Improved Bounds on the Spectral Gap above Frustration Free Ground States of Quantum Spin Chains *

WOLFGANG L SPITZER[1] **and SHANNON STARR**[2]
[1] *Department of Mathematics, University of California, Davis, Davis, CA 95616-8633, USA. email spitzer@math.ucdavis.edu*
[2] *Department of Physics, Princeton University, Princeton, NJ 08544, USA. email sstarr@math.princeton.edu*

(9 December 2002)

Abstract In [20, 21] Nachtergaele obtained explicit lower bounds for the spectral gap above many frustration free quantum spin chains by using the "martingale method". We present simple improvements to his main bounds which allow one to obtain a sharp lower bound for the spectral gap above the spin-1/2 ferromagnetic XXZ chain. As an illustration of the method, we also calculate a lower bound for the spectral gap of the AKLT model, which is about 1/3 the size of the expected gap.

Mathematics Subject Classifications (2000). 82B10, 82C10
Keywords: Martingale method, spectral gap, VBS states, XXZ model, AKLT model

1. Introduction

In this paper, we obtain lower bounds on the spectral gap above frustration free ground states, which are special ground states which exist for certain special quantum spin systems. In the simplest case, a "frustration free" quantum spin chain can be defined as follows: For each site, let $\mathcal{H}_x \equiv \mathbb{C}^d$ be a finite dimensional Hilbert space, such that all $\mathcal{H}_x$ are equivalent to one another. For any finite subset $\Lambda \subset \mathbb{Z}$, one defines $\mathcal{H}_\Lambda = \bigotimes_{x\in\Lambda} \mathcal{H}_x$. Consider an interaction Φ, defined on all finite subsets of $\mathbb{Z}$, such that $\Phi(\Lambda_0) \in \mathcal{B}(\mathcal{H}_{\Lambda_0})$ for all finite Λ_0. Then for any finite subset $\Lambda \subset \mathbb{Z}$, we define

$$H(\Lambda) = \sum_{\Lambda_0 \subset \Lambda} \Phi(\Lambda_0)\,,$$

where, as usual, $\Phi(\Lambda_0) \in \mathcal{B}(\mathcal{H}_{\Lambda_0})$ is extended to $\mathcal{H}_\Lambda$ by tensoring with the identity on $\mathcal{H}_{\Lambda\setminus\Lambda_0}$. This is the set-up for a typical quantum spin system. What distinguishes a frustration free quantum spin system are two assumptions

1. For each Λ_0, $\Phi(\Lambda_0)$ is a positive operator.
2. For each nonempty Λ, the subspace $\mathcal{G}(\Lambda) = \mathrm{Ker}(H(\Lambda))$ satisfies $\dim(\mathcal{G}(\Lambda)) > 0$.

In this case, we call $\mathcal{G}(\Lambda)$ the ground state space, and declare the ground states to be frustration free.

We point out an obvious implication. Since all $\Phi(\Lambda_0)$ are positive,

$$\mathrm{Ker}(H(\Lambda)) = \cap_{\Lambda_0\subset\Lambda} \mathrm{Ker}(\Phi(\Lambda_0))\,.$$

This means that any ground state $\psi \in \mathcal{G}(\Lambda)$ not only minimizes the total energy, i.e. is a minimum energy eigenvector of $H(\Lambda)$, it also minimizes all the local interaction energies. This is to be contrasted with a frustrated spin system.

To mention just a few interesting examples of frustration free systems, there are ferromagnets: the isotropic Heisenberg ferromagnet has a frustration free ground state, as do the anisotropic "kink" ferromagnets known as XXZ models. A frustration free antiferromagnet is the AKLT model, invented by Affleck, Kennedy, Lieb and Tasaki [1], which is an important model for understanding the role of spin-dimension in antiferromagnetic chains. (Haldane [12] conjectured that the antiferromagnetic, isotropic Heisenberg chain has gapless excitations above the ground state if the spin is half-integer but is gapfull if the spin takes an integer value.) There are numerous stochastic models, such that the Markov generators can be viewed as frustration free models. Of these, we just mention one, the Kac model [13], for which the gap has recently been computed exactly by Carlen, Carvalho and Loss [5] using a technique, similar in spirit, but different than the martingale method, which is specific to mean field models. We point this out to indicate that this is still a very active field of research.

Henceforth we will assume that Φ is only a nearest neighbor interaction, and that it is translation invariant. Neither of these requirements is essential.[2] This allows us to write $H_\Lambda = \sum_{x,x+1\in\Lambda} h(x,x+1)$, where $h(x,x+1)$ is a fixed positive operator of just two spins.

2. Main Theorem

The main result of this paper is a method of calculating lower bounds for the spectral gap, by diagonalizing a family of finite dimensional matrices, with dimensions much smaller than the starting Hamiltonians H_Λ, $\Lambda \subset \mathbb{Z}$. There are two situations where this is particularly useful: First, if there is a symmetry which allows explicit diagonalization of these matrices, which is the case for the spin-1/2 XXZ model. Second, for the finitely correlated states (see Section 4) the matrices involved are all of a fixed finite dimension. If the dimension is small enough, the matrices can be diagonalized by hand, which is the case for the AKLT model.

Suppose $\Lambda = [a,b]$ and $\Lambda_0 = [c,d] \subset \Lambda$. Then we define $H(c,d) = H_{\Lambda_0}$, interpreted as an operator on $\mathcal{H}_\Lambda$. We suppress the dependence on Λ from the notation. Similarly we define $\mathcal{G}(c,d) = \operatorname{Ker} H(c,d)$. We define $G(c,d) = \operatorname{Proj}(\mathcal{G}(c,d))$ to be the orthogonal projection onto $\mathcal{G}(c,d)$. The spectral gap, $\gamma(c,d)$, is defined to be the largest number such that

$$H(c,d) \geq \gamma(c,d)(1 - G(c,d))\,.$$

This number is independent of Λ, and depends only on $c - d + 1 = \#[c,d]$. We define $\gamma_N = \min_{2\leq n\leq N} \gamma(1,n)$, which has the useful property of being a nonincreasing function of N. A sensible convention is that $H(a,a) = 0$, $\mathcal{G}(a,a) = \mathcal{H}_a = \mathbb{C}^d$, and $G(a,a) = 1$. However, there is no way to define $\gamma(a,a)$, since both $H(a,a)$ and $1 - G(a,a)$ vanish.

The auxiliary Hamiltonian, whose analysis is easier than that of H_Λ and which gives information on $\gamma(a,b)$, is just the ground state projector restricted to an appropriate

[2] An example of a non-translation invariant frustration free model was treated in [6]. There, the XXZ model is perturbed by a defect, which nevertheless is still frustration free. In that paper, the "martingale method" is applied in a slightly different way because there the spin chain is grown on both edges, while we will grow the spin chain along one edge only.

subspace. Suppose $\Lambda_0 = [-m, n] \subset \Lambda = [a, b]$. Define

$$\varepsilon(m, n) = \sup\{\langle\psi|G(0, n)\psi\rangle : \psi \in \mathcal{G}(-m, 0)\,,\ \psi \in \mathcal{G}(-m, n)^{\perp}\,,\ \|\psi\| = 1\}\,. \tag{2.1}$$

The value of $\varepsilon(m, n)$ is independent of a and b as long as $a \leq -m$ and $b \geq n$, and by translation symmetry the answer would be the same if the interval $[-m, n]$ were replaced by $[x - m, x + n]$ for any x. For $m, n \geq 1$, we define

$$\varepsilon_{m,n} = \sup_{m' \geq m} \varepsilon(m', n)\,,$$

which is monotone decreasing in m, but not necessarily in n.

Our main result is the following:

Theorem 2.1. *Suppose that $\varepsilon_{m,n} < 1/2$, for some $n \geq m \geq 1$. For any $N \geq 2$ and $\psi \in \mathcal{G}(1, N)^{\perp}$, define*

$$\psi_1 = (1 - G(1, N - n))\psi\,,\ \psi_2 = G(1, N - n)\psi\,,$$

whence it follows $\psi = \psi_1 + \psi_2$. (If $N \leq n + 1$ then $\psi_1 = 0$, $\psi_2 = \psi$.) Then

$$\langle\psi|H(1, N)\psi\rangle \geq \gamma_{m+n}(\alpha(\varepsilon_{m,n})\|\psi_1\|^2 + \beta(\varepsilon_{m,n})\|\psi_2\|^2)\,, \tag{2.2}$$

where $\alpha(\varepsilon) = (\sqrt{1 - \varepsilon} - \sqrt{\varepsilon})^2$ and $\beta(\varepsilon) = \sqrt{1 - \varepsilon}(\sqrt{1 - \varepsilon} - \sqrt{\varepsilon})$.

Our improved bound is due to the following simple estimate.

Lemma 2.2. *(a) Suppose H is a positive matrix, $G = \mathrm{Proj}(\mathrm{Ker}(H))$, and γ is the gap; so that $H \geq \gamma(1 - G)$. If ϕ is a vector which is not in the kernel of H – not necessarily orthogonal to the kernel either – then for any ψ*

$$\langle\psi|H\psi\rangle \geq \gamma|\langle\psi|(1 - G)\phi\rangle|^2/\langle\phi|(1 - G)\phi\rangle\,. \tag{2.3}$$

(b) If G is an orthogonal projection, ϕ is a nonzero vector, and ψ is any vector orthogonal to ϕ, then

$$|\langle\psi|G\phi\rangle|^2 \leq \|\psi\|^2\|\phi\|^2\langle G\rangle_{\phi}(1 - \langle G\rangle_{\phi})\,, \tag{2.4}$$

where $\langle G\rangle_{\phi} = \langle\phi|G\phi\rangle/\|\phi\|^2$.

Proof: Both are by Cauchy-Schwarz. First,

$$\begin{aligned} |\langle\psi|(1 - G)\phi\rangle|^2 = |\langle(1 - G)\psi|(1 - G)\phi\rangle|^2 &\leq \|(1 - G)\psi\|^2\|(1 - G)\phi\|^2 \\ &\leq \gamma^{-1}\langle\psi|H\psi\rangle\|(1 - G)\phi\|^2\,. \end{aligned}$$

For the second, choose an orthonormal basis $\phi_1, \phi_2, \ldots$ with $\phi_1 = \phi/\|\phi\|$. Then

$$\begin{aligned} |\langle\psi|G\phi\rangle| \leq \sum_{n\geq 2}|\langle\psi|\phi_n\rangle\langle\phi_n|G\phi\rangle| &\leq \Big(\sum_{n\geq 2}|\langle\psi|\phi_n\rangle|^2\Big)^{1/2}\Big(\sum_{n\geq 2}\langle G\phi|\phi_n\rangle\langle\phi_n|G\phi\rangle\Big)^{1/2} \\ &= \|\psi\|\Big(\sum_{n\geq 2}\langle G\phi|\phi_n\rangle\langle\phi_n|G\phi\rangle\Big)^{1/2}\,. \end{aligned}$$

Moreover,

$$\sum_{n\geq 2}\langle G\phi|\phi_n\rangle\langle\phi_n|G\phi\rangle = \langle G\phi|G\phi\rangle - \langle G\phi|\phi_1\rangle\langle\phi_1|G\phi\rangle = \|G\phi\|^2(1 - \|G\phi\|^2/\|\phi\|^2)\,.$$

∎

Proof: (Of Theorem 2.1) The proof is by induction. The first step is to check that the proposition is trivially satisfied for $N \leq m+n$. Indeed, since $\alpha \leq \beta \leq 1$, we have

$$\gamma(N) \geq \gamma_{m+n} \geq \max(\alpha\gamma_{m+n}, \beta\gamma_{m+n}),$$

by the definition of γ_{m+n}. The next step is the induction step.

For the induction hypothesis we suppose the proposition is true whenever $N \leq N_0$, where $N_0 \geq m+n$, and then prove the proposition is also true for $N = N_0+1$. This means we must show

$$\langle\psi|H(1,N)\psi\rangle \geq \gamma_{m+n}[\alpha\|\psi_1\|^2 + \beta\|\psi_2\|^2] \tag{2.5}$$

where $G(1,N)\psi = 0$, $\psi_1 = [1-G(1,N-n)]\psi$ and $\psi_2 = [G(1,N-n) - G(1,N)]\psi$. It is useful to make a further orthogonal decomposition

$$\psi_{11} = [1-G(1,N-2n)]\psi_1\,,\ \ \psi_{12} = [G(1,N-2n) - G(1,N-n)]\psi_1\,.$$

(We have assumed that $N > 2n$, otherwise we set $\psi_{11} = 0$.) By the induction hypothesis, we conclude

$$\langle\psi_1|H(1,N-n)\psi_1\rangle \geq \gamma_{m+n}[\alpha\|\psi_{11}\|^2 + \beta\|\psi_{12}\|^2]\,.$$

Since $\psi_2 \in \mathcal{G}(1,N-n)$, it is clear that $\langle\psi|H(1,N-n)\psi\rangle = \langle\psi_1|H(1,N-n)\psi_1\rangle$. So to prove equation (2.5), we just need to prove

$$\langle\psi|H(N-n,N)\psi\rangle \geq \gamma_{m+n}(\beta\|\psi_2\|^2 - (\beta-\alpha)\|\psi_{12}\|^2)\,. \tag{2.6}$$

We assume

$$\|\psi_2\|^2 > \Big(1-\frac{\alpha}{\beta}\Big)\|\psi_{12}\|^2 = \frac{\sqrt{\varepsilon}}{\sqrt{1-\varepsilon}}\|\psi_{12}\|^2 \tag{2.7}$$

because otherwise equation (2.6) is automatically satisfied. (Henceforth, we will write ε in place $\varepsilon_{m,n}$.)

To estimate the left hand side of equation (2.6), we use Lemma 2.4. Specifically, by part (a), substituting ψ_2 for ϕ, we have

$$\langle\psi|H(N-n,N)\psi\rangle \geq \gamma_{m+n}\frac{|\langle\psi|[1-G(N-n,N)]\psi_2\rangle|^2}{\langle\psi_2|[1-G(N-n,N)]\psi_2\rangle}\,,$$

where we have used $\gamma_{m+n} \leq \gamma_{n+1}$. We define the number

$$\eta = \langle G(N-n,N)\rangle_{\psi_2} = \langle\psi_2|G(N-n,N)\psi_2\rangle/\|\psi_2\|^2$$

which allows us to rewrite

$$\langle\psi|H(N-n,N)\psi\rangle \geq \gamma_{m+n}|\langle\psi|[1-G(N-n,N)]\psi_2\rangle|^2\Big/(1-\eta)\|\psi_2\|^2\,. \tag{2.8}$$

The denominator is not zero because, by (2.7), $\psi_2 \neq 0$, and by the definition of $\varepsilon_{m,n}$, $\eta \leq \varepsilon_{m,n} < 1/2$.

Writing G in place of $G(N-n,N)$,

$$\langle\psi|[1-G]\psi_2\rangle = (1-\eta)\|\psi_2\|^2 - \langle\psi_{11}|G\psi_2\rangle - \langle\psi_{12}|G\psi_2\rangle\,. \tag{2.9}$$

It is easy to see that

$$\langle\psi_{11}|G\psi_2\rangle = \langle\psi_{11}|G(N-n,N)\psi_2\rangle = 0 \tag{2.10}$$

because ψ_{11} and $G(N-n,N)\psi_2$ are both eigenvectors of the Hermitian operator $G(1,N-2n)$, with different eigenvalue. (Note that $G(N-n,N)$ commutes with $G(1,N-2n)$ because they are localized on disjoint intervals.) By Lemma 2.4 (b),

$$|\langle\psi_{12}|G\psi_2\rangle| \leq \|\psi_{12}\|\cdot\|\psi_2\|\sqrt{\eta(1-\eta)}\,. \tag{2.11}$$

Combining equations (2.9), (2.10) and (2.11), gives

$$\langle\psi|[1-G]\psi_2\rangle \geq \sqrt{1-\eta}\|\psi_2\|(\sqrt{1-\eta}\|\psi_2\| - \sqrt{\eta}\|\psi_{12}\|)\,. \tag{2.12}$$

Using (2.7) and observing that $\frac{\eta}{1-\eta} \leq \frac{\sqrt{\varepsilon}}{\sqrt{1-\varepsilon}}$, the right hand side of (2.12) is positive. Hence

$$|\langle\psi|[1-G]\psi_2\rangle|^2 \geq (1-\eta)\|\psi_2\|^2(\sqrt{1-\eta}\|\psi_2\| - \sqrt{\eta}\|\psi_{12}\|)^2\,. \tag{2.13}$$

Therefore, by (2.8), equation (2.6) would follow from

$$(\sqrt{1-\eta}\|\psi_2\| - \sqrt{\eta}\|\psi_{12}\|)^2 \geq \beta\|\psi_2\|^2 - (\beta-\alpha)\|\psi_{12}\|^2\,. \tag{2.14}$$

The proof is reduced to checking that the bilinear form

$$B(x,y) = (\beta-\alpha+\eta)x^2 - 2\sqrt{\eta(1-\eta)}xy + (1-\eta-\beta)y^2$$

is positive. The conditions for positivity are

$$\beta-\alpha+\eta \geq 0\,, \quad 1-\eta-\beta \geq 0\,, \tag{2.15}$$

$$\eta(1-\eta) \leq (1-\eta-\beta)(\beta-\alpha+\eta)\,. \tag{2.16}$$

They must be satisfied for all η in the range $[0,\varepsilon]$ (where we write ε for $\varepsilon_{m,n}$). Condition (2.16) is equivalent to $(1-\beta)(\beta-\alpha) \geq \eta(2\beta-\alpha)$. By condition (2.15), this is the same as

$$\eta \leq (1-\beta)(\beta-\alpha)/(2\beta-\alpha)$$

for all $\eta \leq \varepsilon$. However the inequality with $\eta = \varepsilon$ subsumes all those with $\eta < \varepsilon$. Again using (2.15), this is equivalent to

$$\alpha \leq \beta(1-2\varepsilon-\beta)/(1-\varepsilon-\beta)\,.$$

The maximum value of α satisfying these constraints occurs when

$$\alpha = (\sqrt{1-\varepsilon} - \sqrt{\varepsilon})^2\,, \quad \beta = 1-\varepsilon-\sqrt{\varepsilon(1-\varepsilon)}\,,$$

which is just what we set out to prove. ∎

The previous theorem can be used to obtain uniform bounds for the spectral gap above all finite volume ground states. The situation for infinite volume ground states is usually more complicated than just taking the limit of the finite volume gaps. An infinite volume ground state is a state ω on the quasilocal observable algebra $\mathcal{A}_0$ which is locally stable in the sense that for any finite volume Λ, and any $X \in \mathcal{A}_\Lambda$, $\omega(X^*\delta(X)) \geq 0$, where δ is the derivation $\delta(X) = \lim_{\Lambda' \nearrow \mathbb{Z}}[H_{\Lambda'}, X]$. For such a ground state, one can define the gap $\gamma(\omega)$ to be the greatest number γ satisfying

$$\omega(\delta(X)^*\delta^2(X)) \geq \gamma\,\omega(\delta(X)^*\delta(X))\,,$$

for any local observable $X \in \mathcal{A}_\Lambda$, with Λ finite. The purpose of taking three powers of δ on the left hand side is to restrict to the positive energy states, since many ground states may be quasilocal perturbations of one another. With this definition, we can state the main corollary.

Corollary 2.3. *Suppose that $\varepsilon_{m,n} < 1/2$, for some $n \geq m \geq 1$. Then there is a uniform lower bound on the finite volume spectral gaps*

$$\inf_{N\geq 2} \gamma(N) \geq \gamma_{m+n}(1 - 2\sqrt{\varepsilon_{m,n}(1-\varepsilon_{m,n})})\,.$$

Moreover, if ω is an infinite volume ground state, obtained as a weak-$*$ limit of finite volume ground states, then $\gamma(\omega)$ obeys the same bound.

Proof: The proof of the first part follows trivially from Theorem 2.1. One simply observes that $\beta(\varepsilon) > \alpha(\varepsilon)$ and $\alpha(\varepsilon) = 1 - 2\sqrt{\varepsilon(1-\varepsilon)}$.

The proof of the second part follows an argument of Koma and Nachtergaele [17], which was stated for the XXZ model. Suppose ω is a weak-$*$ limit of finite volume ground states $\omega_N(\cdots) = \langle \psi_N | \cdots \psi_N \rangle$, where the unit vectors ψ_N are each in $\mathcal{H}_{\Lambda_N}$ for $\Lambda_N \nearrow \mathbb{Z}$. Then for any local $X \in \mathcal{A}_\Lambda$, there exists an N_0 with $X \in \mathcal{A}_{\Lambda_N}$ for all $N \geq N_0$, and $\omega(X) = \lim_{N\to\infty} \omega_N(X)$, the limit starting at $N = N_0$. Now note that $\delta(X)^* \delta^2(X)$ is also local, localized in $\Lambda^{(2)}$ consisting of all points no more than 2 units from Λ. There is also an N_0 such that $\Lambda_N \supset \Lambda^{(2)}$ for all $N > N_0$, and then we evaluate

$$\omega_N(\delta(X)^* \delta^2(X)) = \langle \delta(X)\psi_N | H(\Lambda_N) \delta(X) \psi_N \rangle \,.$$

Let $\psi'_N = \delta(X)\psi_N$. Then if ϕ is any ground state of $H = H(\Lambda_N)$,

$$\langle \phi | \psi'_N \rangle = \langle \phi | (HX - XH) \psi_N \rangle = 0$$

because $H\psi_N = H\phi = 0$. So ψ'_N is orthogonal to the ground state space, hence

$$\omega_N(\delta(X)^* \delta^2(X)) \geq \gamma(N) \, \langle \psi'_N | \psi'_N \rangle = \gamma(N) \, \omega_N(\delta(X)^* \delta(X)) \,.$$

Using the uniform lower bound for the finite volume gap, and taking liminf's proves the claim. ∎

Theorem 2.1 apparently is a stronger result than the last corollary. For example, one can use it to prove that, whenever the lower bound on the infinite volume gap is achieved, then the first excitation is not an "edge state". The only example we know where the lower bound is attained is the XXZ model, and there the first excitation is actually not an eigenvalue but the bottom of a continuous band of spectrum with quasi-momentum 0. This is an important difference between the XXZ model and the AKLT model, where the lowest variational excitations (see [10]) have quasi-momentum equal to π.

3. Example 1 : The XXZ model

The XXZ ferromagnet, with *kink*-inducing boundary fields is a translation invariant, nearest-neighbor model with anisotropic exchange

$$h(n, n+1) = -\operatorname{sech}(\xi) \boldsymbol{S}_n \cdot \boldsymbol{S}_{n+1} - (1 - \operatorname{sech}(\xi)) S^3_n S^3_{n+1} + j \tanh(\xi)(S^3_n - S^3_{n+1}) \,.$$

where $\xi \geq 0$ determines the amount of anisotropy. The isotropic Heisenberg model is $\xi = 0$, while $\xi = \infty$ is the Ising model with kink boundary conditions. For $j = 1/2$, the Hamilton is invariant under the action of the quantum group $\mathrm{SU}_q(2)$, where $q = e^{-\xi}$. In this section we will prove the following proposition.

Proposition 3.1. *Let $\xi > 0$. Then the spectral gap above any infinite volume, weak-$*$ limit of finite volume kink ground states is $\gamma(\xi) = 1 - \operatorname{sech} \xi$.*

The representation of $\mathrm{SU}_q(2)$ is generated by three spin operators

$$S^3(1, N) = \sum_{n=1}^N S^3_n \,, \quad S^+(1, N) = \sum_{n=1}^N e^{-(n-1)\xi} S^+_n \,, \quad S^-(1, N) = \sum_{n=1}^N e^{-(N-n)\xi} S^-_n \,.$$

The first is the usual generator of U(1); the other two are deformations of the spin raising and lowering operators for SU(2). Notably, they are not permutation invariant. For us, the fundamental feature is a decomposition property. Define $S^{3,+,-}(a, b) =$

$\tau^{a-1}(S^{3,+,-}(1,b+1-a))$, where τ is the translation automorphism. Then for any $1 \le n \le N$

$$S^+(1,N) = S^+(1,n) + e^{-n\xi} S^+(n+1,N)\,,$$
$$S^-(1,N) = e^{-(N-n)\xi} S^-(1,n) + S^-(n+1,N)\,.$$

The representation theory $\mathrm{SU}_q(2)$ for $q > 0$ is identical to that of SU(2) in many respects. In particular, the Clebsch-Gordon series is the same. This means the irreducible representations of each dimension, or multiplets, occur with the usual multiplicities in $\mathcal{H}(1,N)$. Particularly, there is a unique maximal dimensional multiplet. As in the isotropic case, it is known [22, 3, 11] that the ground state subspace of the XXZ ferromagnet is precisely this maximal dimensional multiplet. Moreover, it was shown [18] that all the infinite volume ground states are obtained as weak-$*$ limits of these states, which means Proposition 3.1 is particularly relevant. We refer to a generic n-dimensional $\mathrm{SU}_q(2)$ multiplet as $V(n)$.

To prove the proposition, we use Corollary 2.3. Nachtergaele already calculated every $\varepsilon(m,1)$ in [20]. We will repeat his argument, for completeness. This also serves as a slightly simpler preparation for the calculation in the next section. A major motivation for this letter is the fact that Nachtergaele's calculation, with our improved bound, is sufficient to calculate the exact gap.

We now state a lemma which is generally helpful for calculating $\varepsilon(m,n)$.

Lemma 3.2. *The Hermitian operator $K(-m,n) = G(-m,0)G(0,n)G(-m,0)$ has range contained in the Hilbert space $\mathcal{G}(-m,0)\otimes\mathcal{G}(1,n)$. The eigenspace for eigenvalue* 1 *is precisely $\mathcal{G}(-m,n)$. The number $\varepsilon(m,n)$ is the largest eigenvalue less than* 1.

Proof: The first statement is true because $G(0,n)G(-m,0)$ has range contained in $\mathcal{G}(1,n)$, and $[G(-m,0),G(1,n)] = 0$ implies $K(-m,n)$ also has range contained in $\mathcal{G}(1,n)$. Clearly $\mathcal{G}(-m,n)$ is an eigenvalue-1 eigenspace of $K(-m,n)$. If $K(-m,n)\psi = \psi$, then $G(-m,0)\psi = \psi$ and $\langle\psi|G(0,n)\psi\rangle = 1$, which implies $G(0,n)\psi = \psi$. Since H is a nearest-neighbor Hamiltonian, this implies $\psi \in \mathcal{G}(-m,n)$. The last statement follows from the minimax principle. ∎

Proof: (Of Proposition 3.1) In the case of the XXZ model, the space $\mathcal{G}(-m,0) \subset \mathcal{H}(-m,0)$ is the unique $(m+2)$-dimensional multiplet $V(m+2)$. The space $\mathcal{G}(1,1)$ is just $V(2)$. By Clebsch-Gordon, $V(m+2)\otimes V(2) = V(m+3)\oplus V(m+1)$. Since $G(-m,0)$ and $G(0,1)$ are $\mathrm{SU}_q(2)$ invariant, so is $K(-m,1)$. This implies that $K(-m,1) = \lambda_1 \operatorname{Proj}(V(m+3)) + \lambda_2 \operatorname{Proj}(V(m+1))$. Of course, since $V(m+3)$ is the maximal dimensional irreducible multiplet in $\mathcal{H}(-m,1)$, we know $\lambda_1 = 1$, by the lemma. We can restrict to the two-dimensional (invariant) subspace of $\mathcal{G}(-m,0)\otimes\mathcal{G}(1,1)$, which is the intersection with the eigenspace $\{\psi : S^3(-m,1)\psi = m/2\psi\}$. Then it is easy to determine λ_2 from the trace. In the following we will use a nonorthogonal basis, which simplifies the algebra a bit.

Every vector can be written as

$$\psi = (\alpha S^-(-m,0) + \beta e^{-(m+1)\xi} S^-(1,1)|\Omega\rangle$$

for some α, β, where $|\Omega\rangle$ is the all up-spins state. Note $S^-(1,1)$ is the spin lowering operator for the interval $[1,1] = \{1\}$; i.e., it is the spin lowering operator at the site 1. We observe that

$$\psi = (\alpha S^-(-m,-1) + \alpha e^{-m\xi} S^-(0,0) + \beta e^{-(m+1)\xi} S^-(1,1)|\Omega\rangle\,.$$

Since $S^-(0,1) = S^-(0,0) + e^{-\xi}S^-(1,1)$, it is clear that

$$G(0,1)\psi = \Big(\alpha S^-(-m,-1) + \frac{\alpha+\beta e^{-2\xi}}{1+e^{-2\xi}} e^{-m\xi}(S^-(0,0)+e^{-\xi}S^-(1,1))\Big)|\Omega\rangle\,.$$

Similarly, since $S^-(-m,0) = S^-(-m,-1) + e^{-m\xi}S^-(0,0)$, we have

$$\begin{aligned} G(-m,0)G(0,1)\psi = \Big(&(Z+e^{-2m\xi})^{-1}(\alpha Z + e^{-2m\xi}\frac{\alpha+e^{-2\xi}\beta}{1+e^{-2\xi}})S^-(-m,0) \\ &+ \frac{\alpha+e^{-2\xi}\beta}{1+e^{-2\xi}} e^{-(m+1)\xi}S^-(1,1)\Big)|\Omega\rangle \end{aligned}$$

where $Z = \|S^-(-m,-1)|\Omega\rangle\|^2 = (1-e^{-2m\xi})/(1-e^{-2\xi})$. (We are exploiting the fact that $G(0,1)$ and $G(-m,0)$ are rank-one projections restricted to the appropriate subspace.) Hence the eigenvalues are equal to the eigenvalues of the matrix

$$\frac{1}{(Z+e^{-2m\xi})(1+e^{-2\xi})}\begin{bmatrix} Z(1+e^{-2\xi})+e^{-2m\xi} & e^{-2(m+1)\xi} \\ (Z+e^{-2m\xi}) & e^{-2\xi}(Z+e^{-2m\xi}) \end{bmatrix}$$

The trace of this matrix is $1 + e^{-2\xi}Z/((Z+e^{-2m\xi})(1+e^{-2\xi}))$. Since one of the eigenvalues is 1, that means the other eigenvalue is

$$\varepsilon(m,1) = \frac{e^{-2\xi}Z}{(Z+e^{-2m\xi})(1+e^{-2\xi})} = \frac{e^{-2\xi}}{1+e^{-2\xi}}\cdot\frac{1-e^{-2m\xi}}{1-e^{-2(m+1)\xi}}\,.$$

This is increasing with m. So, $\varepsilon_{1,1} = e^{-2\xi}/(1+e^{-2\xi}) < 1/2$ for $\xi > 0$. Also, $\gamma_2 = 1$. So $\gamma_2(1-2\sqrt{\varepsilon_{1,1}(1-\varepsilon_{1,1})}\,) = 1 - \operatorname{sech}\xi$.

By Corollary 2.3 this gives the lower bound for the gap of Proposition 3.1. But it is also easy to obtain this as an upper bound for the gap above the all-up-spins state, using a sequence of variational wave function approximating a long spin-wave state of quasimomentum zero. Since the all-up-spin state is a weak-$*$ limit of the kink ground states upon translation to the left, this upper bound also works for the gap above the kink states. ∎

Note: *The exact gap of the spin-*$1/2$ *XXZ chain was previously determined rigorously, by different means, by Koma and Nachtergaele* [17].

4. Example 2 : The AKLT model

In this section we will show how to calculate $\varepsilon(m,n)$ for the AKLT model. The AKLT model is a spin-1, translation invariant, nearest neighbor model with interaction

$$h(n,n+1) = P_2(\boldsymbol{S}_n + \boldsymbol{S}_{n+1})$$

where P_2 is projection onto the spin-2 states in the tensor product of the two spins-1. This model was introduced by Affleck, Kennedy, Lieb and Tasaki, in [1]. They calculated the ground states, proved that the ground states are exponentially clustering, and proved the existence of a positive spectral gap [2]. However, their existence proof gave no explicit lower bound on the gap, which is is a question of physical interest (see e.g. [4]). In [21], Nachtergaele developed a version of the martingale method for quantum spin systems to obtain explicit lower bounds on the gap of many finitely correlated spin chains. Finitely correlated spin chains are a class of frustration free quantum spin chains, including the AKLT model [9]. The martingale method was originally developed by Lu and Yau [19] to obtain the exact scaling of the spectral gap and logarithmic Sobolev inequality constant for certain stochastic spin systems.

The key point of Lu and Yau's proof for the spectral gap was the introduction of a filtration, such that the Gibbs measure is a martingale with respect to that filtration. This is the same idea for the case of quantum spin systems, except in place of the Gibbs measure one uses the ground state. The martingale property guarantees that the decomposition one obtains for the Hilbert space is actually a direct sum decomposition. In retrospect, this same decompositon is also the key element of the existence proof in [2], although they did not call it a martingale decomposition.

There are several versions of the frustration free ground states of the AKLT model. Affleck, Kennedy, Lieb and Tasaki first wrote them down in the Ising basis, as a weighted sum over certain configurations which are "alternating sign rows". These are configurations of $+$'s, $-$'s and 0's such that between any two $+$'s there is a $-$ and between any two $-$'s there is a $+$. Aside from this rule, 0's may occur anywhere. Physicists refer to this as diluted Neél order. On the basis of this, it was conjectured [7] that the AKLT state has a long range order, for an order parameter which itself is a product of arbitrarily many terms. Partly in order to explain this, Kennedy and Tasaki [14] came up with a nonlocal unitary transformation. One of the nice features of their transformation is the fact that it makes the ground states into product states. This simplifies calculations, so we shall use their representation.

There is another representation of the AKLT ground state, which we shall not use, but we would like to mention, which is the quantum Markov chain (known as matrix product representation by physicists). Fannes, Nachtergaele and Werner [9] used this to prove many interesting features of the valence bond solid states, in particular that all the correlations could be modeled on a finite-dimensional space. They called states with this property, which include many generalizations of the original VBS states, finitely correlated states. More recently, physicists Kolezhuk and Mikeska [16] have used this approach to study spin ladders, and derived new examples of frustration free models beyond those which were known for spin chains.

The Kennedy-Tasaki unitary transformation is $U = \prod_{j<k} \exp(i\pi S_j^z S_k^x)$. It makes the four ground states of the AKLT model, which are otherwise somewhat complicated states, into product states $\Phi_k(1,N) = \bigotimes_{n=1}^N (\phi_k)_n$, where $\phi_{1,2} = [|0\rangle \pm \sqrt{2}|+\rangle]/\sqrt{3}$ and $\phi_{3,4} = [|0\rangle \pm \sqrt{2}|-\rangle]/\sqrt{3}$. We denote $\Phi_i(a,b) = \bigotimes_{n=a}^b (\phi_i)_n$. These four states are not orthogonal, but we can calculate the Gram matrix

$$M(a,b) := [\langle \Phi_i(a,b) | \Phi_j(a,b) \rangle]_{i,j=1}^4 = [\langle \phi_i | \phi_j \rangle^{b-a+1}]_{i,j=1}^4 \,.$$

Specifically,

$$M(1,N) = I_4 + 3^{-N} \begin{bmatrix} 0 & (-1)^N & 1 & 1 \\ (-1)^N & 0 & 1 & 1 \\ 1 & 1 & 0 & (-1)^N \\ 1 & 1 & (-1)^N & 0 \end{bmatrix} . \tag{4.17}$$

We denote the inverse as $W(a,b) := M(a,b)^{-1}$, and it equals

$$\begin{aligned} W(1,N) &= (1 + 2(-1)^N 3^{-N} - 3 \cdot 3^{-2N})^{-1} \\ &\quad \times \left(I_4 - 3^{-N} \begin{bmatrix} -(-1)^N 2 & (-1)^N & 1 & 1 \\ (-1)^N & -(-1)^N 2 & 1 & 1 \\ 1 & 1 & -(-1)^N 2 & (-1)^N \\ 1 & 1 & (-1)^N & -(-1)^N 2 \end{bmatrix} \right) . \end{aligned} \tag{4.18}$$

We can write the ground state projector as

$$G(a,b) = \sum_{i,j=1}^{4} \Phi_i(a,b) W(a,b)_{ij} \Phi_j(a,b)^\dagger\,,$$

where $\Phi_j(a,b)^\dagger$ is the covector dual to $\Phi_j(a,b)$. Indeed this operator is clearly Hermitian, and it has the correct range. It is an easy exercise for the reader to check that it behaves correctly on the four ground states.

We will use Lemma 3.2 to calculate $\varepsilon(m,n)$. An important point is that we do not need to use an orthogonal basis to obtain a matrix representation. So, instead, we use the basis

$$\Theta_{ij}(-m,n) = \Phi_i(-m,0) \otimes \Phi_j(1,n)\,.$$

We will calculate the matrix for $G(-m,0)G(0,n)G(-m,0)$ in this basis. Of course $G(-m,0)$ fixes $\Theta_{ij}(-m,n)$. Before calculating $G(0,n)\Theta_{ij}(-m,n)$, we note that for any $a \leq c < b$, $\Phi_i(a,b) = \Phi_i(a,c) \otimes \Phi_i(c+1,b)$. Hence,

$$\begin{aligned} &G(0,n)\Theta_{ij}(-m,n) \\ &= \sum_{kl} \Phi_k(0,n) \otimes \Big(W(0,n)_{kl}\Phi_l(0,n)^\dagger \Phi_i(-m,-1) \otimes \Phi_i(0) \otimes \Phi_j(1,n)\Big) \\ &= \sum_{kl} W(0,n)_{kl} M(0,0)_{li} M(1,n)_{lj} \Phi_i(-m,-1) \otimes \Phi_k(0,n)\,. \end{aligned}$$

Finally,

$$\begin{aligned} &G(-m,0)G(0,n)\Theta_{ij}(-m,n) \\ &= \sum_{rs}\sum_{kl} \Phi_r(-m,0) \otimes \Big(W(-m,0)_{rs} W(0,n)_{kl} M(0,0)_{li} M(1,n)_{lj} \\ &\qquad \Phi_s(-m,0)^\dagger \Phi_i(-m,-1) \otimes \Phi_k(0,0) \otimes \Phi_k(1,n)\Big) \\ &= \sum_{klrs} W(0,n)_{kl} M(0,0)_{li} M(1,n)_{lj} W(-m,0)_{rs} M(0,0)_{sk} M(-m,-1)_{si} \Theta_{rk}\,. \end{aligned}$$

Let $\mathcal{A}$ be the 16×16 matrix with

$$\mathcal{A}(ij,rk) = \sum_{ls} W(0,n)_{kl} M(0,0)_{li} M(1,n)_{lj} W(-m,0)_{rs} M(0,0)_{sk} M(-m,-1)_{si}\,.$$

The top four eigenvalues are 1, corresponding to the four-fold degenerate ground state space $\mathcal{G}(-m,n)$. Therefore, $\varepsilon(m,n)$ is the fifth largest eigenvalue.

To diagonalize $\mathcal{A}$, we observe that while U has obscured the SU(2) symmetry of $\widetilde{H}_N$, it is still there. It's just no longer a "local" symmetry. Each of $\mathcal{G}(-m,0)$ and $\mathcal{G}(1,n)$ is equal to $V(3) \oplus V(1)$. Thus, the space on which $\mathcal{A}$ acts is $(V(3) \oplus V(1))^{\otimes 2} = V(5) \oplus 3V(3) \oplus 2V(1)$, using Clebsch-Gordon. Abusing notation, we refer to $3V(3)$ and $2V(1)$ as the spin-1 and spin-0 subspace. $\mathcal{A}$ is SU(2) symmetric since each $G(a,b)$ is. Using this and Schur's lemma, we know that $V(5)$ is an eigenspace of $\mathcal{A}$ with eigenvalue $\lambda_5 \in \mathbb{Q}$. We also know that the eigenspace of $\mathcal{A}$ for eigenvalue 1 is $\mathcal{G}(-m,n) \equiv V(3) \oplus V(1)$. Let $V(1)'$ be the orthogonal complement of $\mathcal{G}(-m,n) \cap 2V(1)$ in $2V(1)$. Since this is one dimensional, it is an eigenspace for $\mathcal{A}$ with eigenvalue $\lambda_1 \in \mathbb{Q}$. Let W be the orthogonal complement of $\mathcal{G}(-m,n) \cap 3V(3)$ in $3V(3)$. Then $\mathcal{A}$ restricted to this block, and restricted to an eigenspace of S^3, is a 2×2 matrix. There are two eigenvalues $\lambda_{3,\pm}$, which are irrational, but solvable by radicals. So $\mathcal{A}$ is solvable by hand, and the eigenvalues are

$$\begin{aligned}
\lambda_5(m,n) &= \frac{1}{9}\cdot\frac{(1-(-1/3)^m)(1-(-1/3)^n)}{(1-(-1/3)^{m+1})(1-(-1/3)^{n+1})}\,,\\
\lambda_1(m,n) &= \frac{1}{9}\cdot\frac{(1-(-1/3)^{m-1})(1-(-1/3)^{n-1})}{(1-(-1/3)^{m+1})(1-(-1/3)^{n+1})}\,,\\
\lambda_{3\pm}(m,n) &= \frac{1}{9}\cdot\frac{1+(-1/3)^m+(-1/3)^n-19(-1/3)^{m+n+2}\pm 2\sqrt{R(m,n)}}{(1-(-1/3)^{m+1})(1-(-1/3)^{n+1})}\\
R(m,n) &= [(-1/3)^m-(-1/3)^n]^2-128(-1/3)^{2m+2n+4}\\
&\qquad +4(-1/3)^{m+n+2}(1+(-1/3)^m)(1+(-1/3)^n)\,.
\end{aligned}$$

To calculate $\varepsilon_{1,1}$, we set $n=1$ and observe that the maximum eigenvalue occurs for $m=1$, $\lambda_5(1,1)=1/4$. Since $\gamma(2)=1$, Corollary 2.3 gives

$$\inf_{N\geq 2}\gamma(N)\geq 1-\frac{\sqrt{3}}{2}\approx 0.133975\,.$$

If we instead use $\varepsilon_{3,3}$, we must numerically diagonalize the six-site chain to calculate γ_6. We have done this using Lanczos, and the answer is $\gamma_6 = 0.3985(\pm 0.00005)$. Taking $n=3$ we find $\varepsilon(3,3)=\lambda_5(3,3)=(7/20)^2$. So we obtain

$$\inf_{N\geq 2}\gamma(N)\geq 0.137194\,.$$

Several authors have calculated variational upper bounds for the spectral gap above the AKLT ground states. One of these authors was Knabe [15] who not only calculated variational upper bounds, but obtained rigorous lower bounds, too. Fáth and Solyóm [10] showed Knabe's trial states to be equivalent to the single mode approximation states of Arovas, Auerbach and Haldane [4], and themselves gave a third interpretation as hidden domain walls. These all give an upper bound of $\gamma(N)\leq 10/27=0.370\ldots$. Our lower bound is not spectacular, but at least has the correct order. By yet different methods, Fannes, Nachtergaele and Werner [8] obtained the excellent lower bound $\gamma(N)\geq 3/10$.

Acknowledgements

It is a pleasure to thank Bruno Nachtergaele for several fruitful discussions. S.S. is a National Science Foundation Postdoctoral Fellow.

References

[1] I. Affleck, T. Kennedy, E.H. Lieb, and H. Tasaki, *Rigorous results on valence-bond ground states in antiferromagnets*, Phys. Rev. Lett. **59** (1987), no. 7, 799–802.

[2] ———, *Valence bond ground states in isotropic quantum antiferromagnets*, Commun. Math. Phys. **115** (1987), 477–528.

[3] F. C. Alcaraz, S. R. Salinas, and W. F. Wreszinski, *Anisotropic ferromagnetic quantum domains*, Phys. Rev. Lett **75** (1995), 930–933.

[4] D.P. Arovas, A. Auerbach, and F.D.M. Haldane, *Extended Heisenberg models of antiferromagnetism: Analogies to the fractional quantum Hall effect*, Phys. Rev. Lett. **60** (1988), no. 6, 531–534.

[5] E. Carlen, M. Carvalho, and M. Loss, *Determination of the spectral gap for Kac's Master equation and related stochastic evolutions*, Preprint, 2002, arXiv:math-ph/0109003.

[6] P. Contucci, B. Nachtergaele, and W.L. Spitzer, *The ferromagnetic Heisenberg XXZ chain in a pinning field*, Phys. Rev. B **66** (2002), 0644291–13, arXiv:math-ph/0204011.

[7] M. den Nijs and K. Rommelse, *Preroughening transitions in crystal surfaces and valence-bond phases in quantum spin chains*, Phys. Rev. B. **40** (1989), 4709.

[8] M. Fannes, B. Nachtergaele, and R. F. Werner, *Valence bonds states on quantum spin chains as ground states with spectral gap*, J. Phys. A: Math. Gen. **24** (1991), L185–L190.

[9] ———, *Finitely correlated states of quantum spin chains*, Commun. Math. Phys. **144** (1992), 443–490.

[10] G. Fáth and J. Sólyom, *Hidden domain walls as elementary excitations in the spin-1 valence-bond-solid model*, Acta Physica Polonica A **85** (1994), no. 2, 371–375.

[11] C.-T. Gottstein and R. F. Werner, *Ground states of the infinite q-deformed Heisenberg ferromagnet*, Preprint, arXiv:cond-mat/9501123.

[12] F.D.M. Haldane, *Continuum dynamics of the 1-d Heisenberg antiferromagnet: Identification with the $O(3)$ nonlinear sigma model*, Phys. Lett. A **93** (1983), 464–468.

[13] M. Kac, *Foundations of kinetic theory*, Proc. 3rd Berkeley symp. Math. Stat. Prob. (J. Neyman, ed.), vol. 3, Univ. of California, 1956, pp. 171–197.

[14] Tom Kennedy and Hal Tasaki, *Hidden symmetry breaking and the Haldane phase in $s = 1$ quantum spin chains*, Commun. Math. Phys. **147** (1992), 431–484.

[15] S. Knabe, *Energy gaps and elementary excitations for certain VBS-quantum antiferromagnets*, J. Stat. Phys. **52** (1988), 627–638.

[16] A. K. Kolezhuk and H.-J. Mikeska, *Finitely correlated generalized spin ladders*, Internat. J. Modern Phys. B **12** (1998), no. 23, 2325–2348, arXiv:cond-mat/9803176.

[17] T. Koma and B. Nachtergaele, *The spectral gap of the ferromagnetic XXZ chain*, Lett. Math. Phys. **40** (1997), 1–16, arXiv:cond-mat/9512120.

[18] ———, *The complete set of ground states of the ferromagnetic XXZ chains*, Adv. Theor. Math. Phys. **2** (1998), 533–558, arXiv:cond-mat/9709208.

[19] S.L. Lu and H.T. Yau, *Spectral gap and logarithmic Sobolev inequality for Kawasaki and Glauber dynamics*, Commun. Math. Phys. **156** (1993), 399–433.

[20] B. Nachtergaele, *A lower bound for the spectral gap of the ferromagnetic XXZ chain*, Preprint, arXiv:cond-mat/9501098.

[21] ———, *The spectral gap for some spin chains with discrete symmetry breaking*, Commun. Math. Phys. **175** (1996), 565–606.

[22] V. Pasquier and H. Saleur, *Common structures between finite systems and conformal field theories through quantum groups*, Nuclear Physics B **330** (1990), 523–556.