q}, \tilde{\boldsymbol{r}}, \tilde{\boldsymbol{Q}}\}$. For the extremizer, we search the subspace with $(Q_{(a\alpha)(b\beta)}, \tilde{Q}_{(a\alpha)(b\beta)})$ equal to $(Q, \tilde{Q})$ $(a = b, \alpha = \beta)$, $(q_1, -\tilde{q}_1)$ $(a = b, \alpha \neq \beta)$, or $(q_0, -\tilde{q}_0)$ $(a \neq b)$, with $\tilde{r}_a = \tilde{r}$. This is the RS in the present formula of two replica numbers n and ν. Then, we obtain

$$\psi_0(n,\nu,\mu) = \operatorname*{extr}_{\tilde{\Theta}_0}\left\{\ln\int \mathrm{D}y\mathrm{D}w\left(\int \mathrm{D}v\left(\int \mathrm{D}u\, \mathrm{e}^{-\frac{\mu}{2\nu}\left(\sigma_y y-\sqrt{q_0}\,w-\sqrt{q_1-q_0}\,v-\sqrt{Q-q_1}\,u\right)^2}\right)^{\nu}\right)^{n}\right.$$
$$+\frac{1}{2}n\tilde{r}r+\frac{1}{2}n\nu\tilde{Q}Q-\frac{1}{2}n\nu(\nu-1)\tilde{q}_1q_1-\frac{1}{2}n(n-1)\nu^2\tilde{q}_0q_0$$
$$\left.+\frac{1}{\alpha}\ln\int \mathrm{D}z\left(\int \mathrm{D}t\sum_c\left(\operatorname*{Tr}_{x|c}\mathrm{e}^{-\frac{\tilde{r}}{2\nu}c-\frac{\tilde{Q}+\tilde{q}_1}{2}cx^2+t\sqrt{\tilde{q}_1-\tilde{q}_0}\,cx+z\sqrt{\tilde{q}_0}\,cx}\right)^{\nu}\right)^{n}\right\}, \tag{A.5}$$

where $\tilde{\Theta}_0 = \{Q, q_1, q_0, \tilde{r}, \tilde{Q}, \tilde{q}_1, \tilde{q}_0\}$ and $\operatorname{Tr}_{x|c} = \int \mathrm{d}_c x$. We assume that (A.5) is true not only for positive integers (n, ν) but also for real numbers (n, ν). In taking the limits $(n, \nu) \to (0, 0)$, we introduce $\chi = \beta(Q - q_1)$, $q = q_0$, $\hat{r} = \tilde{r}$, $\hat{Q} = \nu(\tilde{Q} + \tilde{q}_1) - \nu^2\tilde{q}_1$, $\hat{\chi} = \nu^2\tilde{q}_1$, and $\hat{q} = \nu^2\tilde{q}_0$, which are assumed to be of the order $O(1)$ in these limits. Following some straightforward calculations, the replica identity is given by

$$\phi_0(\mu) = \lim_{n\to 0}\lim_{\nu\to 0}\frac{1}{n}\psi_0(n,\nu,\mu), \tag{A.6}$$

thus yielding (19).

A.2. The limit $\alpha \to 0$ in the ℓ_0 case

We examine the behavior of the zero point of entropy, ϵ_0, in the large-size limit of the basis matrix, $\alpha \to 0$. The parameter μ corresponding to the zero point ϵ_0, μ_0, can be formally written using (23) and (24), as

$$\mu_0 = -\frac{\hat{\chi}(\mu_0)}{2\phi_0(\mu_0)} = -\frac{1}{2}\hat{\chi}\left\{\frac{1}{2}\ln\frac{1+\chi}{1+\chi+\mu_0\Delta} - \frac{1}{2}\frac{\mu_0(q+\sigma_y^2)}{1+\chi+\mu_0\Delta}\right.$$
$$\left.+\frac{1}{2}\left(\hat{r}r+\hat{Q}Q-\frac{\hat{\chi}}{\mu_0}\chi+\hat{q}q\right)+\frac{1}{\alpha}\int \mathrm{D}z\ln\left(1+\sqrt{\frac{\hat{\chi}+\hat{Q}}{\hat{Q}+\hat{q}}}\,\mathrm{e}^{-\frac{1}{2}\hat{r}+\frac{1}{2}\frac{\hat{q}}{\hat{Q}+\hat{q}}z^2}\right)\right\}^{-1}. \tag{A.7}$$

A numerical calculation indicates the behavior of $\mu_0 \to \infty$ as $\alpha \to 0$, while $\hat{Q}$, $\hat{q}$, Q, q, $\chi \sim O(1)$ are kept finite. We will determine the scalings of the relevant variables for $\alpha \to 0$ so as to agree with these observations. A crucial observation from (21*d*) is that the factor Y should vanish, in order to cancel the vanishing α, yielding

$$\frac{1}{\alpha}Y \propto \frac{\sqrt{\mu_0}}{\alpha}\mathrm{e}^{-\frac{1}{2}\hat{r}} = O(1) \Rightarrow \begin{cases}\hat{r} = \tilde{r} - 2\rho\ln\alpha \\ \mu_0 \propto \alpha^{2-2\rho}\end{cases}, \tag{A.8}$$

where we introduce an exponent ρ controlling the divergence speed of $\hat{r}$ and μ_0. Since we assume the divergence of μ_0, ρ must be larger than unity. The value of ρ is determined

J. Stat. Mech. (2016) 063302

by solving (A.7) in a self-consistent manner. The scaling of the remaining order parameter $\hat{\chi}$ is determined by

$$\hat{\chi} \to \frac{\mu_0(\alpha)}{1+\chi} + \frac{q+\sigma_y^2}{\Delta^2} \to \infty. \tag{A.9}$$

Now, we know all of the scalings of the order parameters, and can reduce (A.7) to the dominant part, as

$$\mu_0 \approx \frac{\frac{\mu_0}{1+\chi} + \frac{q+\sigma_y^2}{\Delta^2}}{2\rho r \ln \alpha + \ln(1+\chi+\mu_0\Delta)}. \tag{A.10}$$

By solving this in the leading scaling, we obtain

$$\rho = \frac{1}{1-r},\ \mu_0 \approx \frac{\mathrm{e}^{(1+\chi)^{-1}}}{\Delta}\alpha^{-2\frac{r}{1-r}} + O(1). \tag{A.11}$$

By inserting (A.9) and (A.11) into (23), we get (47)

Appendix B. Some calculations for the ℓ_1-based methods

B.1. Derivations of f_1 and ϵ_1^{LS}

Based on (45), we introduce

$$\psi_1(n,\nu,\beta,\mu,\kappa) = \frac{1}{M}\ln\left[Z_1^{n-1}(\mu,\kappa|\boldsymbol{y},\boldsymbol{A}) \int \mathrm{d}\boldsymbol{\xi}\, \mathrm{e}^{-\mu(\mathcal{H}_1(\boldsymbol{\xi}|\boldsymbol{y},\boldsymbol{A})+\kappa\|\boldsymbol{\xi}\|_0)} \left(\int \mathrm{d}_{|\boldsymbol{\xi}|_0}\boldsymbol{x}\, \mathrm{e}^{-\frac{\beta}{2}\|\boldsymbol{y}-\boldsymbol{A}(|\boldsymbol{\xi}|_0\circ\boldsymbol{x})\|_2^2} \right)^{\nu} \right]_{\boldsymbol{y},\boldsymbol{A}}. \tag{B.1}$$

By calculating this in the case of positive integers (n,ν), we obtain

$$\psi_1(n,\nu,\beta,\mu,\kappa) = \frac{1}{M}\ln \underset{\{\boldsymbol{\xi}^a\}}{\mathrm{Tr}} \underset{\{\boldsymbol{x}^\alpha\}}{\mathrm{Tr}} \left(\left[\mathrm{e}^{-\frac{\mu}{2}\sum_a\left(y_j-\sum_i A_{ji}\xi_i^a\right)^2 - \frac{\beta}{2}\sum_\alpha\left(y_j-\sum_i A_{ji}|\xi_i^1|_0 x_i^\alpha\right)^2} \right]_{\boldsymbol{y},\boldsymbol{A}} \right)^M, \tag{B.2}$$

where $\mathrm{Tr}_{\{\boldsymbol{\xi}^a\}} = \prod_{a=1}^n \int \mathrm{d}\boldsymbol{\xi}^a \mathrm{e}^{-\mu\left(\lambda\sum_i|\xi_i^a|+\kappa\sum_i|\xi_i^a|_0\right)}$, and $\mathrm{Tr}_{\{\boldsymbol{x}^\alpha\}} = \prod_{\alpha=1}^\nu \int \mathrm{d}_{|\boldsymbol{\xi}^1|_0}\boldsymbol{x}^\alpha$. Let us introduce the variables $s_j'^a = \sum_i A_{ji}\xi_i^a$, $s_j^\alpha = \sum_i A_{ji}|\xi_i^1|x_i^\alpha$, $P_{ab} = \frac{1}{M}\sum_i \xi_i^a\xi_i^b$, $C_{\alpha a} = \frac{1}{M}\sum_i(|\xi_i^1|_0 x_i^\alpha)\xi_i^a$, and $Q_{\alpha\beta} = \frac{1}{M}\sum_i(|\xi_i^1|_0 x_i^\alpha)(|\xi_i^1|_0 x_i^\beta)$. As in the ℓ_0 case, we can rewrite the variables $\{s_j'^a, s_j^\alpha\}$ as random variables from a zero-mean multivariate normal distribution, with the covariances $[s_j'^a s_k'^b]_{\boldsymbol{A}} = \delta_{jk}P_{ab}$, $[s_j'^a s_k^\alpha]_{\boldsymbol{A}} = \delta_{jk}C_{a\alpha}$, and $[s_j^\alpha s_k^\beta]_{\boldsymbol{A}} = \delta_{jk}Q_{\alpha\beta}$. The application of the central limit theorem here is justified by the nonzeroness of compression rate r shown in figure 5(b) derived from (36). Using these variables, we obtain

J. Stat. Mech. (2016) 063302

$$\psi_1(n,\nu,\beta,\mu,\kappa)=\frac{1}{M}\ln\operatorname*{Tr}_{\{\boldsymbol{\xi}^a\}\{\boldsymbol{x}^\alpha\}}\operatorname*{Tr}_{\boldsymbol{P}}\operatorname*{Tr}_{\boldsymbol{C}}\operatorname*{Tr}_{\boldsymbol{Q}}\left(\left[\mathrm{e}^{-\frac{\mu}{2}\sum_a(y_j-s_j'^a)^2-\frac{\beta}{2}\sum_\alpha(y_j-s_j^\alpha)^2}\right]_{y_j,\{s_j'^a,s_j^\alpha\}|\boldsymbol{P},\boldsymbol{C},\boldsymbol{Q}}\right)^M, \quad \text{(B.3)}$$

where $\operatorname*{Tr}_{\boldsymbol{P}}=\prod_{a,b}\int \mathrm{d}P_{ab}\delta(MP_{ab}-\sum_i\xi_i^a\xi_i^b)$, $\operatorname{Tr}_{\boldsymbol{C}}=\prod_{a,\alpha}\int \mathrm{d}C_{\alpha a}\delta\Big(MC_{\alpha a}-\sum_i(|\xi_i^1|_0x_i^\alpha)\xi_i^a\Big)$, and $\operatorname{Tr}_{\boldsymbol{Q}}=\prod_{\alpha,\beta}\int \mathrm{d}Q_{\alpha\beta}\delta\Big(MQ_{\alpha\beta}-\sum_i(|\xi_i^1|_0x_i^\alpha)(|\xi_i^1|_0x_i^\beta)\Big)$. After introducing the Fourier representation of the delta function, the saddle-point method is employed, to obtain

$$\begin{aligned}\psi_1(n,\nu,\beta,\mu,\kappa)=\operatorname*{extr}_{\Theta_1}\Bigg\{&\ln\left[\mathrm{e}^{-\frac{\mu}{2}\sum_a(y-s'^a)^2-\frac{\beta}{2}\sum_\alpha(y-s^\alpha)^2}\right]_{y,\{s'^a,s^\alpha\}|\boldsymbol{P},\boldsymbol{C},\boldsymbol{Q}}\\&+\sum_{a,b}\frac{\tilde{P}_{ab}}{2}P_{ab}+\sum_{a,\alpha}\tilde{C}_{a\alpha}C_{a\alpha}+\sum_{\alpha,\beta}\frac{\tilde{Q}_{\alpha\beta}}{2}Q_{\alpha\beta}\\&+\frac{1}{\alpha}\ln\operatorname*{Tr}_{\{\xi^a\}\{x^\alpha\}}\operatorname{Tr}\mathrm{e}^{-\sum_{a,b}\frac{\tilde{P}_{ab}}{2}\xi^a\xi^b-\sum_{a,\alpha}\tilde{C}_{a\alpha}\xi^a(|\xi^1|_0x^\alpha)-\sum_{\alpha,\beta}\frac{\tilde{Q}_{\alpha\beta}}{2}(|\xi^1|_0x^\alpha)(|\xi^1|_0x^\beta)}\Bigg\},\end{aligned} \quad \text{(B.4)}$$

where $\Theta_1=\{\boldsymbol{P},\boldsymbol{C},\boldsymbol{Q},\tilde{\boldsymbol{P}},\tilde{\boldsymbol{C}},\tilde{\boldsymbol{Q}}\}$. For the extremizer, we search the subspace with $(P_{ab},\tilde{P}_{ab})$ equal to $(P,\tilde{P})$ $(a=b)$ or $(p,-\tilde{p})$ $(a\neq b)$; $(C_{a\alpha},\tilde{C}_{a\alpha})$ equal to $(C,-\tilde{C})$ $(a=1)$ or $(c,-\tilde{c})$ $(a\neq 1)$; and $(Q_{\alpha\beta},\tilde{Q}_{\alpha\beta})$ equal to $(Q,\tilde{Q})$ $(\alpha=\beta)$ or $(q,-\tilde{q})$ $(\alpha\neq\beta)$. This is the RS assumption for the present case. Thus, we obtain

$$\begin{aligned}\psi_1(n,\nu,\beta,\mu,\kappa)=\operatorname*{extr}_{\tilde{\Theta}_1^{\mathrm{LS}},\Theta_1}\Bigg\{&\ln\int\mathrm{D}y\mathrm{D}z\mathrm{D}w\left(\int\mathrm{D}v\mathrm{e}^{-\frac{\mu}{2}\left(\sigma_y y-\sqrt{p}\,w-\sqrt{P-p}\,v\right)^2}\right)^{n-1}\\&\times\int\mathrm{D}v\mathrm{e}^{-\frac{\mu}{2}\left(\sigma_y y-\sqrt{p}\,w-\sqrt{P-p}\,v\right)^2}\left(\int\mathrm{D}u\,\mathrm{e}^{-\frac{\beta}{2}\left(\sigma_y y-\frac{c}{\sqrt{p}}w-\frac{C-c}{\sqrt{P-p}}v-\sqrt{q-\frac{(C-c)^2}{P-p}-\frac{c^2}{p}}z-\sqrt{Q-q}\,u\right)^2}\right)^\nu\\&+\frac{1}{2}n\tilde{P}P-\frac{1}{2}n(n-1)\tilde{p}p-\nu\tilde{C}C-(n-1)\nu\tilde{c}c+\frac{1}{2}\nu\tilde{Q}Q-\frac{1}{2}\nu(\nu-1)\tilde{q}q\\&+\frac{1}{\alpha}\ln\int\mathrm{D}z\mathrm{D}w\mathrm{D}v\mathrm{D}u\left(\operatorname*{Tre}_{\xi}^{-\frac{\tilde{P}+\tilde{p}}{2}\xi^2+\left(u\sqrt{\tilde{p}-\tilde{c}}+v\sqrt{\tilde{c}}\right)\xi}\right)^{n-1}\\&\times\operatorname*{Tre}_{\xi}^{-\frac{\tilde{P}+\tilde{p}+\tilde{C}-\tilde{c}}{2}\xi^2+\left(u\sqrt{\tilde{p}-\tilde{c}}+w\sqrt{\tilde{C}-\tilde{c}}+v\sqrt{\tilde{c}}\right)\xi}\\&\times\left(\operatorname*{Tre}_{x}^{-\frac{\tilde{Q}+\tilde{q}+\tilde{C}-\tilde{c}}{2}(|\xi|_0x)^2+\left(z\sqrt{\tilde{q}-\tilde{c}}+w\sqrt{\tilde{C}-\tilde{c}}+v\sqrt{\tilde{c}}\right)|\xi|_0x}\right)^\nu\Bigg\},\end{aligned} \quad \text{(B.5)}$$

where $\tilde{\Theta}_1^{\mathrm{LS}}=\{C,c,Q,q,\tilde{C},\tilde{c},\tilde{Q},\tilde{q}\}$ and $\Theta_1=\{P,p,\tilde{P},\tilde{p}\}$.

J. Stat. Mech. (2016) 063302

The free-energy density f_1 is now derived as

$$f_1(\mu,\kappa) = -\lim_{n\to 0}\lim_{\nu\to 0}\frac{1}{\mu n}\psi_1(n,\nu,\beta,\mu,\kappa) \tag{B.6}$$

$$= \operatorname*{extr}_{\tilde{\Theta}_1}\left\{\frac{1}{2\mu}\ln(1+\mu(P-p)) + \frac{1}{2}\frac{P+\sigma_y^2}{1+\mu(P-p)} - \frac{1}{2\mu}(\tilde{P}P+\tilde{p}p)\right.$$
$$\left. -\frac{1}{\mu\alpha}\int \mathrm{D}v \ln \operatorname*{Tr}_{\xi} e^{-\frac{\tilde{P}+\tilde{p}}{2}\xi^2 + v\sqrt{\tilde{p}}\xi}\right\}. \tag{B.7}$$

In the limit $\mu\to\infty$, we introduce $\chi_p = \mu(P-p)$, $\hat{P} = \mu^{-1}(\tilde{P}+\tilde{p})$, and $\hat{\chi}_p = \mu^{-2}\tilde{p}$, which are assumed to be of the order $O(1)$. Taking the $\mu\to\infty$ limit in (B.7) leads to (34).

On the other hand, in order to evaluate ϵ_1^{LS}, in addition to $\hat{\Theta}_1 = P, \chi_p, \hat{P}, \hat{\chi_p}$, in taking the limit $\mu\to\infty$ we define the parameters $\chi_c = \beta(C-c)$, $\chi_q = \beta(Q-q)$, $\hat{C} = \beta^{-1}(\tilde{C}+\tilde{c})$, $\hat{\chi}_c = \beta^{-2}\tilde{c}$, $\hat{Q} = \beta^{-1}(\tilde{Q}+\tilde{q})$, and $\hat{\chi}_q = \beta^{-2}\tilde{q}$, which are assumed to be of the order $O(1)$. Then, through the formula

$$\epsilon_1^{\mathrm{LS}} = \lim_{\beta\to\infty}\lim_{\mu\to\infty}\lim_{n\to 0}\lim_{\nu\to 0} -\frac{1}{\beta\nu}\psi_1(n,\nu,\beta,\mu,0), \tag{B.8}$$

we obtain (42).

B.2. The limit $\alpha\to 0$ in the ℓ_1 case

The EOSs (35) show that in the limit $\alpha\to 0$ we have

$$\chi_p, \hat{P}, \hat{\chi}_p = O(1). \tag{B.9}$$

From (36) and the asymptotic formula of the complementary error function $\mathrm{erfc}(\cdot)$, we see in the limit $\alpha\to 0$ we have

$$\frac{e^{-\frac{1}{2}\theta^2}}{\alpha\sqrt{2\pi}\,\theta} = O(1), \Rightarrow \theta = O(\sqrt{|\ln\alpha|}) \to \infty, \tag{B.10}$$

which is realized by controlling λ as $O(\sqrt{|\ln\alpha|})$. Using these scalings, and the asymptotic expansion of the complementary error function for large θ in (35*d*), we obtain

$$P = O(\theta^{-2}) = O(|\ln\alpha|^{-1}) \to 0. \tag{B.11}$$

By inserting these scalings into (38), we obtain (48).

The asymptotic form of ϵ_1^{LS} can be similarly obtained. Following some lengthy but straightforward calculations, we obtain

$$\hat{\chi}_q = O(|\ln\alpha|^{-1}) \to 0, \tag{B.12a}$$

$$\hat{Q} = O(1), \tag{B.12b}$$

$$\hat{\chi}_c = O(|\ln \alpha|^{-1}) \to 0, \tag{B.12c}$$

$$\hat{C} = O(1), \tag{B.12d}$$

$$\chi_q = O(1), \tag{B.12e}$$

$$Q = O(|\ln \alpha|^{-1}) \to 0, \tag{B.12f}$$

$$\chi_c = O(1), \tag{B.12g}$$

$$C = O(|\ln \alpha|^{-1}) \to 0. \tag{B.12h}$$

By substituting these scalings into (44*c*), we obtain (49).

References

[1] Pearson K 1901 *Phil. Mag.* **2** 559–72
[2] Eckart C and Young G 1936 *Psychometrika* **1** 211–8
[3] Su X and Khoshgoftaar T M 2009 *Adv. Artif. Intell.* **2009** 421425
[4] Markovsky I 2012 *Low Rank Approximation: Algorithms, Implementation, Applications* (Berlin: Springer)
[5] Olshausen B A and Field D J 1996 *Nature* **381** 607–9
[6] Olshausen B A and Field D J 1997 *Vis. Res.* **37** 3311–25
[7] Olshausen B A and Field D J 2004 *Curr. Opin. Neurobiol.* **14** 481–7
[8] Terashima H and Hosoya H 2009 *Netw.: Comput. Neural Syst.* **20** 253–67
[9] Terashima H and Okada M 2012 *Adv. Neural Inf. Process. Syst.* **25** 2312–20
[10] Terashima H, Hosoya H, Tani T, Ichinohe N and Okada M 2013 *Neurocomputing* **103** 14–21
[11] Donoho D L 2006 *IEEE Trans. Inform. Theory* **52** 1289–306
[12] Candès E J and Tao T 2005 *IEEE Trans. Inform. Theory* **51** 4203–15
[13] Candès E J, Romberg J and Tao T 2006 *IEEE Trans. Inform. Theory* **52** 489–509
[14] Candès E J and Tao T 2006 *IEEE Trans. Inform. Theory* **52** 5406–25
[15] Natarajan B K 1995 *SIAM J. Comput.* **24** 227–34
[16] Davis G, Mallat S and Avellaneda M 1997 *Constr. Approx.* **13** 57–98
[17] Temlyakov V N 1998 *Adv. Comput. Math.* **8** 249–65
[18] Temlyakov V N 1999 *J. Approx. Theory* **98** 117–45
[19] Temlyakov V N 2000 *Adv. Comput. Math.* **12** 213–27
[20] Temlyakov V N 2003 *Found. Comput. Math.* **3** 33–107
[21] Gilbert A C, Muthukrishnan S and Strauss M J 2003 *Proc. 40th Annu. ACM-SIAM Symp. on Discrete Algorithms* pp 243–52
[22] Tropp J A, Gilbert A C, Muthukrishnan S and Strauss M J 2003 *Proc. Int. Conf. on Image Processing* pp I37–40
[23] Tropp J A 2004 *IEEE Trans. Inform. Theory* **50** 2231–42
[24] Donoho D L, Elad M and Temlyakov V N 2006 *IEEE Trans. Inform. Theory* **52** 6–18
[25] Foucart S and Rauhut H 2013 *A Mathematical Introduction to Compressive Sensing* (Berlin: Springer)
[26] Rish I and Grabarnik G 2014 *Sparse Modeling: Theory, Algorithms, and Applications* (Boca Raton, FL: CRC Press)
[27] Nishimori H 2001 *Statistical Physics of Spin Glasses and Information Processing: an Introduction* (Oxford: Oxford University Press)
[28] Dotsenko V 2001 *Introduction to the Replica Theory of Disordered Statistical Systems* (Cambridge: Cambridge University Press)
[29] Swendsen R H and Wang J-S 1986 *Phys. Rev. Lett.* **57** 2607–9
[30] Hukushima K and Nemoto K 1996 *J. Phys. Soc. Japan* **65** 1604–8
[31] Pati Y C, Rezaiifar R and Krishnaprasad P S 1993 *Conf. Record of 27th Asilomar Conf. on Signals, Systems and Computers* pp 40–4

J. Stat. Mech. (2016) 063302

[32] Davis G M, Mallat S G and Zhang Z 1994 *SPIE J. Opt. Eng.* **33** 2183–91
[33] Donoho D L, Maleki A and Montanari A 2009 *Proc. Natl Acad. Sci. USA* **106** 18914–9
[34] Monasson R and O'Kane D 1994 *Europhys. Lett.* **27** 85–90
[35] Tibshirani R 1996 *J. R. Stat. Soc.* B **58** 267–88
[36] Cohen A, Dahmen W and DeVore R 2009 *J. Am. Math. Soc.* **22** 211–31
[37] Das A and Kempe D 2008 *Proc. 40th Annu. ACM Symp. on Theory of Computing* pp 45–54
[38] Das A and Kempe D 2011 *Proc. 28th Int. Conf. on Machine Learning* pp 1057–64
[39] Franz S and Parisi G 1997 *Phys. Rev. Lett.* **79** 2486
[40] Ferrenberg A M and Swendsen R H 1988 *Phys. Rev. Lett.* **61** 2635–8
[41] Skodras A, Christopoulos C and Ebrahimi T 2001 *IEEE Signal Proc. Mag.* **18** 36–58
[42] Candès E J and Wakin M B 2008 *IEEE Signal Proc. Mag.* **25** 21–30
[43] Rubinstein R, Bruckstein A M and Elad M 2010 *Proc. IEEE* **98** 1045–57
[44] Gleichman S and Eldar Y C 2011 *IEEE Trans. Inf. Theory* **57** 6958–75
[45] Sakata A and Kabashima Y 2013 *Europhys. Lett.* **103** 28008

J. Stat. Mech. (2016) 063302