QPEN: Quantum Projection and Quantum Entanglement Enhanced Network for Cross-Lingual Aspect-Based Sentiment Analysis

Xingqiang Zhao,Hai Wan*,Kunxun Qi

School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, P.R.China
zhaoxq23@mail2.sysu.edu.cn, wanhai@mail.sysu.edu.cn, qikx@mail2.sysu.edu.cn

Abstract

Aspect-based sentiment analysis (ABSA) has attracted much attention due to its wide application scenarios. Most previous studies have focused solely on monolingual ABSA, posing a formidable challenge when extending ABSA applications to multilingual scenarios. In this paper, we study upgrading monolingual ABSA to cross-lingual ABSA. Existing methods usually exploit pre-trained cross-lingual language to model cross-lingual ABSA, and enhance the model with translation data. However, the low-resource languages might be under-represented during the pre-training phase, and the translation-enhanced methods heavily rely on the quality of the translation and label projection. Inspired by the observation that quantum entanglement can correlate multiple single systems, we map the monolingual expression to the quantum Hilbert space as a single quantum system, and then utilize quantum entanglement and quantum measurement to achieve cross-lingual ABSA. Specifically, we propose a novel quantum neural model named QPEN (short for *quantum projection and quantum entanglement enhanced network*). It is equipped with a proposed quantum projection module that projects aspects as quantum superposition on a complex-valued Hilbert space. Furthermore, a quantum entanglement module is proposed in QPEN to share language-specific features between different languages without transmission. We conducted simulation experiments on the classical computer, and experimental results on SemEval-2016 dataset demonstrate that our method achieves state-of-the-art performance in terms of F1-scores for five languages.

Introduction

Aspect-based sentiment analysis (ABSA) is a fine-grained task for sentiment analysis. It aims at inferring the sentiment polarities over specific aspects in a sentence (Liu 2012; Hu et al. 2019). With the development of the product reviews and the social media, ABSA has been widely applied in many real-world applications.

However, most existing methods focus merely on the monolingual ABSA task, and can hardly be directly applied to the multi-lingual scenario. In practice, the real-world applications such as E-commerce systems accept the reviews of user in different languages, requiring high-quality multilingual systems for the sentiment analysis. Therefore, the cross-lingual ABSA, which aims to train a ABSA model primarily on the data in one language and then apply it to other languages, has become an essential ingredient in multilingual sentiment analysis systems.

In recent years, cross-lingual pre-trained language models such as the *multilingual BERT (m-BERT)* model (Devlin et al. 2019) and the *XLM-Roberta (XLM-R)* model (Conneau et al. 2020) have become a prevalent paradigm for tackling cross-lingual ABSA. Thanks to the multilingual knowledge learned in the pre-training stage (Wu and Dredze 2019), these pre-trained models are able to be fine-tuned on labeled data in source language (usually English) and then be directly applied to the data in target language. To further investigate the transfer of language-specific knowledge in solving the cross-lingual ABSA problem, an aspect code-switching mechanism with knowledge distillation (Zhang et al. 2021) has been proposed to enhance the cross-lingual alignment for pre-trained cross-lingual language models.

However, it is still challenging to adopt the cross-lingual pre-trained language models to tackle the cross-lingual ABSA task since the low-resource languages might be under-represented during the pre-training phase (Conneau et al. 2020; Pfeiffer et al. 2020). Making use of the translated target language data with projected labels is a plausible way to transfer language-specific knowledge (Li et al. 2020), whereas the performance of such translation-based methods heavily relies on the quality of the translation and label projection. In practice, the task-specific knowledge in the translated data would also be limited if the quality of the projected label is unsatisfactory.

To tackle these challenges, we investigate the problem of cross-lingual ABSA by introducing quantum modules to pre-trained cross-lingual language models, based on the observation that quantum entanglement can correlate multiple single systems in quantum systems. Specifically, we propose a novel quantum cross-lingual ABSA model named QPEN (short for *quantum projection and quantum entanglement enhanced network*). It is equipped with a quantum projection module and a quantum entanglement module, as shown in Figure 1. Given a review context, QPEN first encodes the context to contextualized representation by applying a pre-trained cross-lingual language model such as mBERT or

*Corresponding author

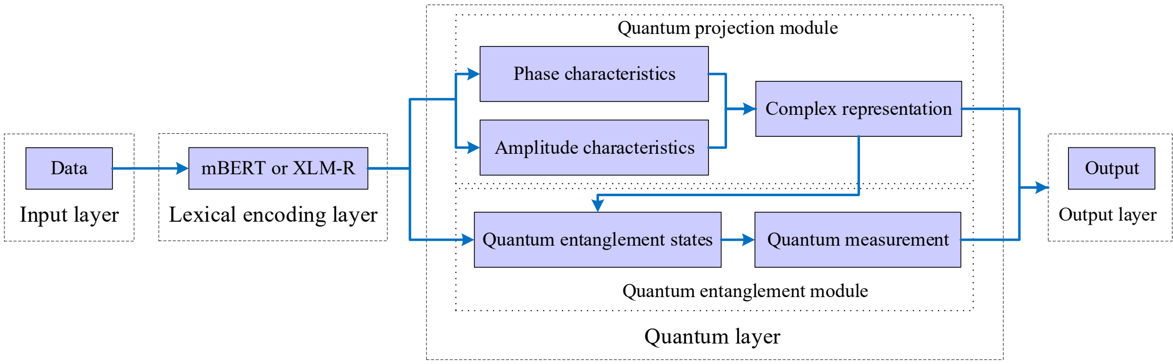

Figure 1: The overall architecture of QPEN. QPEN consists of the input layer, the lexical encoding layer, the quantum layer, and the output layer, of which the quantum layer contains the quantum projection module and the quantum entanglement module.

XLM-R. We propose two quantum modules to enhance the cross-lingual language modelling for cross-lingual ABSA. The first module is the quantum projection module, which projects the contextualized representation to the complex representation in a Hilbert space. With this module, we can formulate each aspect as a quantum superposition state on a complex-valued Hilbert space. In order to enforce the model sharing the language-specific knowledge between different languages, we propose a quantum entanglement module that enables each particle to measure its own quantum state to share the entire information without transmission throughout the quantum system, yielding entanglement representation that captures shared language-specific knowledge between different languages. The complex representation and the entanglement representation are then concatenated for the sequence labeling classification.

We evaluate our model QPEN on the SemEval-2016 dataset (Pontiki et al. 2016). Experimental results demonstrate that the proposed model outperforms state-of-the-art methods by an average absolute gain of 1.03% (resp. 2.31%) in terms of F1-score based on the mBERT (resps. XLM-R) model. The main contributions of this work include:

1. We proposed a quantum neural model named QPEN for cross-lingual ABSA, which is equipped with a quantum projection module and a quantum entanglement module.To the best of our knowledge, it is the first method introducing quantum modules for cross-lingual ABSA.
2. We provide the quantum circuit implementation for the proposed quantum entangled module, which allows our model QPEN to be implemented on a quantum computer.
3. We conduct the simulation experiments on a classical computer using the SemEval-2016 dataset to demonstrate significant improvements achieved by QPEN.

Related Works

Our work shares the same research line with aspect-based sentiment analysis and quantum neural networks.

Aspect-Based Sentiment Analysis. Traditional sentiment analysis tasks (Liu 2012) are sentence-level or document-level oriented. In contrast, ABSA is an entity-level oriented and a more fine-grained task for sentiment analysis. Most recent researches on cross-lingual ABSA mainly focus on its sub-tasks including the cross-lingual aspect term extraction (Lin et al. 2014) and aspect sentiment classification (Lambert 2015; Barnes, Lambert, and Badia 2016). To obtain language knowledge of the target languages, translation systems are used to obtain pseudo-parallel data (Zhou, Wan, and Xiao 2015). A word or phrase alignment algorithm such as fastAlign (Dyer, Chahuneau, and Smith 2013) is then utilized to project the label from the source to the target sentence. Since the performance of such methods heavily depends on the quality of the translation and alignment, different strategies (Klinger and Cimiano 2015; Li et al. 2020) are proposed to further improve the data quality. Another line of work uses the cross-lingual word embeddings trained on large parallel bilingual corpus (Ruder, Vulić, and Søgaard 2019). By switching the word embeddings between different languages, the model can be used in a language-agnostic manner (Barnes, Lambert, and Badia 2016; Akhtar et al. 2018; Wang and Pan 2018; Jebbara and Cimiano 2019). Recently, the transformer-based models pre-trained on large multilingual corpus, such as the *multilingual BERT (mBERT)* model (Devlin et al. 2019) and the *XLM-Roberta (XLM-R)* model (Conneau et al. 2020), have shown significant improvements for various cross-lingual *Natural Language Processing (NLP)* (Martin 2009) tasks. Thanks to the language knowledge learned in the pre-training process, fine-tuning the model on the labeled source language data and directly conducting the inference on the target data can achieve impressive cross-lingual adaptation performance (Wu and Dredze 2019; Pires, Schlinger, and Garrette 2019; Karthikeyan et al. 2020). Some studies further utilize the translation system together with the pre-trained models (Fei, Zhang, and Ji 2020; Hu et al. 2020; Singh et al. 2020; Li et al. 2020). To investigate the importance of language-specific knowledge in solving the cross-lingual ABSA problem, the model ACS (Zhang et al. 2021) was distilled on the unlabeled target language data, which improves the performance to the same level as the supervised method.

Quantum Neural Networks. The quantum method provides new ideas and directions for the development of NLP

(Martin 2009), and some quantum and quantum-inspired neural networks have been designed (Li et al. 2021; Zhao, Hou, and Xu 2022; Yan, Wu, and Yan 2023). Recently, a fundamentally new, quantum cognitively motivated fusion strategy for predicting sentiment judgments was proposed. In particular, utterances were formulated as quantum superposition states of positive and negative sentiment judgments, and unimodal classifiers were formulated as mutually incompatible observables, on a complex-valued Hilbert space with positive-operator valued measures (Gkoumas et al. 2021a). In view of the advantages of *quantum probability (QP)* (Gudder 2014) in modeling such uncertainty, a transparent quantum probabilistic neural model and a QP driven multi-task learning framework have been proposed (Gkoumas et al. 2021b; Liu et al. 2021). Besides, a novel perspective on conversational emotion recognition has been provided by drawing an analogy between the task and a complete span of quantum measurement (Li et al. 2021). To accurately and comprehensively model complicated interactions, a comprehensive framework quantum-like multi-modal network for multi-modal sentiment analysis has been designed, which leverages the mathematical formalism of *quantum theory (QT)* and a *long short-term memory (LSTM)* network (Zhang et al. 2020). In addition, a novel quantum neural network has been proposed for learning combinatorial optimization problems in a supervised manner to achieve better and faster results (Ye, Yan, and Yan 2023), a hybrid quantum-classical generative adversarial network has been designed for the image generation via learning discrete distribution (Zhou et al. 2023), and a quantum convolutional neural network based on variational quantum circuits has been constructed (Gong et al. 2024).

Considering the importance of the language-specific knowledge in solving the cross-lingual ABSA problem, inspired by the quantum technology, we propose a *quantum projection and quantum entanglement enhanced network (QPEN)* for cross-lingual ABSA tasks since the quantum entanglement enables each particle to measure its own quantum state to share the whole information without transmission throughout the quantum system.

Preliminaries on Quantum Theory

We introduce the key concepts of the quantum cognition (Busemeyer and Bruza 2012; Fell et al. 2019), which we exploit to construct the proposed work.

Quantum Projection

Hilbert Space. Quantum cognition exploits an infinite complex-valued vector space, called Hilbert space $\mathcal{H}$, in which the state of a quantum system is represented as a unit-length vector. Different from classical probability, quantum probability events are defined as orthonormal basis states. A projective geometric structure establishes relationships between states vectors and basis states (Hughes 1989; Halmos 2017). The same Hilbert space can be represented by different sets of orthonormal basis states, and the same state can be defined over different sets of orthonormal basis states.

Quantum Superposition. Quantum superposition is one of the fundamental concepts in *quantum mechanics (QM)*(Merzbacher 1998), which describes the uncertainty of a single particle. In the micro world, a particle like a photon can be in multiple mutually exclusive basis states simultaneously with a probability distribution. A general pure state $|\varphi\rangle$ is a vector on the unit sphere, represented by

$$|\varphi\rangle = \omega_1|e_1\rangle + \omega_2|e_2\rangle + \ldots + \omega_n|e_n\rangle \tag{1}$$

where $\{|e_1\rangle, |e_2\rangle, \ldots, |e_n\rangle\}$ are basis states forming an orthogonal basis of the Hilbert Space, and the probability amplitudes $\{\omega_1, \omega_2, \ldots, \omega_n\}$ are complex scalars with $\sum_{i=1}^{n} |\omega_i|^2 = 1$, and $|\cdot|$ is the modulus of a complex number. $|\varphi\rangle$ is a superposition state when it is not identical to a certain basis state $|e_i\rangle$. In particular, in a two-dimensional Hilbert Space $\mathcal{H}_2$ spanned by basis states $|0\rangle$ and $|1\rangle$, a pure state $|\varphi\rangle$ is represented as

$$|\varphi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle \tag{2}$$

where $\theta \in [0, 2\pi]$, $\phi \in [0, 2\pi]$, i is the imaginary number and $i^2 = -1$.

Quantum Measurement. Quantum measurement is another fundamental concept in quantum cognition for calculating quantum probabilities. In quantum mechanics, *projection-valued measure (PVM)* removes a system state from uncertainty to a precise event, by projecting a state to its certain corresponding basis state. In the absence of measurement, there is uncertainty in the state in that it takes all possible measurement values simultaneously. After measurement, the state collapses onto a certain basis state. However, PVMs on subsystems of a larger system cannot be described by a PVM acting on the system itself. *Positive-operator valued measure (POVM)* overcomes this constraint, by associating a positive probability for each measurement outcome, ignoring the post-measurement state (Nielsen and Chuang 2002). That is to say, POVM is a generalization of PVM, providing mixed information of a state for the entire ensemble of all the subsystems (Gkoumas et al. 2021a).

Quantum Entanglement

In quantum mechanics, when several particles interact with each other, the properties of each particle have been integrated into the properties of the whole system, which cannot describe the properties of each particle alone, and can only describe the properties of the whole system, this phenomenon is called quantum entanglement (Horodecki et al. 2009). Quantum entanglement is a phenomenon that occurs purely in quantum systems. In classical mechanics, no similar phenomenon can be found. The most common quantum entanglement state are Bell states (Nielsen and Chuang 2002). Bell states are the maximally quantum entangled states of two qubits, an EPR (Einstein-Podolsky-Rosen) pair is one of the four Bell states.

$$\begin{aligned} |\phi^{\pm}\rangle &= \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle) \\ |\psi^{\pm}\rangle &= \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \end{aligned} \tag{3}$$

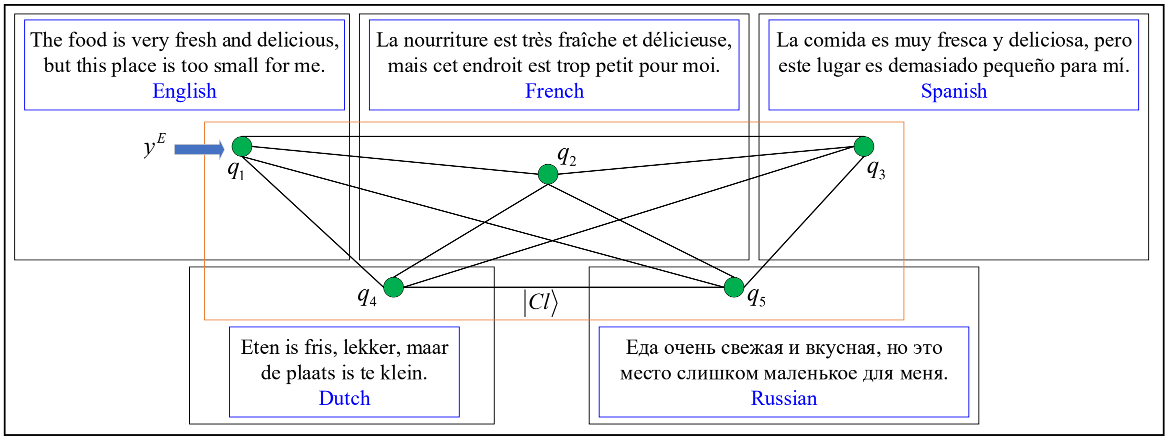

Figure 3: The quantum entangled states for cross-lingual ABSA tasks with five languages, where $|Cl\rangle$ is the quantum entanglement state of five languages and q_i is a single quantum system for a language.

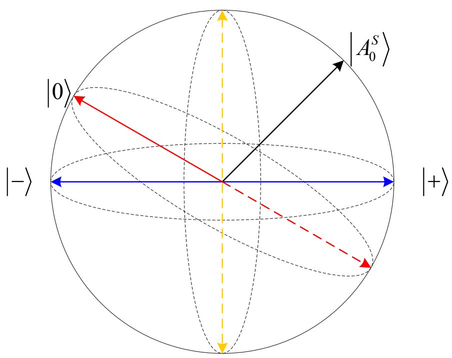

Figure 2: Representations of the aspect in a Hilbert space.

Quantum Projection and Quantum Entanglement Enhanced Network

An overview of QPEN is shown in Figure 1. The cross-lingual ABSA task can be formulated as a sequence labelling problem (Li et al. 2019; He et al. 2019). Given a sentence $x = \{x_i\}_{i=1}^{L}$ with L tokens, the model predicts a label sequence $y = \{y_i\}_{i=1}^{L}$ where $y_i \in Y = \{B, I, E, S\} - \{POS, NEU, NEG\} \bigcup \{O\}$ denotes the aspect boundary and its sentiment polarity for the corresponding token x_i. For example, $y_i = B - POS$ means x_i is the beginning of a positive aspect term. In the cross-lingual transfer setting, we only have the sentence-label pair in the source language S, i.e., $(x^S; y^S) \in D^S$ and aim to predict the label sequence y^T for the sentence x^T in the target language T.

Following method ACS (Zhang et al. 2021), we adopt two cross-lingual pre-trained models, i.e., mBERT (Devlin et al. 2019) and XLM-R (Conneau et al. 2020), as the lexical encoder to compute contextualized representation.

Quantum Projection Module

We formulate every aspect as a mutually incompatible observable on a complex-valued Hilbert space, which is a 3-dimensional vector space $\mathcal{H}_3$ spanned by basis states $\{|+\rangle, |-\rangle, |0\rangle\}$. The basis states $|+\rangle$, $|0\rangle$, and $|-\rangle$ correspond to the positive, the neutral, and the negative sentiments, respectively. We represent an aspect A_0^S as a pure state $|A_0^S\rangle$ on $\mathcal{H}_3$, and the representations of the aspect and the basis states in a Hilbert space are shown in Figure 2. One aspect is represented as a pure state $|A_0^S\rangle$ of positive, neutral, and negative sentiments on a 3-dimensional Hilbert space.

$$|A_0^S\rangle = \alpha|+\rangle + \beta|0\rangle + \gamma|-\rangle, \tag{4}$$

where $|\alpha|^2 + |\beta|^2 + |\gamma|^2 = 1$.

For the ternary sentiment analysis task, each observable is associated with three eigenstates and three eigenvalues, with common eigenvalues of 1, 0, and -1 for the positive, the neutral, and the negative sentiments. Incompatibility falls under different sets of eigenstates $\{|A_0^S, +\rangle, ||A_0^S, 0\rangle, ||A_0^S, -\rangle\}$ defining one aspect basis.

$$\begin{aligned} \langle A_0^S, +|A_0^S, +\rangle &= \langle A_0^S, 0|A_0^S, 0\rangle \\ &= \langle A_0^S, -|A_0^S, -\rangle = 1 \end{aligned} \tag{5}$$

$$\begin{aligned} \langle A_0^S, +|A_0^S, 0\rangle &= \langle A_0^S, +|A_0^S, -\rangle \\ =\langle A_0^S, 0|A_0^S, +\rangle &= \langle A_0^S, 0|A_0^S, -\rangle \\ =\langle A_0^S, -|A_0^S, 0\rangle &= \langle A_0^S, -|A_0^S, +\rangle = 0 \end{aligned} \tag{6}$$

In quantum theory (Nielsen and Chuang 2002), a general observable $\hat{O}$ can be decomposed to its eigenstates $\{|\lambda_i\rangle\}$ of the orthonormal basis as $\hat{O} = \lambda_i|\lambda_i\rangle\langle\lambda_i|$, where eigenvalues $\{\lambda_i\}$ are possible values that a state can take for the corresponding events after quantum measurement. Thus, the aspect observable is

$$\begin{aligned} \hat{A}_0^S = & (+1)|A_0^S, +\rangle\langle A_0^S, +| \\ & + (0)|A_0^S, 0\rangle\langle A_0^S, 0| \\ & + (-1)|A_0^S, -\rangle\langle A_0^S, -| \end{aligned} \tag{7}$$

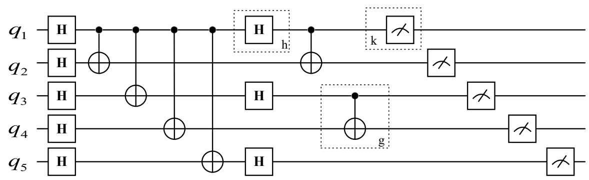

Figure 4: Quantum circuit implementation of the quantum entanglement module, where h, g, k denote the Hadamard gate, the CNOT gate, the measure gate, respectively.

The observable for the final sentiment decision $\hat{S}_0^A$ is

$$\hat{S}_0^A = (+1)|+\rangle\langle+| + (0)|0\rangle\langle0| + (-1)|-\rangle\langle-| \tag{8}$$

Following the projective geometric structure, the measurement probability on an eigenstate equals the projection of the system state onto it. Thus, the sentiment of one aspect is determined by the observable $\hat{S}_0^A$.

Quantum Entanglement Module

Language-specific knowledge is essential for tackling the cross-lingual ABSA task (Zhang et al. 2021). Meanwhile, in QM, when several particles interact with each other, every particle can obtain the information of the whole quantum system by measuring its own quantum state. So we try to create the quantum entanglement between different languages to share the language-specific knowledge, and the quantum entanglement between different languages that we created is shown in Figure 3. Quantum entanglement module can be implemented by combination of CNOT gates and Hadamard gates, and the quantum circuit is shown in Figure 4.

To create the quantum entanglement, we design a quantum entanglement state with five qubits.

$$\begin{aligned} |Cl\rangle &= |00000\rangle + |11111\rangle \\ &= \frac{1}{2}[(|0000\rangle + |1111\rangle) \otimes (|0\rangle + |1\rangle) \\ &\quad + (|0000\rangle - |1111\rangle) \otimes (|0\rangle - |1\rangle)] \\ &= \frac{1}{\sqrt{2}}[(|\phi^+\rangle|\phi^+\rangle + |\phi^-\rangle|\phi^-\rangle) \\ &\qquad \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ &\quad + (|\phi^+\phi^-\rangle + |\phi^-\phi^+\rangle) \\ &\qquad \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)] \\ &= \frac{1}{\sqrt{2}}[(|\phi^+\rangle|\phi^+\rangle + |\phi^-\rangle|\phi^-\rangle) \otimes |+\rangle \\ &\quad + (|\phi^+\phi^-\rangle + |\phi^-\phi^+\rangle) \otimes |-\rangle] \end{aligned} \tag{9}$$

The most common representation of quantum mechanical phenomena are transformation matrices (Miller and Thornton 2006).A qubit can be expressed by a column vector, and the two orthogonal quantum states are $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, respectively. We assign the language-specific knowledge to each particle that each language own in the designed quantum entangled state $|Cl\rangle$, and all particles can obtain the whole language-specific knowledge of other particles in the whole system by measuring their own quantum state, which completes the information sharing of the whole language-specific knowledge.

		EN	FR	SP	DU	RU
Train	No.Sen	2000	1664	2070	1722	3655
	No.Asp	1743	1641	1856	1231	3077
Test	No.Sen	676	668	881	575	1209
	No.Asp	612	650	713	373	949

Table 1: Statistics of the data in each language. No.Sen denotes the number of sentences and No.Asp denotes the number of aspects in each set respectively.

Loss Function

We minimize the following total objective function:

$$L_C = -\sum_{(s,a)\in A_{ll}} \sum_{c\in C} \log p(a) \tag{10}$$

where L_C is a standard cross-entropy loss, A_{ll} contains all sentence-aspect pairs, and C is the collection of distinct sentiment polarities.

Experiments

To evaluate our method QPEN, we transform the quantum state into linear algebraic representation and apply them to the classical cross-lingual ABSA task since the most common representation of quantum mechanical phenomena are transformation matrices (Miller and Thornton 2006). Then we conducted simulation experiments on the SemEval-2016 dataset (Pontiki et al. 2016) with five languages. All the experiments [1] are performed on a workstation with a single machine with 125GB of RAM and 40GB of video memory, two physical CPU with 24 cores Intel(R) Xeon(R) Gold 6248R CPU @ 3.00GHz, and a single GPU (Nvidia A100).

Dataset

We conduct experiments on SemEval-2016 dataset (Pontiki et al. 2016), including real user reviews in English (EN), French (FR), Spanish (SP), Dutch (DU), and Russian (RU) [2]. The data in each language is already split into training and testing sets. We keep the split and further sample 20% data from the training set as the validation set for model selection. Summary data statistics are shown in Table 1, where No.Sen and No.Asp denote the number of sentences and the number of aspects in each set, respectively.

[1] The code and datasets are available at https://github.com/sysulic/QPEN.git

[2] There is one more language data namely Turkish is provided in the SemEval workshop. However, we leave it out in the experiments due to its extremely small testing set (less than 150 sentences)

Methods	mBERT					XLM-R				
	FR	SP	DU	RU	Avg	FR	SP	DU	RU	Avg
ZERO-SHOT	45.60	57.32	42.68	36.01	45.40	56.43	67.10	59.03	56.80	59.84
TRANSLATION-TA	40.76	50.74	47.13	41.67	45.08	47.00	58.10	56.19	50.34	52.91
BILINGUAL-TA	41.00	51.23	49.72	43.67	46.41	49.34	61.87	58.64	52.89	55.69
TRANSLATION-AF	48.03	59.74	49.73	50.17	51.92	57.07	66.61	61.26	59.55	61.12
BILINGUAL-AF	48.05	60.23	49.83	51.24	52.34	57.91	68.04	60.80	60.81	61.89
ACS	49.65	59.99	51.19	52.09	53.23	59.39	67.32	62.83	60.81	62.59
ACS-DISTILL-S	52.23	62.04	52.72	53.00	55.00	61.00	68.93	62.89	60.97	63.45
ACS-DISTILL-M	52.25	62.91	53.40	54.58	55.79	59.90	69.24	63.74	62.02	63.73
QPEN(this work)	**53.27**	**63.84**	**54.61**	**55.36**	**56.97**	**63.21**	**71.59**	**66.16**	**64.52**	**65.79**

Table 2: Experimental results of the cross-lingual ABSA task on SemEval-2016 dataset.

Methods	mBERT					XLM-R				
	FR	SP	DU	RU	Avg	FR	SP	DU	RU	Avg
ACS	49.65	59.99	51.19	52.09	53.23	59.39	67.32	62.83	60.81	62.59
ACS+QP	**50.34**	**61.78**	**52.89**	**53.39**	**54.58**	**61.31**	**69.72**	**64.26**	**62.72**	**64.33**
ACS+QE	**50.61**	**62.07**	**53.08**	**53.47**	**54.83**	**60.56**	**69.05**	**63.97**	**61.83**	**63.72**
QPEN	**53.27**	**63.84**	**54.61**	**55.36**	**56.97**	**63.21**	**71.59**	**66.16**	**64.52**	**65.79**

Table 3: Ablation study results, where QP and QE are the quantum projection and quantum entanglement modules.

Baselines

ZERO-SHOT (Conneau et al. 2020) utilizes the labeled source data to fine-tune the model and directly conduct inference on the target data, which has shown to be a strong baseline for the cross-lingual adaptation (Wu and Dredze 2019). To compare with the previous translation-based method, we adopt the baseline that utilizes the pseudo-labelled data with the Translate-then-Align paradigm (TRANSLATION-TA) (Li et al. 2020) and the combination of the source data with such translated data (BILINGUAL-TA) (Zhang, Zhang, and Fu 2019). ACS is an aspect code-switching mechanism to augment the training data with code-switched bilingual sentences (Zhang et al. 2021).

Implementation Details

We conduct experiments based on two multilingual pre-trained models, the cased multilingual BERT (mBERT) model (Devlin et al. 2019) and the base XLM-Roberta (XLM-R) model (Conneau et al. 2020). Google translate API [3] is used for the translation process. We train the models based on mBERT and XLM-R up to 1500 and 2000 steps respectively and conduct model selection on the last 500 steps. The learning rate is 5e-5 and the range of batch size is $\{8, 12, 16\}$. The best choices are selected by the performance on the source language data. We use a learning rate being 5e-5 and the batch size being 16 for both mBERT model and XLM-R model.We set the epsilon parameter of Adam optimizer to be in the range [2e-5,3e-5].

[3] https://translate.google.com/

Comparison Results

To evaluate QPEN and other typical methods, we use macro-averaged F1-score as the main evaluation metrics, where a prediction will be judged as correct only if both its boundary and sentiment polarity are correct. For all experiments, we report the average F1-scores over 5 runs with different random seeds. The main experimental results are reported in Table 2, where the results of methods ZERO-SHOT, TRANSLATION-TA, and BILINGUAL-TA are from Ref. (Li et al. 2020) while the results of methods TRANSLATION-AF, BILINGUAL-AF, ACS, ACS-DISTILL-S, and ACS-DISTILL-M are from Ref. (Zhang et al. 2021). We found that QPEN improves on both the mBERT and XLM-R models, and the effect is better on model XLM-R than mBERT. Specifically, QPEN outperforms state-of-the-art methods by an average absolute gain of 1.03% in terms of F1-score based on mBERT and 2.31% in terms of F1-score based on XLM-R.Results demonstrate that formulating aspects as quantum superposition states on a complex-valued Hilbert space and sharing language-specific knowledge by quantum entanglement can effectively improve the performance of cross-langual ABSA tasks. The performance difference between QPEN-mBERT (resp. QPEN-XLM-R) and ACS-DISTILL-M-mBERT (resp. ACS-DISTILL-M-XLM-R) is statistically significant with p-value=1.6e-3 (resp. 1.3e-3) < 0.05 by a two-tailed t-test.

Ablation Study

To further investigate the role of quantum projection and quantum entanglement in QPEN, we conduct extensive ablation studies and take method ACS (Zhang et al. 2021) as

Source English Sentence	Target Language	ACS	QPEN
Our server was very helpful and friendly.	Dutch	$(1 \quad \surd)$	$(1 \quad \surd)$
Be prepared to wait, because the place is pretty tiny.	Dutch	$(-1 \quad \surd)$	$(-1 \quad \surd)$
The service was very pleasant and the desert was good.	French	$(1 \quad \surd, \quad 1 \quad \surd)$	$(1 \quad \surd, \quad 1 \quad \surd)$
I choose to go with one of the special, the braised lamb shank in red wine, which was excellent.	French	$(0 \quad \times)$	$(1 \quad \surd)$
I highly recommend Caviar Russe to one who wants delicious top grade caviar and fantastic service.	French	$(0 \quad \times)$	$(1 \quad \surd)$
Slightly above average wines start at $70+ with only one selection listed at $30+.	Dutch	$(1 \quad \times)$	$(1 \quad \times)$

Table 4: Case studies, where "1", "0", and "-1" represent the positive, the neutral, and the negative sentiment, respectively. "$\surd$" denotes the sentiment judgment of aspect is correct while "$\times$" denotes the sentiment judgment is wrong.

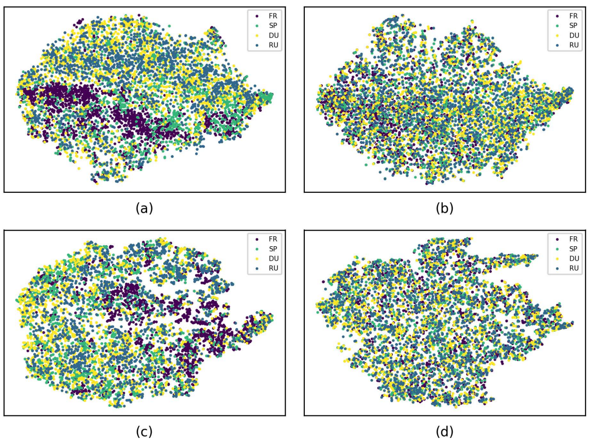

Figure 5: Visual analysis on the test set by adapting t-SNE on the sentence embeddings from QPEN and ACS, where (a) is the sentence embedding plot from ACS with model m-BERT while (b) is the sentence embedding plot from QPEN with model mBERT, and (c) is the sentence embedding plot from ACS with model XLM-R while (d) is the sentence embedding plot from QPEN with model XLM-R, respectively.

the baseline method. The results are reported in Table 3.

Ablation studies show that quantum projection and quantum entanglement improve the performance on models of both mBERT and XLM-R models. Besides, the effect of quantum projection on XLM-R model is better than that on mBERT model, and the effect of quantum entanglement on mBERT model is better than that on XLM-R model.

Case Study and Error Study

Table 4 shows several cases from ACS and QPEN, where "1", "0", and "-1" denote the positive, the neutral, and the negative sentiments, respectively, and "$\surd$" and "$\times$" denote the sentiment judgments of aspect are correct and wrong, respectively. We highlight the aspect words in blue.

Taking the third sample as an example, QPEN is also able to accurately identify the sentiment of each aspect even when there are many aspects in a sentence. Take the fourth sample as an example, the aspect "braised lamb shank in red wine" is a long aspect term, and "shank" is missed with method ACS, so ACS fails. However, QPEN succeeded by employing quantum entanglement to identify the aspect and its sentiment. Besides, in the fifth sample, the aspect has two sub-aspects "caviar" and "service", so ACS fails since it based on the label projection. But QPEN succeeded based on the quantum entanglement and the quantum projection.

Take the sixth sample as an example, both the ACS and QPEN fail to correctly predict the sentence "Slightly above average wines start at $70+ with only one selection listed at $30+.". Due to the difficulty to identify the sentiment of word "wines" in the sentence, and "$70+" and "$30+" are just price comparisons without any sentiment words.

Visual Analysis

We conduct visual analysis on the test set by adapting t-SNE on the sentence embeddings from QPEN and the baseline ACS (Zhang et al. 2021), as shown in Figure 4. In which, (a) is the sentence embedding plot from ACS with the model m-BERT, (b) is the sentence embedding plot from QPEN with the model mBERT, (c) is the sentence embedding plot from ACS with the model XLM-R, and (d) is the sentence embedding plot from QPEN with the model XLM-R, respectively.

Compared with the the baseline method ACS, method QPEN yields more hybridized representation between different languages. This implies that the propose quantum modules encourage to share language-specific knowledge between different languages, resulting in more accurate predictions for the cross-lingual ABSA task. We leave the optimization of our model in the future work with the evolutionary algorithm (Chaturvedi, Su, and Welsch 2021).

Conclusion

In this paper, we have proposed a quantum projection and quantum entanglement enhanced network for cross-lingual aspect-based sentiment analysis tasks, named QPEN. It is equipped with a proposed quantum projection module to project aspects as a quantum superposition state on a complex-valued Hilbert space, and a proposed quantum entanglement module to enforce the model sharing the language-specific between different languages. Extensive experiments on SemEval-2016 dataset show that QPEN is superior to the baseline approaches. Our ablation study, case study and visual analysis further confirm the effectiveness of key components in the proposed QPEN model.

Acknowledgments

This paper was supported by National Natural Science Foundation of China (No. 62276284, 61976232, 51978675), National Key Research and Development Program of China (No. 2021YFA1000504), Guangdong Basic and Applied Basic Research Foundation (No. 2023A1515011470, 2022A1515011355), Guangzhou Science and Technology Project (No. 202201011699), Guizhou Provincial Science and Technology Projects (No. 2022-259), Humanities and Social Science Research Project of Ministry of Education (No. 18YJCZH006), as well as the Fundamental Research Funds for the Central Universities, Sun Yat-sen University (No. 23ptpy31).

References

Akhtar, M. S.; Sawant, P.; Sen, S.; Ekbal, A.; and Bhattacharyya, P. 2018. Solving data sparsity for aspect based sentiment analysis using cross-linguality and multilinguality. Association for Computational Linguistics.

Barnes, J.; Lambert, P.; and Badia, T. 2016. Exploring Distributional Representations and Machine Translation for Aspect-based Cross-lingual Sentiment Classification. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, 1613–1623.

Busemeyer, J. R.; and Bruza, P. D. 2012. *Quantum models of cognition and decision*. Cambridge University Press.

Chaturvedi, I.; Su, C. L.; and Welsch, R. E. 2021. Fuzzy aggregated topology evolution for cognitive multi-tasks. *Cognitive Computation*, 13: 96–107.

Conneau, A.; Khandelwal, K.; Goyal, N.; Chaudhary, V.; Wenzek, G.; Guzmán, F.; Grave, E.; Ott, M.; Zettlemoyer, L.; and Stoyanov, V. 2020. Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 8440–8451.

Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 4171–4186.

Dyer, C.; Chahuneau, V.; and Smith, N. A. 2013. A simple, fast, and effective reparameterization of ibm model 2. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 644–648.

Fei, H.; Zhang, M.; and Ji, D. 2020. Cross-Lingual Semantic Role Labeling with High-Quality Translated Training Corpus. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 7014–7026.

Fell, L.; Dehdashti, S.; Bruza, P.; and Moreira, C. 2019. An experimental protocol to derive and validate a quantum model of decision-making. *arXiv preprint arXiv:1908.07935*.

Gkoumas, D.; Li, Q.; Dehdashti, S.; Melucci, M.; Yu, Y.; and Song, D. 2021a. Quantum cognitively motivated decision fusion for video sentiment analysis. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 827–835.

Gkoumas, D.; Li, Q.; Yu, Y.; and Song, D. 2021b. An entanglement-driven fusion neural network for video sentiment analysis. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, 1736–1742. International Joint Conferences on Artificial Intelligence Organization.

Gong, L.-H.; Pei, J.-J.; Zhang, T.-F.; and Zhou, N.-R. 2024. Quantum convolutional neural network based on variational quantum circuits. *Optics Communications*, 550: 129993.

Gudder, S. P. 2014. *Quantum probability*. Academic Press.

Halmos, P. R. 2017. *Finite-dimensional vector spaces*. Courier Dover Publications.

He, R.; Lee, W. S.; Ng, H. T.; and Dahlmeier, D. 2019. An Interactive Multi-Task Learning Network for End-to-End Aspect-Based Sentiment Analysis. 504–515.

Horodecki, R.; Horodecki, P.; Horodecki, M.; and Horodecki, K. 2009. Quantum entanglement. *Reviews of modern physics*, 81(2): 865.

Hu, J.; Ruder, S.; Siddhant, A.; Neubig, G.; Firat, O.; and Johnson, M. 2020. Xtreme: A massively multilingual multitask benchmark for evaluating cross-lingual generalisation. In *International Conference on Machine Learning*, 4411–4421. PMLR.

Hu, M.; Peng, Y.; Huang, Z.; Li, D.; and Lv, Y. 2019. Open-Domain Targeted Sentiment Analysis via Span-Based Extraction and Classification. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 537–546.

Hughes, R. I. 1989. *The structure and interpretation of quantum mechanics*. Harvard university press.

Jebbara, S.; and Cimiano, P. 2019. Zero-shot cross-lingual opinion target extraction. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2486–2495.

Karthikeyan, K.; Wang, Z.; Mayhew, S.; and Roth, D. 2020. Cross-Lingual Ability of Multilingual BERT: An Empirical Study. In *International Conference on Learning Representations*.

Klinger, R.; and Cimiano, P. 2015. Instance selection improves cross-lingual model training for fine-grained sentiment analysis. In *Proceedings of the Nineteenth Conference on Computational Natural Language Learning*, 153–163.

Lambert, P. 2015. Aspect-level cross-lingual sentiment classification with constrained SMT. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, 781–787.

Li, Q.; Gkoumas, D.; Sordoni, A.; Nie, J.-Y.; and Melucci, M. 2021. Quantum-inspired neural network for conversational emotion recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 13270–13278.

Li, X.; Bing, L.; Zhang, W.; and Lam, W. 2019. Exploiting BERT for End-to-End Aspect-based Sentiment Analysis. 34–41.

Li, X.; Bing, L.; Zhang, W.; Li, Z.; and Lam, W. 2020. Unsupervised cross-lingual adaptation for sequence tagging and beyond. *arXiv preprint arXiv:2010.12405*.

Lin, Z.; Jin, X.; Xu, X.; Wang, W.; Cheng, X.; and Wang, Y. 2014. A cross-lingual joint aspect/sentiment model for sentiment analysis. In *Proceedings of the 23rd ACM international conference on conference on information and knowledge management*, 1089–1098.

Liu, B. 2012. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1): 1–167.

Liu, Y.; Zhang, Y.; Li, Q.; Wang, B.; and Song, D. 2021. What does your smile mean? Jointly detecting multi-modal sarcasm and sentiment using quantum probability. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, 871–880.

Martin, J. H. 2009. *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Pearson/Prentice Hall.

Merzbacher, E. 1998. *Quantum mechanics*. John Wiley & Sons.

Miller, D. M.; and Thornton, M. A. 2006. QMDD: A decision diagram structure for reversible and quantum circuits. In *Proceedings of the 36th International Symposium on Multiple-Valued Logic*, 30–30. IEEE.

Nielsen, M. A.; and Chuang, I. 2002. Quantum computation and quantum information.

Pfeiffer, J.; Vulić, I.; Gurevych, I.; and Ruder, S. 2020. MAD-X: An Adapter-Based Framework for Multi-Task Cross-Lingual Transfer. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 7654–7673.

Pires, T.; Schlinger, E.; and Garrette, D. 2019. How Multilingual is Multilingual BERT? In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 4996–5001.

Pontiki, M.; Galanis, D.; Papageorgiou, H.; Androutsopoulos, I.; Manandhar, S.; Al-Smadi, M.; Al-Ayyoub, M.; Zhao, Y.; Qin, B.; De Clercq, O.; et al. 2016. Semeval-2016 task 5: Aspect based sentiment analysis. In *International workshop on semantic evaluation*, 19–30.

Ruder, S.; Vulić, I.; and Søgaard, A. 2019. A survey of cross-lingual word embedding models. *Journal of Artificial Intelligence Research*, 65: 569–631.

Singh, J.; McCann, B.; Keskar, N. S.; Xiong, C.; and Socher, R. 2020. XLDA: Cross-Lingual Data Augmentation for Natural Language Inference and Question Answering.

Wang, W.; and Pan, S. J. 2018. Transition-based adversarial network for cross-lingual aspect extraction. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, 4475–4481.

Wu, S.; and Dredze, M. 2019. Beto, Bentz, Becas: The Surprising Cross-Lingual Effectiveness of BERT. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 833–844.

Yan, G.; Wu, H.; and Yan, J. 2023. Quantum 3D Graph Learning with Applications to Molecule Embedding.

Ye, X.; Yan, G.; and Yan, J. 2023. Towards Quantum Machine Learning for Constrained Combinatorial Optimization: a Quantum QAP Solver.

Zhang, M.; Zhang, Y.; and Fu, G. 2019. Cross-Lingual Dependency Parsing Using Code-Mixed TreeBank. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 997–1006.

Zhang, W.; He, R.; Peng, H.; Bing, L.; and Lam, W. 2021. Cross-lingual aspect-based sentiment analysis with aspect term code-switching. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 9220–9230.

Zhang, Y.; Song, D.; Li, X.; Zhang, P.; Wang, P.; Rong, L.; Yu, G.; and Wang, B. 2020. A quantum-like multimodal network framework for modeling interaction dynamics in multiparty conversational sentiment analysis. *Information Fusion*, 62: 14–31.

Zhao, Q.; Hou, C.; and Xu, R. 2022. Quantum-Inspired Complex-Valued Language Models for Aspect-Based Sentiment Classification. *Entropy*, 24(5): 621.

Zhou, N.-R.; Zhang, T.-F.; Xie, X.-W.; and Wu, J.-Y. 2023. Hybrid quantum–classical generative adversarial networks for image generation via learning discrete distribution. *Signal Processing: Image Communication*, 110: 116891.

Zhou, X.; Wan, X.; and Xiao, J. 2015. CLOpinionMiner: Opinion target extraction in a cross-language scenario. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(4): 619–630.