

# Supplementary Material for EPAPS

## Quenching the Anisotropic Heisenberg Chain: Exact Solution and Generalized Gibbs Ensemble

### THERMODYNAMIC LIMIT OF THE OVERLAPS

In order to apply the quench action logic we need to compute the thermodynamic limit of the overlap coefficients and in particular their leading extensive part

$$S[\rho] = \lim_{\text{th}} S_{\lambda} = -\lim_{\text{th}} \ln \frac{\langle \Psi_0 | \{\pm \lambda_j\}_{j=1}^{M/2} \rangle}{\|\{\pm \lambda_j\}_{j=1}^{M/2}\|}. \quad (\text{S1})$$

The procedure is to consider the overlap coefficient for a generic finite size Bethe state  $|\{\lambda_j\}_{j=1}^M\rangle$  that in the thermodynamic limit,  $N \rightarrow \infty$  with  $M/N = 1/2$  fixed, flows to a set of distributions  $|\{\lambda_j\}_{j=1}^M\rangle \rightarrow |\rho\rangle$ . Equivalently, in the thermodynamic limit the eigenvalue of a smooth diagonal observable  $\mathcal{A}$  can be recast into a sum of integrals weighted by the distributions  $\rho = \{\rho_n\}_{n=1}^{\infty}$ :

$$\mathcal{A}|\{\lambda_j\}_{j=1}^M\rangle = \left[ \sum_{j=1}^M A_j \right] |\{\lambda_j\}_{j=1}^M\rangle \rightarrow \left[ N \sum_{n=1}^{\infty} \int_{-\pi/2}^{\pi/2} d\lambda \rho_n(\lambda) \tilde{A}_n(\lambda) \right] |\rho\rangle. \quad (\text{S2})$$

By assumption the extensive part of the overlap coefficient  $S[\rho]$  is smooth and does not depend on finite size differences within the set of Bethe states that scale to the same densities  $\rho$ . The number of Bethe states for each set of distributions  $\rho$  is given by the extensive Yang-Yang entropy [1]:  $e^{S_{YY}[\rho]}$ . We are then free to select a representative finite size Bethe state out of all states that scale to the same  $\rho$ . We consider the state  $|\{\lambda_j\}_{j=1}^M\rangle$  consisting of  $2n_s$  strings such that  $2n_s = \sum_{n=1}^{\infty} M_n$ , where  $M_n$  is the number of  $n$ -strings in the state and we choose all  $M_n$  to be even. Note that this is one possible choice for a representative state. Different choices regarding the evenness of the fillings  $\{M_n\}_{n=1}^{\infty}$  lead to different expressions for the exact overlap formula (S4), but are believed [2, 3] to have the same extensive smooth part  $S[\rho]$ .

For any finite size  $N$ , Bethe states are organized in deviated strings. We use the following notation to label the rapidities of such states:

$$\lambda_j \rightarrow \lambda_{\alpha}^{n,a} = \lambda_{\alpha}^n + \frac{i\eta}{2}(n+1-2a) + i\delta_{\alpha}^{n,a}, \quad (\text{S3})$$

where  $a = 1, \dots, n$  and  $\alpha = 1, \dots, M_n$ . The string deviations  $\delta_{\alpha}^{n,a}$  are vanishing in the thermodynamic limit. Though the string hypothesis is not systematically verified around the ground state of the zero-magnetized spin chain [4, 5], it is effectively verified away from the ground state, for example at finite temperatures [6], and by extension in the circumstances discussed here.

The finite size overlap formula between this class of representative states and the Néel state is given in Ref. [7],

$$\frac{\langle \Psi_0 | \{\pm \lambda_j\}_{j=1}^{M/2} \rangle}{\|\{\pm \lambda_j\}_{j=1}^{M/2}\|} = \gamma \times \sqrt{\frac{\det_{M/2}(G_{jk}^+)}{\det_{M/2}(G_{jk}^-)}} \quad \text{with} \quad \gamma = \prod_{j=1}^{M/2} \frac{\sqrt{\tan(\lambda_j + i\eta/2) \tan(\lambda_j - i\eta/2)}}{2 \sin(2\lambda_j)}. \quad (\text{S4})$$

The prefactor  $\gamma$  has to leading order no explicit system size dependence from the string deviations  $\delta \rightarrow 0$ , but is exponentially vanishing when the particle number  $M$  is sent to infinity due to the product over all rapidities. We focus on the ratio of the two determinants, where the matrices are given by

$$G_{(n,\alpha,a),(m,\beta,b)}^{\pm} = \delta_{(n,\alpha,a),(m,\beta,b)} \left( NK_{\eta/2}(\lambda_{\alpha}^{n,a}) - \sum_{(\ell,\gamma,c)} K_{\eta}^+(\lambda_{\alpha}^{n,a}, \lambda_{\gamma}^{\ell,c}) \right) + K_{\eta}^{\pm}(\lambda_{\alpha}^{n,a}, \lambda_{\beta}^{m,b}). \quad (\text{S5})$$

Here,  $K_{\eta}^{\pm}(\lambda, \mu) = K_{\eta}(\lambda - \mu) \pm K_{\eta}(\lambda + \mu)$  and  $K_{\eta}(\lambda) = \frac{\sinh(2\eta)}{\sin(\lambda + i\eta) \sin(\lambda - i\eta)}$ . Divergences in system size as  $1/\delta$  occur in each string block ( $n = m, \alpha = \beta$ ) when  $b = a + 1$  in the term  $K_{\eta}(\lambda_{\alpha}^{n,a} - \lambda_{\alpha}^{n,a+1}) \sim \frac{i}{(\delta_{\alpha}^{n,a+1} - \delta_{\alpha}^{n,a})}$ . On the other hand, for our representative state the terms  $\pm K_{\eta}(\lambda + \mu)$  in  $G^{\pm}$  are never divergent since all string centers are strictly positive.

We conclude that the divergences in  $1/\delta$  in  $\det_{M/2}(G^+)$  cancel exactly the divergences in  $\det_{M/2}(G^-)$ , as they occur in exactly the same form. The same cancellation applies to divergences appearing in  $K_\eta(\lambda - \mu)$  when two rapidities from different strings get close in the thermodynamic limit  $\mu \rightarrow \lambda \pm i\eta + g(N)$  with  $\lim_{\text{th}} g(N) = 0$ . We are then able to take the thermodynamic limit  $\lim_{\text{th}}$  for the overlap coefficients analogously to Ref. [8]. Being non-exponential in system size, the ratio of the two determinants can then be neglected in the contribution to the overlaps that is leading in system size. The thermodynamic overlap coefficients then read as

$$S[\rho] = \lim_{\text{th}} S_\lambda = -\frac{N}{2} \sum_{n=1}^{\infty} \int_0^{\pi/2} d\lambda \rho_n(\lambda) \ln W_n(\lambda), \quad (\text{S6})$$

where

$$W_n(\lambda) = \frac{1}{2^{n+1} \sin^2 2\lambda} \frac{\cosh n\eta - \cos 2\lambda}{\cosh n\eta + \cos 2\lambda} \prod_{j=1}^{\frac{n-1}{2}} \left( \frac{\cosh(2j-1)\eta - \cos 2\lambda}{(\cosh(2j-1)\eta + \cos 2\lambda)(\cosh 4\eta j - \cos 4\lambda)} \right)^2, \quad (\text{S7})$$

if  $n$  odd, and

$$W_n(\lambda) = \frac{\tan^2 \lambda \cosh n\eta - \cos 2\lambda}{2^n \cosh n\eta + \cos 2\lambda} \frac{1}{\prod_{j=1}^{\frac{n}{2}} (\cosh 2(2j-1)\eta - \cos 4\lambda)^2} \prod_{j=1}^{\frac{n-2}{2}} \left( \frac{\cosh 2j\eta - \cos 2\lambda}{\cosh 2j\eta + \cos 2\lambda} \right)^2, \quad (\text{S8})$$

if  $n$  even.

## GTBA EQUATIONS FOR THE NÉEL-TO-XXZ QUENCH

In this section we focus on the derivation of the saddle point state, specified by the set of distribution  $\rho^{\text{sp}}$  obtained by varying the quench action  $S_{QA}[\rho] = 2S[\rho] - \frac{1}{2}S_{YY}[\rho]$  with respect to each root density  $\rho_n(\lambda)$ . Variation of the Yang-Yang entropy is well-known [1]. It should be noted that in front of the Yang-Yang entropy there is an extra factor  $\frac{1}{2}$ . The reason is that only parity-invariant Bethe states contribute and therefore the number of microstates in the ensemble  $\rho$  is the square root of the usual number of microstates. Furthermore, only states in the magnetization sector  $M = N/2$  have non-zero overlap with the initial Néel state. In order to vary with respect to all  $\rho_n(\lambda)$  independently, we need to add a Lagrange multiplier term to the quench action:

$$-h N \left( \sum_{m=1}^{\infty} m \int_{-\pi/2}^{\pi/2} d\lambda \rho_m(\lambda) - \frac{1}{2} \right), \quad (\text{S9})$$

where  $h$  is the Lagrange multiplier. Variation with respect to  $\rho_n(\lambda)$  then leads to the saddle point conditions

$$\ln \eta_n(\lambda) = -2h n - \ln W_n(\lambda) + \sum_{m=1}^{\infty} a_{nm} * \ln(1 + \eta_m^{-1})(\lambda), \quad (\text{S10})$$

where  $n \geq 1$ ,  $\eta_n(\lambda) = \rho_{n,h}(\lambda)/\rho_n(\lambda)$  and the convolution  $*$  is defined by

$$(f * g)(\lambda) = \int_{-\pi/2}^{\pi/2} d\mu f(\lambda - \mu)g(\mu) \quad (\text{S11})$$

The kernels are defined by  $a_{nm}(\lambda) = (1 - \delta_{nm})a_{|n-m|}(\lambda) + 2a_{|n-m|+2}(\lambda) + \dots + 2a_{n+m-2}(\lambda) + a_{n+m}(\lambda)$ , where

$$a_n(\lambda) = \frac{1}{\pi} \frac{\sinh(n\eta)}{\cosh(n\eta) - \cos(2\lambda)}. \quad (\text{S12})$$

The functions  $-2hn - \ln(W_n)$  are called driving terms. For each fixed value of  $h$  the set (S10) of Generalized Thermodynamic Bethe Ansatz (GTBA) equations has a solution in terms of the functions  $\eta_n$ . Substituting this solution into thermodynamic Bethe equations leads to the saddle point distributions  $\rho^{\text{sp}}$ . The parameter  $h$  is fixed by the magnetization condition  $M/N = 1/2$  of the initial Néel state,

$$\sum_{m=1}^{\infty} m \int_{-\pi/2}^{\pi/2} d\lambda \rho_m^{\text{sp}}(\lambda) = \frac{1}{2}. \quad (\text{S13})$$

As for the TBA equations at finite temperature [9], one can recast the GTBA Eqs. (S10) into a factorized form where there is no infinite sum over string lengths. We will now derive this result. We use the Fourier transform conventions

$$\hat{f}_k = \text{FT}[f](k) = \int_{-\pi/2}^{\pi/2} e^{2ik\lambda} f(\lambda) d\lambda, \quad k \in \mathbb{Z}, \quad (\text{S14})$$

$$f(\lambda) = \text{FT}^{-1}[\hat{f}](\lambda) = \frac{1}{\pi} \sum_{k \in \mathbb{Z}} e^{-2ik\lambda} \hat{f}_k. \quad (\text{S15})$$

The Fourier transforms of the kernels are  $\hat{a}_{n,k} = e^{-|k|n\eta}$ , and using the convolution theorem this implies  $a_m * a_n = a_{m+n}$ . A set of identities for the kernels can then be derived [9]

$$(a_0 + a_2) * a_{nm} = a_1 * (a_{n-1,m} + a_{n+1,m}) + (\delta_{n-1,m} + \delta_{n+1,m}) a_1, \quad n \geq 2, m \geq 1, \quad (\text{S16a})$$

and

$$(a_0 + a_2) * a_{1,m} = a_1 * a_{2,m} + a_1 \delta_{2,m}, \quad m \geq 1, \quad (\text{S16b})$$

where we used the convention  $a_0(\lambda) = \delta(\lambda)$ . Convolving the GTBA Eqs. (S10) with  $(a_0 + a_2)$ , the infinite sum can be removed, and we find that

$$(a_0 + a_2) * \ln(\eta_n) = (a_0 + a_2) * g_n - a_1 * (g_{n-1} + g_{n+1}) + a_1 * [\ln(1 + \eta_{n-1}) + \ln(1 + \eta_{n+1})]. \quad (\text{S17})$$

Here, the driving terms of the original GTBA equations are rewritten in a more convenient form with  $g_n(\lambda) = -\ln W_n(\lambda) - 2n \ln 4$ , where

$$g_n = \sum_{l=0}^{n-1} \ln \left[ \frac{s_{n-1-2l} c_{n-1-2l} s_{-n+1+2l} c_{-n+1+2l}}{t_{n-2l} t_{-n+2l}} \right], \quad (\text{S18a})$$

$$t_n = \frac{s_n}{c_n}, \quad s_n(\lambda) = \sin \left( \lambda + \frac{i\eta n}{2} \right), \quad c_n(\lambda) = \cos \left( \lambda + \frac{i\eta n}{2} \right). \quad (\text{S18b})$$

Defining  $g_0(\lambda) = 0$  and  $\eta_0(\lambda) = 0$ , Eq. (S17) holds for  $n \geq 1$ . Let us rewrite the new driving terms  $\tilde{d}_n = (a_0 + a_2) * g_n - a_1 * (g_{n-1} + g_{n+1})$  of Eq. (S17). We first rewrite  $g_n$  such that only positive indices are present:

$$g_n = 2\delta_{n \bmod 2, 1} \ln \left[ s_0^{(2)} \right] + 4 \sum_{l=1}^{\lfloor n/2 \rfloor} \ln \left[ s_{n+1-2l}^{(2)} \right] + 2 \sum_{l=1}^{n-1} \ln \left[ \frac{c_l^{(2)}}{s_l^{(2)}} \right] + \ln \left[ \frac{c_0^{(2)}}{s_0^{(2)}} \right] + \ln \left[ \frac{c_n^{(2)}}{s_n^{(2)}} \right], \quad (\text{S19})$$

where  $s_l^{(2)} = s_l s_{-l}$ ,  $c_l^{(2)} = c_l c_{-l}$  or, explicitly,

$$s_l^{(2)}(\lambda) = \sin \left( \lambda + \frac{i\eta l}{2} \right) \sin \left( \lambda - \frac{i\eta l}{2} \right) = \sin^2(\lambda) + \sinh^2 \left( \frac{\eta l}{2} \right), \quad (\text{S20a})$$

$$c_l^{(2)}(\lambda) = \cos \left( \lambda + \frac{i\eta l}{2} \right) \cos \left( \lambda - \frac{i\eta l}{2} \right) = \cos^2(\lambda) + \sinh^2 \left( \frac{\eta l}{2} \right). \quad (\text{S20b})$$

Now we use that for  $\tilde{a}_\alpha(\lambda) = \frac{1}{2\pi} \frac{\sinh(2\alpha)}{\sin^2(\lambda) + \sinh^2(\alpha)}$  and  $f_\beta(\lambda) = \ln[\sin^2(\lambda) + \sinh^2(\beta)]$  the following relation holds ( $\alpha, \beta > 0$ ):

$$\tilde{a}_\alpha * f_\beta = f_{\alpha+\beta} - 2\alpha. \quad (\text{S21})$$

This implies the identities

$$a_m * \ln \left[ \frac{c_l^{(2)}}{s_l^{(2)}} \right] = \ln \left[ \frac{c_{l+m}^{(2)}}{s_{l+m}^{(2)}} \right], \quad a_l * \ln \left[ \frac{s_0^{(2)}}{s_2^{(2)}} \right] = \ln \left[ \frac{s_l^{(2)}}{s_{l+2}^{(2)}} \right] \quad \text{and} \quad a_l * \ln \left[ \frac{c_0^{(2)}}{c_2^{(2)}} \right] = \ln \left[ \frac{c_l^{(2)}}{c_{l+2}^{(2)}} \right]. \quad (\text{S22})$$

From this we can calculate  $\tilde{d}_{2n}$  and  $\tilde{d}_{2n-1}$  for all  $n \geq 1$  explicitly:

$$\tilde{d}_{2n} = \ln \left[ \frac{c_0^{(2)}}{c_2^{(2)}} \right] - \ln \left[ \frac{s_0^{(2)}}{s_2^{(2)}} \right], \quad \tilde{d}_{2n-1} = \ln \left[ \frac{c_0^{(2)}}{c_2^{(2)}} \right] + \ln \left[ \frac{s_0^{(2)}}{s_2^{(2)}} \right]. \quad (\text{S23})$$

The GTBA equations can be written compactly as

$$(a_0 + a_2) * \ln(\eta_n) = \tilde{d}_n + a_1 * [\ln(1 + \eta_{n-1}) + \ln(1 + \eta_{n+1})], \quad (\text{S24a})$$

where  $n \geq 1$ , the  $\lambda$ -dependence is left implicit and by convention  $\eta_0(\lambda) = 0$  and  $a_0(\lambda) = \delta(\lambda)$ . The driving terms are given by

$$\tilde{d}_n(\lambda) = \ln \left[ \frac{\cos^2(\lambda)}{\cos^2(\lambda) + \sinh^2(\eta)} \right] - (-1)^n \ln \left[ \frac{\sin^2(\lambda)}{\sin^2(\lambda) + \sinh^2(\eta)} \right]. \quad (\text{S24b})$$

Using the convolution theorem once again, one can invert the operation of  $(a_0 + a_2)*$  and bring it to the right hand side of Eq. (S24a). The Fourier transforms of the driving terms are

$$\hat{\tilde{d}}_{n,k} = 2\pi \frac{(1 - e^{-2\eta|k|})}{|k|} \left( \frac{(-1)^n - (-1)^k}{2} \right). \quad (\text{S25})$$

Defining

$$\begin{aligned} \hat{d}_{n,k} &:= \frac{\hat{\tilde{d}}_{n,k}}{\hat{a}_{0,k} + \hat{a}_{2,k}} = 2\pi \frac{\tanh(\eta k)}{k} \left( \frac{(-1)^n - (-1)^k}{2} \right), \\ \hat{s}_k &:= \frac{\hat{a}_{1,k}}{\hat{a}_{0,k} + \hat{a}_{2,k}} = \frac{1}{2 \cosh(k\eta)}, \end{aligned} \quad (\text{S26})$$

the GTBA equations in Fourier space are

$$\text{FT}[\ln(\eta_n)](k) = \hat{d}_{n,k} + \hat{s}_k \left( \text{FT}[\ln(1 + \eta_{n-1})](k) + \text{FT}[\ln(1 + \eta_{n+1})](k) \right). \quad (\text{S27})$$

In real space they are precisely the GTBA equations of the main text.

### NEAREST-NEIGHBOR SPIN-SPIN CORRELATION

In this section we show how to compute the expectation value of the nearest-neighbor spin-spin correlator,  $\langle \sigma_j^z \sigma_{j+1}^z \rangle = 1 + \frac{4}{NJ} \langle \frac{\partial H}{\partial \Delta} \rangle$ , in the thermodynamic limit on a generic, translationally invariant Bethe eigenstate specified by a set of smooth distributions  $\boldsymbol{\rho}$ . This is a direct consequence of the Hellmann-Feynman theorem [10]. Independently, an analogous implementation was recently presented in Ref. [11].

A naive application of the Hellmann-Feynman theorem on the saddle point state leads to a wrong result

$$\langle \boldsymbol{\rho}^{\text{sp}} | \frac{1}{N} \frac{\partial H}{\partial \Delta} | \boldsymbol{\rho}^{\text{sp}} \rangle \neq \frac{\partial}{\partial \Delta} \langle \boldsymbol{\rho}^{\text{sp}} | \frac{1}{N} H | \boldsymbol{\rho}^{\text{sp}} \rangle = -\frac{J}{2}. \quad (\text{S28})$$

This is because the overlaps of two Bethe states with different values of  $\Delta$  are exponentially small in system size, which makes the thermodynamic limit and the derivative with respect to  $\Delta$  noncommuting,

$$\lim_{\delta \rightarrow 0} \lim_{\text{th}} \langle \boldsymbol{\lambda}(\Delta) | \boldsymbol{\lambda}(\Delta + \delta) \rangle \neq \lim_{\text{th}} \lim_{\delta \rightarrow 0} \langle \boldsymbol{\lambda}(\Delta) | \boldsymbol{\lambda}(\Delta + \delta) \rangle. \quad (\text{S29})$$

In order to apply the Hellmann-Feynman theorem we need to take the derivative of the energy eigenvalue of a generic Bethe state at finite size  $N$  and then take the thermodynamic limit. Under the string hypothesis the energy eigenvalue  $\omega_{\boldsymbol{\lambda}}$  becomes a function only of the string centers  $\lambda_{\alpha}^n$  of the Bethe state, since the string deviations vanish exponentially in system size,

$$\omega_{\boldsymbol{\lambda}} = -\pi J \sinh(\eta) \sum_{n,\alpha} a_n(\lambda_{\alpha}^n), \quad (\text{S30})$$

where  $a_n$  is defined in Eq. (S12). We can now apply the Hellmann-Feynman theorem to this finite size state by taking the derivative of the energy with respect to  $\Delta$ :

$$\langle \boldsymbol{\lambda} | \sigma_1^z \sigma_2^z | \boldsymbol{\lambda} \rangle = 1 + \frac{4}{NJ} \frac{d\omega_{\boldsymbol{\lambda}}}{d\Delta} = 1 + \frac{4}{NJ} \frac{1}{\sinh \eta} \frac{d\omega_{\boldsymbol{\lambda}}}{d\eta} = 1 - \frac{4\pi}{N} \sum_{n,\alpha} \left[ \frac{\cosh \eta}{\sinh \eta} a_n(\lambda_{\alpha}^n) + (\partial_{\eta} a_n)(\lambda_{\alpha}^n) + \partial_{\lambda_{\alpha}^n} a_n(\lambda_{\alpha}^n) \frac{d\lambda_{\alpha}^n}{d\eta} \right]. \quad (\text{S31})$$

The quantities  $\frac{d\lambda_\alpha^n}{d\eta}$  are obtained by deriving the reduced Bethe equations for string centers [12, 13]

$$\theta_n(\lambda_\alpha^n) - \frac{1}{N} \sum_{m,\beta} \theta_{nm}(\lambda_\alpha^n - \lambda_\beta^m) = \frac{2\pi}{N} I_\alpha^n \quad (\text{S32})$$

(see [9] for the definitions of  $\theta_n$  and  $\theta_{nm}$ ) with respect to the interaction parameter  $\eta$ ,

$$N \tilde{a}_n(\lambda_\alpha^n) - \sum_{m,\beta} \tilde{a}_{nm}(\lambda_\alpha^n - \lambda_\beta^m) + \sum_{m,\beta} G_{(n,\alpha),(m,\beta)} \frac{d\lambda_\beta^m}{d\eta} = 0. \quad (\text{S33})$$

Here we introduced the reduced Gaudin matrix for string centers [14]

$$G_{(n,\alpha),(m,\beta)} = \delta_{(n,\alpha),(m,\beta)} \left( N a_n(\lambda_\alpha^n) - \sum_{k,\gamma} a_{nk}(\lambda_\alpha^n - \lambda_\gamma^k) \right) + a_{nm}(\lambda_\alpha^n - \lambda_\beta^m) \quad (\text{S34})$$

and the functions  $\tilde{a}_n$  and  $\tilde{a}_{nm}$ ,

$$\tilde{a}_n(\lambda) = \frac{1}{2\pi} \partial_\eta \theta_n(\lambda) = -a_n(\lambda) \frac{n \sin 2\lambda}{2 \sinh n\eta}, \quad (\text{S35})$$

$$\tilde{a}_{nm}(\lambda) = \frac{1}{2\pi} \partial_\eta \theta_{nm} = (1 - \delta_{nm}) \tilde{a}_{|n-m|}(\lambda) + 2\tilde{a}_{|n-m|+2}(\lambda) + \dots + 2\tilde{a}_{n+m-2}(\lambda) + \tilde{a}_{n+m}(\lambda). \quad (\text{S36})$$

We define a set of the auxiliary functions  $\mathbf{h} = \{h_n\}_{n=1}^\infty$  such that

$$h_n(\lambda_\alpha^n) = \sum_{m,\beta} (G^{-1})_{(n,\alpha),(m,\beta)} \left( N \tilde{a}_m(\lambda_\beta^m) - \sum_{k,\gamma} \tilde{a}_{mk}(\lambda_\beta^m - \lambda_\gamma^k) \right). \quad (\text{S37})$$

In the thermodynamic limit, expression (S31) can then be recast as a functional of the distributions  $\boldsymbol{\rho}$  and the auxiliary functions  $\mathbf{h}$ ,

$$\langle \boldsymbol{\rho} | \sigma_1^z \sigma_2^z | \boldsymbol{\rho} \rangle = 1 + \lim_{\text{th}} \frac{4}{NJ} \frac{d\omega_\lambda}{d\Delta} = 1 - 4\pi \sum_{n=1}^\infty \int_{-\pi/2}^{\pi/2} d\lambda \rho_n(\lambda) \left[ \frac{\cosh \eta}{\sinh \eta} a_n(\lambda) + \partial_\eta a_n(\lambda) - (\partial_\lambda a_n(\lambda)) h_n(\lambda) \right], \quad (\text{S38})$$

where the auxiliary functions are determined by a set of linear integral equations

$$\rho_{n,t}(\lambda) h_n(\lambda) + \sum_{m=1}^\infty a_{nm} * (\rho_m h_m)(\lambda) = \tilde{a}_n(\lambda) - \sum_{m=1}^\infty \tilde{a}_{nm} * \rho_m(\lambda). \quad (\text{S39})$$

As in the case of GTBA Eqs. (S10), they can be reduced to

$$(a_0 + a_2) * (\rho_{n,t} h_n) = \left[ (a_0 + a_2) * d_n - a_1 * (d_{n-1} + d_{n+1}) \right] + a_1 * (h_{n-1} \rho_{n-1,h} + h_{n+1} \rho_{n+1,h}), \quad (\text{S40})$$

with  $h_0(\lambda) = d_0(\lambda) = 0$  and the driving terms given by ( $n \geq 1$ )

$$d_n(\lambda) = \tilde{a}_n(\lambda) - \sum_{m=1}^\infty \tilde{a}_{nm} * \rho_m(\lambda). \quad (\text{S41})$$

Analogously to the local conserved charges, the correlator (S38) can be expressed solely in terms of  $\rho_{1,h}$  and the auxiliary function  $h_1$  (see [3]),

$$\begin{aligned} \langle \boldsymbol{\rho} | \sigma_1^z \sigma_2^z | \boldsymbol{\rho} \rangle = 1 + 4 \left\{ - \frac{\cosh \eta}{\sinh \eta} \sum_{k \in \mathbb{Z}} \left( \frac{e^{-|k|\eta} - \hat{\rho}_{1,h}(k)}{2 \cosh k\eta} \right) + \sum_{k \in \mathbb{Z}} |k| \left[ \frac{e^{-|k|\eta}}{2 \cosh k\eta} + \tanh(|k|\eta) \left( \frac{e^{-|k|\eta} - \hat{\rho}_{1,h}(k)}{2 \cosh k\eta} \right) \right] \right. \\ \left. - \pi \int_{-\pi/2}^{\pi/2} d\lambda \rho_{1,h}(\lambda) h_1(\lambda) \frac{\partial}{\partial \lambda} s(\lambda) \right\}, \quad (\text{S42}) \end{aligned}$$

where the function  $s$  is given in the main text and  $\hat{\rho}_{1,h}$  is the Fourier transform, defined in (S14), of  $\rho_{1,h}$ .

## NUMERICAL LINKED-CLUSTER EXPANSIONS

In this section we discuss the numerical linked-cluster expansion (NLCE) results. Such an expansion allows us to obtain the infinite-time average, or the diagonal ensemble (DE), result for the expectation value of spin-spin correlations. This approach was introduced in Ref. [15] to study quenches with initial thermal states. It can be straightforwardly tailored to study quenches with initial ground states as explained in detail, for the particular quenches considered in this work, in Ref. [16]. Here we report the results that are relevant to the discussion in the main text.

The NLCE calculations are done in clusters with up to 18 sites, as in Ref. [15]. We denote as  $\mathcal{O}_l^{\text{DE}}$  the result obtained for an observable  $\mathcal{O}$  when adding the contribution of all clusters with up to  $l$  sites. We should stress that while the convergence of the NLCE calculations improves for increasing values of  $\Delta$  [16], for large  $\Delta$  they do not allow us to discriminate between the results of the quench action (QA) and the GGE calculations, which are very close to each other. They also do not allow us to discriminate between the QA and GGE results as  $\Delta \rightarrow 1$  when the observable considered is  $\sigma_1^z \sigma_2^z$ . In that regime, the QA and GGE results are also too close to each other, becoming identical for  $\Delta = 1$ . As shown in Fig. 3 of the main text, the largest relative differences between the QA and GGE predictions are seen for  $\langle \sigma_1^z \sigma_3^z \rangle$  when  $\Delta \rightarrow 1$ . This is the observable and regime in which we focus our NLCE effort.

In Fig. 1, we show results for  $\langle \sigma_1^z \sigma_3^z \rangle_l^{\text{DE}}$  versus  $l$  for several values of  $\Delta$  between 1 and 3. A few remarks are in order on those results. They can be seen to be converging with increasing  $l$ , the amplitude of the oscillations decreases, but do not quite converge for the cluster sizes accessible to us. As  $\Delta$  decreases from 3, one can see that the convergence of the NLCE initially worsens, the amplitude of the oscillations increases for any two contiguous values of  $l$ , and then improves very close to  $\Delta = 1$ . Finally, it is apparent that the results for  $\langle \sigma_1^z \sigma_3^z \rangle^{\text{DE}}$  decrease with decreasing  $\Delta$ , as expected from the exact calculations discussed in the main text.

We now turn our attention to the regime in which  $\Delta$  is very close to 1. The fact that the convergence of our NLCE calculations improves in that regime is better seen in Fig. 2, where we report results for  $\Delta = 1, 1.01, 1.1, \text{ and } 1.2$ . However, the results of our bare NLCE sums still do not allow us to discriminate between the QA and GGE results for  $\Delta = 1$ , which are depicted as continuous and dashed horizontal lines, respectively. As discussed in NLCE studies of systems in thermal equilibrium [17, 18], whenever NLCE series do not converge to a desired accuracy, one can use resummation techniques to accelerate the convergence of the series and obtain more accurate results. There are two resummation techniques that were explained in detail in Ref. [18], which we have found to improve convergence in our problem. Those are Wynn's and Brezinski's algorithms. They can be applied multiple times (in what we call "cycles") to a series for an observable and are expected to improve convergence with each cycle. However, the application of too many cycles can also lead to numerical instabilities [18].

In a nutshell, every cycle of those algorithms leads to a new series with fewer elements [18]. The last element after each cycle is expected to approach the  $l \rightarrow \infty$  result. The reduction in the number of elements after each cycle limits the number of times each algorithm can be applied to a series. In Wynn's algorithm, each cycle reduces the number

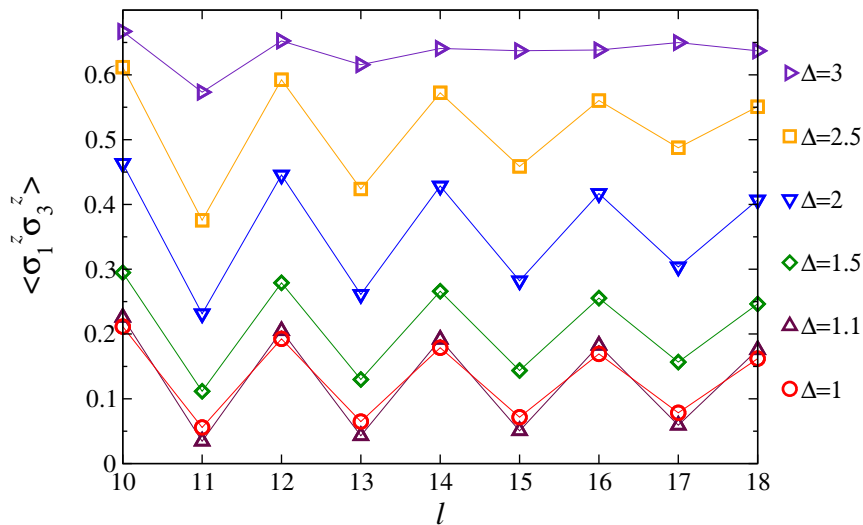


FIG. 1. Next-nearest neighbor correlations  $\langle \sigma_1^z \sigma_3^z \rangle_l^{\text{DE}}$  versus  $l$  for quenches with  $\Delta = 1, 1.1, 1.5, 2, 2.5, \text{ and } 3$ . Results are reported for the last nine orders of the NLCE.

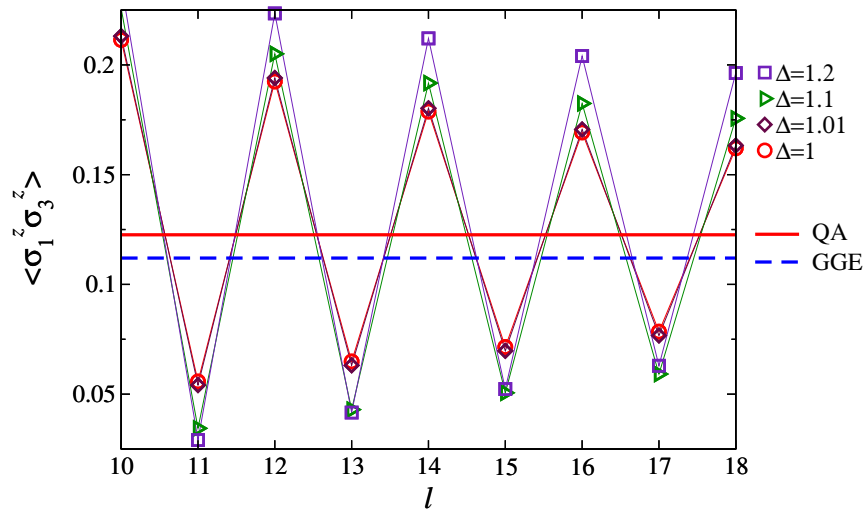


FIG. 2. Next-nearest neighbor correlations  $\langle \sigma_1^z \sigma_3^z \rangle_l^{\text{DE}}$  versus  $l$  for quenches with  $\Delta = 1, 1.01, 1.1,$  and  $1.2$ . Results are reported for the last nine orders of the NLCE. The QA and GGE predictions for  $\Delta = 1$  are reported as continuous and dashed horizontal lines, respectively.

of elements by two, and in Brezinski's algorithm each cycle reduces the number of elements by three. Since our bare series for  $\langle \sigma_1^z \sigma_3^z \rangle^{\text{DE}}$  has 18 elements (we consider clusters with up to 18 sites), the maximal number of cycles we can apply for Wynn's algorithm is 8 and for Brezinski's algorithm is 5.

In Fig. 3 we report the results (the last element) after each cycle of Wynn's and Brezinski's algorithms. As  $\Delta$  approaches 1, we find that the resummations are stable and lead to very close results. This gives us confidence in the robustness of the resummations around the Heisenberg point. Figure 3 shows that, as  $\Delta \rightarrow 1$  and as the number of cycles increases, Brezinski's algorithm appears to converge to a slightly larger value of  $\langle \sigma_1^z \sigma_3^z \rangle^{\text{DE}}$  than the QA prediction, while Wynn's resummations seem to converge to a result slightly below the QA prediction. Both are far from the GGE prediction. Hence, our NLCE results support the correctness of the QA predictions for the outcome of the relaxation dynamics in these systems.

In the main text, we report the results of Brezinski's resummations after one cycle as representative of the outcome of the resummation techniques used. This because those results are almost identical to the ones obtained using Wynn's algorithm after one cycle, and are in between the results obtained after the maximal number of cycles that

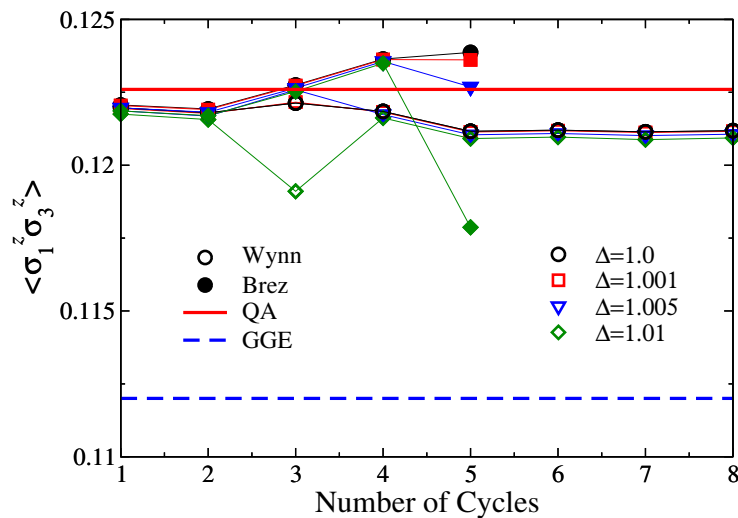


FIG. 3. Next-nearest neighbor correlations  $\langle \sigma_1^z \sigma_3^z \rangle^{\text{DE}}$  after each cycle of Wynn's (empty symbols) and Brezinski's (filled symbols) algorithms. Results are reported for  $\Delta = 1, 1.001, 1.005,$  and  $1.01$ , and are compared to the QA (continuous line) and GGE (dashed line) results for  $\Delta = 1$ .

could be applied in each algorithm. In the main text, we also report an interval of confidence that contains all results obtained within Brezinski's and Wynn's resummations, except for  $\Delta = 0.015$  for which the last cycle of Brezinski's algorithm resulted in  $\langle \sigma_1^z \sigma_3^z \rangle^{\text{DE}} = 0.1346$ , which is out of the interval of confidence reported. We should add that the fluctuations in the values of  $\Delta$  after resummations increase as  $\Delta$  increases for  $\Delta < 2$ . Because of this, and the fact that the QA and GGE results approach each other with increasing  $\Delta$ , we cannot use our NLCE results to discriminate between the QA and GGE predictions as one moves further away from the Heisenberg point.

- 
- [1] V. E. Korepin, N. M. Bogoliubov, and A. G. Izergin, *Quantum Inverse Scattering Method and Correlation Functions* (Cambridge Univ. Press, 1993).
  - [2] J.-S. Caux and F. H. L. Essler, *Phys. Rev. Lett.* **110**, 257203 (2013).
  - [3] B. Wouters, M. Brockmann, J. De Nardis, D. Fioretto, R. Vlijm and J.-S. Caux, *to be published*.
  - [4] F. Woynarovich, *J. Phys. A: Math. Gen.* **15**, 2985 (1982).
  - [5] O. Babelon, H. J. de Vega, and C. M. Viallet, *Nucl. Phys. B* **220**, 13 (1983).
  - [6] A. M. Tselik and P. B. Wiegmann, *Adv. Phys.* **32**, 453 (1983).
  - [7] M. Brockmann, J. D. Nardis, B. Wouters, and J.-S. Caux, *J. Phys. A: Math. Theor.* **47**, 145003 (2014).
  - [8] J. De Nardis, B. Wouters, M. Brockmann, and J.-S. Caux, *Phys. Rev. A* **89**, 033601 (2014).
  - [9] M. Takahashi, *Thermodynamics of one-dimensional solvable models* (Cambridge University Press, Cambridge, 1999).
  - [10] R. P. Feynman, *Phys. Rev.* (1939).
  - [11] M. Mestyan and B. Pozsgay, arXiv:1405.0232.
  - [12] H. Bethe, *Zeit. für Physik* **71**, 205 (1931).
  - [13] M. Takahashi, *Prog. Theor. Phys.* **46**, 401 (1971).
  - [14] J.-S. Caux, R. Hagemans, and J. M. Maillet, *J. Stat. Mech.: Th. Exp.* **2005**, P09003 (2005).
  - [15] M. Rigol, *Phys. Rev. Lett.* **112**, 170601 (2014).
  - [16] M. Rigol, In preparation.
  - [17] M. Rigol, T. Bryant, and R. R. P. Singh, *Phys. Rev. Lett.* **97**, 187202 (2006).
  - [18] M. Rigol, T. Bryant, and R. R. P. Singh, *Phys. Rev. E* **75**, 061118 (2007).