مجلهٔ پژوهش فیزیک ایران، جلد ۲۲، شمارهٔ ۲، تابستان ۱۴۰۱
DOI: 10.47176/ijpr.22.2.01344

سیمای فاز مدل هایزنبرگ: روش یادگیری ماشین

عبدالرضا رسولی کناری[۱] و محمدحسین زارع[۲*]

۱. *دانشکده برق و کامپیوتر، دانشگاه صنعتی قم، قم*

۲. *دانشکده فیزیک، دانشگاه صنعتی قم، قم*

پست الکترونیکی: zare@qut.ac.ir

(دریافت مقاله: ۱۴۰۰/۸/۹ ؛ دریافت نسخهٔ نهایی: ۱۴۰۰/۱۰/۲۱)

چکیده

الگوریتم یادگیری ماشین به عنوان ابزرای قدرتمند، چشم‌انداز خوبی برای مطالعهٔ فازهای مختلف ماده در زمینهٔ فیزیک ماده چگال ترسیم می‌کند. در این مقاله سعی خواهیم کرد با بازبینی در فرمول‌بندی الگوریتم‌های شبکهٔ عصبی عمیق، روشی نو جهت حل مسئلهٔ بهینه سازی سامانه‌های اسپینی، برای بررسی حالت پایهٔ سامانه‌های مغناطیسی زیر دمای کوری، که تقارن دورانی اسپین به صورت خودبه‌خودی شکسته می‌شود، معرفی کنیم. با استفاده از روش شبکهٔ یادگیری عمیق، سیمای فاز کلاسیکی مدل هایزنبرگ همسانگرد شبکهٔ مربعی و شبکهٔ لانه زنبوری را مطالعه کردیم. نتایج به دست آمده با روش یادگیری ماشین، با سیمای فاز کلاسیکی مدل هایزنبرگ که از دیگر روش‌های تحلیلی و محاسباتی پیدا شده، همخوانی کامل دارد. همچنین، در تحقیق حاضر که اساس آن بر یادگیری عمیق است، مزیت بالاتری نسبت به الگوریتم‌های تکاملی، که با چالش اساسی در حل مسئله‌های بهینه‌سازی مواجه‌اند، برخوردار است. بنابراین، توانایی الگوریتم‌های یادگیری ماشین نظیر شبکهٔ عصبی عمیق در حل مسائل فیزیک مادهٔ چگال، استفاده از آن را در مطالعهٔ حالت پایهٔ سامانه‌های مغناطیسی، اجتناب‌ناپذیر می‌سازد.

واژه‌های کلیدی: یادگیری ماشین، شبکهٔ عصبی عمیق، روش بهینه سازی آدام، مدل هایزنبرگ

۱. مقدمه

فیزیکدانان برای بررسی منشأ قوانین طبیعت، از مفاهیم تقارن استفاده می‌کنند. بنابراین مبحث تقارن در فهم طبیعت، نقش مهمی بازی می‌کند. با پیدایش مکانیک کوانتومی در قرن بیستم، استفاده از روش‌های تقارن مبتنی بر نظریهٔ گروه، ابزار خوبی جهت مطالعهٔ سامانه‌های فیزیکی به وجود آمد. در فیزیک کلاسیک، طبق قضیهٔ نوتر، تقارن‌های فیزیک کلاسیک به طور مستقیم، به مشاهده پذیرهای فیزیکی مرتبط می‌شوند. در فیزیک کوانتومی، ناوردایی مشاهده‌پذیرها تحت تبدیلات معین، دلالت بر وجود عملگر یکانی معینی بر روی فضای هیلبرت حالت‌ها دارد. اگر آن تبدیل تقارنی با دینامیک سامانه سازگار باشد، این به معنای جابه‌جاپذیری این عملگر با هامیلتونی است که در این صورت منجر به ساختار چندگانه مشخصی در طیف آن می‌شود. هرچند، سامانه‌هایی وجود دارند که دینامیک آنها تحت تبدیلی ناوردا است، ولی تقارن مورد نظر در طیف یا

مشاهده‌پذیرهای فیزیکی قابل مشاهده نیست که از این، به عنوان شکست خودبه‌خودی تقارن نام برده می‌شود. بنابراین شکست خودبه‌خودی تقارن، زمانی اتفاق می‌افتد که هامیلتونی سامانه تحت تقارن خاصی ناورداست ولی حالت پایهٔ سامانه آن تقارن را ندارد. فازهای مختلف ماده از قبیل ابررسانایی، ابرشارگی و فرومغناطیس و همچنین گذار فازهای ممکن در این سامانه‌ها را می‌توان با استفاده از این پدیده توصیف کرد. مواد مغناطیسی در دماهایی، زیر دمای کوری، خاصیت مغناطیسی پیدا می‌کنند، هرچند برهم‌کنش تبادلی بین ممان‌های مغناطیسی، جهت ارجحی را موجب نمی‌شوند. با توجه به این که از تقارن هامیلتونی نمی‌توان برای تعیین حالت‌های پایهٔ مغناطیسی استفاده کرد، باید از روش‌های شهودی که همراه با حدسیات زیادی است، برای تعیین حالت پایهٔ مغناطیسی استفاده کرد. تعیین حالت پایهٔ مغناطیسی سامانه‌های مغناطیسی ناکام، سخت است؛ زیرا این‌گونه سامانه‌ها، حالت‌های با کمینه موضعی زیادی دارند که انرژی آنها به انرژی حالت کمینهٔ کلی، نزدیک است. این ناکامی مغناطیسی را در سامانه‌هایی که با مدل‌های اسپینی ناکام توصیف می‌شوند، مانند مدل هایزنبرگ شامل برهم‌کنش همسایه‌های اول و دوم روی شبکه‌های مربعی و لانه زنبوری یا در سامانه‌هایی که هندسهٔ شبکهٔ آنها باعث ناکامی می‌شود مانند شبکهٔ مثلثی و شبکهٔ کاگومه، می‌توان جستجو کرد.

در فیزیک مادهٔ چگال، روش‌های محاسباتی زیادی وجود دارند که برای فهم بهتر نتایج تجربی مورد استفاده قرار می‌گیرند. ساختارهای مغناطیسی را می‌توان با حل هامیلتونی اسپینی این گونه سامانه‌ها، با استفاده از روش‌های شبیه سازی عددی نظیر مونت‌کارلو و دینامیک اسپین [۱ و۲]، مورد مطالعه قرار داد. مطالعهٔ دقیق فازهای مغناطیسی متفاوت با استفاده از این روش‌ها خیلی سخت است؛ زیرا حالت‌های مغناطیسی زیادی با انرژی کمینهٔ موضعی متفاوت از انرژی کمینهٔ کلی در این سامانه‌ها وجود دارند. باید توجه کرد که استفاده از این روش برای تعیین حالت پایهٔ سامانه‌های مغناطیسی که هامیلتونی آنها شامل برهم‌کنش‌های تبادلی ناهمسانگرد هستند، به خاطر وجود تعداد بی‌نهایت حالت کمینهٔ موضعی، به مراتب سخت‌تر است.

روش محاسباتی یادگیری ماشین را می‌توان به‌عنوان ابزاری قدرتمند برای تحقیقات فیزیک مدرن نام برد [۳ و ۴]. اخیراً در فیزیک مادهٔ چگال، روش‌های محاسباتی یادگیری ماشین را برای مطالعهٔ فازها و گذار فاز بین حالت‌های مختلف ماده [۵–۱۰]، حالت‌های مغناطیسی کوانتومی [۱۱–۱۷]، سامانهٔ اسپینی کلاسیکی [۱۸–۲۰]، تجزیه و تحلیل نتایج تجربی [۲۱ و ۲۲] و جستجوی مواد جدید [۲۳]، مورد استفاده قرارگرفته‌اند. با استفاده از روش محاسباتی یادگیری ماشین که در مرجع‌های [۱۸–۲۰] ارائه شده‌اند، می‌توان آرایش‌های اسپینی متفاوت را با استفاده از کمینه‌کردن انرژی مغناطیسی، به دست آورد. بنابراین، این نتایج بیانگر این است که روش یادگیری ماشین را می‌توان برای مطالعهٔ سامانه‌های همبستهٔ قوی استفاده کرد و می‌تواند ابزاری مناسب برای حل مسئله‌های سخت و باز در زمینهٔ فیزیک مادهٔ چگال باشد.

در این مقاله سعی داریم که ابتدا روش محاسباتی شبکهٔ یادگیری عمیق را با جزئیات بیان کنیم. سپس به تعیین حالت پایهٔ مغناطیسی مدل اسپینی کلاسیکی هایزنبرگ، که شامل برهم‌کنش تبادلی بین همسایهٔ اول و دوم روی شبکه‌های مربعی و لانه زنبوری است، بپردازیم. در نهایت، سازگاری نتایج روش یادگیری ماشین با نتایجی که قبلاً با روش‌های محاسباتی دیگر به دست آمده است، را بررسی می‌کنیم.

۲. مدل هایزنبرگ

فیزیک ابررساناهای دمای بالا، مانند کوپرایت‌ها، در دو بُعد است، بدین معنی که الکترون‌ها به صورت مؤثر در صفحات دو بُعدی حرکت می‌کنند. بنابراین برای توصیف نظری این‌گونه مواد، از حرکت الکترون‌ها بین صفحات صرف‌نظر می‌شود. در ضمن، اتم‌هایی که نقش اصلی را در فیزیک این دسته از مواد دارند، مانند اتم مس در ابررسانایی کوپرایت‌ها، شبکه‌های دو بُعدی مربعی را تشکیل می‌دهند که اوربیتال d آنها در حال پر شدن است. باید توجه کرد که برهم‌کنش کولنی بین الکترون‌ها،

در اتم‌هایی که اوربیتال d یا f در حال پر شدن دارند، به خاطر گسترش فضایی کمتر در مقایسه با اوربیتال‌های s یا p، بسیار قوی‌تر است. گسترش فضایی کمتر اوربیتال‌های d یا f باعث می‌شود که دو الکترون با اسپین‌های مخالف در فاصلهٔ کمتری نسبت به هم قرار گیرند و در نتیجه دافعهٔ کولنی افزایش می‌یابد. بنابراین از مرتبهٔ بزرگی برهم‌کنش کولنی در عناصر مختلف، می‌توان آنها را به دو دسته، هم‌بستهٔ ضعیف یا هم‌بستهٔ قوی، تقسیم‌بندی کرد. در این مقاله علاقه‌مندیم که روی سامانه‌های هم‌بستهٔ قوی متمرکز شویم. در این‌گونه سامانه‌ها که برهم‌کنش کولنی قوی است، الکترون‌ها به صورت جداگانه، جایگاه‌های شبکه را اشغال می‌کنند که انرژی سامانه را کمینه کنند. به این حالت، حالت نیمه‌پر گفته می‌شود. در این صورت، درجات آزادی بار یخ می‌زنند و فیزیک این دسته از سامانه‌ها، به صورت مؤثر، تنها با درجات آزادی اسپینی توصیف می‌شود.

در این مقاله علاقه‌مندیم مدل هایزنبرگ همسانگرد $S=\frac{1}{2}$ شبکه‌های مربعی و لانه زنبوری را که به صورت زیر تعریف می‌شود مطالعه کنیم:

$$H = J_1 \sum_{\langle ij \rangle} \vec{S}_i \cdot \vec{S}_j + J_2 \sum_{\langle\langle ij \rangle\rangle} \vec{S}_i \cdot \vec{S}_j \,, \qquad (1)$$

که J_1 و J_2 به ترتیب برهم‌کنش‌های تبادلی پادفرومغناطیس بین همسایه‌های اول و دوم روی شبکه‌های مربعی و لانه زنبوری را نشان می‌دهند،(شکل ۱ را ببنید). قابل ذکر است که در این دو مورد، ناکامی مغناطیسی از وجود رقابت بین برهم‌کنش‌های تبادلی بین همسایه‌های اول و دوم روی این دو شبکه ناشی می‌شود؛ زیرا اسپین‌ها تمایل دارند نسبت به اسپین‌های روی جایگاه‌های همسایه، به صورت پادموازی قرار بگیرند که انرژی سامانه کمینه شود. بنابراین یک اسپین که در یک جایگاه قرار دارد نمی‌تواند به طور هم‌زمان، با اسپین‌های همسایه اول و دوم خود به صورت پادموازی جهت‌گیری کند. در این صورت به ازای مرتبهٔ بزرگی خاصی از ثابت‌های جفت‌شدگی، سامانه نمی‌تواند حالت پایهٔ یکتا پیدا کند و حالت‌هایی با انرژی یکسان در این‌گونه سامانه‌ها زیاد می‌شود که در این حالت، می‌گوییم که سامانه در پیدا کردن حالت پایه، ناکام است.

۳. الگوریتم یادگیری ماشین

حالت پایهٔ مدل اسپینی شبکهٔ مربعی یا لانه زنبوری، در اصل یافتن آرایش اسپینی با کمترین انرژی است که این یک مسئلهٔ بهینه سازی خواهد بود. بدین منظور می‌توان از الگوریتم‌های تکاملی مانند ژنتیک (GA)[1]، تبرید شبیه‌سازی شده (SA)[2] و کلونی مورچه‌ها برای حل این‌گونه مسئله‌ها استفاده کرد. روش حل مسئله با استفاده از الگوریتم‌های تکاملی، بدین گونه است که با کاوش و اکتشاف فضای حالت‌های محتمل برای اسپین‌ها، سعی در پیدا کردن آرایش اسپینی است که در آن تابع هدف، در اینجا انرژی حالت پایهٔ هایزنبرگ، به کمینهٔ کلی خود برسد. چالش اساسی در حل مسئله‌های بهینه‌سازی با الگوریتم‌های تکاملی، بزرگ بودن فضای حالت‌های ممکن برای مسئله و همچنین وجود حالت‌های زیاد با کمینه‌های موضعی است که پیدا کردن پاسخ درست، که انرژی به ازای آن کمینه کلی دارد، را خیلی زمان‌بر و سخت می‌کند.

اخیراً روش‌های یادگیری ماشین، به‌ویژه شبکه‌های عصبی عمیق که در آن قدرت پردازش الگوریتم‌های یادگیری ماشین افزایش یافته است، می‌تواند برای حل مسئله‌هایی با فضای حالت‌های بزرگ، مناسب باشد. بنابراین با استفاده از شبکه‌های عصبی عمیق می‌توان چالش جدی در حل مسئله‌های اسپینی با روش‌های الگوریتم تکاملی را برطرف کرد. همچنین، شبکه‌های عصبی عمیق که قدرت پردازش بالاتری دارند را می‌توان برای حل مسائل بهینه‌سازی و یافتن فضای حالت‌های ممکن مسئله، جهت پیدا کردن آرایش اسپینی با انرژی کمینه کلی استفاده کرد. در این مقاله سعی خواهیم کرد با بازبینی در فرمول‌بندی شبکه‌های عصبی عمیق، روشی نو جهت حل مسئلهٔ بهینه سازی سامانه‌های اسپینی، برای پیدا کردن حالت پایه، ارائه دهیم. در این بخش، ابتدا روش شبکه‌های عصبی عمیق را با جزئیات بیشتر توضیح خواهیم داد و هم‌چنین این روش را جهت بررسی حالت پایهٔ سامانه‌های اسپینی، مورد بازبینی قرار می‌دهیم.

۱. Genetic Algorithm

۲. Simulated Annealing

۳. ۱. شبکهٔ عصبی عمیق

روش شبکهٔ عصبی عمیق از سه لایهٔ اصلی از نورون‌ها (۱) لایهٔ ورودی، (۲) لایهٔ پنهان میانی و (۳) لایهٔ خروجی تشکیل شده است. مبنای اصلی کار در روش شبکهٔ عصبی عمیق، یافتن رابطهٔ مؤثر بین داده‌های ورودی و داده‌های خروجی است، که بر اساس این رابطه، بتوان برای هر داده ورودی جدید، خروجی درست را به دست آورد. با شبکهٔ عصبی عمیق می‌توان هرگونه روابط خطی یا غیرخطی را بین داده‌های ورودی و داده‌های خروجی مدل‌سازی کرد. در واقع، در شبکهٔ عصبی عمیق، باید مجموعه پارامتر Θ را طوری انتخاب کرد که برای مجموعه داده‌های آموزشی، رابطهٔ زیر برقرار باشد:

$$Y = f_{\Theta}(X), \quad X = \{X_1, X_2, \ldots, X_n\}, \quad Y = \{Y_1, Y_2, \ldots, Y_n\}, \qquad (۲)$$

که در آن Θ، مجموعه پارامترهای یادگیری شونده شبکهٔ عصبی، X و Y به ترتیب داده‌های آموزشی ورودی و خروجی شبکه و n تعداد نمونه‌های آموزشی هستند. در این روش، با تغییر پارامتر یادگیری، Θ، باید رابطهٔ (۲) به ازای تمام داده‌های آموزشی (ورودی و خروجی) برقرار باشد یا کمترین مقدار خطا را داشته باشد. تابع خطا، که معمولاً در محاسبات از تابع مجذور مربعات استفاده می‌شود، را می‌توان با مقایسهٔ داده‌های خروجی داده‌شده Y و خروجی شبکه با استفاده از رابطهٔ زیر، به دست آورد:

$$L = \sum_i |Y_i^* - Y_i|^2 = \sum_i |f_{\Theta}(X_i) - Y_i|^2, \qquad (۳)$$

که در آن Y و Y^* به ترتیب داده‌های خروجی شبکه و مقدار واقعی در داده‌های آموزشی هستند. اینجا نیاز است که Θ را طوری بیابیم که مقدار تابع خطا (L) کمینه شود. بدین منظور باید از تابع خطا نسبت به پارامتر یادگیری شونده، مشتق بگیریم، $\frac{dL}{d\Theta}$، و در جهت عکس شیب تابع مذکور، به سمت نقطهٔ کمینه حرکت کنیم. لازم به ذکر است: مقادیر پارامترهای یادگیری شونده که با توجه به مقدار مشتق به دست‌آمده اند باید طوری تغییر پیدا کنند که در جهت منفی شیب، به سمت نقطهٔ کمینه نزدیک شوند، تا بهترین مقدار برای این پارامترها به دست بیاید. از این روش به‌عنوان گرادیان کاهشی[۱] نام برده می‌شود.

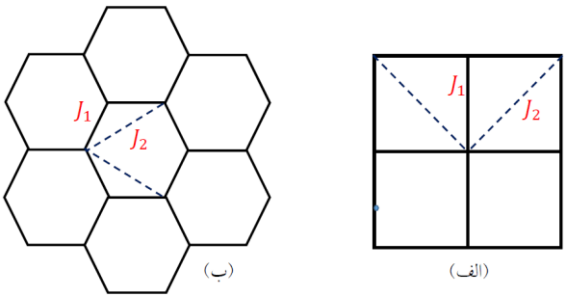
چالش اصلی در استفاده از روش شبکهٔ عصبی عمیق برای حل مسائل بهینه‌سازی، فقدان داده‌های آموزشی مورد نیاز برای یادگیری شبکه است. درواقع، در مسئلهٔ بهینه‌سازی، باید به مجموعه‌ای از جواب‌ها دسترسی داشته باشیم که از آنها به عنوان داده‌های آموزشی، برای یادگیری شبکهٔ عصبی استفاده کنیم. بنابراین با داشتن مجموعه‌ای از این داده‌های آموزشی، که به عنوان داده‌های ورودی و خروجی شبکهٔ عصبی در نظر گرفته می‌شوند، می‌توان به شبکه، آموزش لازم را برای حل مسئله داد. در یک مسئلهٔ بهینه‌سازی، تابع هدف مشخص است و تلاش برای یافتن پاسخ کمینه است، در حالی که در شبکهٔ عصبی عمیق سعی می‌شود با داشتن پاسخ‌های مسئله، تابع مورد نظر تخمین زده شود. باید توجه کرد که در الگوریتم‌های تکاملی با داشتن تابع هدف، با خلق جواب‌های متفاوت، سعی در انتخاب بهترین جواب از میان جواب‌های ممکن هستیم، در حالی که در شبکهٔ عصبی عمیق، با داشتن جواب‌ها، سعی در پیدایش رابطهٔ مؤثر بین داده‌های ورودی و خروجی هستیم. در بخش بعدی، سعی خواهیم کرد با بازنویسی فرمول‌بندی شبکهٔ عصبی عمیق، از توانایی‌های آن برای حل مسئلهٔ بهینه‌سازی سامانه‌های اسپینی، جهت پیدا کردن حالت پایهٔ مغناطیسی، استفاده کنیم.

۳. ۲. شبکهٔ عصبی عمیق بهینه‌ساز برای سامانه‌های اسپینی

همان طور که قبلاً ذکر شد، روش اصلی حل مسئله با شبکهٔ یادگیری عمیق[۲] (DL)، یافتن مجموعه پارامترهای یادگیری شونده است. هرچند، برای یافتن پارامترهای یادگیری شونده، نیاز است مجموعه‌ای از جواب‌های درست را ، به‌عنوان داده‌های آموزشی، به شبکه عصبی عمیق داد. در مورد حالت پایهٔ سامانه‌های اسپینی، که در این مقاله مدنظر است، آرایش‌های

۱. Gradient Descent

۲. Deep Learning



شکل ۱. نمایش شبکه‌های (الف) مربعی و (ب) لانه زنبوری، که J_1 و J_2 به ترتیب برهم‌کنش‌های بین اسپین هر جایگاه با اولین و دومین همسایه‌اش را نشان می‌دهد.

اسپینی درست را که به ازای آن، انرژی سامانه در کمینهٔ کلی خودش قرار دارد، از قبل نداریم. بنابراین داده‌های ورودی و خروجی مورد نیاز برای آموزش شبکهٔ عصبی عمیق، در دسترس نیستند.

برای پیدا کردن حالت پایهٔ مدل هایزنبرگ با استفاده از روش شبکهٔ عصبی، این روش را به گونه‌ای بازنویسی می‌کنیم که آرایش اسپین‌ها به عنوان پارامترهای شبکهٔ عصبی شناخته شوند. درواقع، با بازنویسی این روش، به جای در نظر گرفتن بردار اسپین به‌عنوان نورون‌های ورودی، از آنها به‌عنوان پارامترهای یادگیری شونده استفاده خواهیم کرد. در نتیجه می‌توان با استفاده از شبکهٔ عصبی عمیق، بهترین آرایش با کمترین انرژی را پیدا کرد.

برای حل مسئله با این روش، شبکهٔ با شرایط مرزی دوره‌ای به ابعاد $L_x \times L_y$ در نظر می‌گیریم. در اینجا، اسپین‌ها همانند برداری $\vec{S} = [S_x, S_y, S_z]$ روی جایگاه‌ها مختلف شبکه در نظر گرفته می‌شوند. برای هر آرایش اسپینی، می‌توان هامیلتونی سامانه را به صورت زیر بازنویسی کرد:

$$H = J_1 \sum_{\langle ij \rangle} \vec{S}_i \cdot \vec{S}_j + J_2 \sum_{\langle\langle ij \rangle\rangle} \vec{S}_i \cdot \vec{S}_j = \sum_i \vec{M}_i \cdot \vec{S}_i , \qquad (۴)$$

که در آن

$$\vec{M}_i = J_1 \sum_{j:\langle ij \rangle} \vec{S}_j + J_2 \sum_{j:\langle\langle ij \rangle\rangle} \vec{S}_j$$

توجه داشته باشید که اثرات برهم‌کنش هر اسپین با همسایه‌هاش را به صورت میدان مغناطیسی مؤثری همانند $\vec{M}_i$ در نظر می‌گیریم. بنابراین، انرژی کلاسیکی هر اسپین برابر است با:

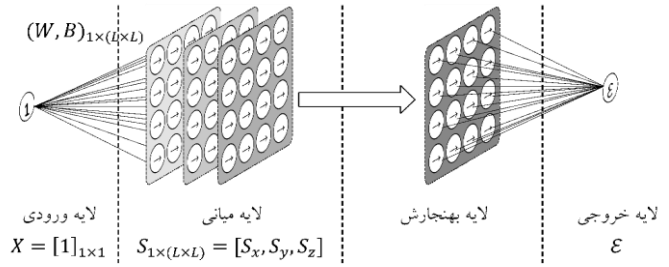
$$\varepsilon = \frac{1}{Q} \sum_i \vec{M}_i \cdot \vec{S}_i , \qquad (۵)$$

که Q تعداد کل اسپین‌های شبکه را نشان می‌دهد. طبق این رابطه، برای کمینه شدن انرژی سامانه، اسپین‌ها روی هر جایگاه شبکه باید به گونه‌ای باشند که در خلاف جهت میدان مؤثر $\vec{M}_i$ که احساس می‌کنند، جهت‌گیری کنند. در اینجا، انرژی میانگین سامانه، به عنوان تابع هدف است که برای پیدا کردن آرایش اسپینی مناسب به ازای ثابت‌های تبادلی مختلف در روش شبکهٔ عصبی عمیق، باید تا حد امکان به مقدار کمینهٔ آرمانیش، نزدیک شود تا تابع خطا به کمترین مقدار خود برسد. در این مقاله، برای سادگی در انجام محاسبات، $J_1 = 1$ در نظر گرفته‌ایم.

اینجا سعی می‌کنیم روش عددی شبکهٔ عصبی عمیق را با جزئیات بیشتر توضیح دهیم. همان طور که در شکل ۲ نشان داده شده است، لایه ورودی شامل ماتریس واحد است که به نورون‌های لایهٔ میانی، مرتبط است. نورون‌های لایهٔ میانی را می‌توان با ضرب ماتریس ورودی در ماتریس وزن $W_{1\times(L\times L)}$ و جمع با ماتریس سوییدهٔ $B_{1\times(L\times L)}$ از طریق رابطهٔ زیر به دست آورد، که همان پارامترهایی هستند که باید توسط شبکه یادگیری شوند.

$$S_{1\times(L\times L)} = X_{1\times 1} \cdot W_{1\times(L\times L)} + B_{1\times(L\times L)} , \qquad (۶)$$

با استفاده از روش گرادیان کاهشی، این مقادیر باید آن قدر تغییر کنند تا بهترین مقدار با کمترین مقدار خطا در نهایت به دست آید که به فرایند فوق، آموزش شبکه می‌گویند. توجه



شکل ۲. طرح کلی شبکهٔ عصبی عمیق، متشکل از سه بخش لایهٔ ورودی، میانی و خروجی.

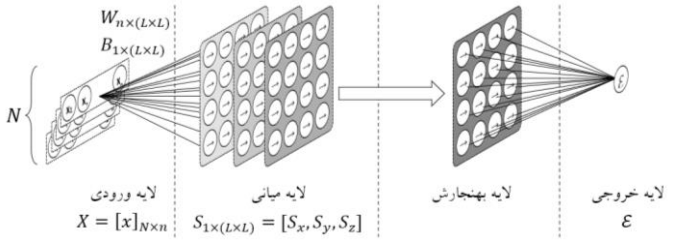
داشته باشید که در شبکهٔ عصبی عمیق، عناصر ماتریس وزن می‌توانند هر مقداری را اختیار کنند. بنابراین ممکن است که مقادیر نهایی به دست آمده برای اسپین‌ها، هر مقدار دلخواهی داشته باشد. برای اعمال شرط سخت، که اندازهٔ اسپین در هر جایگاه باید یک باشد، آنها از یک لایهٔ بهنجارش عبور داده می‌شوند تا بردارهای مربوط به اسپین‌ها، بهنجار شوند. در گام بعدی، با محاسبهٔ انرژی سامانه، مقدار خطا را می‌توان با اختلاف انرژی حاصل با آرمانی‌ترین انرژی سامانه، یعنی $-QSM_{\max}$ (که در آن، $M_{\max}$ بیشینه میدان مغناطیسی مؤثر میانگینی است که هر اسپین می‌تواند احساس کند)، با استفاده از رابطهٔ (۳) به دست آورد و از میزان خطای محاسبه شده، ماتریس‌های وزن و سوییده مربوطه را اصلاح کرد. این فرایند باید تا همگرا شدن این روش به یک مقدار بهینه، ادامه یابد.

یکی از کاستی‌های روش شبکهٔ عصبی عمیق، احتمال بیش‌برازش یا یادگیری بیش‌ازحد شبکه است. با توجه به داده‌های کم آموزشی برای سامانه‌های اسپینی و همچنین زیاد بودن پارامترهای یادگیری، احتمال بیش‌برازش به‌خودی‌خود افزایش می‌یابد. جهت رفع این نقصان، لایهٔ ورودی را بزرگ‌تر خواهیم کرد تا ترکیب احتمالاتی بیشتری جهت پیدا کردن آرایش اسپینی با کمترین انرژی را داشته باشیم. با توجه به بزرگی فضای حالت‌های ممکن برای سامانه‌های اسپینی، تنها با در نظر گرفتن یک نمونهٔ ورودی حتی با ابعاد بزرگ‌تر، منتج به قدرت اکتشافی زیادی برای شبکه نمی‌شود. برای اطمینان از درستی حالت پایهٔ به دست آمده، از تعداد بیشتری نمونهٔ ورودی جهت یادگیری شبکه استفاده می‌کنیم که این امکان را می‌دهد که آرایش‌های اسپینی بیشتری را انتخاب کنیم. بنابراین، بردار ورودی که یک درایه داشت به برداری با n درایه تبدیل می‌شود و همچنین تعداد نمونه‌های ورودی به N نمونه افزایش می‌یابد. برای انجام ضرب ماتریسی در رابطهٔ (۵)، لازم است که ماتریس وزنی همانند $W_{n\times(L\times L)}$ در نظر گرفته شود. هرچند باید دقت کرد که اگر هر مرحلهٔ یادگیری، با در نظر گرفتن تمام نمونه‌ها انجام گیرد، سرعت یادگیری به‌شدت پایین می‌آید. بدین منظور، هر بار از فضای حالت مربوط به ورودی‌های مختلف، به صورت تصادفی چند نمونه انتخاب می‌شود و با میانگیری روی نمونه‌های متفاوت، انرژی محاسبه می‌شود. در شبکهٔ عصبی عمیق، از این به عنوان روش مینی بچ[1] یاد می‌شود. طرح نهایی شبکهٔ عصبی عمیق استفاده شده را در شکل ۳ مشاهده می‌کنید.

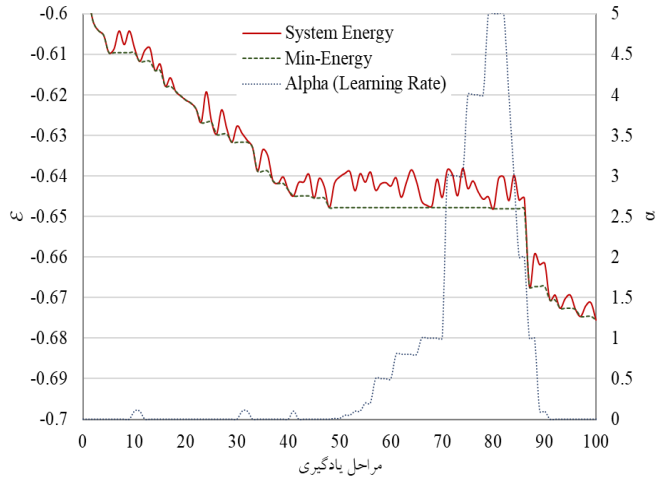
باید دقت کرد که در حل مسئله‌های بهینه‌سازی، بسیار محتمل است که سامانه در کمینه‌های موضعی قرار بگیرد. برای این‌که سامانه بتواند به کمینهٔ کلی خودش برسد، در این مقاله، از دو روش متفاوت استفاده کرده‌ایم. در روش اول، به جای استفاده از روش گرادیان کاهشی جهت یادگیری ماتریس وزن شبکهٔ عصبی، از روش بهینه سازی آدام[2] استفاده کرده‌ایم [۲۴]. در روش آدام، برای انجام محاسبات یادگیری ماتریس وزن، گشتاورهای مرتبهٔ اول و دوم به صورت همزمان، در نظر گرفته می‌شوند. با در نظر گرفتن مراتب بالاتر گشتاور، می‌توان در بسیاری از موارد، سامانه را از قرار گرفتن در کمینه‌های موضعی دور کرد و به کمینهٔ کلی سوق داد. در روش دوم، به جای در نظر گرفتن مقادیر ثابت برای پارامتر نرخ یادگیری آدام (α)،

۱. Mini-Batch

۲. Adam Optimizer



شکل ۳. طرح کلی شبکهٔ عصبی عمیق با داده‌های ورودی بیشتر، جهت رفع احتمال بیش‌برازش.



شکل ۴. تغییر دینامیکی نرخ یادگیری، α، به‌عنوان تابعی از زمان، که دور شدن از کمینهٔ موضعی را نشان می‌دهد.

اجازه می‌دهیم این پارامتر با شرایط لحظه‌ای مسئله تغییر کند. پارامتر نرخ یادگیری آدام جهت کنترل یادگیری شبکهٔ عصبی عمیق استفاده می‌شود. توجه داشته باشید در این روش، با نزدیک شدن تابع هدف به مقدار کمینهٔ کلی، باید پارامتر نرخ یادگیری کاهش یابد تا به سامانه اجازه دهیم که با آهنگ تغییرات کم، به کمینهٔ کلی‌اش نزدیک شود. قابل‌ذکر است که در این حالت، در نظر گرفتن مقدار نرخ یادگیری بزرگ‌تر، سامانه را از کمینهٔ کلی‌اش دور می‌کند. در ضمن، همان طور که در شکل ۴ نشان داده شده است، می‌توان با تغییر پارامتر نرخ یادگیری، سامانه را از کمینهٔ موضعی خارج کرد و به سامانه اجازه داد که در کمینهٔ کلی‌اش قرار گیرد. شکل ۴، یک برش ۱۰۰ مرحله‌ای از یادگیری سیستم شبکهٔ عصبی مورد استفاده را نشان می‌دهد. همان طور که واضح است نمودار تغییرات انرژی سامانه (System_Energy) از مرحلهٔ ۵۰ به بعد، تغییر موثری نمی‌کند و سامانه در کمینهٔ موضعی قرار می‌گیرد و نمودار کمینه انرژی یافت‌شده (Min_Energy) تا این مرحله، ثابت باقی می‌ماند. اگر پارامتر نرخ یادگیری را ثابت در نظر بگیریم، سامانه در کمینهٔ موضعی باقی خواهد ماند. با بهره‌گیری از نرخ یادگیری دینامیکی و افزایش آن، سامانه از مرحلهٔ ۹۰ به بعد، از کمینهٔ موضعی خارج شده و به یک فضای حالت جدید با انرژی پایین‌تر قرار می‌گیرد. نتایج نشان می‌دهد که در نظر گرفتن نرخ یادگیری دینامیک، به سامانه این اجازه را می‌دهد که حالت‌هایی را با انرژی پایین‌تر، نسبت به حالتی که نرخ یادگیری ثابت است، پیدا کند. در ادامه، فرایند یادگیری شبکهٔ عصبی عمیق پیشنهادی را با جزئیات بیشتر توضیح خواهیم داد.

(الف) مقداردهی اولیه:

ابتدا ماتریس ورودی $X_{N\times n}$، ماتریس وزن $W_{n\times(L\times L)}$ و بردار سوییده $B_{1\times(L\times L)}$ با توجه به اندازهٔ شبکهٔ L، بُعد ورودی n و تعداد نمونه‌های در نظر گرفته‌شده برای ورودی N با اعداد تصادفی مقداردهی می‌شوند. قابل ذکر است که اگر اعداد در نظر گرفته شده کاملاً تصادفی باشند طوری‌که میانگین آنها صفر

باشد، سرعت یادگیری بیشتر خواهد شد.

(ب) مرحلهٔ پیش‌رو:

همان گونه که قبلاً ذکر شد، جهت افزایش سرعت یادگیری شبکه، به جای انتخاب همه نمونه‌های ورودی در هر تکرار، تعداد نمونه (R) به صورت تصادفی انتخاب می‌شود و برای محاسبهٔ تابع هدف، روی همه نمونه‌های انتخابی، میانگیری می‌شود. مطابق با رابطهٔ (۵)، آرایش اولیهٔ اسپین‌ها روی شبکه، با ضرب بردار ورودی در ماتریس وزن و جمع آن با ماتریس سوییده به دست می‌آید. برای اعمال شرط سخت، که اندازهٔ اسپین در هر جایگاه باید یک باشد، آنها از یک لایهٔ بهنجارش عبور داده می‌شوند تا بردارهای مربوط به اسپین‌ها، بهنجار شوند. چنانچه انرژی مربوط به آرایش به دست آمده، از انرژی حالت قبلی کمتر باشد، این آرایش به عنوان آرایش جدید انتخاب می‌شود و در غیر این صورت در همان حالت قبلی باقی می‌ماند. احتمال پذیرش را می‌توان با استفاده از اختلاف انرژی حالت‌های قدیم (S_{old}) و حالت جدید (S_{new})، همانند زیر بیان کرد:

$$\Delta\varepsilon = \varepsilon_{new} - \varepsilon_{old}\,,\qquad P(S_{old}\to S_{new}) = \begin{cases} 1 & \Delta\varepsilon < 0 \\ 0 & \Delta\varepsilon > 0 \end{cases}, \tag{۷}$$

(ج) پس‌انتشار خطا و اصلاح پارامترهای ماتریس‌های وزن و سوییده به روش آدام:

مشتق جزئی تابع خطا نسبت به هر یک از پارامترهای یادگیری وزن و سوییده را می‌توان همانند زیر تعریف کرد:

$$\frac{d\Upsilon}{dW} = \frac{d\Upsilon}{d\varepsilon}\frac{d\varepsilon}{dS}\frac{dS}{dW} = \frac{2}{R}\,|\,\varepsilon_{target} - \varepsilon_{net}\,| \times \sum_{i=1}^{L\times L} Q\overrightarrow{M}_i \times \overrightarrow{X},$$
$$\frac{d\Upsilon}{dB} = \frac{d\Upsilon}{d\varepsilon}\frac{d\varepsilon}{dS}\frac{dS}{dB} = \frac{2}{R}\,|\,\varepsilon_{target} - \varepsilon_{net}\,| \times \sum_{i=1}^{L\times L} Q\overrightarrow{M}_i \times \overrightarrow{X}, \tag{۸}$$

که در آن

$$\varepsilon_{net} = \frac{1}{R}\sum_{r=1}^{R}\varepsilon(S_r),\qquad \varepsilon_{target} = -QSM_{max}\,,$$
$$\Upsilon = |\,\varepsilon_{net} - \varepsilon_{target}\,|^2, \tag{۹}$$

با استفاده از رابطهٔ (۹)، بردارهای گشتاور مرتبهٔ اول ($\hat{\rho}$) و گشتاور مرتبهٔ دوم (η)، پارامترهای یادگیری (W, B) در روش آدام از روابط بازگشتی زیر محاسبه می‌شوند:

$$\begin{aligned}
\rho_t^{(W)} &= \gamma_1\cdot\rho_{t-1}^{(W)} + (1-\gamma_1)\frac{d\Upsilon}{dW_t},\\
\eta_t^{(W)} &= \gamma_2\cdot\eta_{t-1}^{(W)} + (1-\gamma_2)\left(\frac{d\Upsilon}{dW_t}\right)^2,\\
\rho^{(W)} &= \frac{\rho_t^{(W)}}{1-\gamma_1^t},\qquad \eta^{(W)} = \frac{\eta_t^{(W)}}{1-\gamma_2^t},\\
\rho_t^{(B)} &= \gamma_1\cdot\rho_{t-1}^{(B)} + (1-\gamma_1)\left(\frac{d\Upsilon}{dB_t}\right),\\
\eta_t^{(B)} &= \gamma_2\cdot\eta_{t-1}^{(B)} + (1-\gamma_2)\left(\frac{d\Upsilon}{dB_t}\right)^2,\\
\rho^{(B)} &= \frac{\rho_t^{(B)}}{1-\gamma_1^t},\qquad \eta^{(B)} = \frac{\eta_t^{(B)}}{1-\gamma_2^t},
\end{aligned} \tag{۱۰}$$

γ_1 و γ_2 به ترتیب، ضرایب فراموشی گرادیان و گشتاور دوم گرادیان هستند و t نمایانگر تکرار در هر مرحله از یادگیری است. t در هر مرحلهٔ یادگیری، یک واحد افزایش می‌یابد. با استفاده از رابطهٔ (۱۰)، می‌توان پارامترهای یادگیری را همانند زیر اصلاح کرد:

$$\begin{aligned}
W_t &= W_{t-1} - \alpha\left(\frac{\rho^{(W)}}{\sqrt{\eta^{(W)}} + \delta}\right),\\
B_t &= B_{t-1} - \alpha\left(\frac{\rho^{(B)}}{\sqrt{\eta^{(B)}} + \delta}\right),
\end{aligned} \tag{۱۱}$$

α پارامتر نرخ یادگیری و δ یک مقدار بی‌نهایت کوچک است که برای رفع تکینگی در محاسبات عددی به مخرج کسر اضافه می‌شود.

(د) اعمال پویای نرخ یادگیری:

بر اساس شرایط لحظه‌ای که سامانه دارد، می‌توان با تغییر پویای نرخ یادگیری بر اساس روابط زیر، سامانه را از باقی ماندن در کمینه‌های موضعی دور کرد.

$$P(\alpha\uparrow) = \begin{cases} 1 & \Delta\varepsilon > 0 \\ 0 & \Delta\varepsilon < 0 \end{cases},\qquad P(\alpha\downarrow) = \begin{cases} 1 & \Delta\varepsilon < 0 \\ 0 & \Delta\varepsilon > 0 \end{cases}, \tag{۱۲}$$

در این رابطه، چنانچه مقدار انرژی تغییر نکند $\Delta\varepsilon = 0$، نرخ یادگیری $\alpha\uparrow$ افزایش می‌یابد تا سامانه از کمینه موضعی خارج شود. به محض کاهش مقدار انرژی و خارج شدن سامانه از

کمینهٔ موضعی $\Delta\varepsilon < 0$، نرخ یادگیری $\alpha \downarrow$ کاهش می‌یابد تا به کمینهٔ کلی نزدیک شود.

(ه) تکرار مراحل یادگیری:

برای رسیدن به حالت پایهٔ درست، مراحل فوق باید تکرار شوند تا یکی از شرایط زیر برقرار شود:

- سامانه به حالتی با مقدار کمینهٔ کلی همگرا شود.
- میزان خطا از یک حد آستانهٔ تعیین‌شده کمتر شود.
- تعداد مراحل تکرار یادگیری از آستانهٔ تعیین‌شده، بیشتر شود.

۴. نتایج

در این بخش، ابتدا با استفاده از روش شبکهٔ عصبی عمیق پیشنهادی، مدل هایزنبرگ شبکهٔ مربعی و شبکهٔ لانه زنبوری مطالعه می‌کنیم و سپس در مقایسه با روش الگوریتم‌های تکاملی، به مزیت‌های روش شبکهٔ عصبی عمیق برای مطالعهٔ سامانه‌های اسپینی، خواهیم پرداخت.

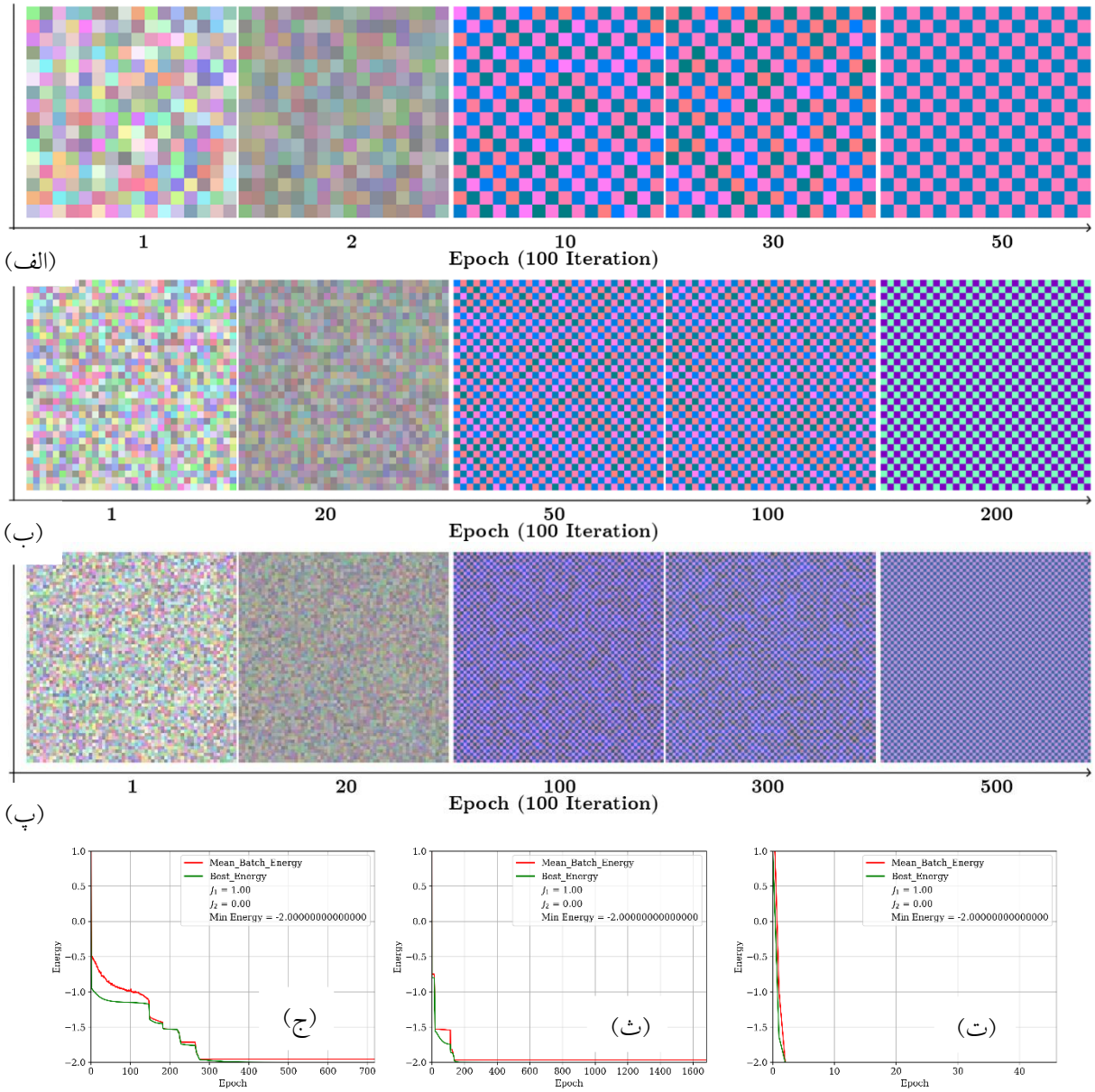
قبل از این‌که به مطالعهٔ سیمای فاز کلی مدل هایزنبرگ همسایهٔ اول و دوم شبکهٔ مربعی بپردازیم، حالت پایهٔ مدل هایزنبرگ شبکهٔ مربعی $J_2 = 0$ را بررسی می‌کنیم. بدین منظور شبکهٔ مربعی ۱۶×۱۶، ۳۲×۳۲ و ۶۴×۶۴ را انتخاب می‌کنیم. برای فهم نسبت به حالت مغناطیسی با کمینه انرژی کلی، می‌توان آرایش اسپینی مربوط به این حالت را در فضای حقیقی رسم کرد. همان طور که در سه ردیف اول شکل ۵ نشان داده شده است آرایش اسپینی در فضای حقیقی به ازای مراحل یادگیری تا نزدیک شدن به آرایش اسپین نهایی (ستون آخر) برای مدل هایزنبرگ شبکهٔ مربعی با اندازه‌های متفاوت ۱۶×۱۶، ۳۲×۳۲ و ۶۴×۶۴ پادفرومغناطیس است. زمانی‌که تعداد مراحل تکرار به عدد اشباع برسد، شکست خودبه‌خودی تقارن اتفاق می‌افتد و آرایش اسپینی با نظم بلند برد پادفرومغناطیس (ستون آخر) پایدار می‌شود. قابل ذکر است که نتیجهٔ به دست آمده برای این حالت، با کارهای قبلی همخوانی کامل دارد [۲۵].

در این مرحله می‌توان گفت سامانهٔ مغناطیسی، کمینهٔ کلی‌اش را پیدا کرده است. برای آرایش‌های اسپینی که بردار موج متناظر با آن متناسب با بردار شبکهٔ وارون است، حالت کمینهٔ انرژی موضعی شامل حوزه‌هایی محدود با نظمی مشابه حالت پایه است که با دیواره‌هایی از هم جدا شده‌اند. باید توجه داشته باشید که در روش شبکهٔ عصبی عمیق، تعداد مراحل یادگیری ثابت در نظر گرفته نشده است و در این روش، با نزدیک شدن به عدد اشباع، آهنگ تغییرات انرژی کاهش می‌یابد تا سامانه، حالت پایه با کمینهٔ کلی را پیدا کند (ردیف آخر شکل ۵ را ببینید). با افزایش اندازهٔ شبکه، تعداد مراحل یادگیری که سامانه به آرایش اسپینی نهایی‌اش برسد، به طور قابل ملاحظه‌ای افزایش می‌یابد.

با این توضیح کلی که چگونه می‌توان با روش عصبی عمیق، حالت پایهٔ سامانه‌های مغناطیسی را پیدا کرد، در ادامه به مطالعهٔ سیمای فاز کلاسیکی مدل هایزنبرگ همسانگرد همسایهٔ اول و دوم شبکهٔ مربعی و شبکهٔ لانه زنبوری خواهیم پرداخت.

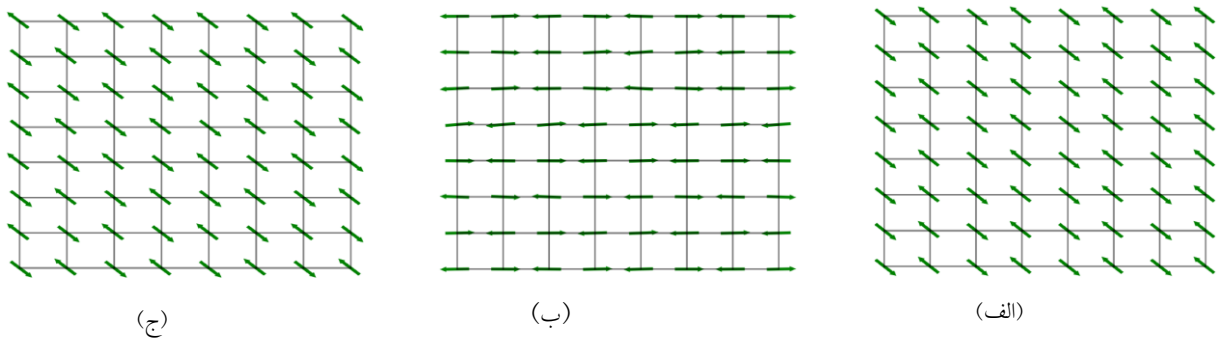
۴. ۱. سیمای فاز کلاسیکی مدل هایزنبرگ شبکهٔ مربعی

آرایش اسپینی به دست آمده با استفاده از روش عددی شبکهٔ عصبی عمیق نشان می‌دهد که با افزایش برهم‌کنش تبادلی همسایهٔ دوم تا $J_2 < \frac{1}{2}$، نظم پادفرومغناطیس هنوز پایدار است که در این آرایش اسپینی، اسپین‌های همسایهٔ هر جایگاه به صورت پادموازی جهت‌گیری کرده‌اند (شکل ۶. الف را ببینید). فاز میانی پیچشی، شکل ۶. ب، فقط به $J_2 = \frac{1}{2}$ پایدار است. محاسبات نشان می‌دهد که حالت پایهٔ کلاسیکی پیچشی، واگنی بی‌نهایت دارد. قابل ذکر است که بردار موج متناظر با نظم پیچشی، ضریب گویایی از بردار شبکهٔ وارون نیست، یعنی بردار موج ناجور $\vec{q}_{inc} \neq \frac{c}{d}G$، که در آن c و d اعداد صحیح هستند. همچنین به ازای $J_2 > \frac{1}{2}$، نظم مغناطیسی هم‌خط با بردار موج $\vec{q} = (\pi, 0)$ یا $\vec{q} = (0, \pi)$ پایدار است. در فاز مغناطیسی هم‌خط، جهت‌گیری اسپین‌ها در یک راستای شبکهٔ مربعی به صورت فرومغناطیس و در راستای دیگر شبکه، آرایش پادفرومغناطیس دارند (شکل ۶. ج را ببینید).

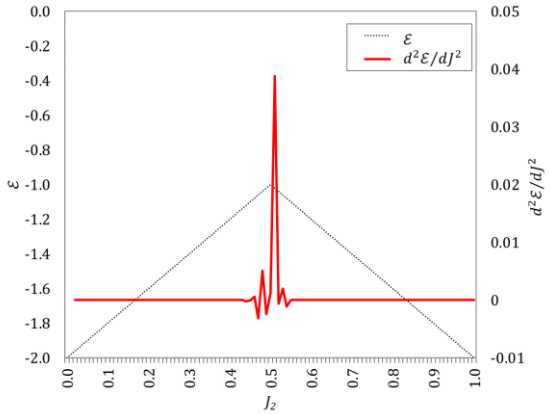


شکل ۵. نمایش آرایش اسپینی حالت $J_۲ = ۰$ به ازای مراحل تکرار متفاوت برای شبکهٔ مربعی با اندازهٔ (الف) ۱۶×۱۶، (ب) ۳۲×۳۲ و (ج) ۶۴×۶۴. رنگ‌بندی در این شکل‌ها، نمایش اسپینی را در فضای حقیقی نمایش می‌دهد. در مراحل یادگیری اولیه، رنگ‌های کاملاً تصادفی روی جایگاه‌های شبکهٔ مربعی انتخاب شده‌اند که بیانگر وجود آرایش اسپینی بی‌نظم است. با افزایش مراحل یادگیری، اسپین‌ها روی شبکهٔ مربعی، آرایش پادفرومغناطیس پیدا می‌کنند. همان طور که در ستون آخر سه ردیف اول مشخص است، واحدهای کوچ با رنگ متفاوت، نمایانگر اسپین‌ها با جهت‌گیری فضایی کاملاً مخالف هم در فضای حقیقی هستند. تغییرات انرژی بر حسب مراحل یادگیری برای شبکهٔ مربعی با اندازهٔ (ت) ۱۶×۱۶، (ث) ۳۲×۳۲ و (ج) ۶۴×۶۴.

انرژی حالت پایه و مشتق مرتبهٔ دوم انرژی حالت پایه مدل هایزنبرگ همسانگرد شبکهٔ مربعی به عنوان تابعی از $J_۲$ در شکل ۷ نشان داده شده است. تکینگی موجود در مشتق مرتبهٔ دوم انرژی به ازای $J_۲ = \frac{۱}{۲}$، بیانگر گذار فاز بین نظم‌های پادفرومغناطیس و هم‌خط است و در ضمن ناپیوستگی مشتق مرتبهٔ اول انرژی به ازای $J_۲ = \frac{۱}{۲}$ نشان دهندهٔ گذار از مرتبهٔ اول



شکل ۶. آرایش اسپینی مدل هایزنبرگ همسانگرد شبکهٔ مربعی به ازای (الف) $J_2 < \frac{1}{2}$، (ب) $J_2 = \frac{1}{2}$ و (ج) $J_2 > \frac{1}{2}$ که با استفاده از روش شبکهٔ عمیق به دست آمده است.



شکل ۷. انرژی حالت پایه (خط نقطه‌چین) و مشتق مرتبهٔ دوم انرژی حالت پایه (خط توپر) مدل هایزنبرگ شبکهٔ مربعی به عنوان تابعی از J_2.

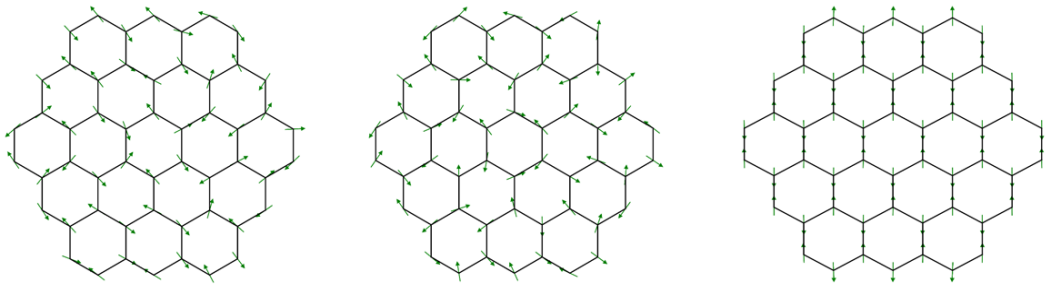
بین این دو نظم مغناطیسی بلند برد است. قابل ذکر است که سیمای فاز کلاسیکی به دست آمده با استفاده از روش شبکهٔ عصبی عمیق برای مدل هایزنبرگ همسانگرد شبکهٔ مربعی، با نتایجی که با روش‌های لاتینجر-تیزا و کمینه سازی وردشی به دست آمده، سازگاری کامل دارد [۲۵].

۴. ۲. سیمای فاز کلاسیکی مدل هایزنبرگ شبکهٔ لانه زنبوری

در این بخش قصد داریم سیمای فاز کلاسیکی مدل هایزنبرگ همسانگرد همسایهٔ اول و دوم شبکهٔ لانه زنبوری را بررسی کنیم. در انجام محاسبات برای مقادیر کوچک J_2، نظم پادفرومغناطیس با بردار موج $\vec{q}=(0,0)$ پایدار است که در هر یاختهٔ واحد، جهت‌گیری اسپین‌ها خلاف جهت هم هستند (شکل ۸. الف را ببینید). برای $J_2=\frac{1}{6}$، گذار فازی از نظم پادفرومغناطیس به نظم پیچشی با بردار موج ناجور وجود خواهد داشت. آرایش‌های اسپینی مربوط به نظم پیچشی شبکهٔ لانه زنبوری به ازای $J_2=0/35$ و $J_2=0/75$ به ترتیب در شکل‌های ۸. ب و ۸. ج نشان داده شده است. به ازای $J_2>\frac{1}{6}$، حالت پایهٔ کلاسیکی، تبهگنی بی‌نهایت دارد یعنی بی‌نهایت آرایش اسپینی با انرژی یکسان وجود دارد. در این حالت، تعدادی نامحدود از حالت‌های پیچشی تبهگن را داریم که حالت پایه می‌تواند از یک رویه انتخاب شود [۲۶].

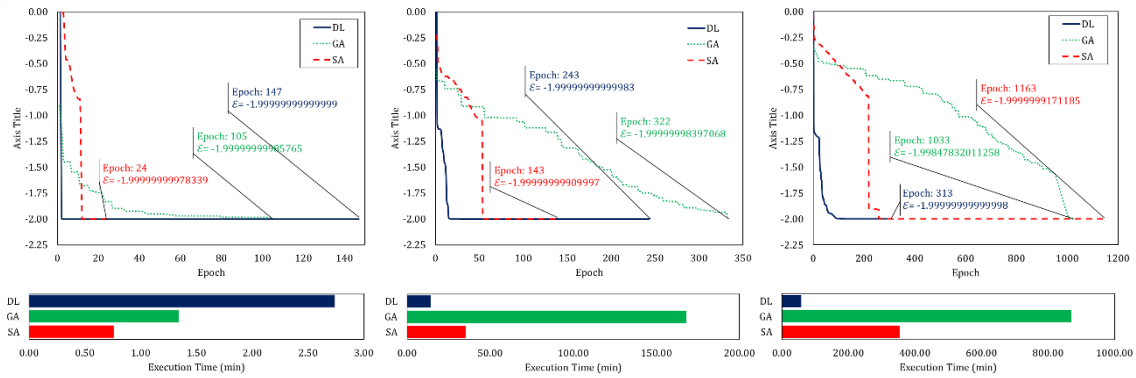
۴. ۳. مقایسهٔ کارایی شبکهٔ یادگیری عمیق با الگوریتم ژنتیک و تبرید شبیه‌سازی شده

در این بخش قصد داریم به مقایسهٔ روش‌های الگوریتم تکاملی نظیر ژنتیک (GA) و تبرید شبیه‌سازی شده (SA) با روش شبکه



(الف) $J_2 = 0.1$ (ب) $J_2 = 0.35$ (ج) $J_2 = 0.75$

شکل ۸. آرایش اسپینی مدل هایزنبرگ همسانگرد شبکهٔ لانه زنبوری به ازای (الف) $J_۲ = ۰/۱$، (ب) $J_۲ = ۰/۳۵$ و (ج) $J_۲ = ۰/۷۵$.



(الف) (ب) (ج)

شکل ۹. محاسبهٔ تغییرات انرژی بر حسب مراحل یادگیری و مدت زمان لازم برای پیدا کردن کمینهٔ کلی مدل هایزنبرگ شبکهٔ مربعی با اندازهٔ (الف) ۴×۴، (ب) ۸×۸ و (ج) ۱۶×۱۶ با استفاده از روش‌های الگوریتم تکاملی نظیر ژنتیک (GA) و تبرید شبیه سازی شده (SA) و روش شبکهٔ یادگیری عمیق (DL).

یادگیری عمیق (DL) بپردازیم. در شکل ۹ (ردیف بالا)، روند یادگیری سامانه شبکهٔ مربعی با اندازه‌های متفاوت (الف) ۴×۴ ، (ب) ۸×۸ (ج) ۱۶×۱۶ برای حالت خاص $J_۲ = ۰$، جهت پیدا کردن حالت پایهٔ نهایی نشان داده شده‌است. شکل ۹ (ردیف پایین) مدت زمانی که طول می‌کشد انرژی سامانه، به کمترین مقدارش همگرا شود را بیان می‌کند که در این حالت، می‌توان ادعا کرد که سامانهٔ اسپینی، کمینهٔ کلی‌اش را پیدا کرده است.

قابل توجه است که برای سامانه‌هایی با اندازهٔ کوچک‌تر، یعنی شبکهٔ مربعی ۴×۴ شکل ۹. الف، روش شبکهٔ یادگیری عمیق، به خاطر پارامترهای زیادی که در این روش جهت کمینه کردن انرژی وجود دارد، نسبت به دو روش دیگر زمان بیشتری نیاز دارد که سامانه بتواند کمینهٔ کلی‌اش را پیدا کند. با افزایش اندازهٔ سامانه، شکل‌های ۹. ب و ج، سامانه با استفاده از روش شبکهٔ عصبی عمیق در تعداد مراحل یادگیری خیلی کمتری می‌تواند حالت کمینهٔ کلی‌اش را پیدا کند. بنابراین نتایج به دست آمده بیانگر این است که روش شبکهٔ عصبی عمیق از نظر بازهٔ زمانی که نیاز است حالت پایهٔ کلاسیکی را پیدا کند، برتری خیلی خوبی نسبت به دیگر روش‌های ذکر شده دارد.

۵. نتیجه‌گیری

با بازبینی در فرمول‌بندی شبکه‌های عصبی عمیق، الگوریتم شبیه‌سازی مناسبی برای محاسبهٔ حالت پایهٔ سامانه‌هایی که زیر دمای کوری شکست تقارن خودبه‌خودی در آنها اتفاق می‌افتد،

معرفی کردیم. در این روش محاسباتی، آرایش اسپین‌ها را به عنوان پارامترهای شبکهٔ عصبی در نظر می‌گیریم، یعنی با در نظر گرفتن بردار اسپین به‌عنوان نورون‌های ورودی، از آنها به عنوان پارامترهای یادگیری شونده استفاده می‌کنیم. در حل مسئله‌های بهینه‌سازی، بسیار محتمل است که سامانه در کمینه‌های موضعی که انرژی بسیار نزدیکی با کمینهٔ کلی دارند، قرار گیرد. برای این‌که سامانه را از قرار گرفتن در کمینه‌های موضعی دور کنیم و به کمینهٔ کلی سوق دهیم، از روش آدام استفاده کرده‌ایم. در این روش، به جای در نظر گرفتن مقادیر ثابت برای پارامتر نرخ یادگیری آدام، اجازه می‌دهیم این پارامتر با شرایط لحظه‌ای مسئله تغییر کند، یعنی با نزدیک شدن تابع هدف به مقدار کمینهٔ کلی، باید پارامتر نرخ یادگیری طوری تغییر یابد، که سامانه بتواند با آهنگ تغییرات کم، به کمینل کلی‌اش نزدیک شود. صحت کارکرد روش شبکهٔ عصبی عمیق پیشنهادی را با مطالعهٔ مدل هایزنبرگ همسانگرد شبکهٔ مربعی و شبکهٔ لانه زنبوری بررسی کردیم که نتایج به دست آمده با نتایج روش‌های الگوریتم تکاملی، کاملاً سازگار است.

مراجع

1. M Seul and D Andelman, *Science* **267** (1995) 476.
2. P Bogdan, E Jonckheere, and S Schirmer. *Chaos, Solitons, & Fractals* **103** (2017) 622.
3. G Carleo, *et al*., *Rev. Mod. Phys*. **91** (2019) 045002.
4. P Mehta, *et al*., *Physics Reports* **810** (2019) 1.
5. L Wang, *Phys. Rev*. B **94** (2016) 195105.
6. P Ponte and R G Melko, *Phys. Rev.* B **96** (2017) 205146.
7. J Carrasquilla and R G Melko, *Nat. Phys.* **13** (2017) 431.
8. K Ch'ng, *et al*., *Phys. Rev.* X **7** (2017) 031038.
9. S J Wetzel, *Phys. Rev.* E **96** (2017) 022140.
10. K Ch'ng, N Vazquez, and E Khatami, *Phys. Rev.* E **97** (2018) 013306.
11. G Carleo and M Troyer, *Science* **355** (2017) 602.
12. Z Cai and J Liu, *Phys. Rev.* B **97** (2018) 035116.
13. J Carrasquilla, *Adv. Phys.* X **5** (2020) 1797528.
14. D L Deng, X Li, and S Das Sarma, *Phys. Rev.* X **7** (2017) 021021.
15. J Hermann, Z Schatzle, and F No´e, *Nature Chem.* **12** (2020) 891.
16. D Pfau, *et al*., *Phys. Rev. Research* **2** (2020) 033429.
17. T Vieijra, *et al*., *Phys. Rev. Lett.* **124** (2020) 097201.
18. K Liu, *et al*., *Phys. Rev. Research* **3** (2021) 023016.
19. N Rao, *et al*., *arXiv: 2102.01103* (2021).
20. H Y Kwon, *et al*., *Phys. Rev.* B **99** (2019) 024423.
21. Y Zhang, *et al*., *Nature* **570** (2019) 484.
22. A Bohrdt, *et al*., *Nature Phys.* **15** (2019) 921.
23. J Schmidt, *et al*., *Npj Comput. Mater.* **5** (2019) 83.
24. D P Kingma and J L Ba, *arXiv:1412.6980* (2014).
25. Z Mortazavizade, H Mosadeq, and M H Zare, *Iran. J. Phys. Res.* **20** (2020) 117.
26. M H Zare, F Fazileh, and F Shahbazi, *Phys. Rev.* B **87** (2013) 224416.