Using the Quantum Probability Ranking Principle to Rank Interdependent Documents

Guido Zuccon, Leif Azzopardi
{guido, leif}@dcs.gla.ac.uk

University of Glasgow, Department of Computing Science,
Glasgow G12 8QQ, UK

Abstract A known limitation of the Probability Ranking Principle (PRP) is that it does not cater for dependence between documents. Recently, the Quantum Probability Ranking Principle (QPRP) has been proposed, which implicitly captures dependencies between documents through "quantum interference". This paper explores whether this new ranking principle leads to improved performance for subtopic retrieval, where novelty and diversity is required. In a thorough empirical investigation, models based on the PRP, as well as other recently proposed ranking strategies for subtopic retrieval (i.e. Maximal Marginal Relevance (MMR) and Portfolio Theory(PT)), are compared against the QPRP. On the given task, it is shown that the QPRP outperforms these other ranking strategies. And unlike MMR and PT, one of the main advantages of the QPRP is that no parameter estimation/tuning is required; making the QPRP both simple and effective. This research demonstrates that the application of quantum theory to problems within information retrieval can lead to significant improvements.

1 Introduction

The Probability Ranking Principle (PRP) is the most widely used and accepted ranking criteria for the retrieval of documents [1]. Although this ranking principle has been shown to be generally applicable, it makes some assumptions which do not always hold [2,3]. When results need to be relevant but also diverse, as is the case in subtopic retrieval, the independence assumption made by the PRP is unrealistic. This is because the PRP neglects relationships between documents at relevance level ignoring the fact that a previous document may already contain similar relevant material [4,5,6,7].

To address this problem, different attempts have been made to formulate a better ranking principle by accounting for the similarity between documents in the ranking process. One such approach is called Maximal Marginal Relevance (MMR) [8]. This was used within a risk minimisation framework to compensate for the PRP's limitations [4]. While recently, Wang and Zhu [5] investigated the PRP's assumption with respect to the certainty in the estimation of a document's probability of relevance. These approaches have been motivated from empirical observations or heuristically adapted to the specific retrieval task, and require a significant amount of parameter tuning to be effective. However, a new ranking principle has been proposed for coping with interdependent document relevance; the Quantum Probability Ranking Principle (QPRP) [9].

The principle is derived from Quantum Probability Theory [10] where interdependent document relevance is captured by "quantum interference". While the theoretical foundations for the principle have been outlined in [9], it has not been empirically tested or validated. Thus, the aim of this paper is to conduct an empirical study comparing models based on the QPRP against models based on the PRP as well as against state-of-the-art non-PRP based models such as MMR [8] and PT [5]. On the subtopic retrieval task of the Interactive TREC Track, the results of this study show that the QPRP consistently outperforms these other models/principles. A distinct advantage that the QPRP has over the other more sophisticated approaches, is that no explicit parameter tuning is required. This work shows for the first time that Quantum Theory can be successfully applied within information retrieval.

The paper continues as follows: Section 2 provides an overview of the subtopic task, and explains how the assumptions of the PRP are too restrictive in this context, before describing the main approaches which aim to account for interdependence between documents. Then, Section 3 presents the QPRP and how it can be applied in order to account for document dependence through "quantum interference". To show the QPRP in action, Section 4 presents the extensive empirical study performed on the subtopic retrieval task, which explicitly focuses on novelty and diversity ranking. Finally, the paper concludes in Section 6, summarising the contribution of this work along with directions of future research.

2 Background & Related Works

The subtopic retrieval task stems from the need of providing a document ranking which covers all the possible different facets (subtopics) relevant to the user's information need. Thus, an IR system aims to maximise the user's satisfaction by retrieving documents which cover all the relevant subtopics in the ranking. Given a test collection, where the subtopics of the relevant documents have been identified, the effectiveness of the retrieval system can be measured in several ways [4,6,7,11]. The three main measures employed are S-recall, S-Mean Reciprocal Rank and S-precision. Subtopics coverage is measured by s-recall at rank k [4], formally defined as

$$s - recall(k) = \frac{| \cup_{i=1}^{k} subtopics(d_i)|}{n_s} \tag{1}$$

where $subtopics(d_i)$ returns the set of subtopics relevant to the topic that are contained in document d_i, and n_s is the number of possible subtopics relevant to the general topic. Intuitively, the fewer documents that have to be examined in order to retrieve all subtopics, the more effective the system. This intuition is also captured by the S-MRR measure, which is defined as the inverse of the rank at which full subtopic coverage is achieved.

For a precision oriented measure, S-precision at r is calculated by taking the ratio between the minimum rank that an optimal system $\mathcal{S}_{opt}$ achieves a S-recall value of r over the corresponding minimum rank the system $\mathcal{S}$ achieves the same S-recall value r [4]. The optimal system reaches recall value of r at the smallest rank k. Formally,

$$s - precision(r) = \frac{minRank(\mathcal{S}_{opt}, r)}{minRank(\mathcal{S}, r)} \tag{2}$$

From the evaluation measures used for subtopic retrieval, it is clear that the intrinsic dependencies between documents need to be considered by the system when ranking. Consequently, models which adhere to the PRP have been shown to result in sub-optimal performance [12].

2.1 The Probability Ranking Principle and its Limitations

The Probability Ranking Principle was first detailed by Robertson [1] in 1977, and its origins stem from initial work performed by Cooper [13] in 1971. Since its inception the PRP has played a vital part in shaping the development of IR models, methods and systems. The intuition underlying the PRP is as follows: in order to obtain the best overall effectiveness, an IR system should rank documents in descending order of their probability of relevance to the user's information need. It has been shown that adhering to this principle guarantees an optimal ranking [1]. In [14], Gordon also shows that the PRP maximises a suitably defined utility function.

The PRP assumes a probability distribution over the documents, which represents the space of events. In [14], the probabilities associated to documents represent the chances that a user is satisfied by observing those documents in response to his/her information need. Such probabilities are approximated by the probability of a document being relevant to the query; and, in practice, they are estimated from statistics extracted from the document and the collection. However, a key assumption made by the PRP is that the probability of relevance of a document is independent from the relevance of other documents (i.e. independence assumption). Consequently, the PRP dictates that at each rank the IR system should select document d such that:

$$d = \underset{d_i \in \mathcal{RE} \setminus RA}{\arg\max} \; P(d_i) \tag{3}$$

with RA being the list of documents that have been ranked, d_i a document belonging to the set of retrieved documents ($\mathcal{RE}$) for a query but not ranked yet, and $P(.)$ being the probability of a document being relevant to the information need. However, when the independence assumption is not upheld, the PRP provides a suboptimal ranking [12].

In the case of subtopic retrieval, there is the explicit requirement of preferring relevant and novel information over redundant. Thus, the documents previously retrieved will influence what documents should be retrieved next. Under the

PRP, however, if two documents have a high probability of relevance, but cover the same topics, they will both ranked in high positions. This has to be avoided in subtopic retrieval, since the two documents have the same relevant content and thus no novel information is conveyed to the user if both are retrieved at high ranks.

2.2 Beyond PRP: Attempts to include document dependence

In the last decade, several attempts have been made to either model or include interdependent document relevance in the ranking process, in particular to cope with interactive information retrieval and subtopic retrieval. For example, the PRP is extended to interactive IR and framed within a situation-based framework in [15]. Under the **Interactive PRP**, users move between situations and the independent relevance assumption is substituted by a weaker condition within each situation. The ranking principle is then derived by the optimum ordering of the choices presented in each situation. As we shall see the QPRP differs from this approach because the quantum probability framework naturally encodes dependent relevance in the interference term[1]; and so, it considers dependence at the document level not at the situation level.

Maximal Marginal Relevance & Risk Minimization: The problem of a document ranking exploiting diversity amongst documents has been heuristically tackled in [8], where a technique called Maximal Marginal Relevance (MMR) was proposed. Document ranking is obtained by balancing the score of similarity between document and query, e.g. the probability of relevance, and a diversity score between the candidate document and all the documents ranked at earlier positions. A successful framework for coping with subtopics retrieval is based on risk minimization [4], where documents are ranked in increasing value of expected risk. In particular, language models are employed to represent documents and queries, while a loss function is used to model users preferences. The preference for retrieving documents that are both relevant and novel is encoded in the MMR function. Documents are selected following the objective function:

$$d = \arg\max \big(value_R(\theta_i; \theta_Q)(1 - c - value_N(\theta_I; \theta_1, \ldots, \theta_{i-1}))\big) \tag{4}$$

where $value_R(\theta_i; \theta_Q)$ is the query likelihood estimated using language models, c represents the relative cost of seeing a non -relevant document compared with seeing a relevant but non-novel document, and $value_N(\theta_I; \theta_1, \ldots, \theta_{i-1})$ is the estimated novelty coefficient.

Portfolio Theory: Risk is also combined with the document relevance estimation in the Portfolio Theory (PT) approach recently proposed in [5]. The intuition behind the PT model for IR is that a measure of uncertainty (variance) is associated to each estimation of document relevance; when ranking documents the IR system should maximize relevance in the ranking while minimizing variance. The ranking criterion proposed by the Portfolio paradigm differs from the

[1] Interference is discussed in Section 3

PRP because the rank of a document is not just function of the estimated probability of relevance of the document itself. Instead, the document's probability of relevance is combined with an additive term which synthesises the risk inclination of the user, and the uncertainty (variance) associated with the probability estimation, along with the correlation between the candidate document and documents ranked previously. In particular, documents are ranked according to:

$$d = \arg\max \big(P(d_i) - bw_{d_i}\delta^2_{d_i} - 2b \sum_{d_k \in RA} w_{d_k}\delta_{d_i}\delta_{d_k}\rho_{d_i,d_k}\big) \qquad (5)$$

where b encodes the risk propensity of the user, RA is the list of documents already ranked, $\delta^2_{d_i}$ is the variance associated to the probability estimation of document d_i, w_{d_i} is a weight inversely proportional to the rank position which express the importance of the rank position itself, and ρ_{d_i,d_k} is the correlation between document d_i and document d_k. Intuitively, the PT's ranking function is affected by the probability of relevance, the variance associated to the probability estimation and the correlation between candidate documents and documents already ranked.

In summary, the ranking functions suggested by the approaches for subtopic retrieval considered in this paper, i.e. MMR and PT, have two components. The first is the probability of relevance of a document with respect to user's information need, and it is in common with the PRP approach. However, the role of the second component in the ranking functions is to encode the degree of novelty/diversity of the candidate document with respect to the ones already ranked. In the QPRP, instead, although apparently it reflects the same schema as MMR and PT, relevance and novelty/diversity estimation are mixed together in the interference term. Moreover, MMR and PT have been inspired by empirical observations and require significant effort in parameter estimation. Conversely, the QPRP is derived from Quantum Probability Theory and does not contain any parameters which need explicit tuning.

3 The Quantum Probability Ranking Principle

Using Quantum Theory within IR was originally proposed by van Rijsbergen [16], and has been subsequently developed in a number of ways [17,18,19,20,21]. Here we consider the Quantum Probability Ranking Principle, that has been recently proposed by Zuccon et al. [9]. The QPRP is derived through the application of quantum probability to the problem of document ranking[2]. The resultant of this work was the following formulation: when ranking documents, the IR system has to maximise the total satisfaction of the user given the document ranking, achievable by maximising the total probability of the ranking. Using the quantum law of total probability, the resultant ranking strategy impose to select at each rank position a document d such that:

$$d = \arg\max \big(P(d_i) + \sum_{d_x \in RA} I_{d_x,d_i}\big) \qquad (6)$$

[2] For a full derivation of the principle, we refer the reader to [9].

where RA is the list of documents already ranked and I_{d_x,d_i} is the "quantum interference" between documents d_x and d_i. The pseudo-code of the QPRP algorithm is sketched in Algorithm 1. The intuition underlying this paradigm is that documents in a ranking share relationships at relevance level, i.e. they interfere with each other. For example, [3,4] showed that the user is more likely to be satisfied by documents addressing his information need in different aspects than documents with the same content. Then, it might be sensible to model documents expressing diverse information as having higher degree of interference than documents that are similar. For the same reason, documents containing novel information might highly interfere with documents ranked in previous positions. Even contrary information might be captured by the interference term: documents containing content contrary to the one presented at the previous rank position might trigger a revision of user's beliefs about the topic.

In summary, interference appears to capture the dependencies in documents' relevance judgements. The QPRP suggests that documents ranked until position $n-1$ interfere with the degree of relevance of the document ranked at position n. However, while the QPRP has been proposed, no experimental work has been performed which validates whether the Quantum based principle provides a better ranking or not. It is the aim of this paper to empirically explore the QPRP in the context of subtopic retrieval. In the following we detail how the QPRP formally differs from the PRP and we provide an outline of the estimation of quantum interference for subtopic retrieval.

Algorithm 1 The ranking strategy of QPRP

$\mathcal{RE} = \{\text{retrieved documents}\}$
$RA[k] = 0, \forall k : 1 \leq k \leq |\mathcal{RE}|$
comment: vector RA will contain the document ranking
$p = 1$
while $\mathcal{RE} \neq \varnothing$ **do**
  $d = \arg\max \left(P(d_i) + \sum_{d_x \in \mathcal{RA}} I_{d_x,d_i}\right)$, with $d_i \in \mathcal{RE}$
  $\mathcal{RE} = \mathcal{RE} \setminus \{d\}$
  $RA[p] = d$
  $p = p + 1$
end while
return RA

3.1 Estimating Probabilities and Interference

Ranking according to the QPRP is quite simple. Firstly, the QPRP uses the same probability estimates as the PRP (i.e the probability of relevance of a document $P(d_i)$). Next, is the interference component, which we shall explain in detail.

In a Quantum Probability Theory the law of total probability is different from standard probability theory (i.e. Kolmogorovian). In particular, the interference component might manifest. While in the PRP the law of total probability, i.e. the total probability associated with the ranking composed by document A and document B, is treated as the sum of the probabilities of the single, independent events (assumption of independent document relevance), the situation in the Quantum framework is different.

The QPRP assumes that underlying the relevance probability distribution there is a primitive concept of a complex amplitude distribution. This assumption follows from Quantum Probability Theory, and is the key point which differentiates the QPRP approach from the traditional PRP. What it means is that no assumption of independence between document relevance is made. The intuition is that a complex number $\phi_i = x_i + jy_i$ (called the amplitude) is associated to each event (i.e. document). Complex probability amplitudes ϕ_i and real[3] probabilities p_i are linked by the relationship $p_i = |\phi_i|^2 = (\sqrt{x_i^2 + y_i^2})^2$ where $|.|$ is the modulus of a complex number.

In this case, the total amplitude of a set of events relates to the sum of amplitudes associated to such events, similarly to what happens in the PRP with probabilities. When deriving the total probability of these events from the amplitudes, the amplitudes themselves are first summed and then their modulus is calculated, leading mathematically to the presence of an additional component, the interference, other than the square of the modulus of each amplitude. This is intuitive if we consider the polar form a complex number, $\phi_i = |\phi_i|(cos\psi_i + jsin\psi_i)$, where $j = \sqrt{-1}$ is the imaginary unit. Then,

$$\begin{aligned} p_{AB} = |\phi_{AB}|^2 = |\phi_A + \phi_B|^2 = \overline{(\phi_A + \phi_B)}(\phi_A + \phi_B) = \\ |\phi_A|^2 + |\phi_B|^2 + \overline{\phi_A}\phi_B + \phi_A\overline{\phi_B} = \\ |\phi_A|^2 + |\phi_B|^2 + 2\,|\phi_A|\,|\phi_B| \cos(\psi_A - \psi_B) = \\ p_A + p_B + 2\sqrt{p_A}\sqrt{p_B}\cos\theta_{AB}. \end{aligned} \tag{7}$$

where $\overline{\phi_A}$ indicates the complex conjugate of ϕ_A and θ_{AB} is the difference between the phases ψ_A and ψ_B. The interference is a real number, since both modulus and cosine are real-valued functions. However, the interference might not be zero: this is the case when the amplitudes do not have orthogonal phases. The difference between PRP and QPRP then resides in the interference component. If the difference of amplitudes's phases is an odd multiple of $\frac{\pi}{2}$, then the probability of joint events is simply the sum of the probabilities of the events, i.e. the sum of the squared amplitudes. When this is not true, the interference term is different from zero. In which case, the total probability obtained using Kolmogorovian probability theory assuming independence between events differs from the total probability employing Quantum Probability Theory. The difference is given by the additive interference term. The interference term might assume a positive value (i.e. constructive interference) or a negative one (i.e. destructive interference). In fact the interference is a function of the modulus of the amplitudes, which is always a positive real number, and of the amplitude's phase. In particular, the interference depends upon the cosine of the amplitude's phase difference.

In the QPRP, the total probability at each cutoff of the ranking is a function of the probabilities associated to the single documents and the interference

[3] meaning belonging to the field of the real numbers.

between each pair of documents contained in the ranking. The maximization of the total probability depends upon the document's probabilities and their interference (see eq. 6). It has been shown that finding the solution of this maximisation problem can be reduced to a minimum set covering problem, which is NP-hard [4]. However, by using a greedy algorithm the complexity is significantly reduced, as it is the case for the MMR and PT approaches.

3.1.1 The QPRP in Action: Interference in Subtopic Retrieval While the QPRP has been proposed, no concrete instantiation has been developed and tested. The main issue is the estimation of the interference term or, equivalently, the estimation of the amplitudes's phase. In this subsection, we outline how to estimate the interference for the task of subtopic retrieval, and then evaluate the instantiation in the remainder of the paper.

The main idea is to capture document interdependence through the interference component. In particular, relationships can be encoded in the phase, while the square roots of the estimated probabilities of relevance act as modulation component. The presence of relevance probabilities guarantees that the interference for documents that are diverse from the ones previously ranked but not (estimated) relevant, i.e. their probability is ~ 0, results null. Vice versa, if the documents involved in the interference have high probability of being relevant, then their contribution will be high. However, the interference also depends, both in sign and modulo, from the cosine of the phase difference between the amplitudes. Being able to derive this component directly from the amplitude distribution would mean being able to generate a complex amplitude distribution from real text statistics: the feasibility of this idea is still under investigation. However, we can try to estimate the phase difference between the amplitudes associated with documents. The estimation of this component depends upon the particular retrieval task. For example, in the subtopic retrieval task constructive interference (positive) might be used to model the interaction between documents covering different facets of the topic, while vice versa destructive interference might occur between documents covering the same subtopics. The converse situation, i.e. constructive interference to model coverage of the same topic while destructive models topical dissimilarity, seems feasible for encoding interdependent document relevance in ad-hoc retrieval task, acting similarly to an iterative implicit feedback mechanism: this is not covered in the present work, but will be topic of further investigations. In the subtopic retrieval scenario, we assume that redundant relevant documents destructively interfere, while documents conveying relevant but novel information generate constructive interference. The implementation details for the estimation of interference are discussed in Section 4.cterized by the correspondent Okapi weight.

4 Empirical Study

The aim of this empirical investigation is two fold:

1. test wether accounting for interdependent document relevance delivers a better document ranking for the subtopic retrieval task (i.e. PRP versus non-PRP), and,

2. compare the ranking strategy based on the QPRP against the classical (MMR) and state of the art (PT) techniques for subtopic retrieval.

To this aim, we conducted the following empirical investigation using the TREC subtopic retrieval track. This uses the documents from the Financial Times of London contained in TREC 6,7 and 8 collections and 20 ad-hoc retrieval topics from the TREC interactive tracks which are composed of subtopics, sometimes referred to as aspects. The collection was indexed using Lemur[4] where standard stop words were removed and Porter stemming applied. For each of the TREC Topics, we used the title of the topic to generate queries. We applied stopping and stemming to both kinds of queries.

The baseline method for the experiments was BM25 as it upholds the Probability Ranking Principle. The more competitive baselines employed were MMR and PT, and these were compared against the QPRP (see below) . For MMR, PT and QPRP methods, the top n documents retrieved by the BM25 baselines were re-ranked accordingly, where we tried n of 100, 200 and 1000. The normalised BM25 score was then used by these methods as the probability of relevance. While other ranking function may have been used, BM25 is a robust baseline previously used and delivers similar performance to Language Models [5]. For each of these methods, a kernel is required to compute the degree of dissimilarity/interdependence between documents. In the experiments reported here we used Pearson's correlation between the weighted term vectors associated to the documents as the kernel. The weighting schema was BM25. This kernel was previously used in [5]. While there are other choices of kernels, like the cosine similarity measures, in a set of preliminary experiments we found that the Pearson's correlation achieved the best results, across all strategies.

Maximal Marginal Relevance: We tested the MMR approach varying the value of the hyper-parameter c in the range $[0, 1]$ by decimal steps. For c equal one, the MMR strategies delivers the same rank as the PRP approach, while for a value of c equal zero the relevance score is discarded in favour of the dissimilarity score. The best results found are reported in the following section.

Portfolio Theory: To compute PT we need the variance associated to the probability estimated provided by the normalised BM25 scores. The variance (indicated with δ^2) becomes an adjunctive parameter of the PT ranking strategy. In [5], they suggest using a constant variance. We investigated the optimal value of the variance in combination with the value of the parameter b that encodes the risk propensity of a user. We considered values of b in the range $[1, 10]$ with unitary increments and values of δ^2 in the range $[10^{-10}, 10^{-1}]$. Here, we report the best results obtained by the possible combinations of parameters given the grid search of b by δ^2. Finally, the correlation ρ between pairs of documents is computed employing Pearson's correlation as described above, while the weight w associated to each rank position r is given by $\frac{1}{log_2 r}$, as in [5].

Quantum Probability Ranking Principle: The implementation of the QPRP ranking strategy does not require any parameter setting/tuning procedure. Without a method to estimate the complex probability amplitudes, we

[4] http://www.lemurproject.org

resort to an approximation of the phase by using Pearson's correlation. By using Pearson's correlation, it also enables us to fairly compare PT against QPRP. Interference between document d_i and d_x is then approximated using $-\sqrt{P(d_i)}\sqrt{P(d_x)}\rho_{d_i,d_x}$, where $P(.)$ is the estimation of the probability of relevance of a document and ρ is the Pearson's correlation between the two documents' term vectors.

5 Results

In Tables 1, 2 and 3 we report the results obtained by the different ranking strategies. In each table, the best results for MMR and PT are reported with respect to S-r@5 for Table 1, S-p@0.1 for Table 2 and S-MRR@100% for Table 3; and the the best performance overall for each measure is highlighted in bold. Note that we only report some of the results for $n = 100$ and $n = 200$ documents, but similar trends were witnessed at different levels on n. From these tables of results, the following points are of interest:

1. first note that in Tables 1, 2 and 3 the results for PT and MMR are generally higher than the equivalent listed in [5], both in absolute terms and percentage increase over the PRP. This is mainly due to the fact that we optimised each method specifically, making them very strong and competitive baselines. But in practice the performance of these methods is likely to be slightly lower;
2. MMR and PT improves upon the PRP at early levels of S-precision and S-recall, but fail to consistently outperform the PRP across all levels;
3. the QPRP improves upon PRP baselines for all levels of S-precision and S-recall. Furthermore, the QPRP outperforms MMR and PT across most levels;
4. the QPRP consistently outperforms other strategies across all topics when considering S-MRR@100%. This means on each topic the QPRP returns complete coverage of all subtopics at a rank lower than all the other strategies;
5. since the topics set is small, performing significance tests would not be appropriate [22, pages 178–180]. However, the QPRP delivers consistently better performance over the PRP, and also outperforms that state of the art methods. Also, as no parameter tuning is required for the QPRP, it represents a very attractive alternative to PT and MMR.

6 Conclusions & Future Works

In this paper we have explored how the QPRP can be applied in the setting of subtopic retrieval; which specifically requires models to account for interdependent document relevance. The QPRP naturally encodes the interdependence through quantum interference. The new ranking strategy has been empirically compared against the PRP and state-of-the-art ranking approaches. We have

Models	***S-r@5***	**S-r@10**	**S-r@20**	**S-r@50**	**S-MRR@100**	**S-p@.1**	**S-p@.2**	**S-p@.5**
PRP	0.2466	0.3900	0.4962	0.6034	0.0086	0.3968	0.3062	0.1941
MMR	0.2697 (+8.56%)	0.3540 (-10.19%)	0.4795 (-3.49%)	0.6032 (-0.02%)	0.097 (+11.39%)	0.4203 (+5.59%)	0.2876 (-6.48%)	0.1964 (+1.17%)
PT	0.2791 (+11.63%)	0.3654 (-6.75%)	0.4444 (-11.66%)	0.5494 (-9.82%)	0.0130 (+33.47%)	**0.4587** (+13.48%)	0.2915 (-5.05%)	0.1769 (-9.73%)
QPRP	**0.3093** (+20.25%)	**0.4063** (+3.99%)	**0.5026** (+1.27%)	**0.6186** (+2.45%)	**0.0177** (+51.31%)	0.4237 (+6.33%)	**0.3446** (+11.14%)	**0.2362** (+17.81%)

Table 1: Subtopic retrieval performance when $n = 200$: where PT ($b = 4$, $\delta^2 = 10^{-5}$) and MMR ($c = 0.5$) are optimized for S-r@5 measure.

Models	***S-p@.1***	**S-p@.2**	**S-p@.5**	**S-p@1**	**S-MRR@100**	**S-r@5**	**S-r@10**	**S-r@50**
PRP	0.3968	0.3062	0.1920	0.0101	0.0071	0.2466	0.3900	0.6034
MMR	0.4501 (+11.83%)	0.3162 (+3.15%)	0.1948 (+1.46%)	0.0106 (+5.27%)	0.0073 (+1.89%)	0.2690 (+8.29%)	0.3784 (-3.08%)	0.6077 (+0.70%)
PT	**0.4807** (+17.45%)	0.2992 (-2.32%)	0.1857 (-3.38%)	0.0086 (-17.10%)	0.0108 (+33.80%)	0.2791 (+11.63%)	0.3622 (-7.69%)	0.5929 (-1.76%)
QPRP	0.4237 (+6.33%)	**0.3452** (+11.29%)	**0.2338** (+17.90%)	**0.01167** (+13.43%)	**0.01621** (+55.61%)	**0.3093** (+20.25%)	**0.4063** (+3.99%)	**0.6150** (+1.89%)

Table 2: Subtopic retrieval performance when $n = 100$: where PT ($b = 3$, $\delta^2 = 10^{-5}$) and MMR ($c = 0.9$) are optimised for S-p@.1 measure.

Models	**S-MRR@25%**	**S-MRR@50%**	**S-MRR@75%**	***S-MRR@100%***
PRP	0.2316	0.1056	0.0707	0.0086
MMR	0.2201 (-5.22%)	0.1135 (+6.96%)	0.0705 (-0.28%)	0.0097 (+11.39%)
PT	0.2131 (-8.68%)	0.1098 (+3.82%)	0.0674 (-4.89%)	0.0154 (+43.92%)
QPRP	**0.2322** (+0.25%)	**0.1355** (+22.06%)	**0.0716** (+1.25%)	**0.0177** (+51.31%)

Table 3: The Subtopic Minimum Reciprocal Rank for various levels of coverage (25%-100%) for $n = 200$. The results for PT ($b = 6$,$\delta^2 = 10^{-5}$) and MMR ($c = 0.5$) are optimised on S-MRR@100%.

shown that accounting for documents dependencies at relevance level delivers a better ranking for subtopic retrieval. Also, the results of our empirical investigation have shown that the QPRP consistently outperforms previous approaches, i.e. MMR and PT, but with the additional advantage that no tedious parameter tuning is required. This research demonstrates that the use of Quantum Probability Theory to model processes within information retrieval can lead to substantial improvements. Future investigations will consider:

1. alternative estimations of the interference;
2. how to derive a complex amplitude distribution from the document corpus;
3. the relationships between interference in the quantum probability framework and conditional probabilities in Kolmogorovian probability theory;
4. test the QPRP employing alternative collections for subtopic retrieval, and;
5. how to apply the QPRP paradigm to other retrieval tasks, e.g. ad-hoc retrieval.

Acknowledgments: We would like to thank Keith van Rijsbergen for his support and mentoring. This work has been conducted as part of the EPSRC Renaissance project (EP/F014384/1).

References

1. Robertson, S.E.: The probability ranking principle in IR. Journal of Documentation **33** (1977) 294–304
2. Eisenberg, M., Berry, C.: Order effects: A study of the possible influence of presentation order on user judgments of document relevance. JASIS **39**(5) (May 2007) 293–300
3. Chen, H., Karger, D.R.: Less is more: probabilistic models for retrieving fewer relevant documents. In: SIGIR '06. (2006) 429–436
4. Zhai, C.X., Cohen, W.W., Lafferty, J.: Beyond independent relevance: methods and evaluation metrics for subtopic retrieval. In: SIGIR '03. (2003) 10–17
5. Wang, J., Zhu, J.: Portfolio theory of information retrieval. In: SIGIR '09. (2009) 115–122
6. Clarke, C.L., Kolla, M., Vechtomova, O.: An effectiveness measure for ambiguous and underspecified queries. In: ICTIR '09. (2009) 188–199
7. Agrawal, R., Gollapudi, S., Halverson, A., Ieong, S.: Diversifying search results. In: WSDM '09. (2009) 5–14
8. Carbonell, J., Goldstein, J.: The use of MMR, diversity-based reranking for reordering documents and producing summaries. In: SIGIR '98. (1998) 335–336
9. Zuccon, G., Azzopardi, L., van Rijsbergen, K.: The quantum probability ranking principle for information retrieval. In: ICTIR '09. (2009) 232–240
10. Khrennikov, A.: Interpretations of probabilities. Walter de Gruyter (2009)
11. Clarke, C.L., Kolla, M., Cormack, G.V., Vechtomova, O., Ashkan, A., Büttcher, S., MacKinnon, I.: Novelty and diversity in information retrieval evaluation. In: SIGIR '08. (2008) 659–666
12. Gordon, M.D., Lenk, P.: When is the probability ranking principle suboptimal. JASIS **43**(1) (1999) 1–14
13. Cooper, W.S.: The inadequacy of probability of usefulness as a ranking criterion for retrieval system output. Unpublished working paper (1971)
14. Gordon, M.D., Lenk, P.: A utility theoretic examination of the probability ranking principle in information retrieval. JASIS **42**(10) (1999) 703–714
15. Fuhr, N.: A probability ranking principle for interactive information retrieval. JIR **12**(3) (June 2008) 251–265
16. van Rijsbergen, C.J.: The Geometry of Information Retrieval. Cambridge University Press (2004)
17. Melucci, M.: A basis for information retrieval in context. ACM TOIS **26**(3) (June 2008) 1–41
18. Piwowarski, B., Lalmas, M.: A quantum-based model for interactive information retrieval. In: ICTIR '09. (2009) 224–231
19. Hou, Y., Song, D.: Characterizing pure high-order entanglements in lexical semantic spaces via information geometry. In: QI 2009. (2009) 237–250
20. Flender, C., Kitto, K., Bruza, P.: Beyond ontology in information systems. In: QI 2009. (2009) 276–288
21. Huertas-Rosero, A.F., Azzopardi, L., van Rijsbergen, C.J.: Eraser lattices and semantic contents. In: QI 2009. (2009) 266–275
22. van Rijsbergen, C.J.: Information Retrieval, 2nd Ed. Butterworth (1979)