2 Criptografia: do clássico ao quântico

2.1 Introdução

A idéia de criptografia é sem dúvida alguma muito antiga. A partir do momento em que o ser humano passou a ter informações que não eram de domínio público, algum tipo de codificação de informação passou a ser requerida. No passado, a grande maioria das aplicações foi militar, com um dos exemplos mais famosos sendo a máquina *enigma* criada pelos alemães na 2ª Guerra Mundial.

Recentemente, especialmente com a proliferação da *Internet* e a crescente divulgação de nossas informações *on-line*, como, por exemplo, preenchimento de cadastros em diversos *websites*, sem falar no comércio eletrônico e movimentações bancárias que não param de crescer, o interesse por métodos criptográficos mais sofisticados vem aumentando na mesma proporção. Não é incomum ouvirmos falar de prisões de quadrilhas de *hackers*, ou de roubo de informações sigilosas pela *Internet*.

2.2 Porquê a criptografia clássica é vulnerável.

Existem dois esquemas principais de criptografia clássica modernos, o público e o privado. Como os nomes indicam, no esquema público parte da informação da chave é revelada publicamente, enquanto no privado é assumido que a Alice e o Bob já possuem a chave (que foi partilhada de alguma forma no passado).

O esquema privado de criptografia originalmente criado por Vernam, foi provado como 100% seguro por Shannon [2]. Esse esquema é conhecido pelo nome de "*one-time pad"* pois a chave utilizada é de uso único de forma a garantir a segurança da mensagem. A idéia é a seguinte: Uma chave binária totalmente aleatória e do mesmo comprimento da mensagem a ser enviada é criada por Alice. Essa mensagem é enviada a Bob através de um meio que é 100 % seguro, ou seja, sabe-se que a Eva não pode interceptar a chave. Para codificar a mensagem Alice

realiza uma operação XOR (OU-exclusivo) da chave com a mensagem a ser enviada. Essa operação também é conhecida como adição módulo 2. A mensagem é então enviada e Bob realiza a mesma operação XOR entre a mensagem e a chave (idêntica a de Alice) que ele possui. A segunda operação XOR faz com que a mensagem seja recuperada (Figura 2).

Para que o sistema seja totalmente seguro, alguns requisitos básicos são necessários. Em relação à chave, ela tem que ser totalmente aleatória (para não ter nenhuma semelhança com a mensagem) e só pode ser usada uma vez. Caso ela seja reutilizada, a espiã Eva pode obter um ganho de informação parcial sobre a mensagem. Nesse caso Eva pode reduzir sua ignorância em relação à nova mensagem; suponha duas mensagens m_1 e m_2 codificadas pela mesma chave k, logo as duas mensagens codificadas serão:

$$c_1 = k \oplus m_1 \tag{2.1}$$

$$c_2 = k \oplus m_2 \tag{2.2}$$

aonde $\oplus$ representa a adição módulo 2.

Como a chave é a mesma para as duas mensagens, Eva pode simplesmente fazer a adição módulo 2 das duas mensagens (ela tem acesso a elas, pois são transmitidas num canal passível de ser monitorado) c_1 e c_2 para obter $m_1 \oplus m_2$ [2] obtendo informação parcial sobre a mensagem. Isso pode não ser aceitável dependendo da aplicação.

A razão da segurança do "*one-time pad*" é que, como a mensagem codificada é formada a partir de uma operação XOR da mensagem a ser transmitida com a chave aleatória, a mensagem torna-se igualmente aleatória, não contendo nenhuma informação a não ser para quem também possua a chave. Por essa razão é impossível para a Eva obter qualquer informação sobre a mensagem original desde que a chave tenha o mesmo comprimento da mensagem, seja aleatória e não tenha sido utilizada anteriormente.

No entanto esse sistema possui uma grande dificuldade operacional: como entregar a chave a Bob de forma 100% segura. Carros-forte podem ser utilizados,

porém são uma opção muito cara e demorada. Nenhum canal clássico de transmissão pode ser utilizado, pois é passivo de monitoração.

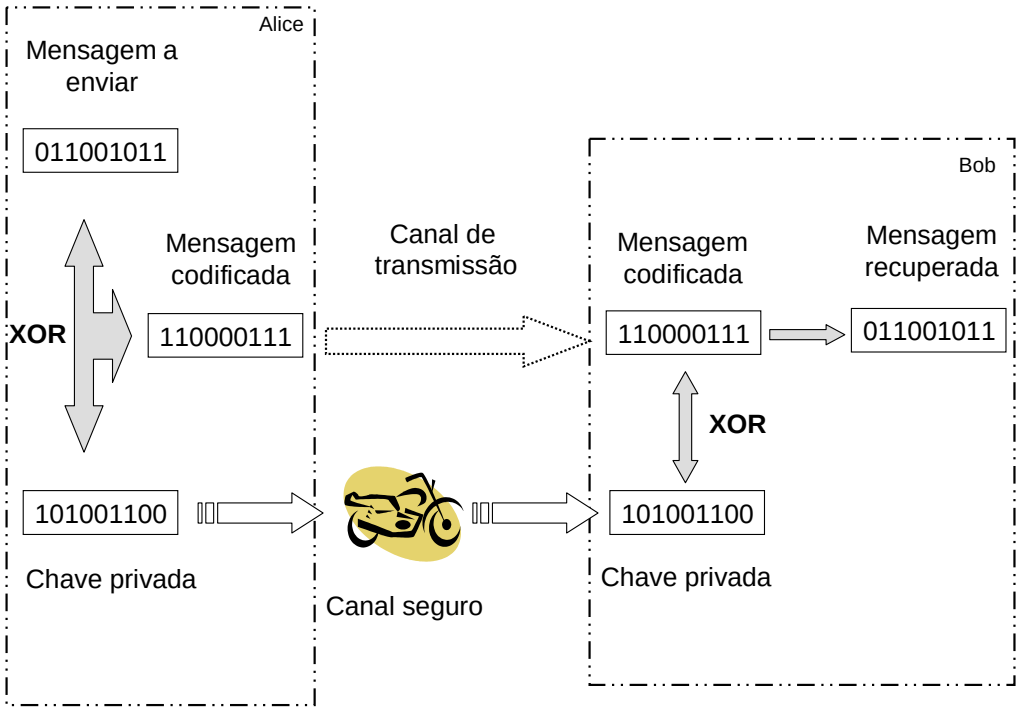

Figura 2: Idéia do "*one-time pad*". A mensagem a enviar é codificada com uma chave privada criada por Alice através de uma operação XOR, isto é, bits iguais retornam 0 enquanto bits diferentes retornam 1. O canal seguro aqui é representado através de um *courier* para tentar tornar a ação de Eva o mais difícil possível.

O esquema de criptografia clássica pública é de longe o mais utilizado pelo fato de oferecer segurança razoável e ser muito mais fácil de ser implementado [2]. A idéia básica dele é ilustrada na figura 3. Bob cria uma chave privada e uma chave pública (não confundir com o esquema anterior) que é representada através de uma caixa aberta na figura 3. A essa caixa está associada uma única chave privada que fica em poder de Bob. Ele produz várias dessas caixas e as envia para quem quiser ouvir. Uma das pessoas que receberam essa caixa é Alice que é quem justamente quer enviar a mensagem a Bob. Ao receber a caixa Alice coloca a mensagem dentro dela (a codifica) e a tranca. O princípio desse esquema está nesse fato, pois ao trancá-la ninguém exceto Bob (que possui a chave privada que gerou a caixa) consegue abri-la, nem mesmo Alice. Ela então envia a caixa

fechada de volta a Bob que consegue abri-la (decodificá-la) e ler a mensagem enviada por Alice.

A segurança desse esquema baseia-se na dificuldade de Eva conseguir abrir a caixa depois que ela tenha sido fechada, isto é, conseguir decodificar a mensagem. É muito fácil para Bob abrir a caixa pois ele possui a chave privada, porém é muito difícil para Eva conseguir decodificar a mensagem sem possuir a chave. No entanto isto não é impossível. Matematicamente a idéia desse esquema é de que é fácil computar a função *f(x)* tendo o valor de *x*, mas é muito difícil fazer a conta reversa, achar *x* a partir de *f(x)*. Em termos computacionais, "muito difícil", significa que o cálculo é de complexidade exponencial, ou seja, à medida que aumenta-se o número de bits da chave o problema torna-se exponencialmente mais complexo.

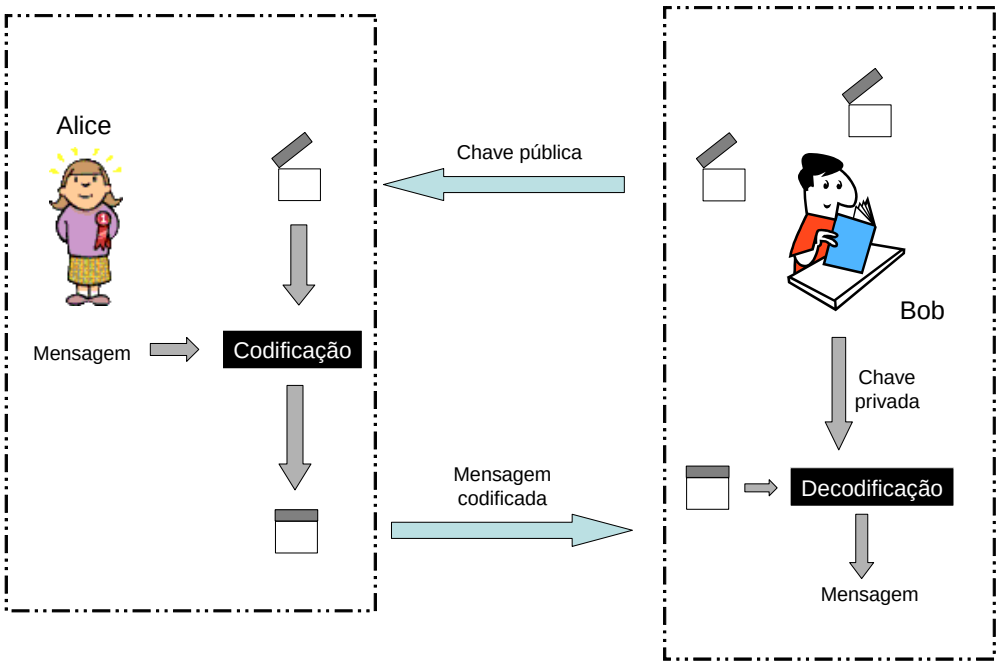

Figura 3: Princípio de funcionamento do esquema de criptografia clássica pública.

O esquema de criptografia pública mais utilizado atualmente é o RSA [6]. O cálculo em que ele se baseia é a fatoração de números extensos. Bob escolhe dois números primos grandes e calcula o seu produto $N = pq$. Ele obtém aleatoriamente uma chave de encriptação e baseada em p e q. Finalmente ele computa uma chave única de decriptação d que é guardada com ele. Ele então

revela N e e publicamente. A partir de N e e Alice pode codificar a mensagem enviando-a para Bob que ao recebê-la utilizará d para decodificá-la.

O problema do RSA é que ele assume que é muito difícil fatorar um número muito grande utilizando um computador clássico (fato que nunca foi provado matematicamente), sem contar que já existe um algoritmo quântico que faz isso em poucos segundos. Infelizmente para nós, no dia em que o primeiro computador quântico for à venda, todos os esquemas públicos de criptografia baseados no RSA serão automaticamente invalidados. Sem contar na possibilidade remota, mas diferente de zero, de que alguém possa descobrir um algoritmo clássico que possa resolver o problema da fatoração muito rapidamente. Como podemos resolver esse impasse? A resposta está no "*one-time pad*", pois sabemos que ele foi provado como matematicamente seguro. O problema está na distribuição da chave. É aí que finalmente, entra a mecânica quântica.

2.3 A teoria quântica entra em cena

Um belo dia, alguém teve uma idéia [8]. Por que não codificar a chave em *qubits* e enviá-los através de uma fibra óptica para o Bob? Dessa maneira se um espião tentasse medí-los, erros seriam inevitavelmente inseridos na transmissão e o espião seria descoberto. É importante lembrar que se Eva fizesse isso num canal clássico ela nunca seria descoberta, pois informação clássica pode ser copiada a vontade sem que haja perda ou alteração da mesma. Como já vimos, o mesmo não ocorre com a informação quântica.

Além disso, um estado quântico desconhecido não pode ser clonado. Suponha que qualquer estado quântico possa ser clonado. A cópia de dois estados quânticos ortogonais $|0\rangle$ e $|1\rangle$ (que são a base de estados da máquina copiadora) pode ser escrita como [6]:

$$|0\rangle \otimes |u\rangle \rightarrow |0\rangle \otimes |0\rangle \otimes |v_0\rangle \tag{2.3}$$

$$|1\rangle \otimes |u\rangle \rightarrow |1\rangle \otimes |1\rangle \otimes |v_1\rangle \tag{2.4}$$

aonde $|u\rangle$ é o estado inicial da máquina copiadora (independente dos estados de entrada $|0\rangle$ ou $|1\rangle$ pois é assumido que a máquina não possui nenhum conhecimento sobre os estados), $|v_0\rangle$ e $|v_1\rangle$ são os estados finais do sistema excluindo o original e a cópia e ⊗ representa produto tensorial. Para dois estados ortogonais, a cópia é bem sucedida. Fazendo agora o mesmo procedimento para o estado $a|0\rangle + b|1\rangle$ que é uma superposição linear dos dois estados ortogonais $|0\rangle$ e $|1\rangle$:

$$\left(a|0\rangle + b|1\rangle\right)\otimes|u\rangle \rightarrow a|0\rangle\otimes|0\rangle\otimes|v_0\rangle + b|1\rangle\otimes|1\rangle\otimes|v_1\rangle \tag{2.5}$$

obtemos (2.5) pela linearidade da mecânica quântica. No entanto para uma clonagem quântica bem sucedida deveríamos obter:

$$\left(a|0\rangle + b|1\rangle\right)\otimes\left(a|0\rangle + b|1\rangle\right)\otimes|v`\rangle \tag{2.6}$$

que é um estado produto. Como as equações 2.5 e 2.6 são diferentes, vemos que uma máquina copiadora quântica não existe, com exceção do caso com estados ortogonais.

É então, fundamental, frisar aqui novamente a idéia da criptografia quântica. Ela não será utilizada para codificar a mensagem. Ela até poderia, mas o espião só seria descoberto após a transmissão de parte da mensagem, havendo vazamento de informação. A aplicação natural então é no "*one-time pad*" para resolver o problema de distribuição da chave. A criptografia quântica, então, ganhou um termo na literatura em inglês que é o QKD (*Quantum Key Distribution* ou Distribuição de Chaves Quânticas).

2.3.1 O protocolo BB84

Como foi mencionado anteriormente a idéia da criptografia quântica é proteger a mensagem através da detecção de um possível espião. A mensagem não é protegida, impedindo que Eva tenha acesso a ela. A proteção é feita através do fato de que a mensagem será alterada pela Eva e a taxa de erro do sistema se elevará fazendo com que Alice e Bob sejam alertados da presença de Eva. Um

protocolo desenvolvido por Bennett e Brassard em 1984 [8] é utilizado para realizar esse processo.

Para a descrição do protocolo será também usada a primeira forma de codificação dos *qubits* numa fibra óptica: a polarização dos fótons. Os fótons são as entidades quânticas mais apropriadas para a transmissão de *qubits*, pois são naturalmente transmitidos através de uma fibra óptica. Para o armazenamento de *qubits*, aplicação fundamental em computação quântica o ideal é utilizar um outro meio como as maneiras comentadas na seção 1.3. apesar de algumas portas lógicas serem implementadas com fótons.

Será utilizada, inicialmente, a polarização linear para a descrição. Escolhemos estados ortogonais para a codificação dos bits, por exemplo $|\rightarrow\rangle$ para bit 0 enquanto que $|\uparrow\rangle$ para bit 1. Dessa maneira Bob consegue separar os dois *qubits* utilizando um divisor de feixe (*beamsplitter)* de polarização (PBS). Para simplificar a explicação vamos supor que a criação de fótons únicos não é um problema. Iremos lidar com o caso mais realista posteriormente.

Alice gera uma chave aleatória, a codifica em fótons com polarizações ortogonais e os envia para Bob que os mede (Figura 4) utilizando dois detectores avalanche para fótons únicos (SPADs). Esse primeiro esquema serve apenas para ilustrar como é feita a transmissão dos *qubits*, mas ele não oferece proteção contra um ataque feito pela Eva por uma simples razão: como a transmissão é feita em uma base ortogonal, e supondo que ela alinhou seu divisor de feixes de polarização de maneira idêntica ao de Bob, ela consegue medir os fótons sem introduzir erros, afinal ela está *usando a mesma base que Bob e Alice*.

Como acabamos de ver, um esquema utilizando uma base ortogonal é passível de ser atacado sem ser detectado. Como podemos resolver o problema? A resposta está em utilizar valores não-ortogonais para a polarização dos fótons. Porém isso também criará uma questão. Se uma base não-ortogonal for utilizada, por exemplo $|\uparrow\rangle$ e $|\nearrow\rangle$ (0^o e 45^o), Eva não conseguirá distinguir os dois estados, o mesmo acontecendo com Bob. É impossível distinguir dois estados quânticos não-ortogonais pois não sabemos sobre qual vetor da base o estado medido irá se projetar. O resultado obtido dependerá das probabilidades (por exemplo a^2 e b^2 no caso de sistemas quânticos de dois estados como a Eq. 1.1) de se obter um dos resultados. Uma conseqüência dessa indistinguibilidade na distinção de dois

estados não-ortogonais é o teorema de não-clonagem quântica apresentado anteriormente. Curiosamente um esquema utilizando somente uma base ortogonal foi proposto, no entanto esse sistema utiliza uma separação temporal para realizar a transmissão, tornando-o potencialmente seguro [13].

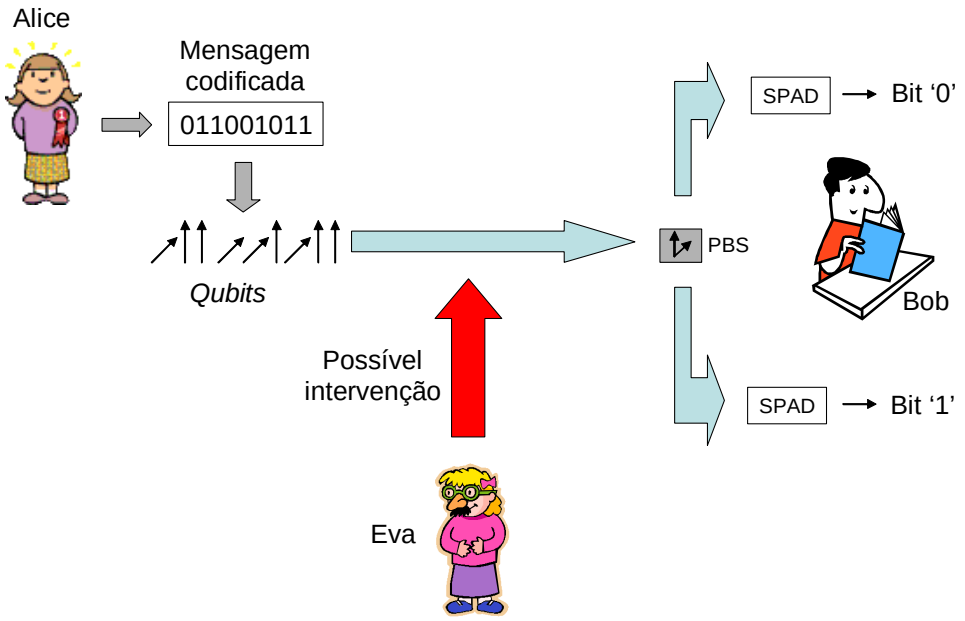

Figura 4: Esquema de uma possível transmissão de *qubits* codificados por Alice utilizando uma base ortogonal de polarização. Esse esquema é vulnerável a um ataque de Eva.

A solução para esse impasse é dada pelo protocolo BB84, utilizando duas bases ortogonais A e B mas que não sejam ortogonais entre si. O exemplo mais utilizado é $|\uparrow\rangle$ e $|\rightarrow\rangle$ para a base A e $|\nearrow\rangle$ e $|\searrow\rangle$ para a base B. Alice escolhe aleatoriamente quais bases irá utilizar com 50 % de chance para cada uma. Bob irá fazer o mesmo, logo, em apenas metade das medições as bases irão concordar. Somente nesses casos eles podem ter certeza dos resultados. Quando as bases forem diferentes o fóton tem probabilidade ½ de ir para qualquer SPAD. Nesses casos os resultados devem ser descartados gerando uma perda de 50 % no sistema.

No entanto essa mudança aleatória entre as duas bases feitas por Alice e Bob é o que dará a segurança ao sistema. Note que agora a Eva, assim como Bob não sabe que base a Alice escolheu para transmitir. Logo, Eva também deve escolher as suas bases para medição, aleatoriamente.

Para completar o sistema ainda falta um detalhe. Alice e Bob compartilham um canal clássico de comunicação cujo único requisito é que ele não pode ser alterado, mas pode ser monitorado (Figura 5). Esse canal tem diversas funções na transmissão, como realizar o processo conhecido como reconciliação de bases entre Alice e Bob (verificação para quais *qubits* as bases foram as mesmas e para quais não) e checar a presença de Eva. Ela agora tem que tentar adivinhar as bases que a Alice utilizou, mas isso é impossível pois a escolha das bases é aleatória. Se Eva tentar medir todos os fótons ela irá introduzir uma taxa de erro de 25% no sistema [3]. Se essa taxa for medida por Alice e Bob, eles sabem que alguém está tentando medir os fótons para tentar obter um ganho de informação. Eles devem então descartar a chave e tentar novamente, ou tentar um outro meio (como um outro canal quântico).

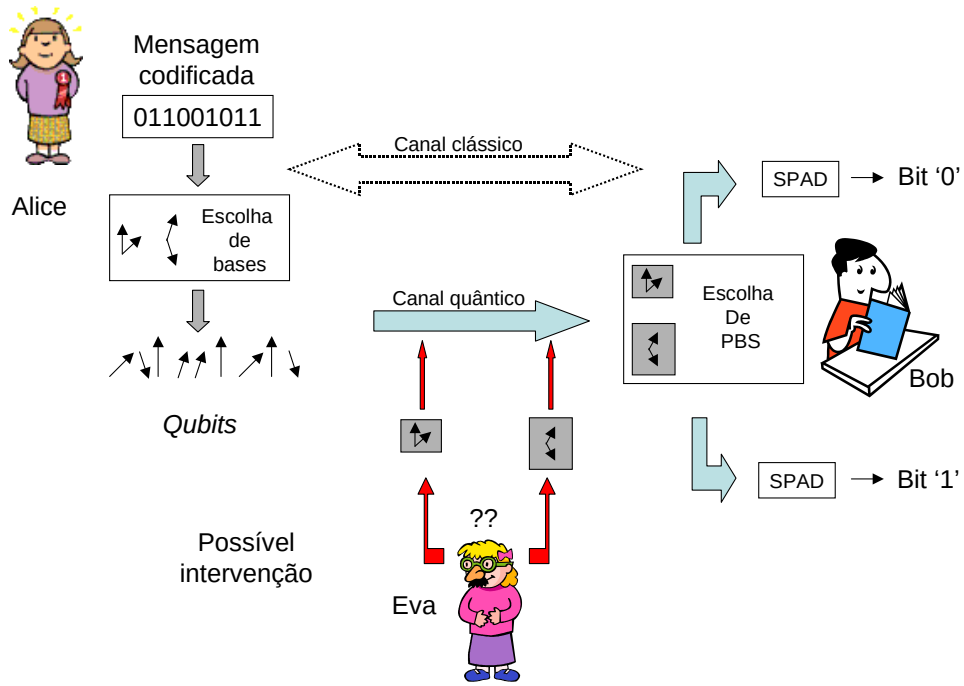

Figura 5: Implementação do BB84. Alice faz uma escolha aleatória das bases em que irá transmitir os *qubits* através do canal quântico enquanto Bob faz uma outra escolha (também aleatória) de qual base ele irá medir os *qubits*. A reconciliação é feita entre eles utilizando o canal clássico. Através desse esquema Eva introduz erros na transmissão que podem ser detectados.

O protocolo passo-a-passo funciona da seguinte maneira:

1. Alice gera a chave a ser utilizada para codificar a mensagem utilizando a idéia do "*one-time pad*". Para cada bit gerado da mensagem é feita uma escolha aleatória de qual das duas bases A e B será utilizada. Para a base A temos $|\rightarrow\rangle$ para o bit 0 e $|\uparrow\rangle$ para o bit 1, enquanto que na base B $|\nearrow\rangle$ representa o bit 1 e $|\searrow\rangle$ o bit 0. O sistema todo é sincronizado e dessa forma Alice anota quais bases foram utilizadas para cada bit para todos os instantes de tempo do relógio do sistema.
2. Bob faz uma escolha aleatória entre as bases A e B para medir cada bit e também anota qual base foi utilizada para cada fóton detectado e em qual SPAD a detecção ocorreu. Haverão muitos casos em que nenhum fóton será detectado num sistema real e nesses casos Bob deverá anotar em quais instantes isso ocorreu também. Nesse momento o conjunto de todos os bits recebidos por Bob é chamado de *raw key*.
3. Após a transmissão de um número suficiente de *qubits* a transmissão é interrompida e Bob avisa Alice em quais instantes de tempo seus SPADs receberam fótons e as bases utilizadas por ele. Em contrapartida ela o avisa quais bases ela usou nesses instantes. Repara-se que ele não revela *qual SPAD foi disparado para esses qubits*. Eles descartam todos os bits da mensagem para os quais bases diferentes foram utilizadas. Essa versão da chave é chamada de *sifted key*.
4. O próximo passo consiste em testar a presença de Eva, consistindo em simplesmente sacrificar parte do conjunto de bits e publicamente calcular a taxa de erro. Se a taxa de erro for de 25% eles sabem que ou Eva tentou interceptar todos os fótons e os reenviou, ou o canal está anormalmente ruidoso. Eles devem descartar todos os fótons e tentar uma nova transmissão.
5. Eles precisam verificar se a chave que eles possuem é idêntica, afinal problemas na transmissão e na detecção levam a erros nos bits. É feito um processo de correção de erros utilizando o canal público da seguinte forma: O conjunto de bits que eles possuem é particionado

em blocos de tamanho *k* de maneira que em cada bloco é altamente improvável que exista mais de um erro. Para cada bloco, Alice e Bob comparam a sua paridade publicamente (a paridade de um bloco é 1 se o número de 1s nesse bloco é par, se o tipo de paridade calculada for par e o contrário se for ímpar). Se as paridades do mesmo bloco para Alice e Bob forem iguais esse bloco é inicialmente aceito como correto. Se as paridades forem diferentes uma busca binária será aplicada ao bloco, ou seja o bloco será seccionado em partes menores e as paridades serão novamente computadas até que o erro seja isolado. Com o intuito de prevenir que Eva obtenha um ganho de informação através dessa discussão pública, Alice e Bob devem descartar o último bit de cada bloco ou sub-bloco cuja paridade foi anunciada. O problema do teste de paridade é que ele só detecta um número ímpar de erros, ou seja, se dois erros estiverem presentes em um bloco, eles não serão detectados. Para conseguir corrigir isso é feita uma permutação aleatória dos bits da *sifted key* diversas vezes além de variação no número de bits de cada bloco seccionado. Isso é feito inúmeras vezes de forma que a probabilidade de que a chave ainda contenha erros seja muito pequena.

6. Finalmente é realizada a amplificação da privacidade, algoritmo utilizado cuja idéia é reduzir a zero a informação de Eva. Sabemos que se ela tentar medir todos os fótons enviados por Alice ela será facilmente detectada. No entanto, se apenas uma pequena fração dos *qubits* for interceptada será gerada uma pequena elevação da taxa de erro que pode ser facilmente confundida pelas perdas no canal. Dessa forma Eva obteve um ganho parcial de informação sobre a chave sem que Alice e Bob percebessem. Esse processo final serve para reduzir a zero a informação que Eva possui sobre a chave. Através da aplicação de uma função matemática da classe de funções *hash* Alice e Bob irão diminuir o conjunto de bits que possuem para torná-los numa chave absolutamente segura [9].

A Tabela 1 fornece um resumo na forma de um exemplo dos passos do protocolo BB84 até a formação da *sifted key*.

Bit a ser enviado	0	1	1	0	1	1	0	0	0	1	0
Escolha de base de Alice	A	B	B	A	B	A	B	B	A	A	B
Fóton enviado	$\|\rightarrow\rangle$	$\|\nearrow\rangle$	$\|\nearrow\rangle$	$\|\rightarrow\rangle$	$\|\nearrow\rangle$	$\|\uparrow\rangle$	$\|\searrow\rangle$	$\|\searrow\rangle$	$\|\rightarrow\rangle$	$\|\uparrow\rangle$	$\|\searrow\rangle$
Escolha de base de Bob	B	A	B	B	A	A	B	B	B	A	B
Bits lidos por Bob (*Raw Key)*	0	0	1	1	1	1	0	0	1	1	0
Bases coincidentes?			√			√	√	√		√	√
Sifted Key			1			1	0	0		1	0

Tabela 1: *Sifted Key* gerada a partir de um conjunto de bits a ser enviado após a verificação das bases coincidentes.

2.3.2 O protocolo EPR

Uma segunda possibilidade para realizar a distribuição quântica de chaves é a utilização de pares de fótons emaranhados. Como foi mencionado anteriormente esses pares possuem uma correlação muito forte em alguma propriedade intrínseca ao serem medidos. Continuaremos utilizando o exemplo de polarização, ou seja, dependendo em qual fonte o par é originado eles revelam possuir a mesma polarização ao serem medidos ou polarizações opostas. Um par EPR pode ser representado pelo singleto:

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle_A|\rightarrow\rangle_B - |\rightarrow\rangle_A|\uparrow\rangle_B\right) \tag{2.3}$$

Os subscritos A e B em (2.3) indicam os fótons A e B do par respectivamente. É interessante notar que não conseguimos escrever esses dois fótons de forma individual, ou seja, eles estão "presos" nesse estado produto. Matematicamente é isso o que indica o emaranhamento entre as partículas. Pares EPR já foram e continuam sendo o foco de discussões intensas nas áreas de filosofia e física, justamente devido a essa correlação quântica.

A fonte de pares emaranhados situa-se entre Alice e Bob (não necessariamente na metade da distância) sendo que um dos fótons do par é enviado a ela e o outro a ele. Da mesma maneira que no protocolo BB84 na seção

anterior ambos escolhem aleatoriamente bases para fazerem as suas medidas. Se eles escolherem a mesma base, seus resultados serão perfeitamente anti-correlacionados e basta que um deles inverta o valor do *qubit* medido, para formar a chave. Quando as bases forem diferentes o *qubit* medido assumirá um valor que não tem significado, então eles descartam o seu valor.

Uma vantagem de se utilizar pares EPR é que existe um teste adicional que pode ser feito para testar a presença de Eva, que é o teste da violação da desigualdade de Bell. Essa violação ocorre quando comparamos os resultados de medidas das polarizações de cada um dos fótons do par. As projeções dos estados de polarização dos fótons são correlacionadas qualquer que seja a base utilizada na medida, desde que elas sejam iguais dos dois lados. Tudo se passa, como se a medida em um deles alterasse o estado do outro. O mecanismo que causa este fenômeno é pouco compreendido, mas de alguma forma é possível alterar o resultado obtido por Bob variando-se a forma de se medir em Alice e vice-versa. A violação da desigualdade de Bell foi verificada experimentalmente por Aspect [4]. No entanto é impossível utilizar esse esquema para transmissão de dados numa velocidade mais rápida do que a luz, pois quem variou o PBS tem que avisar ao outro como o variou (de quanto é o ângulo α) utilizando um canal clássico (limitado pela teoria da relatividade). Caso não haja a sinalização clássica a pessoa do outro lado da linha não tem como distinguir a informação, afinal tudo que ela está vendo são fótons disparando os SPADs de forma aparentemente aleatória. Diz-se aparentemente, pois a informação clássica está ali, ela apenas é inacessível.

O protocolo EPR pode ser ligeiramente modificado para testar para a desigualdade de Bell. Esse teste pode ser feito através da inserção de uma terceira base [3]. A desvantagem é que a probabilidade de Alice e Bob usarem as mesmas bases cai de 1/2 para 2/9. Num sistema perfeito, na ausência de Eva, a violação da desigualdade será máxima [3]. Caso Eva tenha interceptado os fótons ela irá reduzir esse valor, sendo então detectada.

Nesse sistema a chave é formada por uma escolha aleatória de bases feita por Alice e Bob, sendo uma chave aleatória também. A formação da chave como no protocolo BB84 só ocorre depois do processo de transmissão após a reconciliação de bases, verificação da presença de Eva, correção de erro e

amplificação da privacidade. O esquema de transmissão é representado na figura 6.

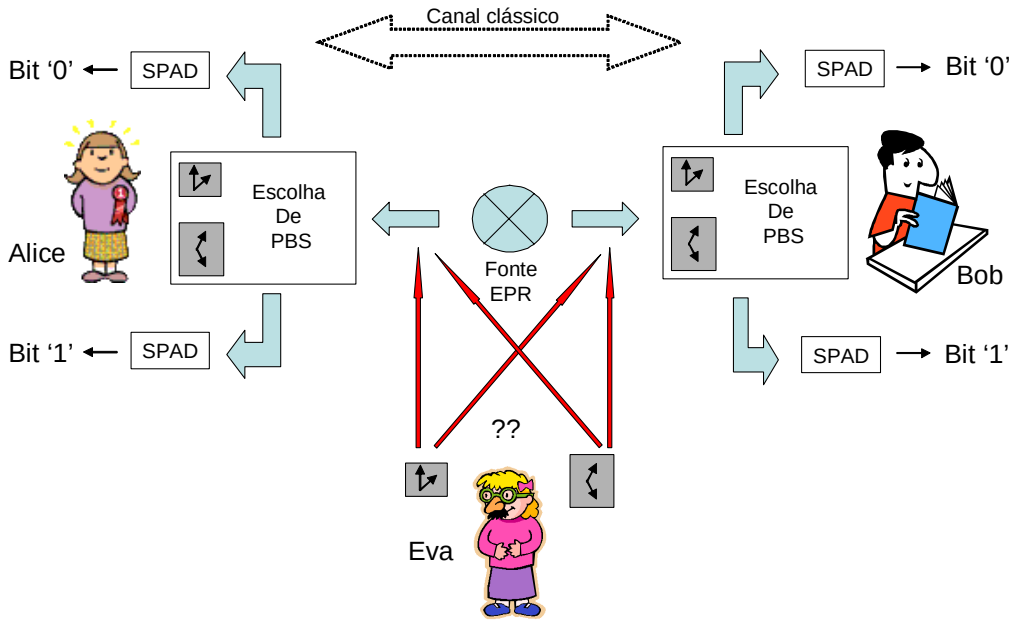

Figura 6: Sistema de criptografia quântica utilizando pares EPR para a transmissão dos *qubits*. Mesmo na situação mais favorável para Eva, em que ela possui o controle da fonte EPR, o esquema ainda é seguro, pois a desigualdade de Bell não será violada de forma máxima nesse caso.

2.3.3 Codificação em fase

O sistema baseado em codificação em polarização apresenta um sério problema para transmissão em fibras ópticas a longas distâncias. Um efeito denominado de Dispersão de Modos de Polarização (PMD) [11] faz com que a polarização da luz varie aleatoriamente ao longo de uma fibra. Isso é causado devido a imperfeições no processo de fabricação e cabeamento que fazem com que o índice de refração não seja uniforme para componentes ortogonais do sinal óptico ao longo da fibra, fenômeno conhecido como birrefringência. Além de fatores intrínsecos a fibra como a fabricação, a birrefringência depende também

de agentes extrínsecos como temperatura e pressão. Em um enlace de fibra esses fatores fazem com que a birrefringência varie aleatoriamente.

Para um sistema de QKD utilizando codificação de polarização existem duas soluções, a utilização de fibras que mantém a polarização da luz constante (PMF) ou soluções ativas de compensação de polarização. A primeira opção é economicamente inviável pois a idéia da transmissão de *qubits* utilizando fótons é aproveitar a extensa malha óptica existente que consiste maciçamente de fibras padrão e de dispersão deslocada (DS). A segunda opção é tecnicamente possível, entretanto, é de alto grau de complexidade especialmente lidando no regime quântico.

Foi então feita uma nova proposta utilizando codificação de fase [12]. A idéia é ilustrada na figura 7. Nós temos basicamente um interferômetro *Mach-Zender* conectando Alice e Bob com um modulador de fase em cada braço. Cada um dos moduladores é controlado por Alice e Bob respectivamente.

Os braços do interferômetro possuem exatamente o mesmo comprimento. Se nenhum deslocamento de fase for feito por Alice ou por Bob, o sinal óptico dividido no primeiro acoplador bidirecional chegará ao segundo exatamente com a mesma fase, disparando o SPAD 1. Se a diferença de fase (imposta pelos deslocadores de fase ϕ_A e ϕ_B) for de π, então o SPAD 2 será disparado. Finalmente se a diferença de fase for de $\pi/2$ a intensidade do sinal será dividida por dois entre os dois detectores. No regime quântico, isto é, fóton a fóton, haverá 50% de probabilidade de qualquer um deles ser disparado.

A grande dificuldade deste sistema é manter a estabilidade mecânica do interferômetro ao longo de quilômetros de distância (o que é praticamente impossível), pois uma pequena variação no comprimento (da ordem de 1μm) de um braço em relação ao outro fará com que a relação de fase seja destruída.

Para tornar essa codificação prática foi proposta uma variação desse esquema: O interferômetro é colapsado em dois "meio-interferômetros" desbalanceados situados em Alice e Bob, sendo que agora somente uma fibra óptica faz a ligação entre os dois (Figura 8) [12]. Nesse caso a estabilidade mecânica dos interferômetros é facilmente realizada estabilizando-se a temperatura.

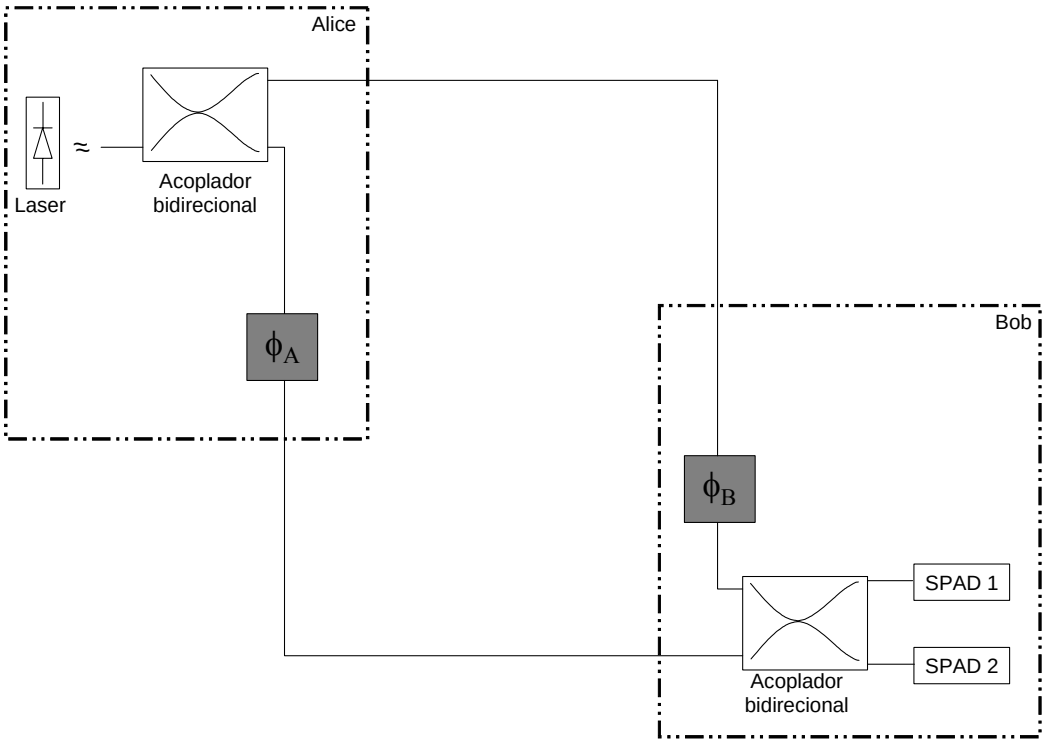

Figura 7: Idéia inicial de um esquema para QKD utilizando codificação de fase. O sistema consiste de um interferômetro Mach-Zender com moduladores de fase ϕ_A e ϕ_B pertencentes a Alice e Bob respectivamente.

Com este novo sistema temos quatro possibilidades de caminho para um fóton. Ele pode escolher o caminho curto-curto, longo-longo, curto-longo e longo-curto. Isso irá gerar três picos distintos separados temporalmente na saída do "meio-interferômetro" de Bob. Se a diferença de caminho de cada "meio-interferômetro" for longa o suficiente (o seu desbalanceamento), não haverá superposição dos três picos. Nota-se que os únicos caminhos que são de nosso interesse são o curto-longo e o longo-curto, pois são os caminhos em que haverá interferência gerada pelos moduladores de fase ϕ_A e ϕ_B. O principal requerimento desse sistema é de que a diferença entre os desbalanceamentos dos dois interferômetros não seja maior do que uma fração de um comprimento de onda. Como foi mencionado acima isso pode ser realizado utilizando-se compartimentos com a temperatura controlada para os "meio-interferômetros".

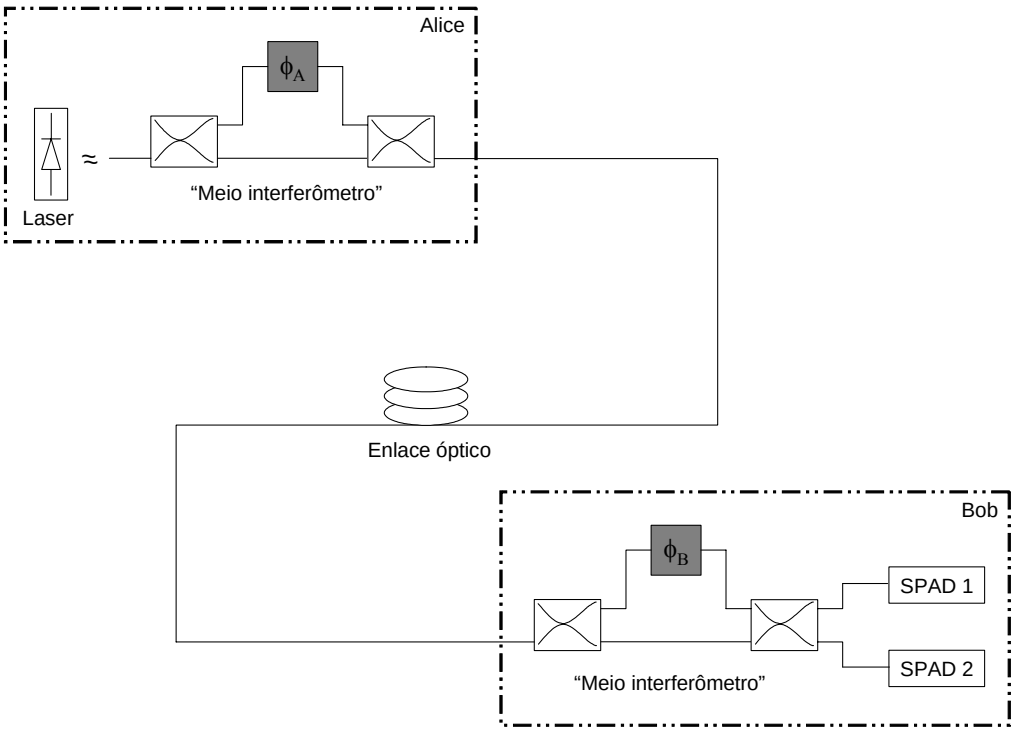

Figura 8: Esquema modificado do sistema para transmissão de *qubits* utilizando codificação de fase composta de dois "meio-interferômetros".

Como o protocolo BB84 pode então ser implementado utilizando a codificação de fase? Alice e Bob podem variar a fase do fóton de 0 a 2π utilizando os moduladores de fase ϕ_A e ϕ_B. Alice então novamente utilizará duas bases para o sistema, só que dessa vez utilizará 0 e π para a base A e $\pi/2$ e $3\pi/2$ para a base B novamente escolhidos aleatoriamente. Bob utilizará dois valores para a detecção, 0 e $\pi/2$, aleatoriamente (Tabela 2). Dependendo da escolha de base utilizada por Alice e do valor utilizado por Bob para o seu modulador de fase, a diferença de fase na chegada do fóton que percorreu os caminhos curto-longo e longo-curto será de 0, $\pi/2$, π ou $3\pi/2$. Para os casos em que a diferença de fase é de $\pi/2$ e $3\pi/2$ haverá igual probabilidade de o fóton ser detectado em qualquer um dos SPADs. Esses casos correspondem a situação em que Alice e Bob utilizaram bases diferentes e devem ser descartados.

Base	Bit a ser enviado	Alice	Bob	$\phi_A - \phi_B$	Bit detectado
A	0	0	0	0	0
A	0	0	$\pi/2$	$3\pi/2$	?
A	1	π	0	π	1
A	1	π	$\pi/2$	$\pi/2$	?
B	0	$\pi/2$	0	$\pi/2$	?
B	0	$\pi/2$	$\pi/2$	0	0
B	1	$3\pi/2$	0	$3\pi/2$	?
B	1	$3\pi/2$	$\pi/2$	π	1

Tabela 2: Esquema de utilização do protocolo BB84 com codificação de fase. "?" corresponde a situação em que existe 50% de probabilidade de que cada SPAD seja disparado. No caso clássico, corresponde a metade da intensidade em cada detector. O *qubit* só pode ser interpretado como correto nos casos em que a defasagem é igual a 0 ou π.

Para os casos em que a diferença de fase é de 0 ou π o fóton tem 100% de probabilidade (no caso ideal) de ser detectado no SPAD 1 ou 2. Esses valores deverão ser guardados para a formação da chave durante o processo de reconciliação que é idêntico ao realizado no caso da codificação em polarização, incluindo os processos de detecção de Eva, correção de erro e amplificação de privacidade. A codificação em fase utilizando BB84 é tão segura quanto a codificação em polarização com o mesmo protocolo. Para o caso da fase Eva novamente não sabe qual base Alice usou pra a transmissão do *qubit*. Mais uma vez, Eva é obrigada a fazer a mesma escolha aleatória de Bob e irá inserir erros na transmissão (25%, se ela interceptar todos os fótons) que Bob detectará mais tarde durante a reconciliação.

Temos então um sistema que funciona como o baseado em polarização com a vantagem adicional de ser insensível à variação da polarização ao longo de fibras ópticas. Na prática, especialmente para distribuição de chaves longas, a codificação de fase requer compensação ativa. A primeira razão para isso é o controle de temperatura para os "meio-interferômetros". Esse controle deve ser muito preciso, pois pequenas variações fazem com que o caminho óptico varie nos "meio-interferômetros" podendo fazer com que os fótons provenientes dos diferentes caminhos se sobreponham no tempo, destruindo a interferência do pico central. Como todo o sistema é sincronizado, essas variações também podem causar *jitter* de tempo nos SPADs. Como veremos adiante, o tempo de chegada dos fótons nos SPADs deve ser bem definido e conhecido. Além disso se

moduladores de fase insensíveis a polarização não forem utilizados haverá variação na probabilidade de se detectar um fóton (ou no caso clássico, variação da intensidade da luz detectada em cada fotodetector).

2.3.4 O protocolo B92

No mesmo trabalho [12], Bennett propôs a possibilidade de realização da transmissão de *qubits* utilizando apenas *quaisquer dois estados não ortogonais.* Este fato é derivado do princípio da mecânica quântica sobre a ambigüidade de se distinguir dois estados não ortogonais. Este protocolo funcionaria da seguinte maneira: Alice passaria a transmitir somente dois estados, 0 e π com Bob fazendo a sua medida também nesses mesmos dois estados. Somente o SPAD 1 será colocado na saída do "meio-interferômetro" de Bob.

Toda vez que eles escolherem o mesmo valor para a medição (0,0 ou π,π) ocorrerá interferência construtiva e o SPAD 1 disparará acusando uma contagem. Quando os valores escolhidos não forem iguais (0,π ou π,0) ocorrerá interferência destrutiva e nenhuma contagem poderá ser detectada. No entanto, esse caso em que as bases discordam não podem ser utilizados para a transmissão de *qubits* pois como será visto no capítulo seguinte, os SPADs possuem ruído que pode gerar contagens espúrias. Normalmente chamadas de contagens de escuro (*dark counts*) elas dão origem a contagens falsas, pois o detector é disparado sem a presença de fótons. Voltando ao caso em que os valores escolhidos por Alice e Bob são diferentes, não podemos confiar na premissa de que o SPAD 1 não vai disparar, pois ele pode retornar uma contagem falsa devido ao ruído e será impossível saber se o valor é falso ou não. Logo durante a reconciliação as bases diferentes serão descartadas.

A reconciliação será feita da seguinte forma para o B92: Bob avisará a Alice *somente os instantes de tempo em que seu SPAD disparou*. O valor escolhido por Bob será mantido em segredo por ele. Toda a vez que o detector tiver disparado, *a menos de erros na transmissão e contagens de escuro*, Bob saberá que Alice utilizou o mesmo valor que ele para o modulador de fase. Associando o valor 0 ao bit 0 e π ao bit 1 (ou vice-versa) Bob tem como saber qual *qubit* foi transmitido. O mesmo é válido para a Alice, pois como ela foi

avisada por Bob em quais instantes ele detectou uma contagem, ela sabe qual valor ele escolheu para ϕ_B, e esse valor será então utilizado para a formação da chave. Todos os instantes em que não é detectada uma contagem Alice e Bob descartam os valores utilizados para os moduladores de fase.

De uma forma mais geral a medida realizada no B92 é conhecida na mecânica quântica como POVM ou Medição do Valor do Operador Positivo. Ela é utilizada na distinção de estados não-ortogonais e funciona da seguinte maneira: ao realizarmos uma medida para a identificação de um estado ela pode retornar os autovalores 0 e 1. Ao recebermos 0, o resultado é considerado inconclusivo e não conhecemos nada sobre o estado. No entanto ao recebermos 1, *temos certeza* de que o estado é aquele medido [16].

Entretanto, existe um problema com esse protocolo tal como ele foi apresentado. Ele é vulnerável a ataques de Eva, pois ela pode fazer a mesma medida que Bob faz. Quando o SPAD dela disparar ela saberá qual foi o valor utilizado por Alice e enviar um outro fóton correspondente para Bob tentando mascarar sua presença. No caso de um resultado inconclusivo (SPAD não dispara) ela pode bloquear o fóton, fazendo Bob acreditar que esse fóton perdido se deve na realidade a perdas no canal, fazendo com que Eva passe despercebida. Uma solução para esse problema é através da utilização de um pulso forte (clássico) de referência acompanhando o *qubit*. Esse pulso complica a situação de Eva consideravelmente, pois agora se ela suprimir o fóton no caso de uma medida inconclusiva, o pulso clássico (que também é utilizado como referência) chega a Bob e o avisa de que deveria haver um *qubit* ali. A outra situação é ainda pior, se Eva suprimir ambos o *qubit* e o pulso de referência Bob simplesmente descartará esse *qubit* pois o pulso clássico não foi recebido. A utilização do pulso clássico de referência torna o B92 seguro contra ataques [16].

2.3.5 Sistemas “Plug and Play”

Com o intuito de resolver o problema da compensação ativa requerida para a codificação de fase foi proposto um sistema que utiliza uma auto-compensação, ou *“Plug and Play”* como foi chamado [3], [14]. O sistema é baseado no uso de um espelho de Faraday (um espelho com uma placa de onda de $\lambda/4$ na frente).

Pode ser mostrado que o uso desse espelho no fim de uma fibra faz com que o estado de polarização de entrada na fibra seja sempre ortogonal ao da saída da mesma fibra, (o pulso que reflete no espelho e volta à entrada) independentemente de qualquer birrefringência na fibra [3]. O único requerimento é que a variação da birrefringência no tempo seja longa em relação ao tempo requerido para a luz realizar a viagem de ida e volta. Dessa forma o sistema pode ser auto-compensado de forma passiva.

Resumidamente o sistema funciona da seguinte forma: a idéia é que o pulso de luz seja proveniente de Bob (Figura 9). Uma fonte laser produz um pulso de

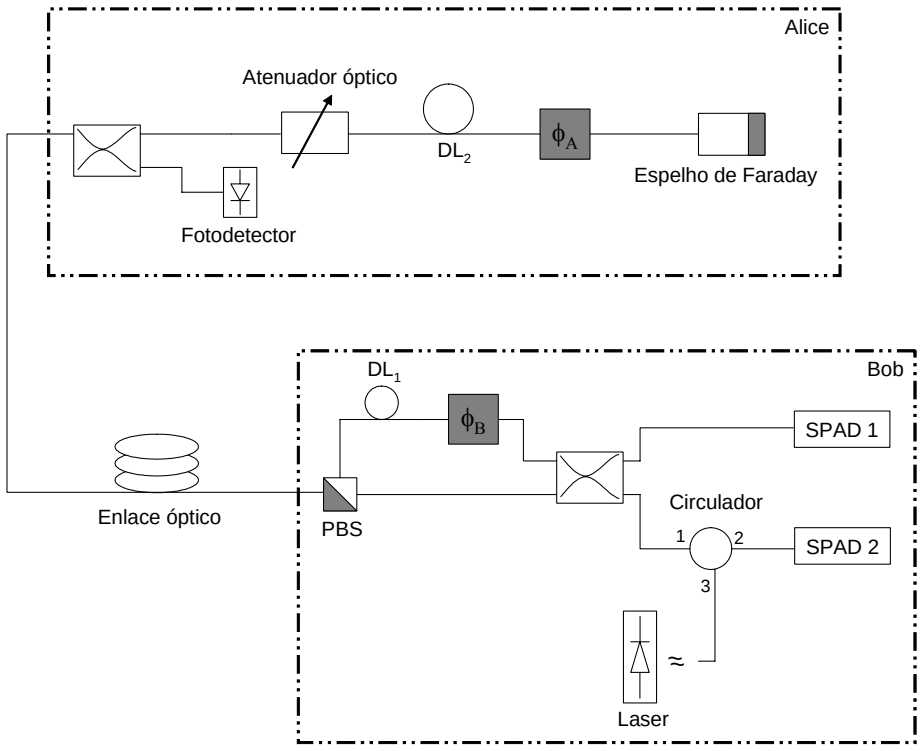

Figura 9: Sistema *"Plug and Play"* para a transmissão de *qubits* utilizando a codificação de fase.

alta intensidade e curto que é guiado por um circulador até um acoplador bidirecional. O pulso é dividido igualmente entre dois caminhos, um curto e um longo. O caminho longo contém o modulador de fase ϕ_B e a linha de atraso DL_1 cuja função é separar temporalmente os dois pulsos. O atraso dado é da ordem de 200ns e o modulador ϕ_B não é ativado ainda. Logo, dois pulsos (P1, referente ao

caminho curto, e P2 que seguiu o caminho longo) distintos são enviados através do enlace óptico. A função do divisor de feixes de polarização PBS será explicada mais adiante.

Os dois pulsos P1 e P2 atravessam o enlace e chegam ao acoplador bidirecional na entrada de Alice. P1 é dividido simetricamente com uma metade chegando ao fotodetector clássico situado numa das saídas do acoplador, e a outra metade seguindo para o atenuador óptico. O fotodetector clássico é utilizado para fornecer uma referência de tempo para o sistema e também é utilizado como proteção contra ataques do tipo "Cavalo de Tróia" [3]. A outra metade segue por um atenuador óptico por outra linha de atraso DL_2, pelo modulador de fase ϕ_A (também desativado) e chega ao espelho de Faraday. O pulso P2 seguirá o mesmo caminho.

Tudo isso é feito para que a compensação do sistema possa ser realizada, pois é agora que irá começar a transmissão dos *qubits* propriamente dita. Afinal até aqui estamos lidando com pulsos de luz clássicos. O pulso refletido pelo espelho de Faraday passa por ϕ_A que agora sim é ativado e aplica o valor escolhido por Alice, mas somente em P1. Após ser inferido o valor do *qubit* o pulso (ainda clássico) passa novamente pelo atenuador óptico. O valor da atenuação aplicada é calculada de tal forma, que ao ser atenuado pela segunda vez, o número médio de fótons por pulso não seja maior do que uma fração de um fóton.

Adiantando um pouco do que será visto no próximo capítulo, e um fato de extrema importância que não foi mencionado ainda é *como gerar fótons unitários*? Como a distribuição de fótons de uma fonte laser é poissoniana a maneira existente de se fazer isso atualmente é atenuar fortemente os pulsos de luz clássicos de forma a se ter somente uma probabilidade de encontrar μ fótons por pulsos. Normalmente por razões de segurança do sistema, $\mu < 1$.

Os dois pulsos (agora quânticos) seguem de volta a Alice. Como eles foram refletidos por um espelho de Faraday, suas polarizações na chegada no PBS de Bob são ortogonais aos valores nas suas saídas anteriormente. Dessa forma eles seguirão por caminhos opostos aos que tomaram na saída, com P1 tomando o caminho longo e P2 o caminho curto. Bob agora infere seu valor escolhido no pulso P1 que interfere com P2 no acoplador bidirecional. Novamente, dependendo

dos valores escolhidos, o SPAD 1 ou o 2 é disparado, ou eles possuem igual probabilidade de acusar uma contagem. Ao último caso corresponde uma escolha de bases conflitante, valor que será descartado durante a reconciliação. Os outros casos corresponderão a uma contagem válida que será guardada para a formação da chave mais tarde. O protocolo BB84 pode então ser aplicado nesse sistema.

A desvantagem do uso desse sistema é que os pulsos passam duas vezes pela mesma fibra e numa delas sua intensidade é forte, gerando espalhamento Rayleigh que pode gerar fótons que induzam contagens falsas. O espalhamento Rayleigh é um fenômeno natural que ocorre em todas as fibras ópticas. Ele é causado pela falta de homogeneidade na estrutura da sílica que compõe a fibra, gerando espalhamento da luz [15]. Na realidade, o limite físico inferior de atenuação em uma fibra é dado por esse espalhamento. Uma pequena fração (≈ 1%) da luz espalhada é recapturada pela fibra na direção contrária à de propagação. Esses fótons podem induzir contagens falsas nos SPADs elevando a taxa de erro do sistema. Essa é a função da linha de atraso DL_2, pois ela evita que pulsos viajando para e de Bob, não estejam presentes no enlace simultaneamente, da seguinte forma: Bob envia pulsos na forma de um trem. Eles são guardados na linha de atraso DL_2, que é simplesmente uma bobina de fibra óptica, ou seja, bem longa. Bob espera até que todo o trem de pulsos enviado por ele esteja de volta para que o próximo trem possa ser enviado. Apesar dessa técnica resolver o problema do espalhamento, ela limita a taxa de repetição do sistema que tem que obedecer ao tempo que o trem leva para ir e voltar, incluindo a passagem pela linha de atraso DL_2.

2.3.6 Codificação por freqüência

A codificação por freqüência utiliza as bandas laterais criadas em uma portadora óptica em Alice, através de uma modulação (fase ou amplitude) gerada a partir de uma portadora de RF (da ordem de GHz). Esse sinal óptico gerado é transmitido pelo enlace até Bob, onde novamente é modulado utilizando a mesma freqüência de RF com sincronismo de fase. Alice e Bob contam com moduladores de fase atuantes no sinal de RF. Dependendo da diferença de fase imposta nesses

moduladores (que será repassada às bandas laterais) interferência destrutiva ou construtiva será gerada nas bandas laterais. Esse sistema pode ser utilizado para a criptografia quântica [3]. Como é o enfoque desse trabalho, a codificação por freqüência será vista com muito mais detalhes no capítulo 4.