Statistical Mechanics of Compressed Sensing: Supplementary Material

February 23, 2010

1 Statistical Mechanics Framework

Here we analyze in more detail the statistical mechanics of the Gibbs distribution $P_G(\mathbf{u}) = \frac{1}{Z}e^{-\beta E(\mathbf{u})}$, where

$$E(\mathbf{u}) = \frac{\lambda}{2T}\mathbf{u}^T\mathbf{A}^T\mathbf{A}\mathbf{u} + \sum_{i=1}^{T} V(u_i + x_i^0), \tag{1}$$

and

$$Z = \int \prod_{i=1}^{T} du_i \, e^{-\beta E(\mathbf{u})}. \tag{2}$$

We wish to compute the average free energy $-\beta \bar{F} = \langle\langle \ln Z \rangle\rangle$, where $\langle\langle \cdot \rangle\rangle$ is an average over the quenched disorder $\mathbf{A}$ and $\mathbf{x}^0$. To do so, we use the replica method, which relies on the identity $\ln \mathrm{Z} = \lim_{n\to 0} \frac{Z^n - 1}{n}$. Z^n can be written as an integral over n replicated variables u_i^a, $a = 1, \ldots n$,

$$\langle\langle Z^n \rangle\rangle = \left\langle\!\left\langle \int \prod_{a=1}^{n} \prod_{i=1}^{T} du_i^a \, \mathrm{e}^{\sum_{a=1}^{n} -\frac{\beta\lambda}{2} \sum_{\mu=1}^{n} \sum_{ij=1}^{T} u_i^a A_{\mu i} A_{\mu j} u_j^a - \beta \sum_{i=1}^{T} V(u_i^a + x_i^0)} \right\rangle\!\right\rangle. \tag{3}$$

We first average over $\mathbf{A}$ and later average over $\mathbf{x}^0$. Now (3) depends on $\mathbf{A}$ only through the variables $b_\mu^a \equiv \frac{1}{\sqrt{T}} \sum_{i=1}^{T} A_{\mu i} u_i^a$, which are jointly gaussian distributed with zero mean and covariance $\langle \delta b_\mu^a \, \delta b_\nu^b \rangle = Q_{ab}\delta_{\mu\nu}$, where $Q_{ab} \equiv \frac{1}{T}\sum_{i=1}^{T} u_i^a u_i^b$. Introducing delta functions $\delta[\sum_{i=1}^{T} u_i^a u_i^b - T Q_{ab}] = \int d\,\hat{Q}_{ab}\, e^{\,i\hat{Q}_{ab}(\sum_{i=1}^{T} u_i^a u_i^b - T Q_{ab})}$ to decouple the integrals over u_i^a and b_μ^a, and performing the gaussian integral over b_μ^a yields,

$$\langle\langle Z^n \rangle\rangle = \left\langle\!\left\langle \int \prod_{a,b=1}^{n} \prod_{i=1}^{T} dQ_{ab}\hat{Q}_{ab} \, \mathrm{e}^{T[i\sum_{ab}\hat{Q}_{ab}Q_{ab} + \frac{1}{T}\sum_{i=1}^{T} \ln \mathcal{Z}_i - \frac{\alpha}{2}\mathrm{Tr}\,\log(I + \beta\lambda Q)]} \right\rangle\!\right\rangle, \tag{4}$$

where $\alpha \equiv \frac{N}{T}$ and

$$\mathcal{Z}_i = \int \prod_{a=1}^{n} du^a \, \mathrm{e}^{-i\sum_{ab} u^a \hat{Q}_{ab} u^b - \beta \sum_a V(u^a + x_i^0)}. \tag{5}$$

The remaining integrals over Q_{ab} and $\hat{Q}_{ab}$ in (4) can be done via the saddle point approximation. We work with a replica symmetric (RS) ansatz for the saddle point: $Q_{ab} = (Q_1 - Q_0)\delta_{ab} + Q_0$ and $\hat{Q}_{ab} = (\hat{q}_1 - \hat{q}_0)\delta_{ab} + \hat{q}_0$. Inserting this ansatz into the above equations, taking the $n \to 0$ limit, and averaging over $\mathbf{x}^0$ yields,

$$-\beta\bar{F} = \mathrm{extr}_{(\hat{q}_1 Q_1 \hat{q}_0 Q_0)} \left\{ i(\hat{q}_1 Q_1 - \hat{q}_0 Q_0) + \langle\langle \ln \zeta \rangle\rangle_{z,x^0} - \frac{\alpha}{2}\left(\frac{\beta\lambda}{1 + \beta\lambda\Delta Q} + \ln\left(1 + \beta\lambda\Delta Q\right) \right) \right\}, \tag{6}$$

where

$$\zeta = \int du\, e^{-i(\hat{q}_1-\hat{q}_0)u^2+\sqrt{-2i\hat{q}_0}zu-\beta V(u+x^0)}, \tag{7}$$

and $\Delta Q \equiv Q_1 - Q_0$. Here z is a zero mean unit variance gaussian variable introduced to decouple the replicas under the RS assumption in (5) via the identity

$$e^{-i\hat{q}_0(\sum_a u^a)^2} = \int \mathcal{D}z\; e^{\sqrt{-2i\hat{q}_0}\,z\sum_a u_a}, \tag{8}$$

where $\mathcal{D}z$ is the standard gaussian measure. Furthermore, the double angular average $\langle\langle\,\cdot\,\rangle\rangle_{z,x^0}$ involves integrating over the Gaussian distribution of z and over the distribution of x^0, i.e.

$$\langle\langle\, F(z,x^0)\,\rangle\rangle_{z,x^0} \equiv (1-f)\int \mathcal{D}z\; F(z,0) + f\int \mathcal{D}z\; dx^0\, P(x^0)F(z,x^0). \tag{9}$$

Extremizing (6) yields the saddle point equations

$$-i\hat{q}_0 = \frac{\alpha}{2}\frac{(\beta\lambda)^2 Q_0}{(1+\beta\lambda\Delta Q)^2} \tag{10}$$

$$i(\hat{q}_1-\hat{q}_0) = \frac{\alpha}{2}\frac{\beta\lambda}{1+\beta\lambda\Delta Q} \tag{11}$$

$$Q_0 = \langle\langle\,\langle u\rangle^2\,\rangle\rangle_{z,x^0} \tag{12}$$

$$\Delta Q = \langle\langle\,\langle \delta u^2\rangle\,\rangle\rangle_{z,x^0}, \tag{13}$$

where the mean $\langle u\rangle$ and variance $\langle \delta u^2\rangle$ of u are computed with respect to the distribution

$$P(u) = \frac{1}{\zeta} e^{-i(\hat{q}_1-\hat{q}_0)u^2+\sqrt{-2i\hat{q}_0}zu-\beta V(u+x^0)}. \tag{14}$$

Substituting (10) and (11) into (14) and completing the square of the quadratic part of the exponent in u then yields the result that $P(u) \propto \exp(-H_{eff})$ with the effective Hamiltonian

$$H_{eff} = \frac{\alpha\beta\lambda}{2(1+\beta\lambda\Delta Q)}\left(u - z\sqrt{Q_0/\alpha}\right)^2 + \beta V(u+x^0). \tag{15}$$

This corresponds to Eq. 3 of the main paper, which is specialized to the case $V(x) = |x|$. The remaining order parameter equations (12) and (13) correspond to Eq. 4 and 5 of the main paper.

At their extremal values, the order parameters Q_1 and Q_0 have simple interpretations in terms of the original Gibbs distribution P_G defined above. Q_1 is the typical normalized length of the residual vector $\mathbf{u}$ when drawn from P_G: $Q_1 = \frac{1}{T}\langle \mathbf{u}\cdot\mathbf{u}\rangle_{P_G}$. Thus as $\lambda,\beta\to\infty$, Q_1 is the squared L_2 reconstruction error $\frac{1}{T}||\hat{\mathbf{x}}-\mathbf{x}^0||_2^2$ in the constrained optimization problem, defined in Eq. 1 of the main paper, for any typical realization of $\mathbf{A}$ and $\mathbf{x}^0$. Similarly, Q_0 is the typical normalized correlation between two residuals $\mathbf{u}^a$ and $\mathbf{u}^b$ each drawn independently from P_G : $Q_0 = \frac{1}{T}\langle \mathbf{u}^a\cdot\mathbf{u}^b\rangle_{P_G}$. Thus as $\lambda,\beta\to\infty$, Q_0 reveals information about the geometry of the solution space to Eq. 1 of the main paper.

Finally, the distribution $P(u)$ itself is simply related to the distribution of the components of the reconstruction $\hat{x}_i$ obtained by solving Eq. 1 in the main paper; the prediction of the RS theory for the distribution of $\hat{x}_i$, conditioned on a value of the true signal $x_i^0 = x^0$ is given by

$$P(\hat{x}_i = x\,|\,x_i^0 = x^0) = \lim_{\lambda,\beta\to\infty}\int \mathcal{D}z\; P(u), \tag{16}$$

where $u = x - x^0$.

2 Perfect Reconstruction Regimes

As described in the main paper, we search for solutions to (10)-(13) which reflect perfect reconstruction performance by first taking the $\lambda \to \infty$ limit, and then making the ansatz

$$\Delta Q = \frac{\alpha}{\beta^2} \Delta q \tag{17}$$

and

$$Q_0 = \frac{\alpha}{\beta^2} q_0, \tag{18}$$

where both Δq and q_0 are $O(1)$ as $\beta \to \infty$. As we see below, the region in the $\alpha - f$ plane in which such solutions can exist depends on the domain of the signal $\mathbf{x}^0$. We search for such solutions in the case $V(x) = |x|$.

2.1 Nonnegative signal reconstruction

We first consider nonnegative signals in which $x_i^0 \geq 0$. In this case, we enforce a nonnegativity constraint on the allowed reconstruction, so that the constrained optimization in Eq. 1 of the main paper is only over $\mathbf{x}$ such that $x_i \geq 0$ for all $i = 1, \dots T$. In the statistical mechanics framework, any integral over the residual u then ranges from $-x^0$ to $+\infty$. The effective Hamiltonian in (15) then becomes quadratic. Under the above ansatz, it is, up to irrelevant additive constants,

$$H_{eff} = \frac{1}{2\Delta q} \left(\beta u - [z\sqrt{q_0} - \Delta q]\right)^2. \tag{19}$$

Now in computing the mean and variance of u, since u scales as $1/\beta$, the nonnegativity constraint, $u \geq -x^0$, is negligible unless $x^0 = 0$, in which case, the constraint implies that $u \geq 0$. Thus, for $x^0 > 0$ the mean and variance of u are

$$\langle u \rangle = (z\sqrt{q_0} - \Delta q)/\beta \tag{20}$$

$$\langle (\delta u)^2 \rangle = \Delta q/\beta^2, \tag{21}$$

while for $x^0 = 0$ the corresponding averages are,

$$\langle u \rangle = (\sqrt{\Delta q}/\beta)\, G_1 \tag{22}$$

$$\langle (\delta u)^2 \rangle = \Delta q/\beta^2\, (G_2 - G_1^2), \tag{23}$$

where

$$G_m = \frac{\int_{-v} \mathcal{D}t\, (t+v)^m}{H(-v)}, \tag{24}$$

where $v = (z\sqrt{q_0} - \Delta q)/\sqrt{\Delta q}$ and $H(x) = \int_x^\infty \mathcal{D}z$. Combining these regimes, the order parameter equations (12) and (13) become

$$\alpha q_0 = (1-f)\Delta q \big\langle\big\langle\, G_1^2 \,\big\rangle\big\rangle_z + f\left(q_0 + \Delta q^2\right) \tag{25}$$

$$\alpha = (1-f) \big\langle\big\langle\, G_2 - G_1^2 \,\big\rangle\big\rangle_z + f. \tag{26}$$

Because the thermal averages $\langle u^m \rangle$, in the $\beta \to \infty$ limit, do not depend on x_0 for nonzero values of x_0, these equations do not depend on the distribution of nonzeros $P(x^0)$. This immediately implies that the region of the $\alpha - f$ plane in which perfect reconstruction is possible via Eq. 1 of the main paper, which coincides with the region in which (25) and (26) have solutions, is independent of the distribution of nonzeros $P(x^0)$ in $\mathbf{x}^0$. Numerically, we have checked that these equations admit finite solutions for any f and $\alpha > \alpha_c(f)$. As α decreases towards the phase boundary $\alpha_c(f)$, Δq and q_0 diverge with $\Delta q/q_0^2 \equiv x^2$, remaining $O(1)$. Taking this limit in (25) and (26) yields the following equations for the phase boundary

$$\alpha = (1-f)\int_x^\infty \mathcal{D}z\,(z-x)^2 + f(1+x^2) \tag{27}$$

$$\alpha = (1-f)H(x) + f. \tag{28}$$

2.2 General signal reconstruction

Here we consider signals in which x_i^0 can take any real value. Any integral over the residual u then ranges from $-\infty$ to $+\infty$. Following a similar analysis as in the case of nonnegative signals, we obtain the same form of equations for the order parameters as in Eqs. (25) and (26), except that now,

$$G_m = \frac{\int_{-\infty}^{-v^-} \mathcal{D}t\,(t+v^-)^m + \int_{-v^+}^{\infty} \mathcal{D}t\,(t+v^+)^m}{H(v^-)+H(-v^+)}, \tag{29}$$

and $v^\pm = (z\sqrt{q_0} \pm \Delta q)/\sqrt{\Delta q}$. Again, numerically, we have checked that these equations admit finite solutions for any f and $\alpha > \alpha_c(f)$ where α_c is now the solution to equations 6 and 7 in the main paper, obtained by taking the limit $\Delta q, q_0 \to \infty$ with $\Delta q/q_0^2 \equiv x^2$, remaining $O(1)$ in (25) and (26) with G_m given by (29).

2.3 Compressed sensing without sparsity optimization

Here we return to the case of nonnegative signals to find a new class of solutions to the order parameter equations (10)-(13) which reflect the existence of a phase in which the solution space to the constraints in Eq. 1 of the main paper condenses to a single point $\hat{\mathbf{x}} = \mathbf{x}^0$. We examine the limit of large β but without the potential term, V in (1) or equivalently keeping β finite and taking $\lambda \to \infty$. We then search for solutions in which

$$Q_0 = \frac{1}{\beta\lambda} q_0 \tag{30}$$

$$\Delta Q_0 = \frac{1}{\beta\lambda}\Delta q, \tag{31}$$

which leads to the following equations for the auxillary order parameters

$$-i\hat{q}_0 = \beta\lambda \frac{\alpha q_0}{2(1+\Delta Q)^2} \tag{32}$$

and

$$i(\hat{q}_1 - \hat{q}_0) = \beta\lambda \frac{\alpha}{2(1+\Delta Q)}. \tag{33}$$

The effective Hamiltonian of u is then,

$$H_{eff} = \frac{\alpha}{2(1+\Delta q)}(u\sqrt{\beta\lambda} - z\sqrt{q_0/\alpha})^2, \tag{34}$$

implying that u scales as $1/\sqrt{\beta\lambda}$. Repeating an analysis similar to that above, we obtain

$$\alpha q_0 = (1-f)\langle\langle\, G_1^2 \,\rangle\rangle_z + f q_0 \tag{35}$$

$$\alpha\frac{\Delta q}{1+\Delta q} = (1-f)\langle\langle\, G_2 - G_1^2 \,\rangle\rangle_z + f, \tag{36}$$

where G_m is given by (24), with $v = z\sqrt{q_0}$.

Now $\langle\langle\, G_1^2 \,\rangle\rangle_z$ monotonically approaches $q_0/2$ from above as $q_0 \to \infty$. This implies that $\alpha > (1+f)/2$ is a necessary condition for the existence of a solution to (35). Numerically, we have checked that it is also sufficient. Furthermore, $\langle\langle\, G_2 - G_1^2 \,\rangle\rangle_z$ montonically approaches 1/2 from below as $q_0 \to \infty$. This means the RHS of (36) is less than $(1+f)/2$, but Δq can still be chosen to solve (36) even if $\alpha > (1+f)/2$. As α approaches $(1+f)/2$ from above, both q_0 and Δq approach infinity, marking a phase transition to a nonzero error.

3 Beyond the Phase Transition: Into the Error Regime

Here we search for solutions to (10)-(13) which describe the performance of compressed sensing in the error phase. Again we focus on the case of L_1 minimization so that $V(x) = |x|$. As in the case of perfect reconstruction, we first take the $\lambda \to \infty$ limit in (15), thereby eliminating λ from the Hamiltonian. However, unlike the case of perfect reconstruction, Eqs. (17) and (18), in the present case, we choose the scaling of the order parameters in the large β regime to be,

$$\Delta Q = \frac{\alpha}{\beta}\Delta q \tag{37}$$

and

$$Q_0 = \alpha q_0, \tag{38}$$

where both Δq and q_0 are $O(1)$. Here, Q_0, which measures the mean square deviation of the reconstructed signal from the original one, is nonzero even at zero temperature, and the thermal fluctuations scale linearly with $1/\beta$. Below we search for solutions of this form to describe the error behavior of compressed sensing as a function of α and f, discussing the cases of nonnegative and general signal reconstruction in parallel.

3.1 Order parameter equations

Under the above ansatz, in the case of nonnegative signals, the effective Hamiltonian for u is

$$H_{eff} = \frac{\beta}{2\Delta q}(u - [\sqrt{q_0}z - \Delta q])^2. \tag{39}$$

Since in the limit of $\beta \to \infty$ u itself is $O(1)$, while its fluctuations are small, the integrals over u are dominated by the global minimum of H_{eff}. The minimum of H_{eff} in the allowed region $u \geq -x^0$, occurs at the interior point $\sqrt{q_0}z - \Delta q$ if $z \geq z^+ \equiv \frac{-x^0+\Delta q}{\sqrt{q_0}}$. Otherwise, it occurs at the boundary $-x^0$. Furthermore, when the minimum is in the interior, the variance of the fluctuations in u is $\langle(\delta u)^2\rangle = \frac{\Delta q}{\beta}$. However, when it is at the boundary, the fluctuations are $O(1/\beta^2)$ and can be neglected. In summary, for nonnegativity constraint we have

$$\langle u \rangle = \begin{cases} \sqrt{q_0}z - \Delta q & \text{if } z \geq z^+ \\ -x^0 & \text{if } z < z^+ \end{cases} \tag{40}$$

and

$$\langle(\delta u)^2\rangle = \begin{cases} \frac{\Delta q}{\beta} & \text{if } z \geq z^+ \\ 0 & \text{if } z < z^+. \end{cases} \tag{41}$$

Similar considerations lead to corresponding results for general signal reconstruction where u is unconstrained by x^0:

$$\langle u\rangle = \begin{cases} \sqrt{q_0}z - \Delta q & \text{if } z \geq z^+ \\ (x^0)^2 & \text{if } z^- \leq z \leq z^+ \\ \sqrt{q_0}z + \Delta q & \text{if } z < z^- \end{cases} \tag{42}$$

and

$$\langle(\delta u)^2\rangle = \begin{cases} \frac{\Delta q}{\beta} & \text{if } z > z^+ \text{ or } z < z^- \\ 0 & \text{if } z^- \leq z \leq z^+, \end{cases} \tag{43}$$

where $z^\pm = \frac{-x^0 \pm \Delta q}{\sqrt{q_0}}$.

Inserting these results into (12) and (13) then gives equations for the order parameters. However, unlike the corresponding equations in the zero error phase, the equations in the error phase depend on the distribution of nonzeros $P(x^0)$. Explicitly, for signals with a nonnegativity constraint, the order parameter equations become

$$\alpha q_0 = (1-f)\int_{\frac{\Delta q}{\sqrt{q_0}}}^{\infty} \mathcal{D}z\,(\sqrt{q_0}z - \Delta q)^2 + f\left\langle\!\!\left\langle \left(\int_{z_+}^{\infty} \mathcal{D}z\,(\sqrt{q_0}z - \Delta q)^2 + (x^0)^2 H(-z^+)\right)\right\rangle\!\!\right\rangle_{x^0} \tag{44}$$

$$\alpha = (1-f)H(\frac{\Delta q}{\sqrt{q_0}}) + f\langle\langle\, H(z^+)\,\rangle\rangle_{x^0}. \tag{45}$$

For unconstrained signals, the corresponding equations are

$$\alpha q_0 = 2(1-f)\int_{\frac{\Delta q}{\sqrt{q_0}}}^{\infty} \mathcal{D}z\,(\sqrt{q_0}z - \Delta q)^2$$

$$+ f\left\langle\!\!\left\langle \int_{z^+}^{\infty} \mathcal{D}z\,(\sqrt{q_0}z - \Delta q)^2 + \int_{-\infty}^{z^-} \mathcal{D}z\,(\sqrt{q_0}z + \Delta q)^2 + (x^0)^2(H(z^-) - H(z^+))\right\rangle\!\!\right\rangle_{x^0} \tag{46}$$

$$\alpha = 2(1-f)H(\frac{\Delta q}{\sqrt{q_0}}) + f\langle\langle\,(H(z^+) + H(-z^-))\,\rangle\rangle_{x^0}. \tag{47}$$

Here, unlike in (12) and (13), the average over x^0 is only over the distribution of nonzero x^0.

Now as $\alpha \to 0$, it can be shown in both cases that $\frac{\Delta q}{\sqrt{q_0}}$ diverges as $\sqrt{\log 1/\alpha}$, while q_0 diverges as $1/\alpha$. The form of this divergence is independent of the distribution of nonzeros. Expanding the above equations in the limit of large $\frac{\Delta q}{\sqrt{q_0}}$ and q_0 then yields the asymptotic result discussed in the main paper: $Q_0 = \alpha q_0 = f(1 + \log(1/\alpha))$. On the other hand, as $\alpha \to \alpha_c$ from below, both q_0 and Δq approach 0 with $\frac{\Delta q}{\sqrt{q_0}}$ remaining $O(1)$. Expanding the above equations in this limit yields the results in the main paper describing the rise of the error near the phase transition between perfect and imperfect reconstruction, and its dependence on the signal statistics.

3.2 Distribution of $\hat{\mathbf{x}}_i$

As noted in (16), the replica method allows us to compute not only the typical reconstruction error, but also the conditional distribution of a reconstruction component $\hat{x}_i$ given the corresponding value of the true signal x_i^0. We denote this distribution $P(x|x^o)$. Given (16) and the scaling of the order parameters in (37) and (38), for any nonnegative signal reconstruction, we have

$$P(x|x^0) = \frac{1}{\sqrt{2\pi q_0}} \exp(-\frac{(x - x_0 + \Delta q)^2}{2q_0}) + H(-z^+)\delta(x). \tag{48}$$

where q_0 and Δq are solutions to (44) and (45). This distribution is simply a gaussian with mean $x^0 - \Delta q$ and variance q_0, truncated at the origin, plus a delta function at the origin whose probability mass is exactly the mass lost by the gaussian due to truncation. Thus we see two sources of error: the reconstruction is

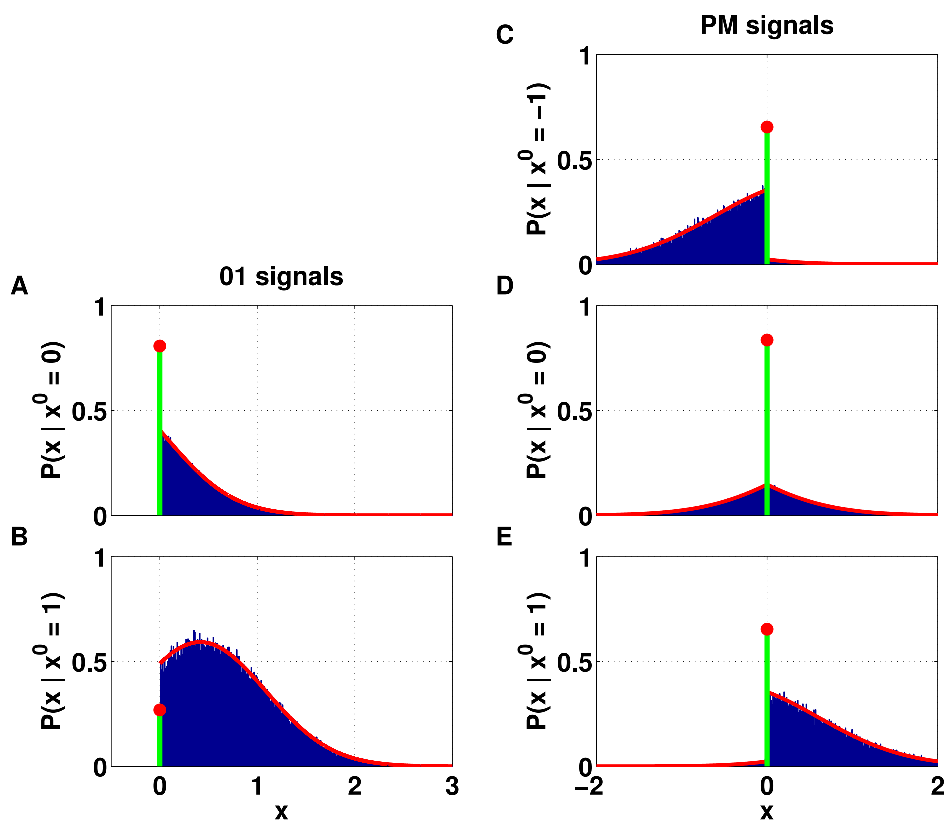

Figure 1: Distribution of the signal reconstruction. (A) and (B) The blue histogram is the conditional distribution of nonzero elements obtained from solving Eq. 1 of the main paper, with positivity constraint, 2000 times, while the height of the green bar represents the average fraction of components $\hat{x}_i$ that are zero, again conditioned on x_i^0. Here $T = 500$, $f = 0.2$ and $\alpha = 0.3$. This yields, through (44) and (45), $q_0 = 0.453$ and $\Delta q = 0.585$. The red curve is the theoretically predicted distribution of nonzeros in (48), while the red dot is the theoretically predicted height of the delta function at $x = 0$ in (48). (C,D,E) Similar results in the case of PM signals with no positivity constraint. Here $f = \alpha = 0.2$, yielding $q_0 = 1.06$ and $\Delta q = 1.43$ through (46) and (47), and the theoretically predicted distribution arises from equation (49).

biased towards the origin by an amount Δq, and suffers a variance q_0 about its mean value. Figure 1A shows examples of the match between theory and simulations for the reconstruction distribution in the case of 01 signals. Similarly, for any general signal reconstruction, we find

$$P(x|x^0) = \frac{1}{\sqrt{2\pi q_0}} \exp(-\frac{(x - x_0 + \mathrm{sgn}(x)\Delta q)^2}{2q_0}) + (H(z^-) - H(z^+))\delta(x), \tag{49}$$

where q_0 and Δq are solutions to (46) and (47). Thus if $x > 0$, the distribution behaves like a gaussian with mean $x^0 - \Delta q$ and variance q_0, while if $x < 0$, it behaves like a gaussian with mean $x^0 + \Delta q$ and the same variance. In both cases, the bias of magnitude Δq acts towards the origin. The remaining probability mass required to normalize the distribution arises from the delta function at the origin. Figure 1B shows examples of the match between theory and simulations for the reconstruction distribution in the case of PM signals.

3.3 L_p norms of the reconstruction.

Given the reconstruction distribution $P(x|x^0)$, it is straightforward to compute various norms of the reconstruction. The average normalized L_p norm of the reconstructed signal $\hat{\mathbf{x}}$, defined as $L_p(\hat{\mathbf{x}}) \equiv \frac{1}{T}\sum_{i=1}^{T} |\hat{x}_i|^p$, is given by the theory as

$$L_p(\hat{\mathbf{x}}) = \langle\langle \int dx\, x^p P(x|x^0) \rangle\rangle_{x^0}.$$

As discussed in the main paper, the theory makes a simple prediction for the L_0 norm of $\hat{x}$ in the error regime: $L_0(\hat{\mathbf{x}}) = \alpha$ independent of α. This result arises as follows. The contribution to the variance of u (in either (41) for nonnegative signals, or (43) for general signals) comes from the regime where $x = u + x^0$ is non-zero, in which case it is equal to $\Delta q/\beta = \Delta Q/\alpha$. Otherwise, thermal fluctuations in u are negligible. Hence, (13) can be written, for both nonnegative and general signals, in the form

$$\Delta Q = \frac{\Delta Q}{\alpha} \Big\langle\Big\langle \int_{x\neq 0} dx\, P(x|x_0) \Big\rangle\Big\rangle_{x^0},$$

from which it follows that $L_0(\hat{\mathbf{x}}) = \langle\langle \int_{x\neq 0} dx\, P(x|x_0) \rangle\rangle_{x^0} = \alpha$.

4 A null model for sparse regression

Here we analyze the statistical mechanics of a Gibbs distribution with energy function

$$E(\mathbf{x}) = -\frac{\lambda}{2T}(\mathbf{y} - \mathbf{A}\mathbf{x})^T(\mathbf{y} - \mathbf{A}\mathbf{x}) + \sum_{i=1}^{T} |x_i|, \tag{50}$$

which is relevant to the case of sparse regression as discussed in the main paper. Here, $\mathbf{A}$ a random NxT matrix as before, and y_μ are i.i.d gaussian distributed with zero mean and variance 1. We perform a replica analysis similar to that in section 1 to compute the typical properties of the Gibbs distribution $P_G(\mathbf{u}) = \frac{1}{Z}e^{-\beta E(\mathbf{u})}$. The replicated partition function Z^n now depends on the quenched disorder $\mathbf{A}$ and $\mathbf{y}$ only through the variables $b_\mu^a \equiv y_\mu - \frac{1}{\sqrt{T}}\sum_{i=1}^{T} A_{\mu i} x_i^a$, which are jointly gaussian distributed with zero mean and covariance $\langle \delta b_\mu^a\, \delta b_\nu^b \rangle = (1 + Q_{ab})\delta_{\mu\nu}$, where $Q_{ab} \equiv \frac{1}{T}\sum_{i=1}^{T} x_i^a x_i^b$. By modifying the calculation in section 1, we obtain the following order parameter equations

$$-i\hat{q}_0 = \frac{\alpha}{2}\frac{(\beta\lambda)^2(Q_0+1)}{(1+\beta\lambda\Delta Q)^2} \tag{51}$$

$$i(\hat{q}_1-\hat{q}_0) = \frac{\alpha}{2}\frac{\beta\lambda}{1+\beta\lambda\Delta Q} \tag{52}$$

$$Q_0 = \langle\langle\, \langle x\rangle^2\, \rangle\rangle_z \tag{53}$$

$$\Delta Q = \langle\langle\, \langle \delta x^2\rangle\, \rangle\rangle_z, \tag{54}$$

where the mean $\langle x\rangle$ and variance $\langle \delta x^2\rangle$ of x are computed with respect to the distribution

$$P(x) = \frac{1}{\zeta} e^{-i(\hat{q}_1-\hat{q}_0)x^2+\sqrt{-2i\hat{q}_0}zx-\beta|x|)}, \tag{55}$$

and

$$\zeta = \int dx\, e^{-i(\hat{q}_1-\hat{q}_0)x^2+\sqrt{-2i\hat{q}_0}zx-\beta|x|}. \tag{56}$$

At their extremal values, the order parameters now have the following interpretations with respect to the original Gibbs distribution P_G. Q_1 is the typical normalized L_2 norm of $\mathbf{x}$ when drawn from P_G: $Q_1 = \frac{1}{T}\langle \mathbf{x}\cdot\mathbf{x}\rangle_{P_G}$, and Q_0 is the typical normalized correlation between two vectors $\mathbf{x}^a$ and $\mathbf{x}^b$ each drawn independently from $P_G : Q_0 = \frac{1}{T}\langle \mathbf{x}^a\cdot\mathbf{x}^b\rangle_{P_G}$.

Just as in CS in the error regime, we first take $\lambda\to\infty$ and then search for solutions in which x is $O(1)$ with small fluctuations $O(1/\beta)$ in the low temperature $\beta\to\infty$ limit. This motivates the following ansatz for the scaling of the order parameters

$$\Delta Q = \frac{\alpha}{\beta}\Delta q \tag{57}$$

and

$$Q_0 = \alpha q_0 - 1. \tag{58}$$

Under this ansatz, $P(x)$ in (55) becomes a Gibbs distribution with effective Hamiltonian

$$H_{eff} = \frac{\beta}{2\Delta q}\left((x-\sqrt{q_0}z)^2 + 2\Delta q|x|\right). \tag{59}$$

In the $\beta\to\infty$ limit, the integrals over x are then dominated by the global minimum of H_{eff}. Furthermore the fluctuations in x are negligible when this minimum is $x=0$. Overall, we have

$$\langle u\rangle = \begin{cases} \sqrt{q_0}z-\Delta q & \text{if } z\geq z^+ \\ 0 & \text{if } z^-\leq z\leq z^+ \\ \sqrt{q_0}z+\Delta q & \text{if } z<z^- \end{cases} \tag{60}$$

and

$$\langle(\delta u)^2\rangle = \begin{cases} \frac{\Delta q}{\beta} & \text{if } z>z^+ \text{ or } z<z^- \\ 0 & \text{if } z^-\leq z\leq z^+, \end{cases} \tag{61}$$

where $z^\pm = \frac{\pm\Delta q}{\sqrt{q_0}}$. Combining this information, (53) and (54) become

$$\alpha q_0 = 2\int_{z^+}^{\infty}\mathcal{D}z\,(\sqrt{q_0}z-\Delta q)^2+1 \tag{62}$$

$$\alpha = H(z^+)+H(-z^-). \tag{63}$$

Furthermore the distribution of the regression components $\hat{x}_i$ is given by

$$P(\hat{x}) = \frac{1}{\sqrt{2\pi q_0}} \exp(-\frac{(\hat{x} + \mathrm{sgn}(x)\Delta q)^2}{2q_0}) + (H(z^-) - H(z^+))\delta(x), \tag{64}$$

where q_0 and Δq are solutions to (62) and (63). From this, we compute the typical normalized L_p norm of $\hat{\mathbf{x}}$ as $L_p(\hat{\mathbf{x}}) = \int d\hat{x}\, \hat{x}^p P(\hat{x})$. This quantity is plotted for $p = 0$ and $p = 1$ in Fig. 4 of the main paper.

5 Analysis of nonconvex compressed sensing: L_p minimization.

In the main paper, and above, we have focused on L_1 minimization, a convex problem for which the RS analysis is valid. Here we extend our RS analysis to the case of L_p minimization, so that the sparsity cost function in (1) is $V(x) = |x|^p$, for $0 \leq p < 1$. This is a nonconvex problem, and so the RS analysis is not guaranteed to be correct. Nevertheless, for completeness, we work out the predictions of the RS theory in this case, and reveal the inconsistency between these predictions and numerical experiments. For simplicity, we focus only on the case of nonnegative signals. The case of general signals can be handled similarly.

5.1 Perfect Reconstruction Regime

We first search for solutions to (10)-(13) which reflect perfect reconstruction. As before, we first take the $\lambda \to \infty$ limit, but now the scaling of order parameters in (17) and (18) generalizes to

$$\Delta Q = \frac{\alpha}{\beta^{2/p}} \Delta q \tag{65}$$

and

$$Q_0 = \frac{\alpha}{\beta^{2/p}} q_0, \tag{66}$$

where both Δq and q_0 are $O(1)$ as $\beta \to \infty$. The effective Hamiltonian in (15) then becomes

$$H_{eff} = \frac{1}{2\Delta q} \left(\beta^{1/p} u - z\sqrt{q_0}\right)^2 + \beta |u + x^0|^p. \tag{67}$$

Now u scales as $\beta^{-1/p}$, and the nonnegativity constraint, $u \geq -x^0$, is negligible unless $x^0 = 0$, in which case, the constraint implies that $u \geq 0$. Furthermore, for $x^0 > 0$, the second term in H_{eff} can be neglected in the $\beta \to \infty$ limit, yielding

$$\langle u \rangle = (z\sqrt{q_0})/\beta^{1/p} \tag{68}$$

$$\langle (\delta u)^2 \rangle = \Delta q / \beta^{2/p}, \tag{69}$$

when $x^0 > 0$. When $x^0 = 0$ the corresponding averages are,

$$\langle u \rangle = G_1/\beta^{1/p} \tag{70}$$

$$\langle (\delta u)^2 \rangle = (G_2 - G_1^2)/\beta^{2/p}\,, \tag{71}$$

where G_1 and G_2 are the first and second moments of the distribution

$$G(t) \propto \exp\left(-\frac{1}{2\Delta q}(t - z\sqrt{q_0})^2 - t^p\right),$$

defined on the domain $t \geq 0$. Combining these regimes, the order parameter equations (12) and (13) become

$$(\alpha - f)\, q_0 = (1-f)\langle\langle\, G_1^2 \,\rangle\rangle_z, \tag{72}$$

$$(\alpha - f)\, \Delta q = (1-f)\langle\langle\, G_2 - G_1^2 \,\rangle\rangle_z. \tag{73}$$

It is clear that $\alpha > f$ is a necessary condition for these equations to have solutions. We have checked numerically that this is also a sufficient condition. We found solutions for all $\alpha > f$, and as α approaches f from above, q_0 and Δq diverge.

Thus the RS theory predicts that $\mathbf{x}^0$ is the global minimum of the constrained L_p optimization problem (Eq. 1 of the main paper with $V(x) = |x|^p$) for all $\alpha > f$. However, we have checked that this is not the case. Although we are not aware of any algorithm that is guaranteed to find the global minimum of the

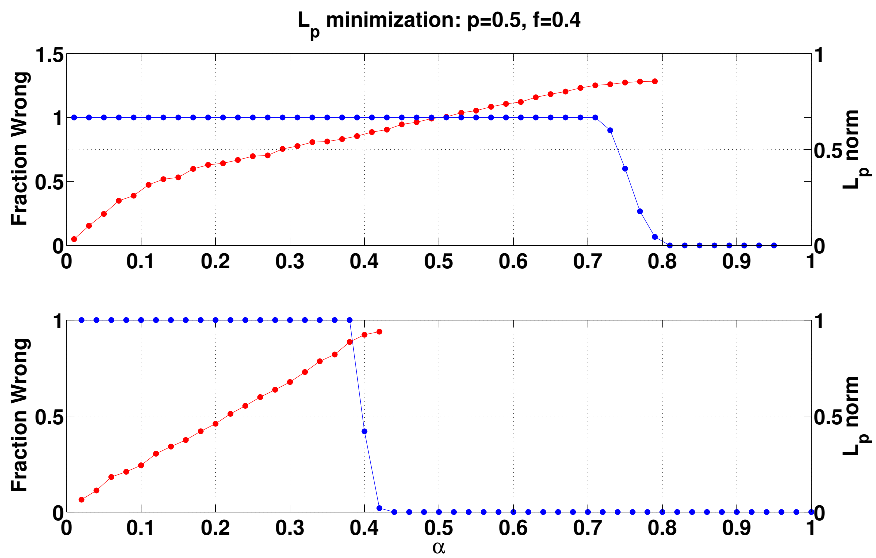

Figure 2: Numerical results on L_p minimization using IRLS. The IRLS algorithm was run 100 times each for various values of α. For each trial, $\mathbf{x}^0$ was a 01 signal with a random fraction $f = 0.4$ of its $T = 500$ entries set to 1. We chose $p = 0.5$. In the top panel, the IRLS algorithm is initialized randomly. The blue curve is the fraction of times IRLS *in*correctly recovered $\mathbf{x}^0$. A transition from perfect to imperfect reconstruction occurs near $\alpha = 0.75$. The red curve is the L_p norm of the reconstruction $\hat{\mathbf{x}}$ offered by IRLS, divided by the L_p norm of $\mathbf{x}^0$, which is f, averaged *only* over the incorrect reconstructions. It is clear that IRLS is doing a poor job in minimizing L_p, since in the range $0.5 < \alpha < 0.8$, IRLS returns reconstructions whose L_p norm is higher than that of $\mathbf{x}^0$. However, despite its poor performance, IRLS manages to find reconstructions whose L_p norm is less than that of $\mathbf{x}^0$ for $\alpha < 0.5$. This contradicts the prediction of the RS theory, which suggests that $\mathbf{x}^0$ is the global minimum of L_p minimization for $\alpha > f = 0.4$. The bottom panel is identical to the top, except now we initialize IRLS with the true solution $\mathbf{x}^0$. For $\alpha > f = 0.4$, $\mathbf{x}^0$ is typically a fixed point of IRLS, indicating its local stability in this regime. For $\alpha < f$, this local stability disappears and IRLS finds solutions with smaller L_p norm than $\mathbf{x}^0$. For both panels, we observe similar behavior for other values of p and f.

L_p problem, we have applied a specific algorithm, iterative reweighted least squares (IRLS), that can find local minima, starting from a given initial condition. We have found that IRLS, when initialized randomly, typically yields solutions $\hat{\mathbf{x}}$ whose L_p norm is *smaller* than that of $\mathbf{x}^0$ when α is close, but still a finite distance above f (see Supp. Fig. 2 top). Thus $\mathbf{x}^0$ is typically not the global minimum of L_p for all $\alpha > f$, and so the perfect reconstruction solution predicted by the RS analysis must be unstable, at least for α close to f.

We did however check numerically that $\mathbf{x}^0$ is a local minimum of the L_p minimization problem for all $\alpha > f$. We did this by initializing the IRLS algorithm at $\mathbf{x}^0$ itself, as opposed to initializing it randomly (see Supp. Fig. 2 bottom). We found that for $\alpha > f$, $\mathbf{x}^0$ is typically a fixed point of the IRLS algorithm, indicating the local stability of $\mathbf{x}^0$ in the L_p potential in this regime of f and α. Thus remarkably, the zero error RS solution, while unstable, nevertheless correctly predicts the regime in which $\mathbf{x}^0$ is a metastable state.

5.2 The Error Regime

Here we describe the predictions of the RS theory for L_p minimization in the error regime. In this case, the scaling of the order parameters is identical to that used in the L_1 case (Eqs. (37) and (38)), yielding an effective Hamiltonian

$$H_{eff} = \frac{\beta}{\Delta q}\left[\frac{1}{2}\left(x - x^0 - z\sqrt{q_0}\right)^2 + \Delta q |x|^p\right]. \tag{74}$$

Here we have made the change of variables $x = u + x^0$, and so x is subject to the constraint $x \geq 0$. Again, since in the limit of $\beta \to \infty$, x itself is $O(1)$, while its fluctuations are small, the integrals over x are dominated by the global minimum x_m of H_{eff}. The condition for a local extremum of H_{eff} in the interior of the domain $x > 0$ is given by

$$x + p\,\Delta q\, x^{p-1} = x^0 + z\sqrt{q_0}. \tag{75}$$

Now, the minimum of the LHS of (75) occurs when $x = x_b \equiv [p(1-p)\Delta q]^{1/(2-p)}$. Thus (75) will have solutions only when $x^0 + z\sqrt{q_0} > x_b + p\,\Delta q\, x_b^{p-1}$. In this case, (75) has precisely two solutions, with the larger one, which is greater than x_b, corresponding to a local minimum of H_{eff}, and the smaller one corresponding to a local maximum. In order for this local minimum to be the global minimum, it must still be smaller than H_{eff} at the boundary $x = 0$, yielding the added condition,

$$H_{eff}(x_m) < H_{eff}(0) \qquad \Longrightarrow \qquad x_m > x_* \equiv [2(1-p)\Delta q]^{1/(2-p)}. \tag{76}$$

If this condition is not satisfied, then $x_m = 0$ is the global minimum of H_{eff}. In summary, we have x_m is the larger solution to (75) when $z > z_* \equiv (x_* + p\,\Delta q\, x_*^{p-1} - x^0)/\sqrt{q_0}$, and otherwise, $x_m = 0$.

It is now straightforward to compute the mean and variance of u for any given z, but to obtain the order parameter equations (12) and (13) we must still integrate over z. To perform this integral, it is useful to first make a change of variables from z to $x_m(z)$. For $z > z_*$, $x_m > x_*$ and x_m obeys (75), yielding the Jacobian

$$\frac{\partial x_m}{\partial z} = \frac{\sqrt{q_0}}{1 - p\,(1-p)\,\Delta q\, x_m^{p-2}}. \tag{77}$$

The distribution G on x_m induced by the gaussian distribution on z is then

$$G(x_m) = \theta(x_m - x^*)\frac{1 - p\,(1-p)\,\Delta q\, x_m^{p-2}}{\sqrt{2\pi q_0}} e^{-\frac{1}{2q_0}(x_m + p\,\Delta q\, x_m^{p-1} - x^0)^2} + \delta(x_m)(1 - H(z_*)), \tag{78}$$

where θ is the usual step function. Then expressing the mean and variance of $u = x - x^0$ in terms of x_m, the order parameter equations (12) and (13) become

$$\alpha q_0 = \left\langle\left\langle \int dx_m G(x_m)(x_m - x^0)^2 \right\rangle\right\rangle_{x^0} \tag{79}$$

$$\alpha = \left\langle\left\langle \int dx_m G(x_m) \frac{1}{(\theta(x_m - x^*)(1 - p\,(1-p)\,\Delta q\, x_m^{p-2})} \right\rangle\right\rangle_{x^0}, \tag{80}$$

where z_* and x_* are defined above. Also, the average over x^0 includes both the zero and nonzero components. Furthermore, in terms of the order parameters, the distribution of the reconstruction x, given the true value x^0 is simply $P(x|x^0) = G(x)$ with $G(x)$ defined in (78).

Thus the RS theory predicts a gap in the distribution whose width is x_*. The existence of this gap is a generic prediction of the RS theory; the distribution of x is induced by the gaussian distribution of z, under the map $z \to x_m(z)$ where x_m is the global minimum of H_{eff} in (74). Two terms contribute to H_{eff}: the first term involving z is an effective quadratic cost arising from the quenched disorder in the constraint of Eq. 1 of the main paper, while the second term arises from the L_p norm. The second term alone favors the global minimum to be at 0. However, because $p < 1$, this term rises sharply at 0, and so, as z varies, the quadratic cost is incapable of shifting the global minimum away from 0 by an amount less than x_*.

While the existence of a gap is an intriguing prediction, we see no evidence for it numerically. We have applied IRLS to L_p minimization and compared the reconstruction distribution to the theoretical predictions of the RS theory (equations (78) to (80)) (see Supp. Fig. 3). At values of α, f and p for which the RS theory predicts a large gap, we see none in the IRLS simulations. Based on this observation, it is likely the above equations yield unstable solutions, just as in the perfect reconstruction regime for L_p minimization.

Given that the key reason for the existence of a gap arises from the fact that the quenched disorder only yields a quadratic contribution to H_{eff}, we initiated a preliminary analysis of the theory at the level of 1RSB. However we find that under the simplest choice of scaling for the order parameters, the 1RSB theory also generically predicts a gap in the reconstruction distribution. It may be the case that alternate scaling regimes exist at 1RSB, or further replica symmetry breaking is required, or that 1RSB is correct and IRLS is not finding solutions near the global minimum that would exhibit a gap in the reconstruction distribution. We leave the exploration of these issues for future work.

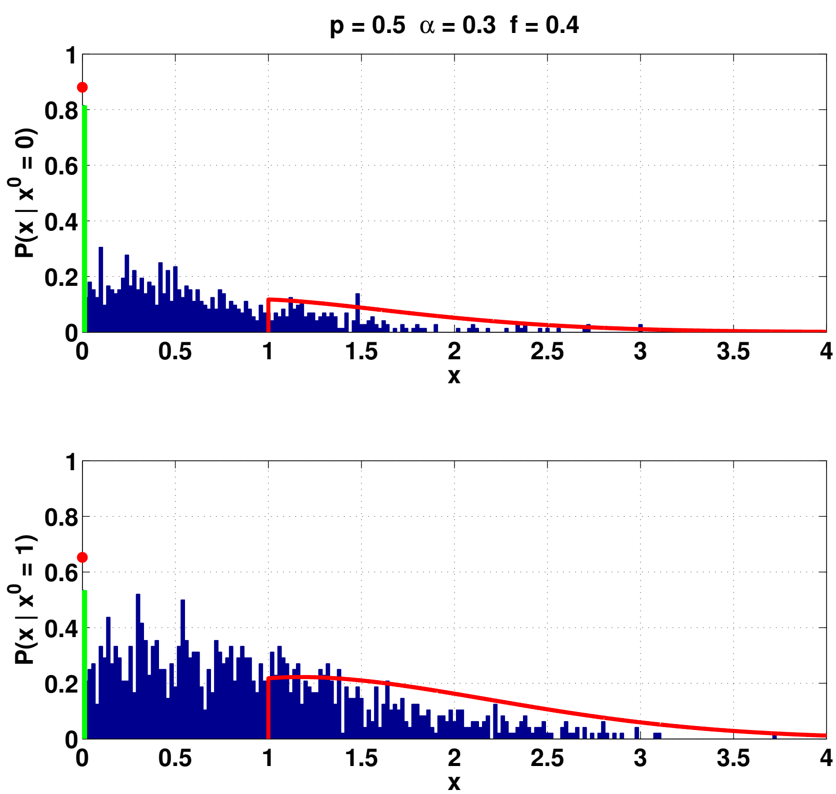

Figure 3: Mismatch between RS theory and IRLS simulations. The IRLS algorithm was run 100 times, to estimate the distribution of the reconstructed components conditioned on the true values (blue and green, same format as in Supp. Fig. 1). The red curves are the predictions of the RS theory, obtained by solving (79) and (80). Qualitatively similar behavior is seen at other values of f, p and α in the failure regime.