Sparsely-Spread CDMA - a statistical mechanics based analysis

Jack Raymond and David Saad

Neural Computation Research Group, Aston University, Aston Triangle, Birmingham, B4 7EJ

E-mail: jack.raymond@physics.org

Abstract.

Sparse Code Division Multiple Access (CDMA), a variation on the standard CDMA method in which the spreading (signature) matrix contains only a relatively small number of non-zero elements, is presented and analysed using methods of statistical physics. The analysis provides results on the performance of maximum likelihood decoding for sparse spreading codes in the large system limit. We present new results for both cases of regular and irregular spreading matrices for the binary additive white gaussian noise channel (BIAWGN) with a comparison to the canonical (dense) random spreading code.

PACS numbers: 64.60.Cn, 75.10.Nr, 84.40.Ua, 89.70.+c

AMS classification scheme numbers: 68P30,82B44,94A12,94A14

1. Background

The subject of multiuser communications is one of great interest from both theoretical and engineering perspectives [1]. Code Division Multiple Access (CDMA) is a particular method for allowing multiple users to access channel resources in an efficient and robust manner, and plays an important role in the current preferred standards for allocating channel resources in wireless communications. CDMA utilises channel resources highly efficiently by allowing many users to transmit on much of the bandwidth simultaneously, each transmission being encoded with a user specific signature code. Disentangling the information in the channel is possible by using the properties of these codes and much of the focus in CDMA is on developing efficient codes and decoding methods.

In this paper we study a variant of the original method, sparse CDMA, where the spreading matrix contains only a relatively small number of non-zero elements as was originally studied and motivated in [2] . While the straightforward application of sparse CDMA techniques to uplink multiple access communication is rather limited, as it is difficult to synchronise the sparse transmissions from the various users, the method can be highly useful for frequency and time hopping. In frequency-hopping code division multiple access (FH-CDMA), one repeatedly switches frequencies during radio transmission, often to minimize the effectiveness of interception or jamming of telecommunications. At any given time step, each user occupies a small (finite) number of the (infinite) M-ary frequency-shift-keying (MFSK) chip/carrier pairs (with gain G, the total number of chip-frequency pairs is MG.) Hops between available frequencies can be either random or preplanned and take place after the transmission of data on a narrow frequency band. In time-hopping (TH-)CDMA, a pseudo-noise sequence defines the transmission moment for the various users, which can be viewed as sparse CDMA when used in an ultra-wideband impulse communication system. In this case the sparse time-hopping sequences reduces collisions between transmissions.

Recently it has been noted that certain types of structurally disordered codes, in particular low-density codes, may have particularly valuable features and are also particularly suitable for analysis by methods developed within the framework of statistical mechanics of disordered systems [3, 4]. For the purposes of the current analysis the replica method is used, a powerful method known to give exact results for many systems. Supplementary analysis and interpretation are presented from a statistical physics perspective.

This work follows the seminal paper of Tanaka [5], and more recent extensions [6], in utilising the replica analysis for randomly spread CDMA with discrete inputs, which established many of the properties of random densely-spread CDMA with respect to several different detectors including Maximum A Posteriori (MAP), Marginal Posterior Maximiser (MPM) and minimum mean square-error (MMSE). Sparsely-spread CDMA differs from the conventional CDMA, based on dense spreading sequences, in that any user only transmits to a small number of chips (by comparison to transmission on all chips in the case of dense CDMA).

The feasibility of sparse CDMA for transmitting information was recently demonstrated [2] and several results for the case of a Gaussian approximation to the correlations (called effective medium approximation) amongst replicas was established, for the case of real (Gaussian distributed) input symbols. In a separate recent study, based on the belief propagation inference algorithm and a binary input prior distribution, sparse CDMA has also been considered as a route to proving results in the densely spread CDMA [7]. In addition this study demonstrated the existence of a *waterfall* phenomenon comparable to the dense code for a subset of ensembles. The waterfall phenomenon is observed in decoding techniques, where there is a dynamical transition between two statistically distinct solutions as the noise parameter is varied. Finally we note a number of pertinent studies concerning the effectiveness of belief propagation as an MPM decoding method [8, 9, 10, 11], and in combining sparse encoding (LDPC) methods with CDMA [12]. Many of these papers however consider the "extreme dilution" regime - in which the number of chip contributions is large but not $O(N)$.

The theoretical work regarding sparsely spread CDMA remained lacking in certain respects. As pointed out in [2], spreading codes with Poisson distributed number of non-zero elements, per chip and across users, are systematically failing in that each user has some probability of not contributing to any chips (transmitting no information). Even in the "partly regular" code [7] ensemble (where each user transmits on the same number of chips) some chips have no contributors owing to the Poisson distribution in chip connectivity, consequently the bandwidth is not effectively utilised. We circumvent this problem by introducing constraints to prevent this, namely taking regular signature codes constrained such that both the number of users per chip and chips per user take fixed integer values. Furthermore we present analytic and numerical analysis without resort to Gaussian approximations of any quantities. Using new tools from statistical mechanics we are able to cast greater light on the nature of the binary prior transmission process. Notably the nature of the decoding state space and relative performance of sparse ensembles versus dense ones across a range of noise levels; and importantly, the question of how the phase coexistence found by Tanaka [5] extends to sparse ensembles, especially close to the transition points determined for the dense ensemble. Phase coexistence is here meant in the same sense as the water-ice phase coexistence, where two stable solutions (decodings) coexist for some parameterisation.

Within this paper we demonstrate the superiority of regular sparsely spread CDMA code over densely spread codes in certain respects, for example, the anticipated bit error rate arising in decoding is improved in the high noise regime and the phase coexistence behaviour is less pervasive. Furthermore, to utilise belief propagation for such an ensemble is certain to be significantly faster and less computationally demanding [13], which has clear power-consumption implications. Other practical issues of implementation, the most basic being non-synchronisation and power control, require detailed study and may make fully harnessing these advantages more complex and application dependent.

The paper is organised as follows: In section 2 we will introduce the general framework and notation used, while the methodology used for the various codes will be presented in

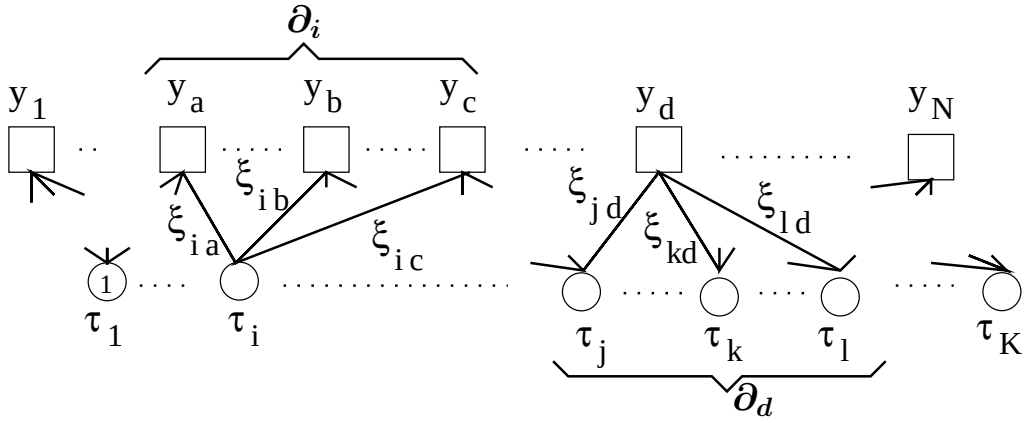

Figure 1. A bi-partite graph is useful for visually realising a problem. A user node i at the bottom interacts through its set of neighbouring factor nodes (∂i) to which it connects. The factor nodes are determined through a similar neighborhood. The interaction at each factor is conditioned on interaction strengths ξ and y_μ (which is an implicit function of the noise ω_μ, and neighbouring input bits $\boldsymbol{b}$ and interaction strengths ξ), assuming a uniform prior on the bits. The statistical mechanics reconstruction problem associates dynamical variables $\boldsymbol{\tau}$ to the user nodes that interact through the factors. The thermodynamical equilibrium state of this system then describes the theoretical performance of optimal detectors.

section 3. The main results for the various codes will be presented in section 4 followed by concluding remarks in section 5.

2. The model

We consider a standard model of CDMA consisting of K users transmitting in a bit interval of N chips. We assume a model with perfect power control and synchronisation, and consider only the single bit interval. In our case the received signal is described by

$$\boldsymbol{y} = \sum_{k=1}^{K} [\boldsymbol{s}_k b_k] + \boldsymbol{\omega} \ , \tag{1}$$

where the vector components describe the values for distinct chips: $\boldsymbol{s}_k$ is the spreading code for user k, $b_k = \pm 1$ is the bit sent by user k (binary input symbols), and $\boldsymbol{\omega}$ is the noise vector. Appropriate normalisation of the power is through the definition of the signature matrix ($\mathbf{s}$). It is possible to include a user or chip specific amplitude variation, which may be due to fading or imperfect power control. We consider a model without these effects. The spreading codes are sparse so that in expectation only C of the elements in vector $\boldsymbol{s}_k$ are non-zero. If, with knowledge of the signature matrix in use, we assume the signal has been subject to additive white Gaussian channel noise of variance σ_0^2/β, where σ_0^2 is the variance of the true channel noise $\langle \omega^2 \rangle$, we can write the likelihood of the transmitted bits

$\boldsymbol{\tau}$ (unknowns given the particular instance) using Bayes Theorem

$$P(\boldsymbol{\tau}|\boldsymbol{y}) = \prod_{\mu=1}^{N} \left[\frac{\sqrt{\beta}}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-\beta}{2\sigma_0^2} \left(\sum_{k=1}^{K} [s_{\mu k}(b_k - \tau_k)] + \omega_\mu \right)^2 \right) \right] P(\boldsymbol{\tau}) \ , \tag{2}$$

and from this define bit error rate, mutual information, and other quantities. The statistical mechanics approach from here is to define a Hamiltonian and partition function from which the various statistics relating to this probability distribution may be determined - and hence all the usual information theory measures. A suitable choice for the Hamiltonian is

$$\mathcal{H}(\boldsymbol{\tau}) = \sum_{\mu=1}^{N} \frac{1}{2\sigma_0^2} \left(\sum_{k=1}^{K} [s_{\mu k}(b_k - \tau_k)] + \omega_\mu \right)^2 + \sum_{k=1}^{K} h_k \tau_k \ . \tag{3}$$

We can here identify $\boldsymbol{\tau}$ as the dynamical variables in the reconstruction problem (dependence shown explicitly). The other quenched variables (parameters), describing the instance of the disorder, are the signature matrix ($\mathbf{s}$), noise ($\boldsymbol{\omega}$) and the inputs ($\boldsymbol{b}$). The variables h_k describe our prior beliefs about the inputs (the rates of each user), and to consider the case of all rates being equal, or some simple statistical correlation is possible within this framework. However, we assume that all users have the same rate, and so $h_k = H$, where maximal rate corresponds to unbiased bits ($H = 0$). The properties of such a system may be reflected in a factor (Tanner) graph, a bipartite graph in which users and chips are represented by nodes (see figure 1).

The calculation we undertake is specific to the case of the thermodynamic limit in which the number of chips $N \to \infty$ whilst the load $\alpha = K/N$ is fixed. Note that α is termed β in many CDMA papers, here we reserve β to mean the "inverse temperature" in a statistical mechanics sense (which here defines different classes of Bayes optimal detectors based on alternative prior beliefs). In all ensembles we may identify the parameter L as the mean number of contributions to each chip, and C as the mean number of contributions per user. As such the following also holds

$$\alpha = \frac{K}{N} = \frac{L}{C} \ . \tag{4}$$

The case in which α is greater than 1 will be called oversaturated, since more than one bit is being transmitted per chip.

The calculations presented henceforth are specific to the case of memoryless noise, drawn from a single distribution. Defining normalised spreading codes such that $\sum_k \boldsymbol{s}_k.\boldsymbol{s}_k = N$ the model signal to noise ratio over a chip interval is becomes

$$SNR = \frac{1}{\sigma_0^2} \tag{5}$$

We henceforth refer to "power spectral density" (PSD) as a measure of the system noise, this is $SNR/2$ - the factor two being connected with physical considerations in implementing the model.

2.1. Code Ensembles

We consider several code ensembles we call irregular, partly regular and regular, which differ in the constraints placed open the signature matrix. The probability distribution

$$P(\mathbf{s}) = \mathcal{N}\left(\prod_{\mu}\left\langle \delta(\sum_{k} s_{\mu k} - \tilde{L})\right\rangle_{P(\tilde{L})}\right)\left(\prod_{k}\left\langle \delta(\sum_{\mu} s_{\mu k} - \tilde{C})\right\rangle_{P(\tilde{C})}\right)\prod_{\mu}\prod_{k} P(s_{\mu,k}) \ , \quad (6)$$

where $\mathcal{N}$ is a normalising constant, $P(\tilde{L})$ is a probability distribution of mean L, $P(\tilde{C})$ is a probability distribution of mean C, and $P(s_{\mu k})$ is the marginal probability distribution which is common to all ensembles

$$P(s_{\mu k}) = (1 - \frac{C}{N})\delta(s_{\mu k}) + \frac{C}{N}\delta(s_{\mu k} - \xi) \ . \quad (7)$$

The interaction strength ξ, is drawn randomly from a single distribution with zero measure at $\xi = 0$, and finite moments, in any instance of a code

$$\phi(\xi) = P(s_{\mu k} = \xi | s_{\mu k} \neq 0) \ . \quad (8)$$

Unlike the dense case the details of this distribution will effect results (though only marginally as reported in [2]). We here investigate the case of Binary Phase Shift Keying (BPSK) which corresponds to a uniform distribution on $\{-\frac{1}{\sqrt{L}}, \frac{1}{\sqrt{L}}\}$, though the analytic results presented are applicable to any distribution of mean square $= 1/L$. Note that disorder in ξ is not a necessity, the case $\xi = 1/\sqrt{L}$ also allows decoding in sparse ensembles.

The case where $P(\tilde{L})$ and $P(\tilde{C})$ are Poissonian identifies the irregular ensemble - where the connections between chips and users are independently distributed. The second distribution called partly regular has $P(\tilde{C}) = \delta_{C,\tilde{C}}$, in which the chip connectivity is again Poisson distributed with mean L, but each user contributes to exactly C chips. This prevents the systematic failure inherent in the irregular ensemble since therein an extensive number of users fail to transmit on any chips. If in addition to the aforementioned constraint all chips receive exactly L contributions, $P(\tilde{L}) = \delta_{L,\tilde{L}}$, the ensemble is called regular. Regular chip connectivity amongst other things prevents the systematic inefficiency due to leaving some chips unaccessed by any users. The case of Poissonian distributions is that in which there is no global control. In many engineering applications constraining users individually (non-Poissonian $P(\tilde{C})$) is practical, whereas coordination between users (non-Poissonian $P(\tilde{L})$) is difficult. The practicalities of implementing the different ensembles we consider are application specific: the advantages inherent in distributing channel resources more evenly amongst users may be lost to practical implentation problems.

3. Methodology

3.1. Information Capacity Lower Bound

In several places the inferiority of the irregular ensembles (partly regular and fully Poissonian) is suggested based on the knowledge that codes which leave a portion of the

bandwidth unutilised cannot be optimal. This annealed approximation here corresponds to a detector reconstructing bits based only on the value of a single chip, so will be tight only in some special cases.

The spectral efficiency is defined as the mutual information between the received signal and reconstructed bits per chip. When averaged over code ensembles this may be identified in connection with the self averaged entropy density

$$\left\langle \frac{1}{N} I(\boldsymbol{\tau};\boldsymbol{y}) \right\rangle = \frac{1}{N} \langle \langle \log_2 P(\boldsymbol{\tau}|\boldsymbol{y}) \rangle_{P_0(\boldsymbol{\tau}|\boldsymbol{y})} + \langle \log_2 P(\boldsymbol{y}) \rangle_{P_0(\boldsymbol{y})} \rangle = \alpha - s/\ln 2 \ , \quad (9)$$

where the subscript zero indicates that the true (generative), rather than model (2), probability distribution and α comes from the entropy of the transmitted bits K divided by the number of chips. The outer average is over the instance of the codes (assumed self averaging). As is common in statistical physics we are able to place a bound on the free energy and thence the entropy by considering an annealed average. In this case factorising the partition sum as single chip components

$$\mathcal{F} = -\lim_{N\to\infty} \frac{1}{N\beta} \langle \ln Z \rangle_I \geq -\lim_{N\to\infty} \frac{1}{N\beta} \sum_{\mu} \langle \ln Z_\mu \rangle_I \ .$$

The average is taken with respect to many variables, but only those in the neighbourhood of μ are relevant. The difference between the ensembles is then independent of the user connectivity. One can carry out the averages over noise and code ensembles numerically. It is necessary also to evaluate the factorised energy to calculate the entropy, and hence a lower bound for spectral efficiency. Note that this result then suggests a scale for the minimal load (α) a channel might be able to bear given L and ω, as well as a comparison between the different ensembles.

The dependency of chip connectivity distribution may be calculated to leading order in the special case that the channel noise $\sigma_0^2 \to 0$. Note that in this case the minimum energy state of $1/(2\beta)$ is always attainable when sites are factorised and hence is the achieved value, and thus the free energy is affine to the spectral efficiency and independent of ensemble type. The entropy on the other hand is not a simple constant due to degeneracy in the ground state dependent on $P(\tilde{L})$

$$\mathcal{F} \geq \left\langle \left\langle -\frac{1}{\beta} \ln \left(\exp\left\{ -\frac{\beta\omega^2}{2\sigma^2} \right\} \sum_{\tau_i | i=1..\tilde{L}} \delta(\sum_{i=1}^{\tilde{L}} \xi_i (b_i - \tau_i)) \right) \right\rangle_{\{b_i,\xi_i\}|i=1..\tilde{L}} \right\rangle_{P(\tilde{L})} . \quad (10)$$

It is straightforward to separate this free energy to give an entropic and energetic part. The entropic part may be written for the case of uniform BPSK modulation as (for a given chip of connectivity $\tilde{L}$)

$$s(\tilde{L}) = \sum_{p=0}^{\tilde{L}} \frac{1}{2^{\tilde{L}}} \binom{\tilde{L}}{p} \ln \left(\sum_{i}^{\min(p,\tilde{L}-p)} \binom{\tilde{L}-p}{i} \binom{p}{i} i! \right) \quad (11)$$

This prooves to be a concave function, so that the ensemble with $\tilde{L} = L$ (regular) has greater spectral efficiency than any other ensemble (and hence minimal entropy). However,

the free energy is asymptotically linear in $\tilde{L}$ suggesting that for large L the difference between the ensembles investigated becomes negligable.

It is possible to consider the opposite limit $\sigma_0^2 \to \infty$ perturbatively. We found that to the leading 3 orders in $\frac{1}{\sigma_0}$ were identical for all code ensembles of the same mean chip connectivity. Thus a simple annealed analysis suggests favourable performance for the case of regular chip connectivity in the low noise, small L regime.

3.2. Replica Method Outline

We determine the properties of the above model including correlations due to the full interaction structure through a statistical mechanics framework, using the replica method. Note nevertheless that within the assumptions made here identical expressions may be obtained by using the cavity method, which is more intuitive on some levels. From the expression of the Hamiltonian (3) we may identify a free energy and partition function as:

$$f = -\frac{1}{N\beta} \ln Z \qquad\qquad Z = \mathrm{Tr}_{\boldsymbol{\tau}} \exp\left(-\beta \mathcal{H}(\boldsymbol{\tau})\right) \ .$$

To progress we make use of the anticipated *self-averaging* properties of the system. The assumption being that in the large system limit any two randomly selected instances will, with high probability, have indistinguishable statistical properties. This assumption has firm foundation in several related problems [14], and is furthermore intuitive after some reflection. If this assumption is true then the statistics of any particular instance can be described completely by the free energy averaged over all instances of the disorder. We are thus interested in the quantity

$$\mathcal{F} = \langle f \rangle = -\lim_{N\to\infty} \frac{1}{N\beta} \langle \ln Z \rangle_I \ , \tag{12}$$

where the angled brackets represent the weighted averages over I (the instances).

Taking the average over disorder is a difficult problem except in special cases, which is why we make use of the replica identity

$$\langle \ln Z \rangle_I = \lim_{n\to 0} \frac{\partial}{\partial n} \langle Z^n \rangle_I \ . \tag{13}$$

We can model the system now as one of interacting replicas, where Z^n is decomposed as a product of an integer number of partition functions with conditionally independent (given the instance of the disorder) dynamical variables. The discreteness of replicas is essential in the first part of the calculation, but a continuation to the real numbers is required in taking $n \to 0$ – this is a notorious assumption, which rigorous mathematics can not yet justify. However, we shall assume validity and since the methodology for the sparse structures is well established [15, 14] we omit our particular details. The final functional form for the free energy is determined from

$$\langle Z^n \rangle = \int \prod_{\boldsymbol{\sigma},b} \left[dP(\boldsymbol{\sigma},b) dP(\boldsymbol{\sigma},b) \right] \exp\{ \ln \mathcal{N} + N(G_1(n) + G_2(n) + G_3(n)) \} \tag{14}$$

$$G_1(n) = \ln \left\{ \int \left[\prod_\alpha \frac{\mathrm{d}\lambda_\alpha}{\sqrt{2\pi}} \right] \exp\left\{ -\sum_\alpha \lambda_\alpha^2/2 \right\} \left\langle \exp\left\{ \frac{\mathrm{i}\sqrt{\beta}\omega}{\sigma_0} \sum_\alpha \lambda_\alpha \right\} \right\rangle_\omega \right. \tag{15}$$

$$\times \left\langle \frac{e^{-L}L^{\tilde{L}}}{\tilde{L}!} \left(\sum_{b,\boldsymbol{\sigma}} P(\boldsymbol{\sigma},b) \left\langle \exp\left\{ \frac{\mathrm{i}\sqrt{\beta}\xi}{\sigma_0} \sum_{\alpha} \lambda_\alpha (b-\tau_\alpha) \right\} \right\rangle_{\xi} \right) \right\rangle_{P(\tilde{L})} \Bigg\} \quad (16)$$

$$G_2(n) = \sum_{\boldsymbol{\sigma},b} P(\boldsymbol{\sigma},b)\hat{P}(\boldsymbol{\sigma},b) \quad (17)$$

$$G_3(n) = \alpha \ln \left\langle \sum_{\boldsymbol{\tau}} \exp\left\{ \beta H \sum_{\alpha} \tau \right\} \left\langle \frac{1}{(-\alpha)^{\tilde{C}}\tilde{C}!} \left(\hat{P}(\boldsymbol{\sigma},b) \right)^{\tilde{C}} \right\rangle_{P(\tilde{C})} \right\rangle_{b} \quad (18)$$

where $\mathcal{N}$ is a constant due to normalising the ensembles (6). This expression may be evaluated at the saddle point to give an expression for the free energy. In the term (18) we account for the cases in which the marginalised probability distribution $P(b)$ and assumed marginal probability distribution (described by H) are asymmetric. In the case of maximal rate which we will consider, the b average is trivial and $H = 0$. Provided that in addition either the noise or interaction strength distributions are symmetric then it is possible to remove the b dependence in the order parameters, since the symmetry $P(b,\boldsymbol{\sigma}) = P(-b,-\boldsymbol{\sigma})$ and $\hat{P}(b,\boldsymbol{\sigma}) = \hat{P}(-b,-\boldsymbol{\sigma})$ leaves the free energy invariant.

3.3. Replica Symmetric Equations

The concise form for our equations is attained using the assumption of replica symmetry (RS). This amounts to the assumption that the correlations amongst replicas are all identical, and determined by a unique shared distribution. The validity of this assumption may be self consistently tested (section 3.5). This assumption differs from that used by Yoshida and Tanaka [2] where the correlations are described by only a handful of parameters rather than a distribution once Replica Symmetry is assumed – this approach may therefore miss some of the detailed structure although it is easier to handle numerically. The order parameter in our case is given by

$$\sum_{b} P(b,b\boldsymbol{\tau}) = \int d\pi(x) \frac{1}{2^n} \prod_{\alpha} (1 + b\tau_\alpha x) \ ; \quad (19)$$

$$\sum_{b} P(b,b\boldsymbol{\tau}) = \hat{q} \int d\hat{\pi}(\hat{x}) \frac{1}{2^n} \prod_{\alpha} (1 + b\tau_\alpha \hat{x}) \ ; \quad (20)$$

$$(21)$$

where $\hat{q}$ is a variational normalisation constant and $\pi, \hat{\pi}$ are normalised distributions on the interval $[-1,1]$. From here onwards we may consider the case in which the bit variables τ_α and interaction strengths ξ are gauged to b.

Using Laplace's method, this gives the following expression for the (RS) free energy at the saddle point

$$\mathcal{F}_{RS} = -\frac{1}{\beta} \mathrm{Extr}_{\pi,\hat{\pi}} \frac{\partial}{\partial n} \left(\mathcal{G}_{1,RS}(\tilde{L})(n) + \mathcal{G}_{2,RS}(n) + \mathcal{G}_{3,RS}(\tilde{C})(n) \right) \quad (22)$$

where

$$\frac{\partial}{\partial n}|_{n=0}\mathcal{G}_{1,RS}(n) \doteq \left(-L\ln 2 + \left\langle \int \prod_{l=1}^{\tilde{L}} [\mathrm{d}\pi(x_l)] \left\langle \ln \mathrm{Tr}_{\{\tau_l=\pm 1\}} \chi_{\tilde{L}}(\boldsymbol{\tau};\{\xi\},\omega)\right\rangle_{\omega,\{\xi\}} \right\rangle_{P(\tilde{L})}\right) \quad (23a)$$

$$\chi_{\tilde{L}}(\boldsymbol{\tau};\{\xi\},\omega) = \exp\left(-\frac{\beta}{2\sigma^2}\left(\omega + \sum_{l=1}^{\tilde{L}}(1-\tau_l)\xi_l\right)^2\right)\prod_{l=1}^{\tilde{L}}(1+\tau_l x_l)\ ; \quad (23b)$$

$$\frac{\partial}{\partial n}|_{n=0}\mathcal{G}_{2,RS}(n) = -L\int \mathrm{d}\pi(x_c)\mathrm{d}\hat{\pi}(\hat{x}_c)\ln(1+x\hat{x}_c)\ ; \quad (23c)$$

$$\frac{\partial}{\partial n}|_{n=0}\mathcal{G}_{3,RS}(n) = \alpha\left\langle \int \prod_{c=1}^{\tilde{C}}[\mathrm{d}\hat{\pi}(\hat{x}_c)] \ln\left(\prod_{c=1}^{\tilde{C}}(1+\hat{x}_c)+\prod_{c=1}^{\tilde{C}}(1-\hat{x}_c)\right)\right\rangle_{P(\tilde{C})}\ . \quad (23d)$$

and the saddle point value for $\hat{w}$ $(= L)$ has been introduced. The averages over $\tilde{L}$ and $\tilde{C}$ encapsulate the differences amongst the ensembles.

Equation (23b) describes the interaction at a single chip in the factor graph (figure 1) of connectivity $\tilde{L}$. The parameter ξ_l and variable τ are the interaction strengths, and reconstructed bits respectively, both gauged to the transmitted bit, while ω is the instance of the chip noise.

The order variational distributions $\{\pi, \hat{\pi}\}$ are chosen so as to extremise (22). The self consistent equations attained by the saddle point method are:

$$\hat{\pi}(\hat{x}) = \left\langle \int \prod_{l=1}^{\tilde{L}}[\mathrm{d}\pi(x_l)] \left\langle \delta\left(\hat{x} - \frac{\mathrm{Tr}_{\{\tau_l=\pm1\}}\ \tau_{\tilde{L}+1}\ \bar{\chi}_{\tilde{L}}(\boldsymbol{\tau};\{\xi\},\omega)}{\mathrm{Tr}_{\{\tau_l=\pm1\}}\ \bar{\chi}_{\tilde{L}}(\boldsymbol{\tau};\{\xi\},\omega)}\right)\right\rangle_{\{\xi\},\omega}\right\rangle_{P_e(\tilde{L})} \quad (24a)$$

$$\bar{\chi}_{\tilde{L}}(\boldsymbol{\tau};\{\xi\},\omega) = \exp\left(-\frac{\beta}{2\sigma^2}\left(\omega + \sum_{l=1}^{\tilde{L}+1}(1-\tau_l)\xi_l\right)^2\right)\prod_{l=1}^{\tilde{L}}(1+\tau_l x_l) \quad (24b)$$

$$\pi(x) = \left\langle \int \prod_{c=1}^{\tilde{C}}[\mathrm{d}\hat{\pi}(\hat{x}_c)]\,\delta\left(x - \frac{\prod_{c=1}^{\tilde{C}}(1+\hat{x}_c) - \prod_{c=1}^{\tilde{C}}(1-\hat{x}_c)}{\prod_{c=1}^{\tilde{C}}(1+\hat{x}_c) + \prod_{c=1}^{\tilde{C}}(1-\hat{x}_c)}\right)\right\rangle_{P_e(\tilde{C})}\ . \quad (24c)$$

The variables $P_e(\tilde{L})$ and $P_e(\tilde{C})$ are here the excess degree distributions of the particular ensemble (6). For regularly constrained ensembles the chip and user excesses are $L-1$ and $C-1$ respectively. For Poissonian distributions the excess degree distribution is the same as for the full distribution.

The other quantities of interest may be determined from the probability distribution for the overlap of reconstructed and sent variables $m_k = \langle \tau_k \rangle$,

$$P(m) = \frac{1}{K}\sum_{k=1}^{K}\delta_{m_k,m} \propto \left\langle \int \prod_{c=1}^{\tilde{C}}[\mathrm{d}\hat{\pi}(\hat{x}_c)]\,\delta\left(m - \frac{\prod_{c=1}^{\tilde{C}}(1+\hat{x}_c) - \prod_{c=1}^{\tilde{C}}(1-\hat{x}_c)}{\prod_{c=1}^{\tilde{C}}(1+\hat{x}_c) + \prod_{c=1}^{\tilde{C}}(1-\hat{x}_c)}\right)\right\rangle_{P(\tilde{C})}, \quad (25)$$

and the entropy density may be calculated from the free energy density by use of the relation

$$s = \beta(e-f)\ , \quad (26)$$

where e is the energy.

3.4. Population Dynamics

Analysis of these equations is primarily constrained by the nature of equations (24*a*-24*c*). No exact solutions are apparent, and perturbative regimes about the ferromagnetic solution (which is only a solution for zero noise) are difficult to handle. Consequently we use population dynamics [16] – representing the distributions $\{\pi(x), \hat{\pi}(\hat{x})\}$ by finite populations (histograms) and iterating this distribution until convergence. In each field update $\{x_i, \hat{x}_i\}$, we consider only a single random sample for each integral, or average, in (24*a*-24*c*), and update the entire population ($\{x_i\}$ or $\{\hat{x}_i\}$) sequentially. It is hoped, and observed, that the large number of fields sampled in this way converges towards the true distribution to leading order in N, the number of fields in the population.

There is a strong analogy between the population dynamics algorithm and that of message passing on a particular instance of the graph. The iteration of the histogram implicit in (24*a*-24*c*) is equivalent to the propagation of a population of "cavity" magnetisations between factor (a) and user (i) nodes, which may be written as the self consistent equations:

$$x_{i\to a} = \frac{1}{\mathcal{N}_x}\left(\prod_{c\in\partial i\setminus a}(1+\hat{x}_c) - \prod_{c\in\partial i\setminus a}(1-\hat{x}_c)\right) ; \tag{27a}$$

$$\hat{x}_{a\to i} = \frac{1}{\mathcal{N}_{\hat{x}}}\mathrm{Tr}_{\{\tau_l=\pm 1\}}\tau_L \exp\left(-\frac{\beta}{2\sigma^2}\left(\omega + \sum_{l\in\partial a\setminus i}(1-\tau_l)\xi_l\right)^2\right) \times \prod_{l\in\partial a\setminus i}(1+\tau_l x_{l\to a}) ; \tag{27b}$$

where $\mathcal{N}_{x,\hat{x}}$ are the relevant normalisations, and the abbreviation ∂y indicates the set of nodes connected to y. In population dynamics the notion of a particular graph with labelled edges is absent however, and the only the distribution of the two types of magnetisations are relevant.

3.5. Stability Analysis

To test the stability of the obtained solutions we consider both the appearance of non-negative entropy, and a stability parameter defined through a consideration of the fluctuation dissipation theorem. The first criteria that the entropy be non-negative is based on the fact that physically viable solutions in discrete systems must have non-negative entropy so that any solution found not meeting this criteria must be based on bad premises; replica symmetry is a likely source.

The stability parameter λ is defined in connection with the cavity method for spin glasses [17] and tests local linear stability of the solutions. It is equivalent to testing the local stability of belief propagation equations as proposed in [18]. A necessary condition for the stability of the *RS* solution is that the "spin glass" susceptibility does not diverge. This condition ensures that fields are not strongly correlated – if they were then a single pair of distributions (W,$\hat{W}$) would not be sufficient to describe the replica correlations.

The spin glass susceptibility when averaged over instances may be defined [19]

$$\zeta = \sum_{d=0}^{\infty} X^d \left\langle \langle \tau_0 \tau_d \rangle_c^2 \right\rangle \tag{28}$$

where d is the distance between two nodes in the factor graph, the inner average denotes the connected correlation function between these nodes, X^d describes the typical number of variables at distance d, and the outer average is over instances of the disorder (self-averaging part). This quantity is not divergent provided that

$$\lambda = \ln \left[\lim_{d \to \infty} X \left\langle \langle \tau_0 \tau_d \rangle_c^2 \right\rangle^{\frac{1}{d}} \right] \tag{29}$$

is negative, since this indicates an asympoptically exponential decrease in the terms of (28) and hence convergence of the sum. In the thermodynamic limit the connected correlation function is dominated by a single direct path which may be decomposed as a chain of local linear susceptibilities:

$$\langle \tau_0 \tau_d \rangle_c \propto \prod_{(i,j)} \frac{\partial x_{i \to a}}{\partial \hat{x}_{b \to i}} \frac{\partial \hat{x}_{b \to i}}{\partial x_{j \to b}} \tag{30}$$

where (i,j) indicate the set of variables on the shortest path between nodes 0 and d in a particular instance of the graph ($27b$). This representation allows us to construct an estimation for λ numerically based on principles similar to population dynamics (later described).

4. Results

Results are presented here for the canonical case of Binary Phase Shift Keying (BPSK) where $\xi_l \in \{1, -1\}$ with equal probability. Furthermore we assume an AWGN model for the true noise ω (of variance σ_0^2). For evaluation purposes we assume the channel noise level is known precisely, so that $\beta = 1$, employing the Nishimori temperature [3]. This guarantees that the RS solution is thermodynamically dominant. Furthermore the energy takes a constant value at the Nishimori temperature and hence the entropy is affine to the free energy. Where of interest we plot the comparable statistics for the single user Gaussian channel (SUG), and the densely spread ensemble, each with MPM decoders – equivalent to maximum likelihood for individual bits.

For population dynamics two parallel populations are initialised either uniformly at random, or in the ferromagnetic state. These two populations are known to converge towards the unique solution, where one exists, from opposite directions, and so we can use their convergence as a criteria for halting the algorithm and testing for the appearance of multiple solutions. In the case where they converge to different solutions we can identify the solution converged to from the ferromagnetic initial state as a *good* solution - in the sense that it reconstructs well, and that arrived at from random initial state as a *bad* solution. In the equivalent BP propagation we cannot choose initial conditions equivalent to ferromagnetic – knowing the exact solution would of course makes the decoding redundant. We therefore expect the properties of the *bad* solution to be those

realisable by belief propagation (though clever algorithms may be able to escape to the good solution under some circumstances).

In order to approximate the stability parameter λ one also randomly initialises two histograms $\{V(v_i), \hat{V}(\hat{v}_i)\}$ with components updated in parallel with W or $\hat{W}$ respectively. If magnetisations $\{x_{i_1}..x_{i_k}\} \in W$ are used to create magnetisation $x_a \in \hat{W}$ then the corresponding update in $\hat{V}$ is

$$\hat{v}_a = \sum_j v_{i_j} \left(\frac{\partial \hat{x}_a}{\partial x_{i_j}} \right)^2 \tag{31}$$

and with similar assignments for the field update of W

$$v_i = \sum_j \hat{v}_{i_j} \left(\frac{\partial x_i}{\partial \hat{x}_{a_j}} \right)^2 \tag{32}$$

renormalising the populations at each step. The renormalisation factor for each population yield (dependent) estimations of λ. Estimation requires convergence in all histograms, and a sufficiently large population for numerical accuracy.

Computer resources however restrict the cases studied in detail to an intermediate PSD regime, and small L. In particular, the problem at low PSD, is the Gaussian noise average, which is poorly estimated, while at high PSD a majority of the histogram is concentrated at magnetisations $x, \hat{x} \approx 1$ not allowing sufficient resolution in the rest of the histogram.

Several different measures are calculated from the converged order parameter, indicating the performance of sparsely-spread CDMA. Using the converged histograms for the fields we are able to determine the quantities free energy, energy and a histogram for the probability distribution, from discretisations of the previously presented equations (24*a*-24*c*). Using the probability distribution we are also able to approximate the decoding bit error rate

$$P_b = \int \mathrm{d}P(x) \frac{1 - \mathrm{sign}(x)}{2} \; ; \tag{33}$$

multi-user efficiency

$$MuE = \frac{1}{SNR} \left[\mathrm{erfc}^{-1}(P_b) \right]^2 \; ; \tag{34}$$

and mutual information between sent and reconstructed bits per chip, $I(\boldsymbol{b}; \boldsymbol{\tau})/N$ (taking a factorised form given the RS assumption)

$$MI = \alpha \left(1 - \int \mathrm{d}P(m) \sum_\tau \frac{1 + \tau m}{2} \log_2 \frac{1 + \tau m}{2} \right) \; . \tag{35}$$

The spectral efficiency is the capacity $I(\boldsymbol{\tau}; \boldsymbol{y})$ per chip, which is affine to the entropy (and the free energy at the Nishimori temperature)

$$\nu = \alpha - s/\ln 2 \; . \tag{36}$$

Negative entropy can be identified when the measured spectral efficiency exceeds the load, and thermodynamic transition points correspond to points of coincident spectral efficiency.

Figure 2 demonstrates some general properties of the regular ensemble in which C:L = 3:3. Equations (24*a*-24*c*) were iterated using population dynamics and the relevant properties were calculated using the obtained solutions; the data presented is averaged over 100 runs and error-bars are omitted for brevity. Figure 2(a) shows the bit error rate in regular and Poissonian codes. The sparse codes demonstrate similar trends to the dense case except the irregular code, which show weaker performance in general, and in particular at high PSD in particular. Detailed trends can be seen in figure 2(b) that shows the multiuser efficiency. Codes with regular user connectivity show superior performance with respect to the dense case at low PSD. Figure 2(c) shows similar trends in the spectral efficiency and mutual information; the effect of the disconnected (user) component is clear in the fact that the irregular code fails to reach capacity at high noise levels. In general it appears the chip connectivity distribution is not critical in changing the trends present, unlike the user connectivity distribution. It was found in these cases (and all cases with unique solutions for given PSD), that the algorithm converged to non-negative entropy values and to a stability measure fluctuating about a value less than 0, as shown in figure 2(d). These points would indicate the suitability of the RS assumption.

The outperformance of dense codes by sparse ensembles with regular user connectivity is new to our knowledge. It also appears that Poissonian chip connectivity outperforms regular connectivity at low PSD. These facts indicate that it may be the effect of locality in the graph, the instance of noise experienced is not homogeneous for all users and this allows effective reconstruction in some neighbourhoods. The effect of these neighbourhoods must mitigate for the poor performance in other neighbourhoods with high instances of the noise. In the dense graph the noise level experienced is homogeneous, in the regular ensemble the core is well connected but with a notion of neighbourhood, and in the Poissonian chip connectivity case there are many neighbourhoods which are relatively isolated. This effect would be most pronounced in disconnected ensembles – TH/FH CDMA, which is why these are effective at rebuffing jamming and other types of interference. This explanation is further supported by experiments with increasing density at fixed α for all ensembles (figure 3).

Figure 3 indicates the effect of increasing density whilst maintaining α in the case of the regular code. As density is increased the statistics of the sparse codes approach that of the dense channel in all ensembles tested. For the irregular ensemble performance increases monotonically with density at all PSD. The rapid convergence to the dense case performance was elsewhere observed for partly regular ensembles, and ensembles based on a Gaussian prior input [2, 7]. At all densities the RS assumption was found to be validated in the stability parameter.

Figure 4 indicates the effect of channel load α on performance. We first explain results for codes in which only a single solution was found (without phase coexistence). For small values of the load a monotonic increase in the bit error rate, and capacity are observed as α is increased with C constant, as shown in figures 4(a) and 4(b), respectively. This matches the trend in the dense case, the dense code becoming superior in performance to the sparse codes as PSD increases. We found that for all ensembles there existed regimes

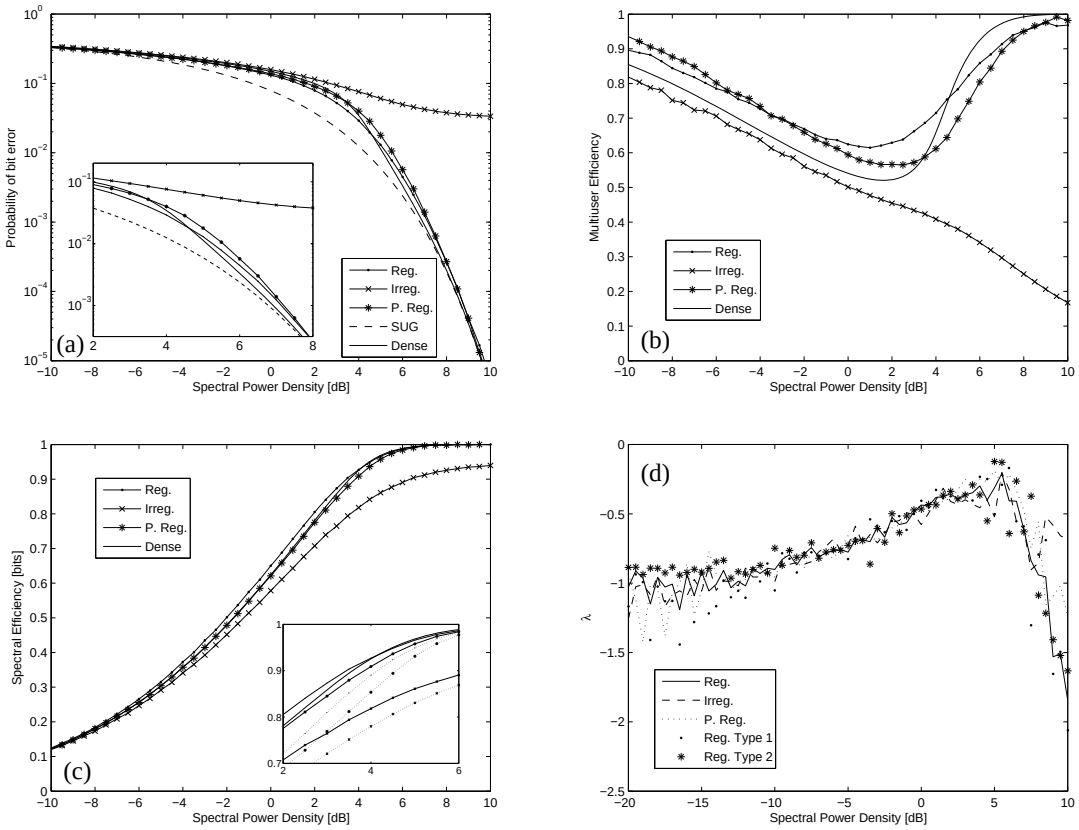

Figure 2. Performance of the sparse CDMA configuration $C : L = 3 : 3$; all data presented on the basis of 100 runs, error bars are omitted and are typically small in subfigures (a)-(c) the smoothness of the curves being characteristic of this level (numerical accuracy was excellent only at intermediate PSDs). (a) The bit error rate is limited by the disconnected component in the case of irregular codes, otherwise trends match the dense case, lower bounded by the SUG. (b) Multiuser efficiency indicates the regular user connectivity codes outperform the dense case below some PSD. (c) The spectral efficiency [——] demonstrates similar trends, the entropy being positive. The gap between the mutual information [· · · · · ·] and spectral efficiency is everywhere small and especially so at small and large PSD, indicating little information loss in the decoding process. (d) The two markers show the mean results for the two different stability estimates in the algorithm for the regular code. There are systematic errors at small PSD, and convergence is good only at intermediate PSD. The lines represent the average of these quantities for each ensemble – all ensembles show a cusp at some PSD, for $3 : 3$ codes the various ensembles shows very similar trends, indicating local stability everywhere.

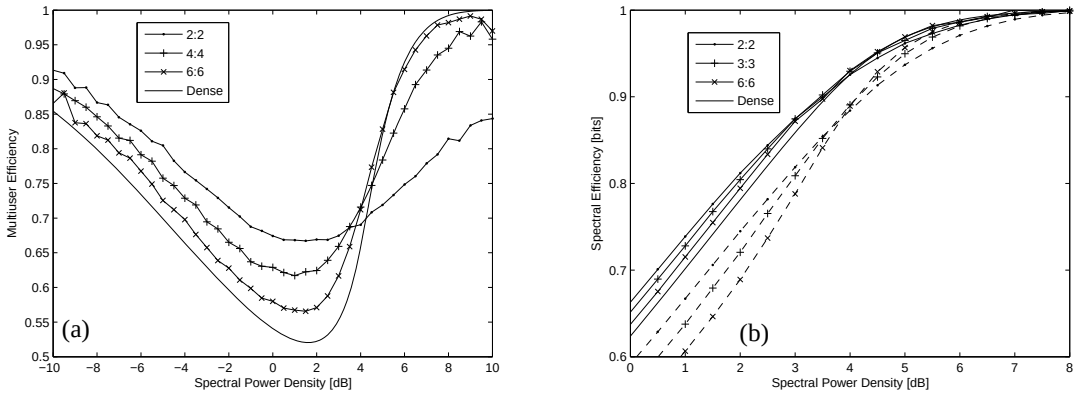

Figure 3. The effect of increasing density for the regular ensemble: (a) Multiuser efficiency, (b) spectral efficiency [——] and mutual information [– – –]. Data presented on the basis of 10 runs, error bars are omitted but of a size comparable with the smoothness of the curves. The performance of sparse codes rapidly approaches that of the dense code everywhere. The PSD threshold beyond which the dense code outperforms the sparse code is fairly stable.

with $\alpha > 1.49$ for which only a single stable solution existed, although the equivalent dense systems has two stable solutions at some PSD [5]. In all single valued regimes we observed positive entropy, and a negative stability parameter. However in cases of large α many features became more pronounced close to the dense case coexistence regime: notably the cusp in the stability parameter, gap between MI and ν and the gradient in Pb.

4.1. Phase Coexistence Regimes

In order to investigate the phase coexistence regime we investigated the states arrived at from random and ferromagnetic initial conditions. Separate heuristic convergence criteria were found for the histograms, and these seemed to work well for the good solution. For the bad solution (as shown figure 4) we simply present results after a fixed number of histogram updates (500) as all convergence criteria tested appeared either too stringent, to require experimentally inaccessible timescales, or did not capture the asymptotic values for important quantities like entropy.

The code we use to demonstrate the phase coexistence properties is the 6:3 regular code. Phase coexistence was found to exist above some PSD. In the case of the $6:3$ code we identified the onset of the bimodal distribution by noticing a divergence in the convergence time (the time for the ferromagnetic and random histograms to converge to a common distribution). The time for this to occur, in a heuristically chosen statistic and accuracy, is plotted in figure 4(d). By a naive linear regression across 3 decades we found a power law exponent of 0.59 and a transition point of $PSD = 10.23dB$, but cannot provide a goodness of fit measure to this data. This would represent the point at which at least two stable solutions co-exist (phase coexistence).

Only two stable solutions were found in the region beyond this critical point and

upto 12dB which we infer to be viable RS solutions. These appear statistically similar to those that exist in the equivalent dense system. The first is the good solution arrived at only from a near ferromagnetic initial histogram condition. This solution has only slightly worse performance than the high performance branch of the dense ensemble. The histogram is strongly peaked near a value 1, and it is suspected that insufficient resolution in this histogram is the reason for very noisy statistics in the stability parameter (although a noisy trend is negative and increasingly so with PSD). The entropy of this solution is positive.

The second solution is the bad solution arrived at from random initial conditions. This histogram representing this solution is very slow to converge. The reason for this may be that the solution is very close to instability as can be seen in the measured stability parameter. Furthermore the existence of a stable good solution may accentuate non-linear fluctuations in the dynamics. The negative entropy shown in 4(c) is strongly suspected to be a result of incomplete convergence in the bad solution, rather than an effect of standard replica symmetry breaking phenomena. Results are presented with the number of histogram updates set to 500, however this is not sufficient for a many of statistics, including entropy, to become stationary. In most other regimes only about 10 histogram updates were required for convergence.

At $PSD \approx 11dB$ one can observe the entropies of the two solutions crossing. This would indicate a second order thermodynamic transition point between two RS solutions if our reasoning is correct, however for the reasons already stated it is difficult to pin down this point. Beyond $PSD \approx 12dB$ only one stable solution is found from both random and ferromagnetic initial conditions, corresponding statistically to a continuation of the good solution. A state which statistically resembles a continuation of the bad state is occasionally arrived at from both initial conditions, this state had a positive stability parameter and negative entropy – so is suspected not to be a viable state. Thus we predict a second dynamical transition in the region of $12dB$, as might be guessed by comparison with the dense case and observation of the trend in the stability parameter (see figure 4).

Our hypothesis is therefore that the trends in the sparse ensembles match those in the dense ensembles within the coexistence region and RS continues to be valid for each of two distinct solutions. The coexistence region for the sparse codes is however smaller in these ensembles. Since our histogram updates dynamically mirror the properties of a belief propagation algorithm on a random graph we can suspect that the bad solution is the performance attainable with belief propagation decoding in the coexistence region, and that there will be problems with convergence near this region. In the cases investigated the bad solution of the sparse ensemble outperforms the bad solution of the dense ensemble, and vice-versa for the good solution. Thus it is likely to be the dynamical transition point for the dense ensemble would corresponds to the PSD beyond which dense CDMA outperforms sparse CDMA at a particular load.

Finally we present figure 5, which shows the the annealed entropy for sparse ensembles of differing chip connectivity in the infinite PSD (zero noise) limit. This shows that at low densities the regular ensemble bound is superior to the irregular ensembles, while the

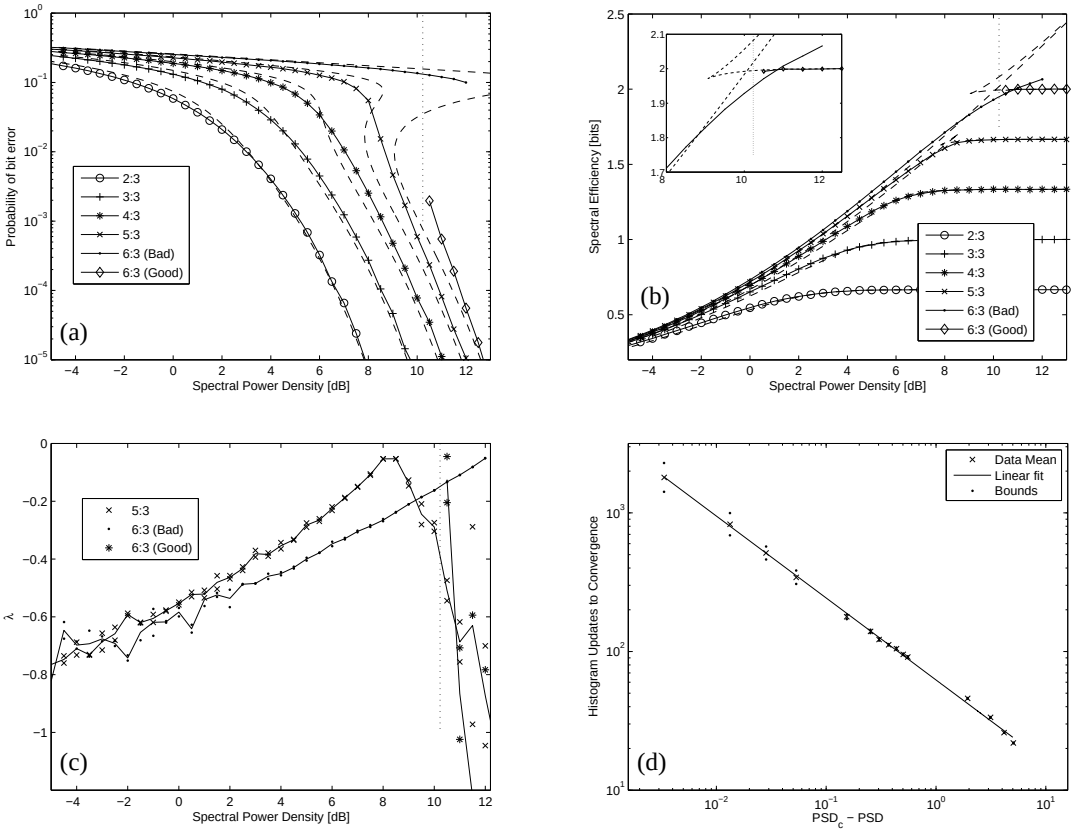

Figure 4. The effect of channel load α on performance for the regular ensemble. Data presented on the basis of 10 runs, error bars omitted but characterised by the smoothness of curves. Dashed lines indicate the dense code analogues. The vertical line indicates the point beyond which $6:3$ random and ferromagnetic initial conditions failed to converge to the same solution, both dynamically stable solutions are shown beyond this point. (a) There is a monotonic increase in bit error rate with the load. (b) As load α is increased there is a monotonic increase in capacity. The spectral efficiency for the "bad" solution exceeds 2 in a small interval (equivalent to negative entropy), this is believed to be a numerical convergence problem. There appears to be a cross over in the entropy of two distinct solutions, indicating a second order phase transition near 11dB. (c) The Stability Parameter was found to be negative for all convergent solutions, indicating the suitability of RS. Where the solution is close to ferromagnetic the stability measure becomes quickly very noisy (as shown for the $5:3$ and $6:3$ codes). (d) Investigation of the the 6:3 code ($\alpha = 2$) indicates a divergence in convergence time as $PSD \rightarrow 10.23dB$ with exponent 0.59 based on a simple linear regression of 15 points (each point is the mean of 10 independent runs). Beyond this point different initial conditions find one of two solutions.

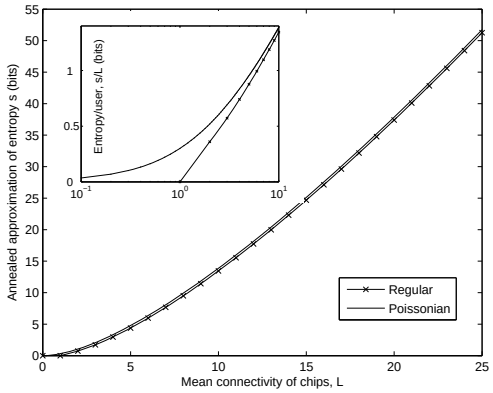

Figure 5. A $PSD \to \infty$ limit annealed approximation to the entropy/chip of different codes. This lower bound differs at leading order between ensembles of regular and Poissonian chip connectivities, but becomes increasingly small as L increases. The curve is asymptotically linear to leading order in L, so difference between Poissonian and regular ensembles are asymptotically negligable.

gap closes (asymptotically to zero) as L increases. This appears consistent with the full numerical results found at high PSD, although regular chip connectivity under performed by comparison with Poisson distributed chip connectivity in the low PSD regime, which was not anticipated in the annealed approximation.

5. Concluding Remarks

Our results demonstrate the feasibility of sparse regular codes for use in CDMA. At moderate PSD it seems the performance of sparse regular codes may be very good. With the replica symmetric assumption apparently valid at practical PSD it is likely that fast algorithms based on belief propagation may be very successful in achieving the theoretical results. Furthermore for lower density sparse codes the problem of the coexistence regime, which limits the performance of practical decoding methods, seems to be less pervasive than for dense ensembles in the over saturated regime.

A direct evaluation of the properties of belief propagation may prove similar results to those shown here. In the absence of replica symmetry breaking states it is normally true that belief propagation performs very well. However, to make best use of the channel resources it may be preferable to implement high load regimes in cases of high PSD, and so overcoming the algorithmic problems arising from phase coexistence is a challenge of practical importance in this case.

Other practical issues in implementation are certainly significant. Similar to the case of dense CDMA there are significant problems relating to multipath, fading and power control, in fact it is known that these effects are more disruptive for the sparse codes, especially regular codes. However, certain situations such as broadcasting (one to many) channels and downlink CDMA, where synchronisation can be assumed, may be practical

points for future implementation. There are practical advantages of the sparse case over dense and orthogonal codes in some regimes. The sparse CDMA method is likely to be particularly useful in frequency-hopping and time-hopping code division multiple access (FH and TH -CDMA) applications where the effect of these practical limitations is less emphasised.

Extensions based on our method to cases without power control or synchronisation have been attempted and are quite difficult. A consideration of priors on the inputs, in particular the effects when sparse CDMA is combined with some encoding method may also be interesting.

Acknowledgments

Support from EVERGROW, IP No. 1935 in FP6 of the EU is gratefully acknowledged. DS would like to thank Ido Kanter for helpful discussions.

Bibliography

[1] S. Verdu. *Multiuser Detection*. Cambridge University Press, New York, NY, USA, 1998.

[2] M. Yoshida and T. Tanaka. Analysis of sparsely-spread cdma via statistical mechanics. In *Proceedings - IEEE International Symposium on Information Theory, 2006.*, pages 2378–2382, 2006.

[3] H. Nishimori. *Statistical Physics of Spin Glasses and Information Processing*. Oxford Science Publications, 2001.

[4] M. Mezard, G. Parisi, and M.A Virasoro. *Spin Glass Theory and Beyond*. World Scientific, 1987.

[5] Toshiyuki Tanaka. A statistical-mechanics approach to large-system analysis of cdma multiuser detectors. *Symposium on Discrete Algorithms (SODA)*, 48(11):2888–2910, 2002.

[6] D. Guo and S. Verdu. *Communications, Information and Network Security*, chapter Multiuser Detection and Statistical Mechanics, pages 229–277. Kluwer Academic Publishers, 2002.

[7] A. Montanari and D. Tse. Analysis of belief propagation for non-linear problems: The example of cdma (or: How to prove tanaka's formula). In *Proceedings IEEE Workshop on Information Theory*, 2006.

[8] Y. Kabashima. A cdma multiuser detection algorithm on the basis of belief propagation. *Jour. Phys. A*, 36(43):11111–11121, 2003.

[9] J.P Neirotti and D. Saad. Improved message passing for inference in densely connected systems. *Europhys. Lett.*, 71(5):866–872, 2005.

[10] A. Montanari, B. Prabhakar, and D. Tse. Belief propagation based multiuser detection. In *Proceedings of the Allerton Conference on Communication, Control and Computing, Monticello, USA*, 2006.

[11] D. Guo and C. Wang. Multiuser detection of sparsely spread cdma. SUBMITTED TO IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS 1, 2007.

[12] T. Tanaka and D. Saad. A statistical-mechanical analysis of coded cdma with regular ldpc codes. In *Proceedings - IEEE International Symposium on Information Theory, 2003.*, page 444, 2003.

[13] D.J. MacKay. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2004.

[14] R. Vicente, D. Saad, and Y. Kabashima. *Advances in Imaging and Electron Physics*, volume 125, chapter Low Density Parity Check Codes - A statistical Physics Perspective, pages 231–353. Academic Press, 2002.

[15] K.Y.M. Wong and D. Sherrington. Graph bipartitioning and spin-glasses on a random network of fixed finite valence. *J. Phys. A*, 20:L793–99, 1987.

[16] M. Mezard and G. Parisi. The bethe lattice spin glass revisited. *Euro. Phys. Jour. B*, 20(2):217–233, 2001.

[17] O. Rivoire, G. Biroli, O.C. Martin, and M. Mzard. Glass models on bethe lattices. *Euro. Phys. J. B*, 37:55–78, 2004.

[18] Y. Kapashima. Propagating beliefs in spin glass models. *J. Phys. Soc. Jpn.*, 72:1645–1649, 2003.

[19] J. Raymond, A. Sportiello, and L. Zdeborova. The phase diagram of random 1-in-3 satisfiability problem. Preprint cond-mat/0702610, 2006.