

3

Análise de Quantizadores Vetoriais das LSF em Codecs de Voz em Ambientes Ruidosos e Redes IP

Em sistemas operando a baixas taxas de bits, a transmissão de forma eficiente dos parâmetros LSF que modelam a envoltória espectral consome uma parcela significativa da taxa de bits total. Por isso, é muito importante que o esquema de quantização das LSF utilizado tenha o melhor desempenho possível levando-se em conta as condições do sistema em que o codec irá operar. Nessa dissertação, os esquemas de QV além de trabalharem com taxas de bits muito baixas, serão afetados por perda de parâmetros LSF que podem ocorrer em redes IP e por ruídos ambientes como o ruído de fábrica, falatório e ruído branco. Veremos que esse ambiente hostil causará, em determinadas situações, um impacto muito negativo na qualidade da voz codificada.

Este capítulo tem por finalidade avaliar, de forma objetiva, o desempenho dos esquemas de quantização vetorial QVPC4, QVPCP2 e QVSM aplicados no Codec descrito na Seção 2.2 e simulados nas condições de rede descritas neste capítulo. Os esquemas serão avaliados através da distorção espectral média e da porcentagem de *outliers*. Também será determinada a medida PESQ - *Perceptual evaluation of speech quality* que é uma predição da percepção de qualidade que seria obtida por um indivíduo, em teste subjetivo de escuta [5], e que foi padronizada pelo ITU-T como Recomendação P.862.

A Seção 3.1 apresenta a base de dados utilizada no treinamento dos QVs e nas simulações, uma breve descrição da medida PESQ e as condições de rede e de ambiente em que ocorrem os experimentos. A Seção 3.2 mostra os resultados das simulações utilizando locuções em ambientes livres de ruído e a Seção 3.3 apresenta os resultados utilizando locuções em ambientes ruidosos. Finalmente, a Seção 3.4 fornece as principais conclusões do capítulo.

3.1

Condições das Simulações

Nas simulações realizadas nesse trabalho, a base de dados usada para treinamento dos QVs foi produzida por 20 locutores masculinos e 20 femininos, onde cada locutor pronunciou 2 conjuntos de dez frases obtidas de listas foneticamente balanceadas para o português falado no Rio de Janeiro [28], resultando em um total de 800 frases. Os desempenhos foram aferidos utilizando-se um conjunto de 60 frases distintas, também foneticamente balanceadas [28], produzidas por 2 locutores masculinos e 2 femininos, gerando uma coleção de 8992 quadros de LSFs, correspondendo a uma seqüência de $8992 \times 21 = 188832$ bits. Para os testes em ambientes ruidosos foram usados somente 3 tipos de ruído, o ruído gaussiano branco gerado artificialmente e 2 tipos de ruído da base de dados *NOISEX-92*: ruído de fábrica e ruído de falatório. Esses 3 tipos foram selecionados pois representam bem os diversos tipos de ambientes ruidosos encontrados pelos usuários diariamente. Foram usados como critérios de avaliação objetiva a distorção espectral média explicitada na equação 2-7 e a distribuição dos *outliers*. São definidas duas classes de *outliers*: porcentagem de LSFs quantizadas com distorção espectral entre 2 e 4 dB e porcentagem de LSFs quantizadas com distorção espectral superior a 4 dB.

Para avaliar a qualidade da voz em codecs operando sobre redes IP utilizamos a recomendação ITU-T P.862 de avaliação perceptiva de qualidade da voz (*PESQ - Perceptual evaluation of speech quality*), que consiste de uma técnica de medição objetiva para estimar a qualidade subjetiva que seria obtida em testes de escuta [6].

O PESQ compara um sinal original com um sinal degradado, resultante da passagem através de um sistema de comunicação. Numa primeira etapa, uma série de atrasos entre o sinal original e degradado são computados e os sinais são alinhados no tempo. No alinhamento de tempo, o sinal degradado é deslocado no tempo para alinhar-se ao sinal original, e os atrasos são contabilizados tanto nos períodos de fala como nos de silêncio. Baseado no conjunto de atrasos encontrados, os sinais original e degradado, já alinhados, são comparados utilizando-se um modelo perceptivo [5], conforme ilustrado Figura 3.1. A chave desse processo é transformar ambos os sinais em uma forma de representação interna, que é análoga à representação psicofísica do sinal no sistema auditivo humano, levando em conta a frequência perceptiva e a intensidade [5][6].

A saída do PESQ utiliza um modelo cognitivo [6] para fornecer

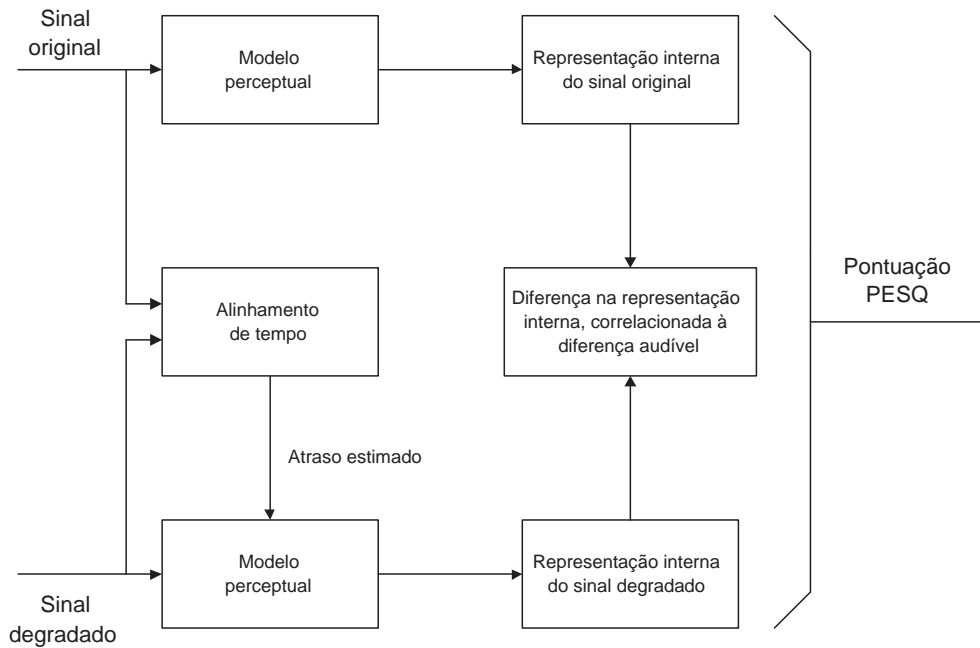


Figura 3.1: Modelo básico do PESQ.

uma predição da percepção de qualidade que seria obtida por uma fala decodificada em um teste de audição subjetivo, como por exemplo MOS (*Mean Opinion Score*). A pontuação do PESQ é mapeada em uma escala tipo MOS, com um limite entre 1,0 e 4,5. Em situações de distorções extremamente altas, o resultado pode ficar abaixo de 1,0, mas isso é muito incomum. Atualmente o PESQ tornou-se um critério bem aceito para medição da qualidade de voz.

Nas simulação da rede IP, usou-se o modelo de perda de quadros descrito na Seção 2.3, esse modelo seguiu as condições de rede consideradas em um trabalho recente [8] que analisa o impacto de perda de pacotes em reconhecedores de voz. Essas condições foram usadas por representarem bem as condições encontradas em uma rede IP, e são fornecidas na Tabela 3.1. As distribuições do comprimento da rajada de perdas de quadros para essas condições são apresentadas na Figura 3.2.

Tabela 3.1: Parâmetros do Modelo de Gilbert usado para simular as condições da Rede com perda de quadros.

$TPQ(\%)$	ppc	B	p	q
0	—	—	0	0
0,6	0,147	1,17	0,005	0,853
9	0,330	1,49	0,066	0,670
28,6	0,500	2,00	0,200	0,500
38,5	0,600	2,50	0,250	0,400

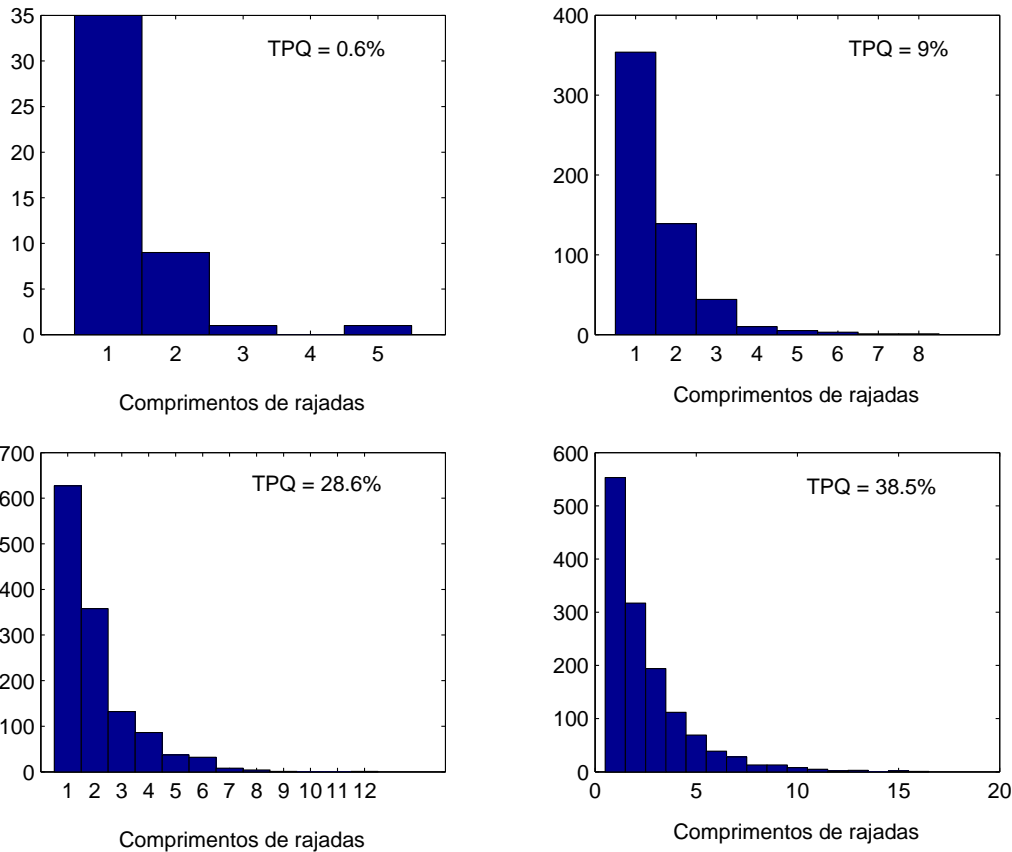


Figura 3.2: Comprimento da rajada de perda de pacotes para os parâmetros de rede apresentados na Tabela 3.1.

3.2

Análise dos Esquemas de QV das LSFs em Ambientes Não-Ruidosos

Nesta seção são apresentados os resultados obtidos utilizando locuções em ambientes não ruidosos, codificadas pelo codec descrito na Seção 2.2 operando em uma rede IP. São testados os esquemas de quantização vetorial QVPC4, QVPCP2 e QVSM, utilizando como medida de desempenho a DE média, o percentual de *outliers* e a recomendação ITU-T P.862 (PESQ).

Os gráficos de distorção espectral *versus* TPQ e porcentagem de *outliers versus* TPQ são mostrados nas Figuras 3.3 e 3.4 respectivamente. Na Figura 3.3, observa-se que o desempenho do QVPC4 é superior a do QVSM para toda a faixa de taxa de perda de quadros considerada e em relação ao QVPCP2 o desempenho é um pouco superior. O melhor desempenho do QVPC4 é confirmado pela Figura 3.4, pois para toda a faixa de interesse, a quantidade de *outliers* com distorção espectral entre 2 e 4 dB do QVPC4 é menor que a dos outros dois quantizadores utilizados nas simulações. Já a quantidade de *outliers* com distorção espectral maior do que 4 dB é comparável para os 3 quantizadores.

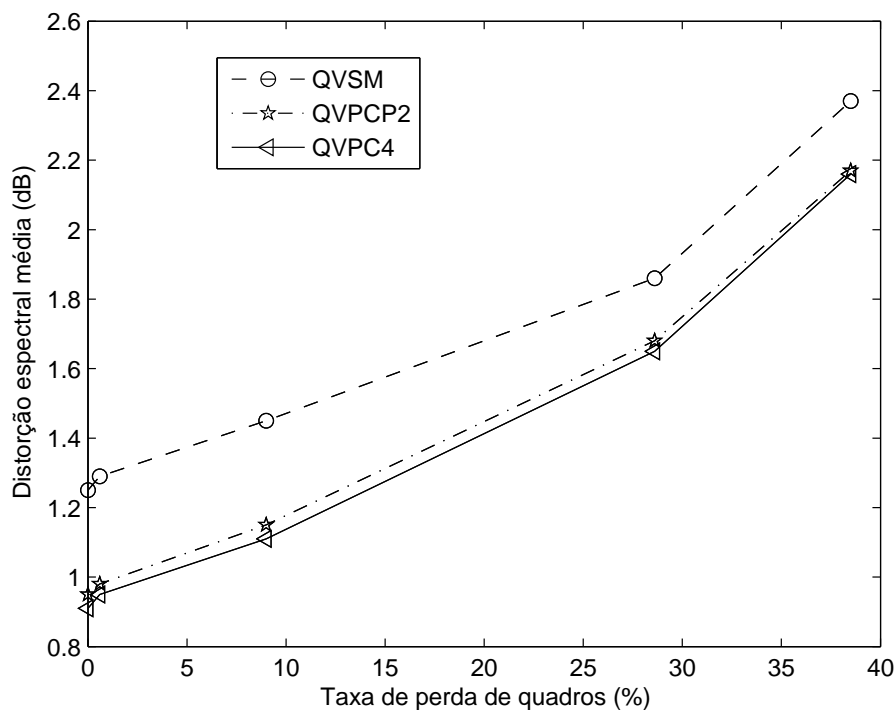


Figura 3.3: Desempenho em termos de DE para TPQ em ambientes não ruidosos.

A Tabela 3.2 e a Figura 3.5 apresentam os resultados em termos do PESQ *versus* TPQ, que mostram a influência da quantização na qualidade da voz em codecs a baixas taxas operando em redes IP. A avaliação feita com o PESQ mostrou que a qualidade da voz utilizando-se os quantizadores vetoriais QVPC4, QVPCP2 e QVSM sofre grande impacto com taxas de perda de pacotes a partir de 28,6%. A essa taxa, começamos a perceber uma maior degradação do sinal de voz. Além disso, o teste com o PESQ confirma que para toda a faixa de taxa de perda de quadro o QVPC4 é o melhor esquema de quantização para codecs a baixas taxas operando

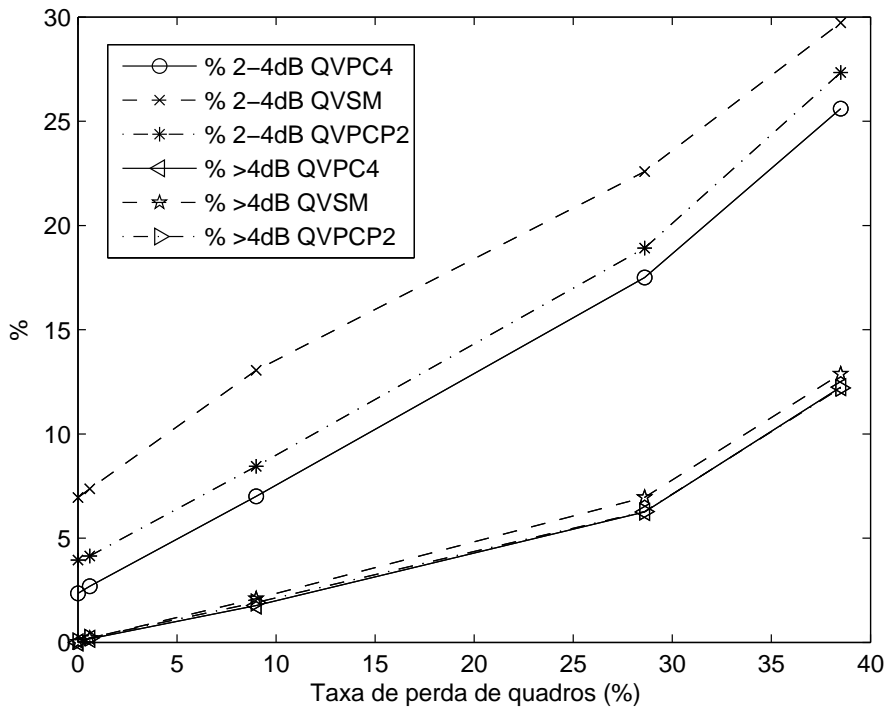


Figura 3.4: Desempenho em termos de *outliers* para TPQ em ambientes não ruidosos.

em redes IP. Entretanto, as diferenças são pequenas quando comparadas ao esquema QVPCP2. Testes de escuta informais validaram os resultados obtidos com o PESQ.

Tabela 3.2: Resultado do teste de qualidade da voz utilizando o PESQ.

TPQ(%)	QVPC4	QVPCP2	QVSM
0	2,747	2,716	2,431
0,6	2,731	2,703	2,437
9	2,710	2,700	2,415
28,6	2,599	2,584	2,350
38,5	2,364	2,346	2,270

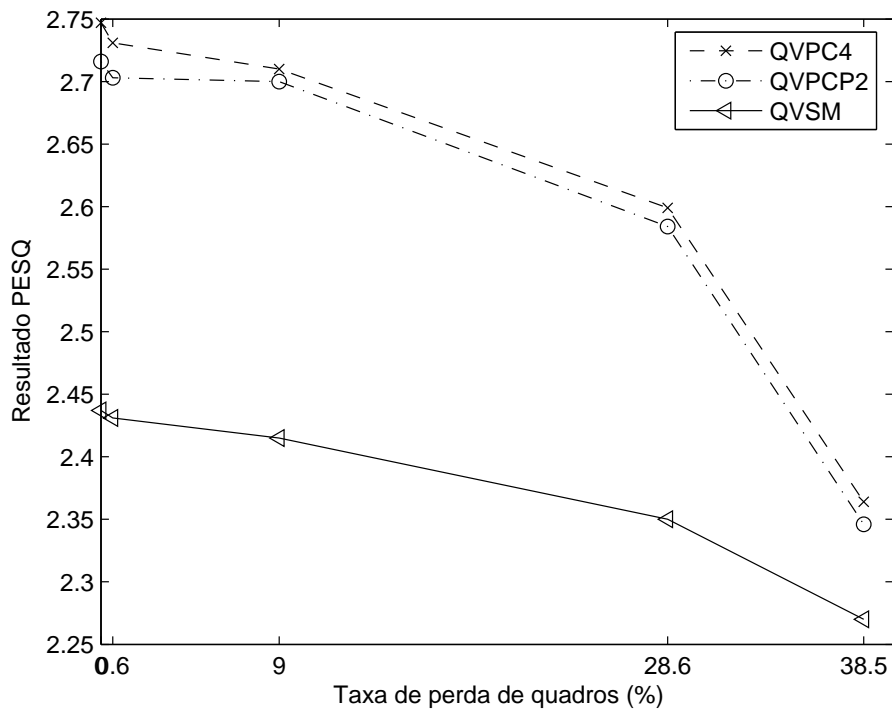


Figura 3.5: Resultado do teste de qualidade da voz utilizando o PESQ *versus* TPQ (%) em ambientes não ruidosos.

3.3

3.3 Análise dos Esquemas de QV das LSF em Ambientes Ruidosos

Nessa seção são apresentados os resultados das simulações utilizando frases em ambientes ruidosos. Os ruídos utilizados foram: ruído de fábrica, ruído de falatório e ruído branco. Novamente, os desempenhos dos QVs foram avaliados através da DE média, do percentual de *outliers* e do PESQ.

3.3.1

3.3.1 Desempenho em termos de Distorção Espectral e Percentual de *Outliers*

Para facilitar a análise do desempenho dos quantizadores vetoriais em termos de DE e percentual de *outliers*, os gráficos apresentarão a DE e *outliers* em função da RSR para os diferentes tipos de ruído considerados. Essa análise é realizada para as seguintes taxas de perda de quadros: 0%, 9%, 28,6% e 38,5%. É importante salientar, que os gráficos relativos à TPQ de 0,6% não são apresentados, pois esta taxa não causou mudanças significativas no desempenho dos QVs quando comparados à TPQ = 0%.

As Figuras 3.6, 3.7 e 3.8 mostram o desempenho dos quantizadores vetoriais (QVPC4, QVPCP2 e QVSM) em termos de Distorção Espectral em ambientes com ruído de fábrica, falatório e ruído branco, respectivamente, e em redes IP sem perda de quadros. Verifica-se através dos gráficos, que o QVPC4 e QVPCP2 obtiveram resultados muito melhores que o QVSM em toda a faixa de variação da RSR ($-5dB$ a $30dB$) nos três tipos de ruídos utilizados. Observamos ainda um pequeno ganho do QVPC4 em relação ao QVPCP2. Nota-se que à medida que a RSR diminui a DE média aumenta e, para o QVSM esse aumento é mais acentuado.

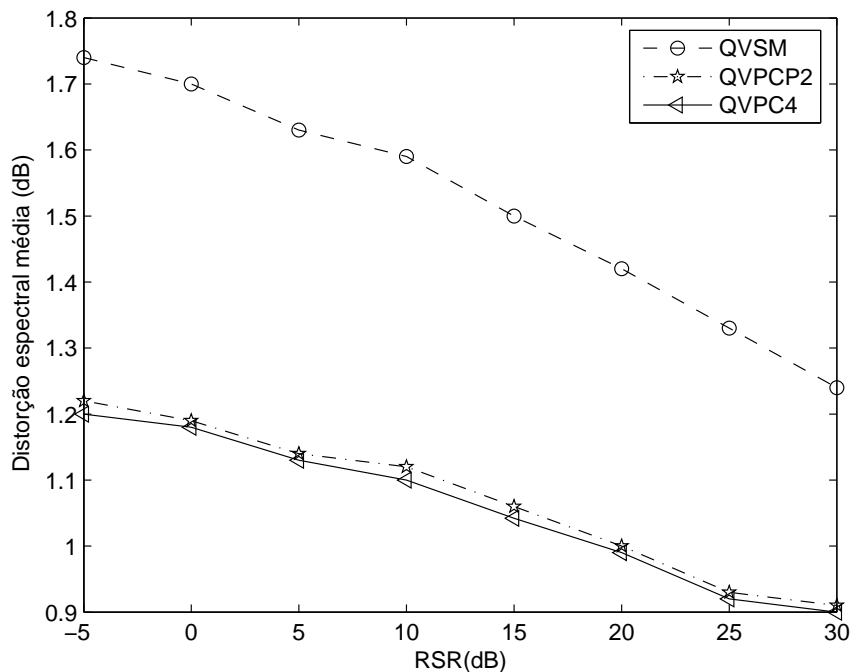


Figura 3.6: Desempenho em termos de DE em diferentes RSR (Ruído de fábrica) e TPQ = 0%.

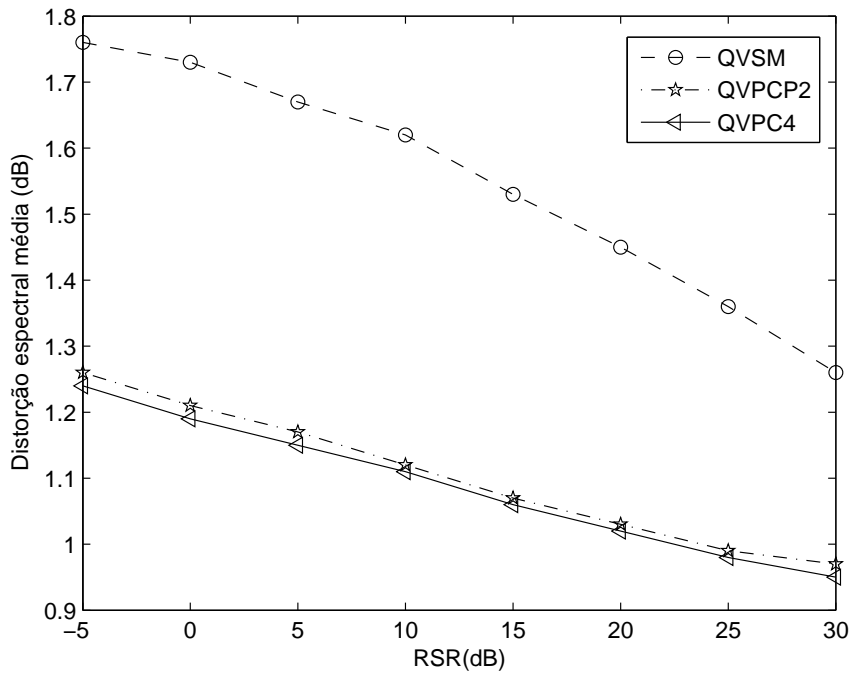


Figura 3.7: Desempenho em termos de DE em diferentes RSR (Ruído de falatório) e TPQ = 0%.

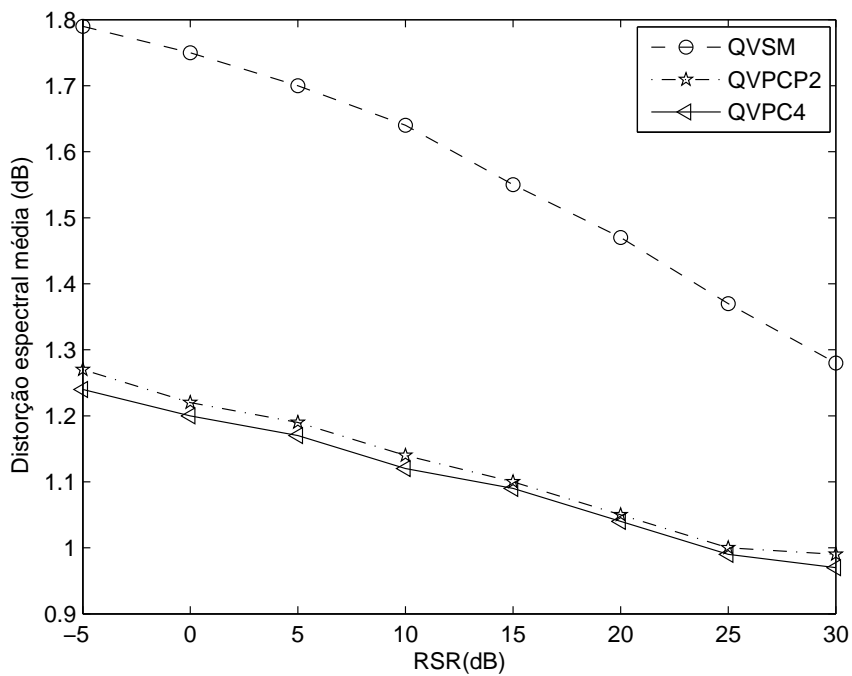


Figura 3.8: Desempenho em termos de DE em diferentes RSR (Ruído branco) e TPQ = 0%.

Os gráficos 3.9, 3.10 3.11, mostram o desempenho em termos de *outliers* nas diferentes RSR. A diminuição da RSR provoca um aumento gradativo de *outliers* em todos os esquemas de quantização vetorial testados nos diversos tipos de ruído. Verifica-se, ainda, que em termos de *outliers* com distorção espectral entre 2 e 4 dB o desempenho do QVPC4 supera os outros esquemas de quantização vetorial nos diversos ambientes ruidosos e com TPQ = 0%. Entretanto, a quantidade de *outliers* com distorção espectral maior do que 4 dB é comparável para os 3 quantizadores.

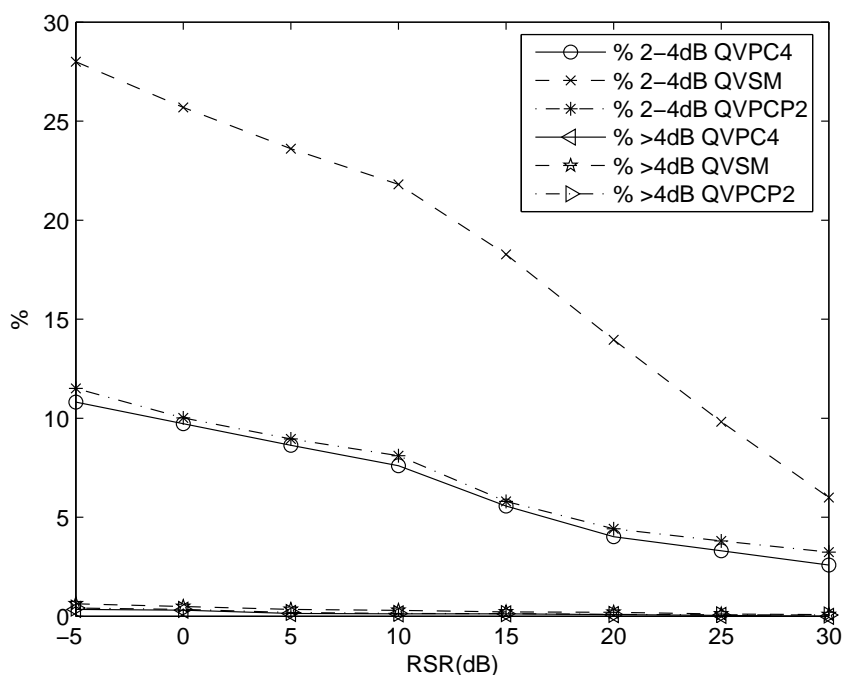


Figura 3.9: Desempenho em termos de *outliers* em diferentes RSR (Ruído de fábrica) e TPQ = 0%.

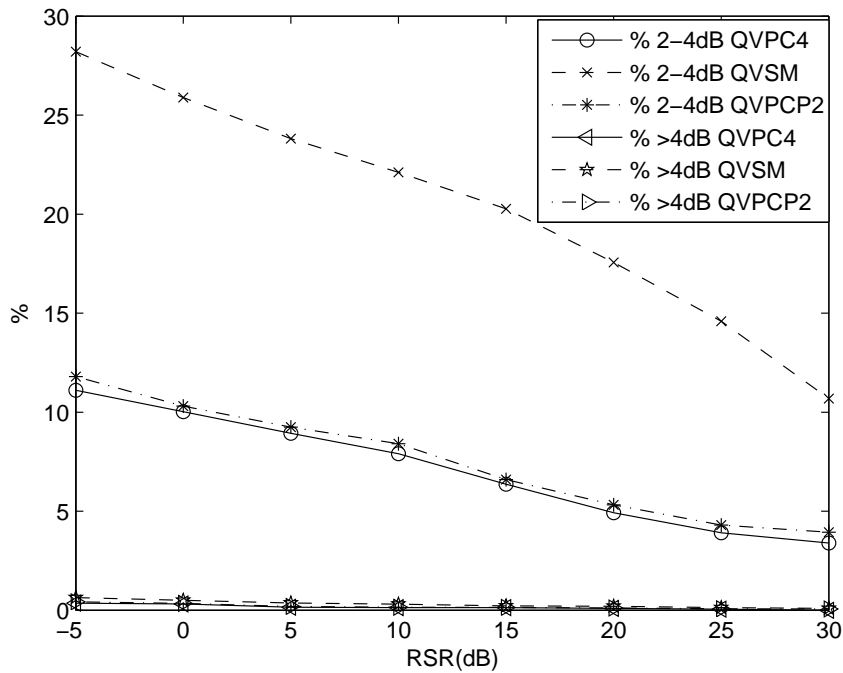


Figura 3.10: Desempenho em termos de *outliers* em diferentes RSR (Ruído de falatório) e TPQ = 0%.

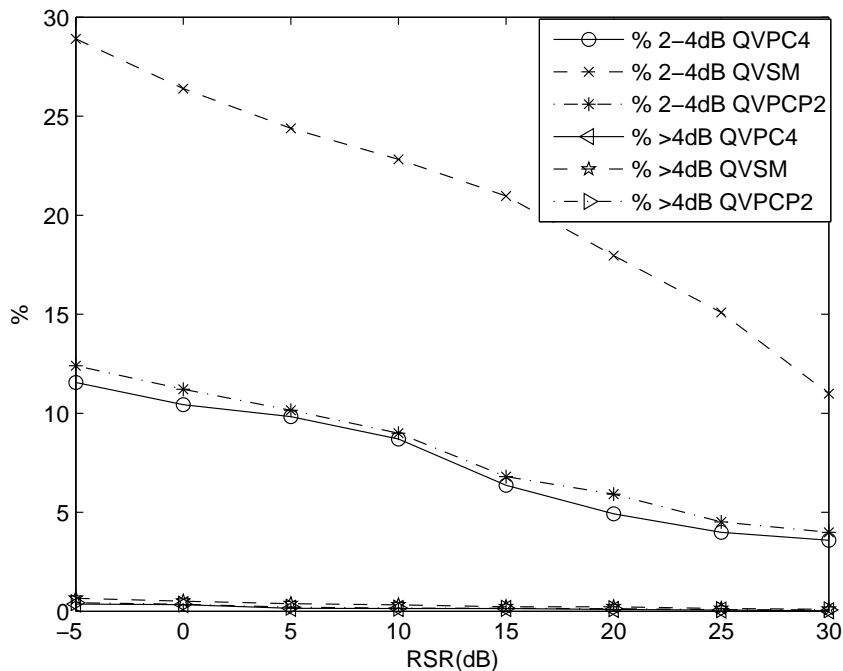


Figura 3.11: Desempenho em termos de *outliers* em diferentes RSR (Ruído de branco) e TPQ = 0%.

Através dos experimentos nota-se que mesmo o codec operando em uma rede IP com taxa de perda de quadros igual a 9% o desempenho dos quantizadores vetoriais, QVPC4 e QVPCP2, é satisfatório, pois apresentam DE com um valor máximo de 1,4dB para uma RSR = -5dB, e esse valor da DE diminui com o aumento da RSR. Isto pode ser visto nas Figuras 3.12, 3.13 e 3.14, onde temos o desempenho dos QVs em termos de Distorção Espectral nos diversos ambientes ruidosos, a uma taxa de perda de quadros igual a 9%. A essa taxa, os QVs têm um desempenho inferior quando comparados com o seu desempenho em um sistema sem perda de quadros. Entretanto, pode-se notar que essa queda de desempenho não é muito significativa. A maior causa de perda de desempenho em termos de DE nas condições de simulação considerada, continua sendo a RSR baixa. É importante destacar, que com a TPQ = 9% o desempenho do esquema QVPC4 continua sendo superior para toda faixa de RSR analisada.

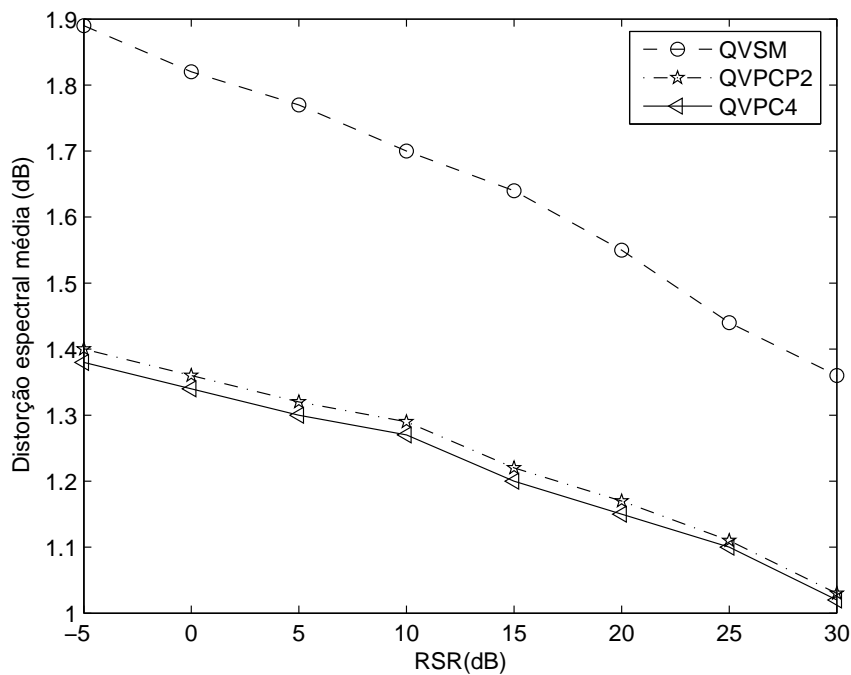


Figura 3.12: Desempenho em termos de DE em diferentes RSR (Ruído de fábrica) e TPQ = 9%.

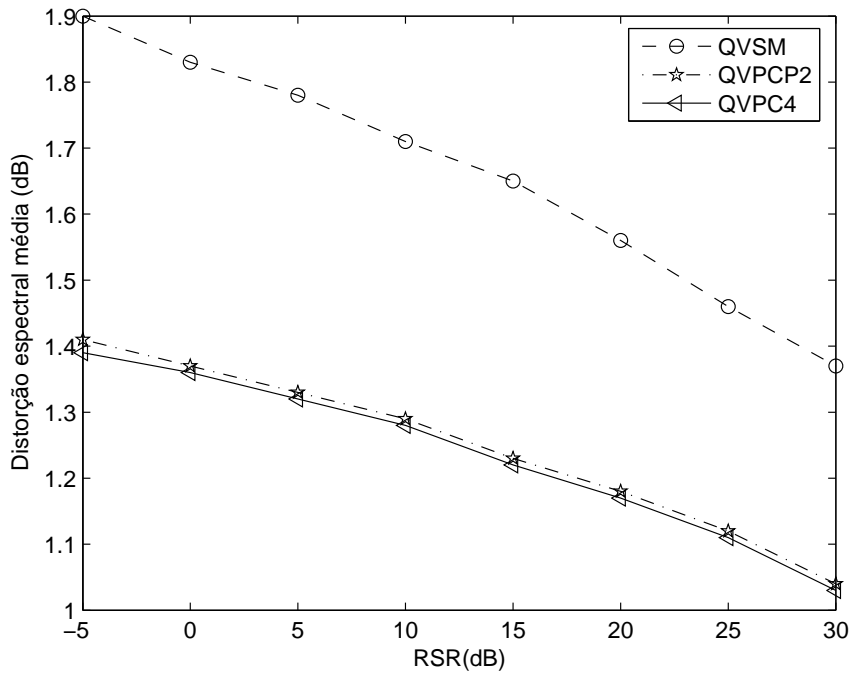


Figura 3.13: Desempenho em termos de DE em diferentes RSR (Ruído de falatório) e TPQ = 9%.

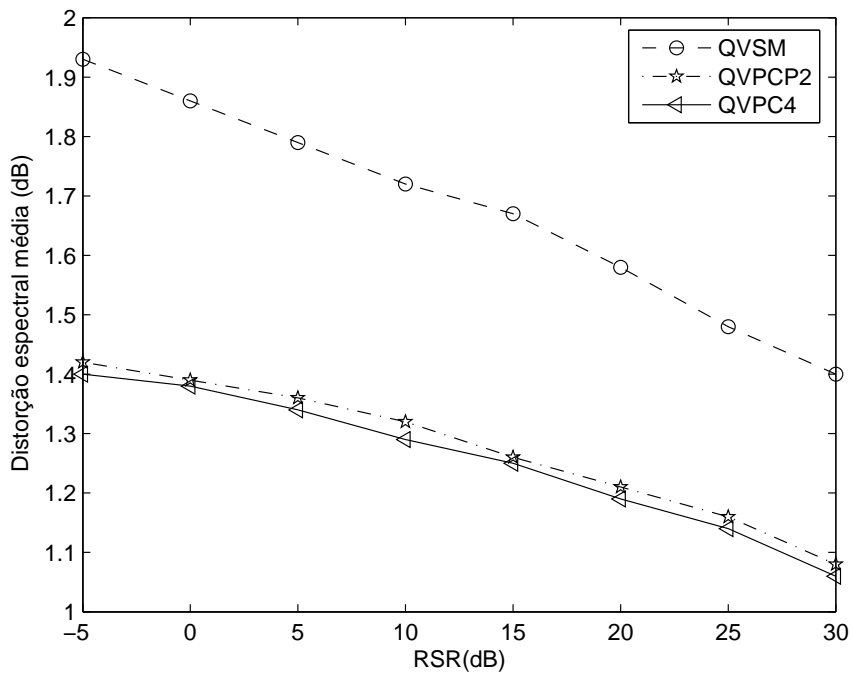


Figura 3.14: Desempenho em termos de DE em diferentes RSR (Ruído de branco) e TPQ = 9%.

O percentual de *outliers* com distorção espectral entre 2 e 4 dB e distorção espectral superior a 4 dB, para frases codificadas em presença de ruído (fábrica, falatório e branco) e com uma TPQ = 9% pode ser observado nas Figuras 3.15, 3.16 3.17. Aqui também verifica-se o melhor desempenho do esquema QVPC4 para toda faixa de RSR analisada para os 3 tipos de ruído utilizados, e o desempenho bastante inferior do QVSM. A essa taxa de perda de quadros, já é possível observar uma sensível queda de desempenho em relação ao percentual de *outliers* em todos os esquemas de quantização vetorial testados. Pode-se observar um grande aumento do percentual de *outliers* com distorção espectral superior a 4 dB. Esse aumento a partir dessa taxa, provavelmente afetará a qualidade da voz codificada. O impacto desse aumento de *outliers* na qualidade da voz, será analisado mais adiante utilizando a medida PESQ descrita no início desse capítulo.

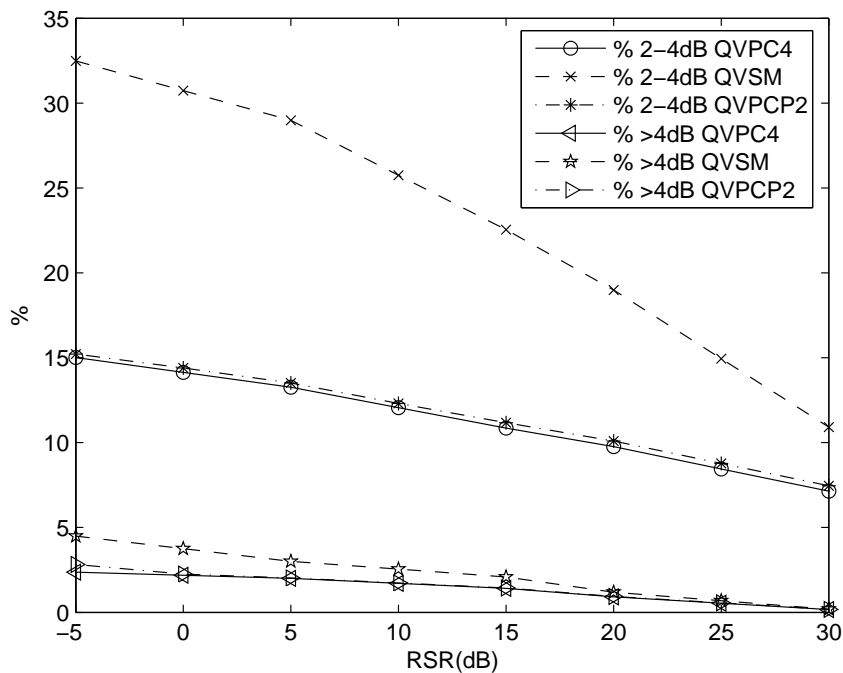


Figura 3.15: Desempenho em termos de *outliers* em diferentes RSR (Ruído de fábrica) e TPQ = 9%.

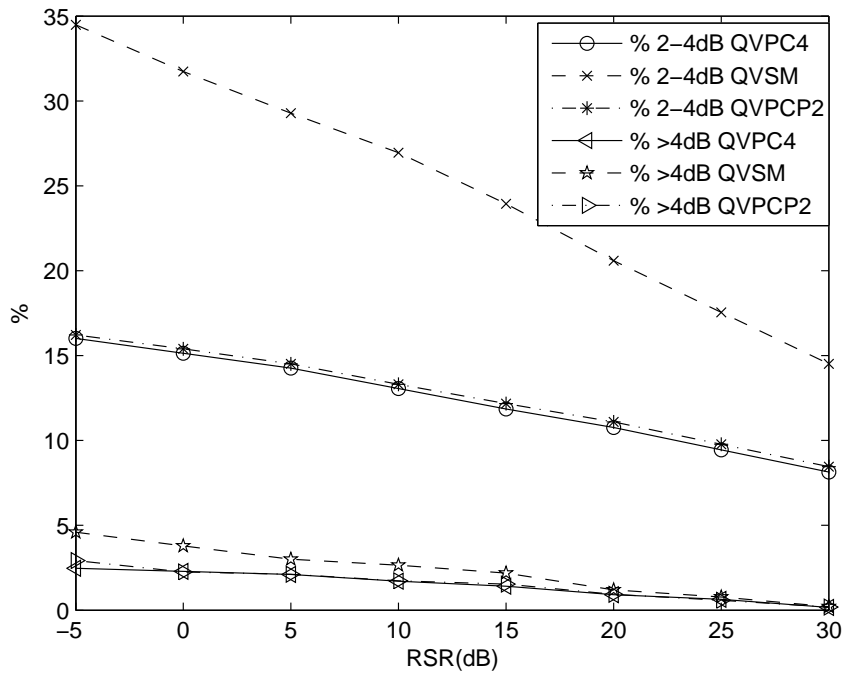


Figura 3.16: Desempenho em termos de *outliers* em diferentes RSR (Ruído de falatório) e TPQ = 9%.

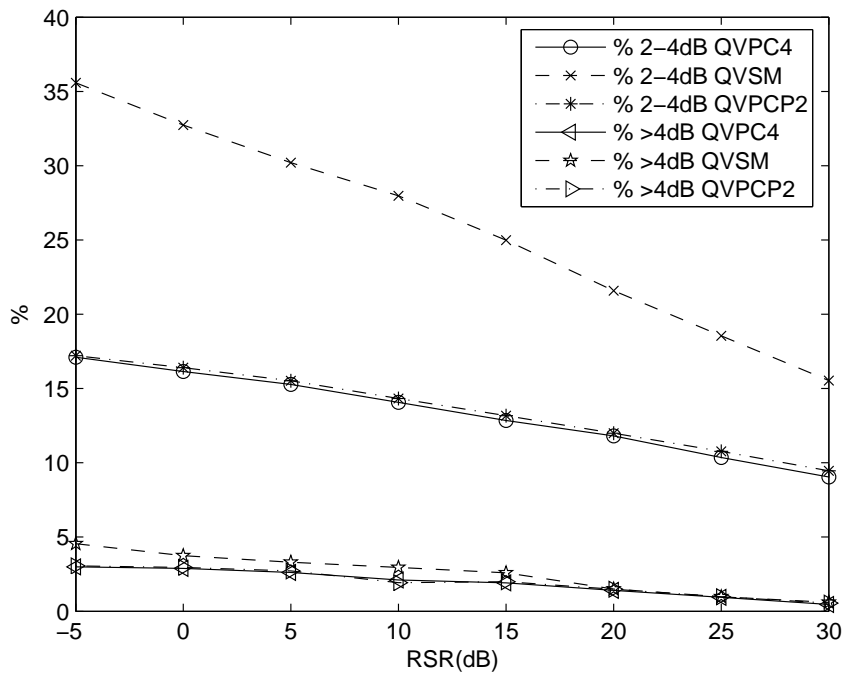


Figura 3.17: Desempenho em termos de *outliers* em diferentes RSR (Ruído de branco) e TPQ = 9%.

Em redes IP, como discutido no Capítulo 2, podem ocorrer taxas de perda de quadros altas e com perdas em rajadas. As Figuras 3.18, 3.19 e 3.20 mostram o desempenho dos quantizadores vetoriais em termos de DE, simulados em uma rede com uma alta taxa de perda de quadros ($TPQ = 28,6\%$). Nessa taxa, observa-se uma queda de desempenho em termos de DE bastante acentuada nos 3 esquemas de quantização vetorial. Essa queda é agravada, também, por uma RSR baixa. O esquema QVPC4 continua obtendo um desempenho muito superior em relação ao QVSM e um desempenho ligeiramente superior ao QVPC2.

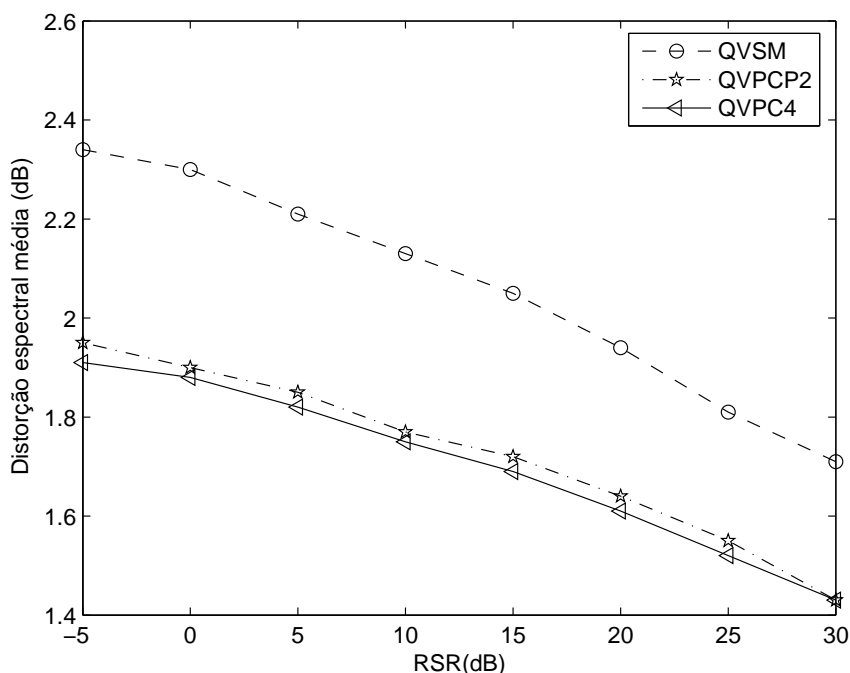


Figura 3.18: Desempenho em termos de DE em diferentes RSR (Ruído de fábrica) e $TPQ = 28,6\%$.

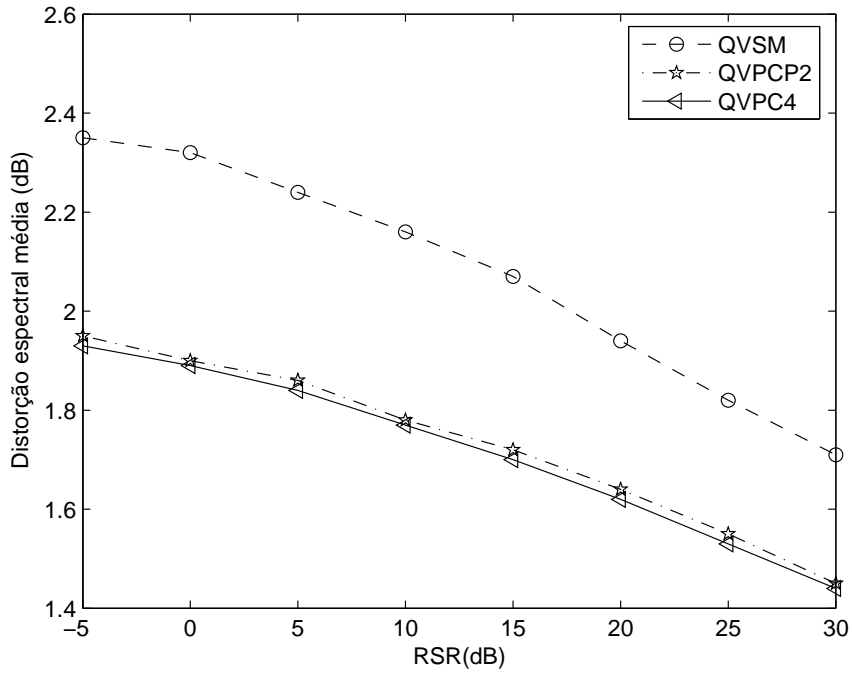


Figura 3.19: Desempenho em termos de DE em diferentes RSR (Ruído de falatório) e TPQ = 28,6%.

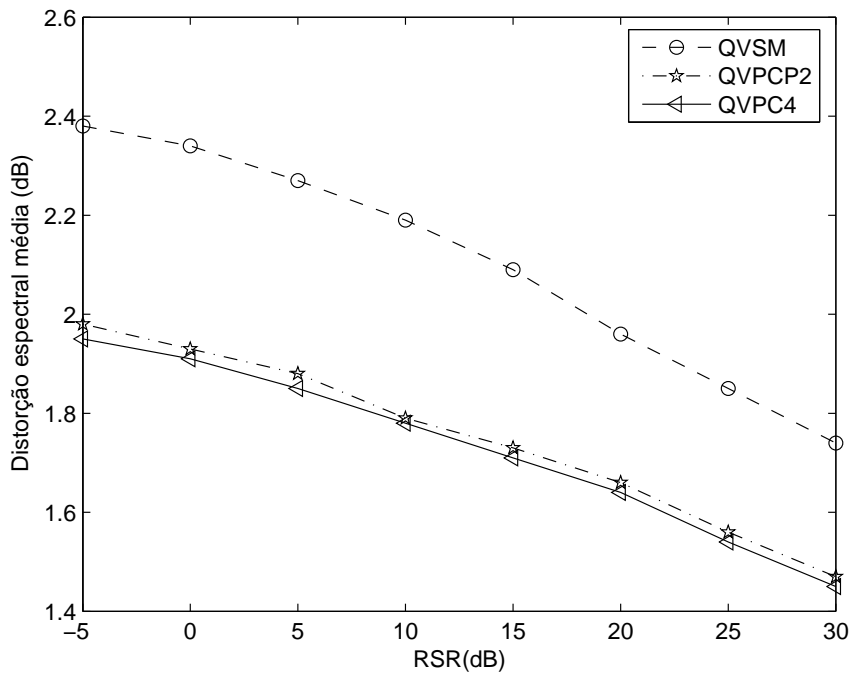


Figura 3.20: Desempenho em termos de DE em diferentes RSR (Ruído de branco) e TPQ = 28,6%.

As Figuras a seguir, 3.21, 3.22 e 3.23, mostram o desempenho bastante ruim dos quantizadores vetoriais, em termos de *outliers*, a uma TPQ = 28,6% . As baixas taxas de RSR também contribuíram para o baixo desempenho dos esquemas testados. Essa alta porcentagem de *outliers*, principalmente com distorções espectrais acima de 4dB, pode prejudicar de maneira significativa a qualidade do sinal codificado.

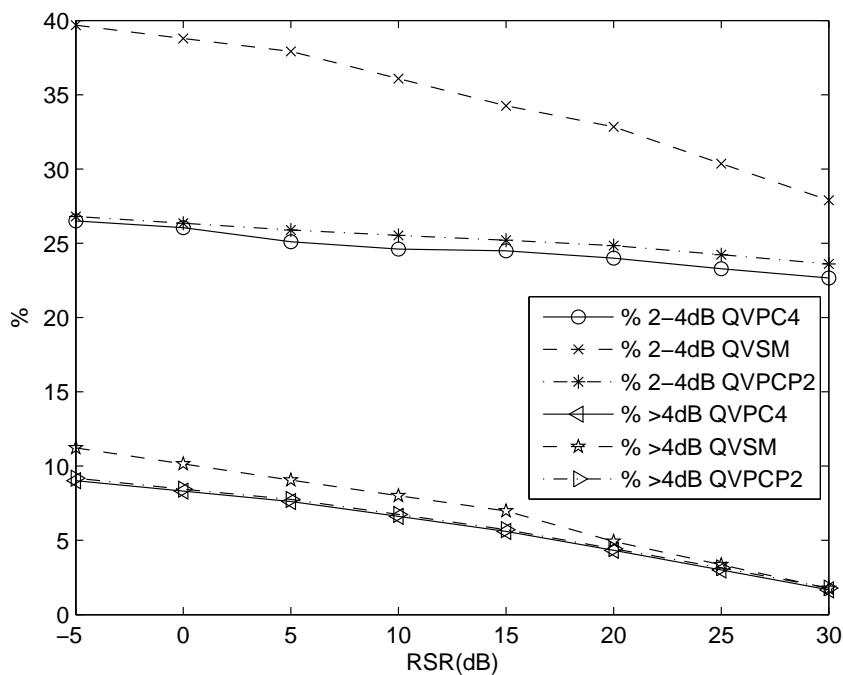


Figura 3.21: Desempenho em termos de *outliers* em diferentes RSR (Ruído de fábrica) e TPQ = 28,6%.

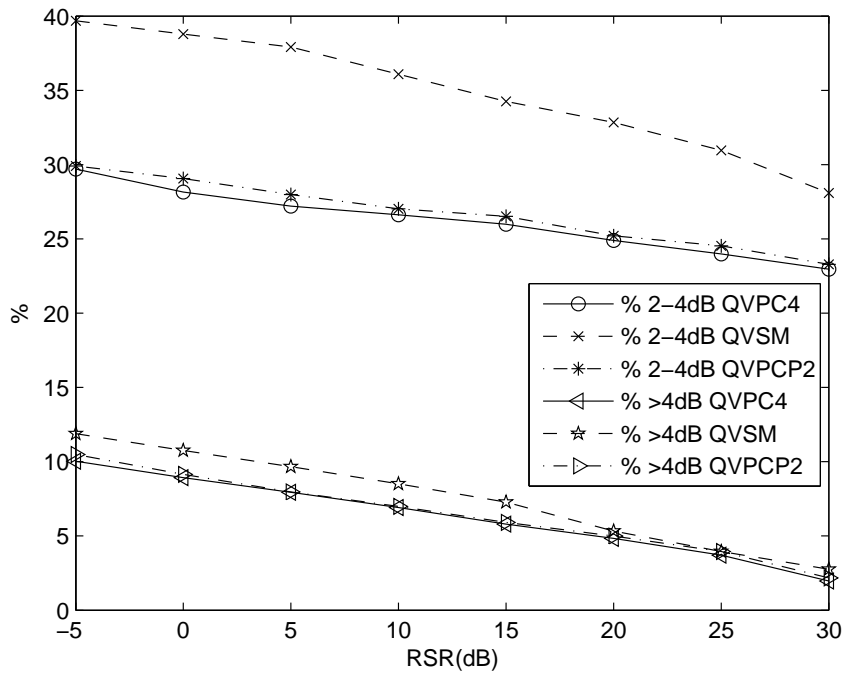


Figura 3.22: Desempenho em termos de *outliers* em diferentes RSR (Ruído de falatório) e TPQ = 28,6%.

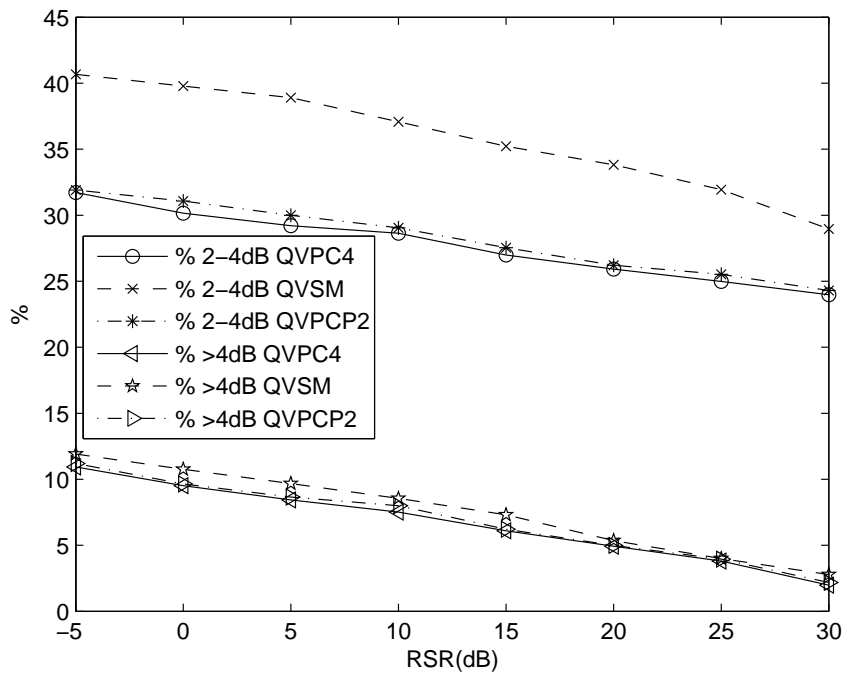


Figura 3.23: Desempenho em termos de *outliers* em diferentes RSR (Ruído de branco) e TPQ = 28,6%.

O cenário mais hostil de perda de quadros usado nessa dissertação foi simulado com uma $TPQ = 38,5\%$. A essa taxa, verifica-se uma enorme perda de desempenho em termos de DE para todos os esquemas de quantização vetorial. Pode-se observar nas Figuras 3.24, 3.25 e 3.26, um desempenho muito fraco dos quantizadores (QVPC4, QVPCP2 e QVSM). Mesmo o QVPC4 que obteve o melhor desempenho em todos os critérios e condições simulados até aqui, apresenta resultados ruins em todos os tipos de ruídos testados, principalmente com RSR baixa (para RSR inferior a $10dB$, a DE é superior a $2,2dB$).

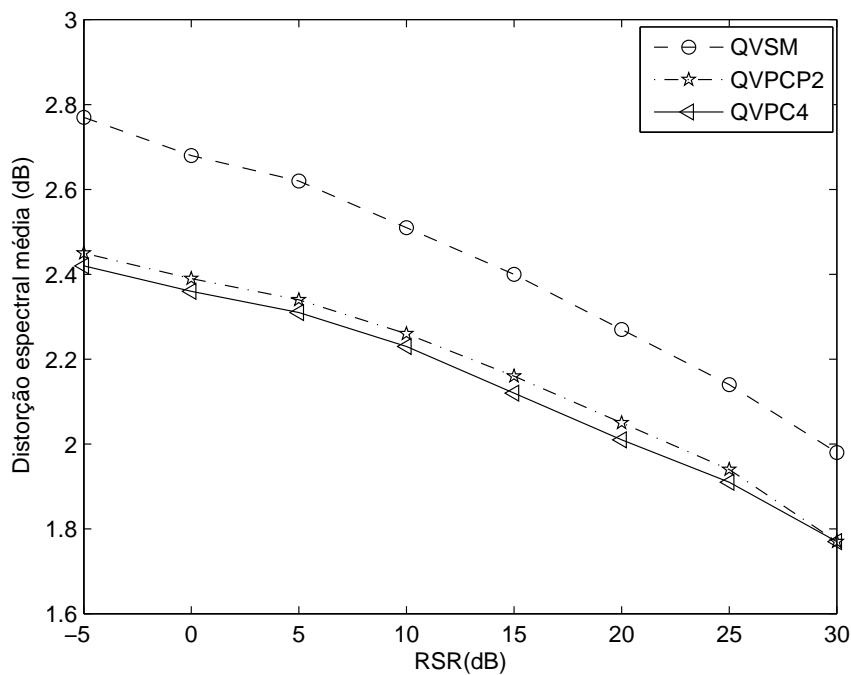


Figura 3.24: Desempenho em termos de DE em diferentes RSR (Ruído de fábrica) e $TPQ = 38,5\%$.

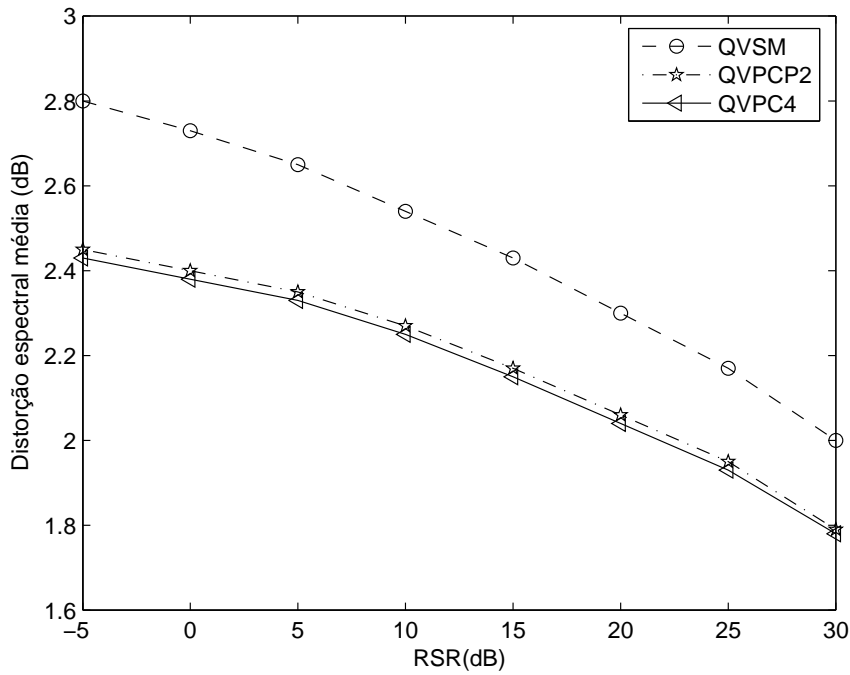


Figura 3.25: Desempenho em termos de DE em diferentes RSR (Ruído de falatório) e TPQ = 38,5%.

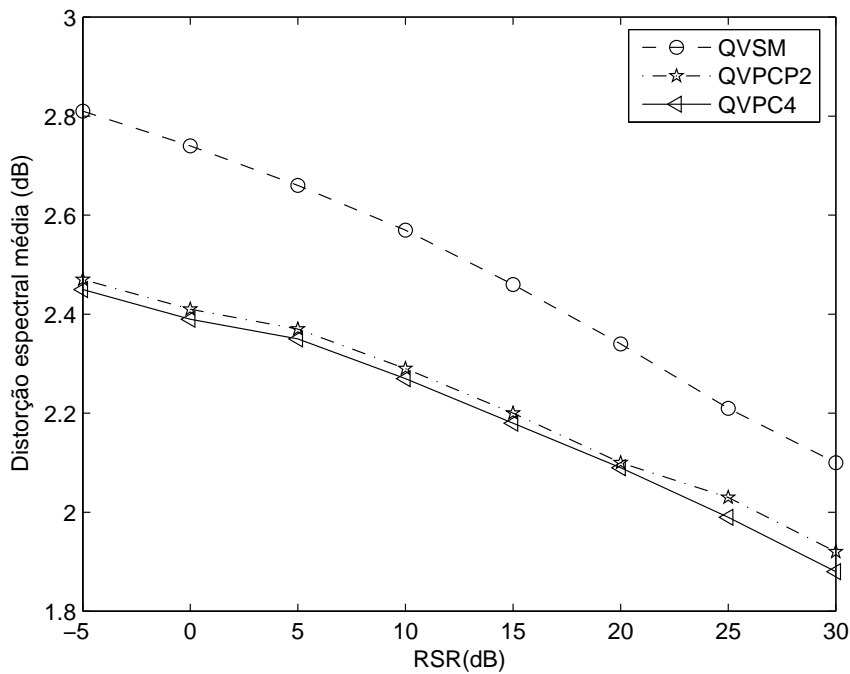


Figura 3.26: Desempenho em termos de DE em diferentes RSR (Ruído de branco) e TPQ = 38,6%.

O desempenho do QVPC4, QVPCP2 e QVSM em relação ao percentual de *outliers* com uma TPQ = 38,5% é mostrado nas Figuras 3.27, 3.28 e 3.29. A essa taxa de perda de quadros, verifica-se um desempenho em relação a percentual de *outliers* muito ruim, para todos os esquemas de QV testados. O codec utilizado nas simulações, trabalhando em uma rede com essa taxa, terá o seu funcionamento comprometido, mesmo usando o QVPC4 que obteve o melhor desempenho para toda faixa de RSR testada. O impacto na qualidade da voz codificada por um codec operando nessas condições, poderá ser melhor observado a seguir, com a utilização da medida PESQ.

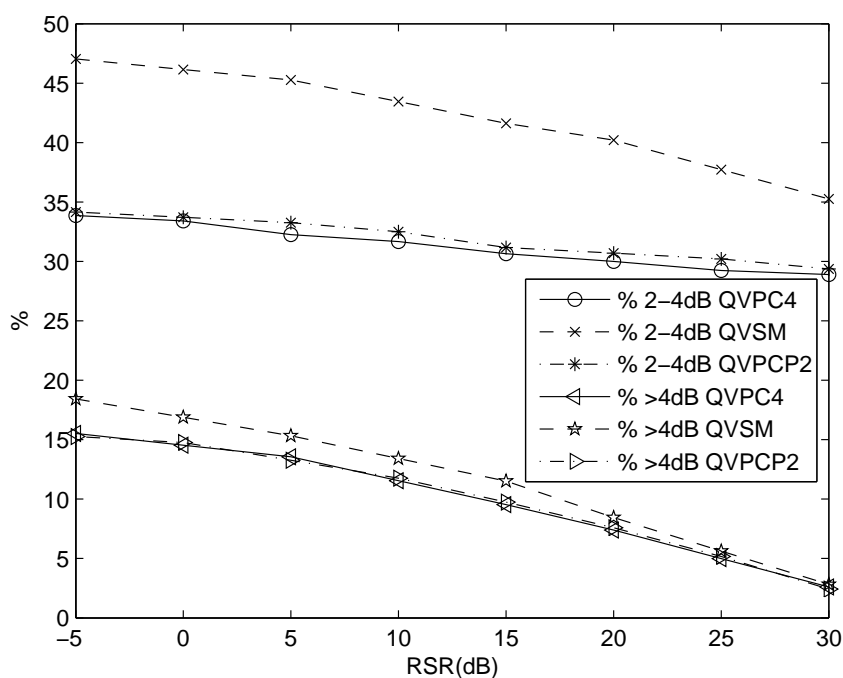


Figura 3.27: Desempenho em termos de *outliers* em diferentes RSR (Ruído de fábrica) e TPQ = 38,5%.

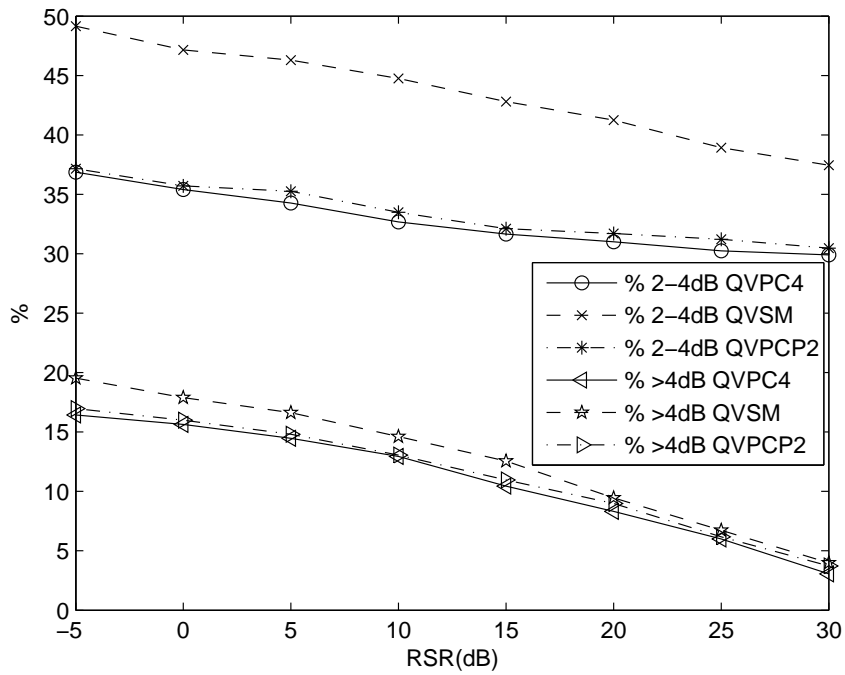


Figura 3.28: Desempenho em termos de *outliers* em diferentes RSR (Ruído de falatório) e TPQ = 38,5%.

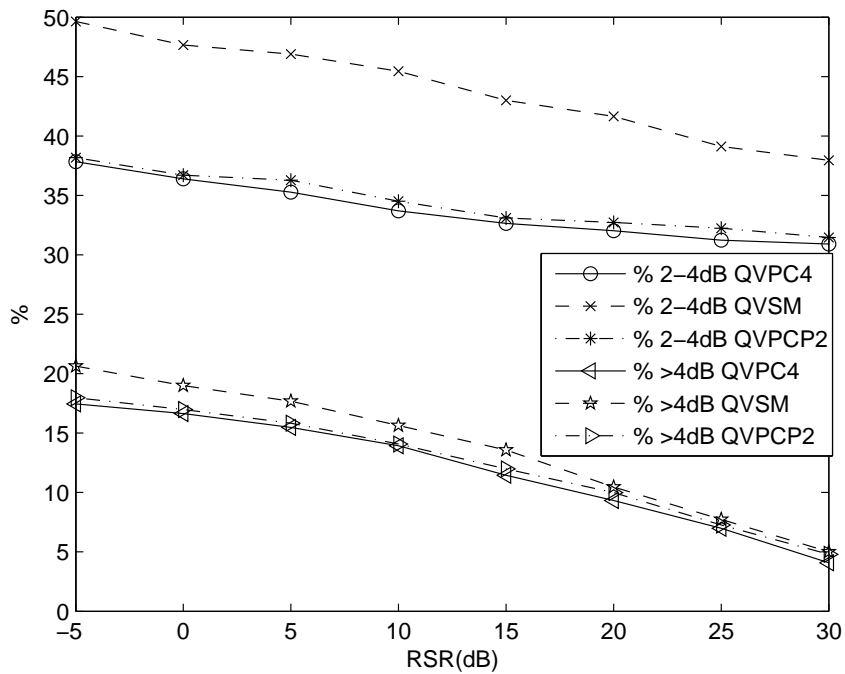


Figura 3.29: Desempenho em termos de *outliers* em diferentes RSR (Ruído de branco) e TPQ = 38,5%.

3.3.2

Desempenho em Termos de Qualidade de Voz Utilizando o PESQ

Esta seção apresenta e discute os resultados obtidos com a utilização da recomendação ITU-T P.862 de avaliação perceptiva de qualidade da voz (PESQ) em vozes codificadas com o Codec descrito na Seção 2.2 operando em uma rede IP e utilizando os esquemas de quantização vetorial QVPC4, QVPCP2 e QVSM. As condições de simulação da Rede IP adotada, seguem o mesmo padrão adotado na Seção 3.3.1, descritas no início deste capítulo. Com o objetivo de facilitar a compreensão e evitar informações desnecessárias, não são apresentados os resultados das simulações obtidos com uma TPQ = 0,6%, uma vez que constatou-se que a essa taxa, os resultados são similares aos resultados obtidos com a TPQ = 0%.

Antes de apresentar os resultados do PESQ para vozes codificadas em ambientes ruidosos e com perda de quadros de LSF, é interessante observar, através da medida PESQ, o impacto causado pelos ruídos de falatório, fábrica e branco em vozes que não sofreram um processo de codificação. Com isso, pode-se avaliar melhor o impacto de ambientes ruidosos na voz codificada. A Figura 3.30 apresenta o resultado PESQ para vozes em ambientes ruidosos antes da codificação. Observa-se que o ruído branco causou a maior degradação na qualidade da voz, enquanto o ruído de fábrica foi o que causou menor impacto, embora as diferenças não sejam grandes . Será analisado mais adiante se isso se confirma para as vozes codificadas. A seguir, são apresentados os resultados das vozes codificadas pelo codec operando em rede IP com diversas TPQ.

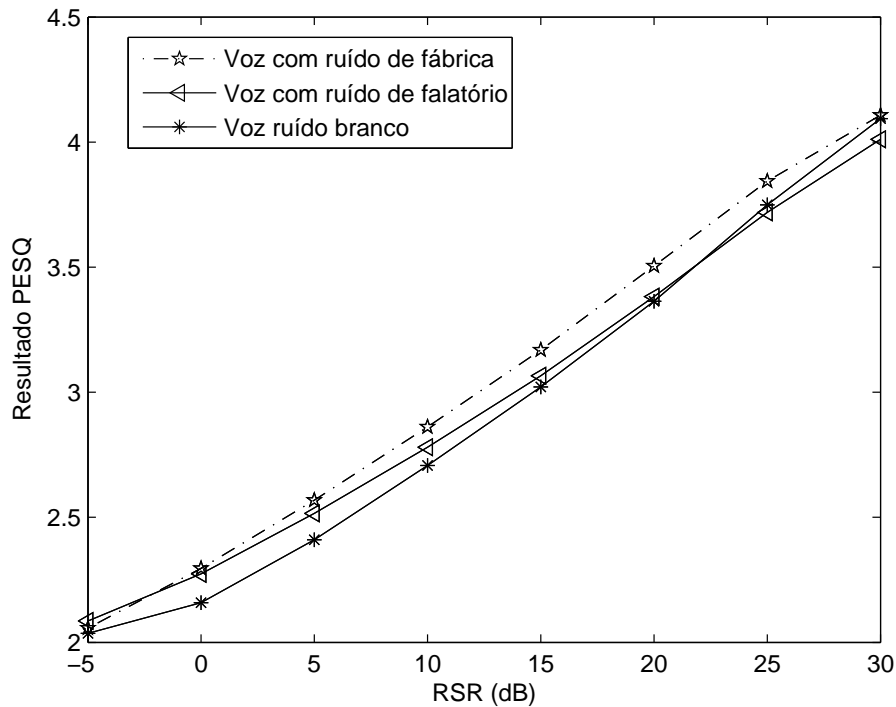


Figura 3.30: Resultado do PESQ em vozes não codificadas em ambientes ruidosos e sem perda de quadros.

A Tabela 3.3 apresenta os resultados do teste de qualidade da voz codificada (utilizando a medida PESQ), com 3 diferentes tipos de ruído, processada por um codec operando em uma rede IP sem perda de quadros e utilizando os esquemas QVPC4, QVPCP2 e QVSM para quantizar os parâmetros LSF. Para RSR baixas, como $-5dB$ e $0dB$ a qualidade do sinal é bastante prejudicada em todos os ambientes ruidosos, independente do quantizador vetorial adotado. Como o codec opera a uma taxa de bits bastante baixa, o que já introduz um ruído de codificação, a simulação em ambientes ruidosos prejudica significativamente o seu desempenho. Entretanto, ao adotar o esquema QVPC4 proposto em [16], obteve-se um desempenho ligeiramente superior ao QVPCP2 em todos os ambientes ruidosos simulados e em toda faixa de RSR analisada. Já o QVSM forneceu um desempenho muito inferior em relação aos outros esquemas.

Ainda é importante ressaltar que para a voz codificada, o ruído de falatório é o mais prejudicial, causando a maior queda de qualidade da voz codificada. Isso pode ser melhor observado na Figura 3.3.2, que apresenta o resultado PESQ para a voz codificada em presença dos 3 tipos de ruído, utilizando o QVPC4 e uma rede IP sem perda de quadros. Esse fato é interessante, pois antes da codificação, como foi visto na Figura 3.30, o ruído

que causa o maior impacto na qualidade da voz é o branco. Isso mostra um melhor desempenho do codec em presença de ruído branco que nos demais ruídos utilizados.

Tabela 3.3: Resultado do teste de qualidade da voz em presença de ruído e com TPQ = 0% utilizando o PESQ.

Ruído	RSR(dB)	-5	0	5	10	15	20	25	30
Falatório	QVPC4	1,833	1,966	2,132	2,275	2,389	2,502	2,590	2,670
	QVPCP2	1,808	1,944	2,112	2,224	2,345	2,466	2,561	2,658
	QVSM	1,727	1,845	2,001	2,098	2,217	2,293	2,347	2,399
Fábrica	QVPC4	1,965	2,112	2,271	2,415	2,536	2,598	2,649	2,696
	QVPCP2	1,953	2,090	2,260	2,402	2,509	2,586	2,638	2,661
	QVSM	1,824	1,946	2,086	2,224	2,306	2,356	2,389	2,392
Branco	QVPC4	2,032	2,155	2,327	2,475	2,597	2,659	2,698	2,733
	QVPCP2	2,022	2,145	2,311	2,461	2,586	2,646	2,680	2,710
	QVSM	1,988	2,032	2,148	2,264	2,353	2,387	2,423	2,430

PUC-Rio - Certificação Digital Nº 0220879/CA

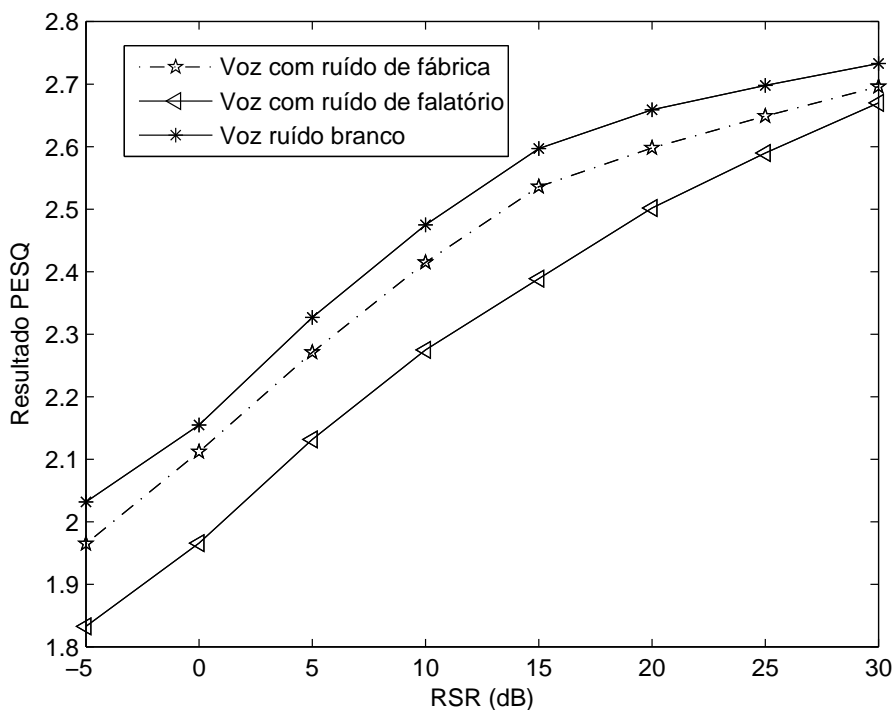


Figura 3.31: Resultado do PESQ com vozes codificadas em ambientes ruidosos utilizando o QVPC4 e sem perda de quadros.

Conforme mostrado na Tabela 3.4, a uma taxa de perda de quadros de 9%, ocorre uma pequena queda na qualidade da voz, com relação a $TPQ = 0\%$, em todas as condições simuladas. O QVPC4 obteve o melhor desempenho em toda faixa de RSR analisada e nos 3 tipos de ruído utilizados. O QVSM foi o que apresentou o pior desempenho. Observa-se, novamente, que o ruído de falatório causou maior degradação na qualidade da voz quando comparado aos outros ruídos, principalmente nas RSR entre $-5dB$ e $5dB$. O ruído branco foi o menos prejudicial, seguido do ruído de fábrica.

Tabela 3.4: Resultado do teste de qualidade da voz em presença de ruído com $TPQ = 9\%$ utilizando o PESQ.

Ruído	$RSR(dB)$	-5	0	5	10	15	20	25	30
Falatório	QVPC4	1,816	1,924	2,093	2,243	2,365	2,481	2,569	2,635
	QVPCP2	1,793	1,912	2,010	2,203	2,342	2,436	2,547	2,629
	QVSM	1,727	1,812	1,960	2,072	2,169	2,277	2,313	2,381
Fábrica	QVPC4	1,942	2,096	2,249	2,380	2,503	2,567	2,617	2,662
	QVPCP2	1,933	2,067	2,248	2,373	2,502	2,564	2,597	2,630
	QVSM	1,800	1,919	2,051	2,164	2,286	2,333	2,362	2,391
Branco	QVPC4	2,026	2,145	2,303	2,459	2,561	2,635	2,670	2,687
	QVPCP2	2,007	2,135	2,298	2,443	2,558	2,619	2,644	2,682
	QVSM	1,980	2,033	2,120	2,238	2,333	2,383	2,400	2,414

Quando ocorrem perda de quadros a taxas mais altas como 28,6% e 38,5% a qualidade da voz começa a cair consideravelmente. Os resultados do PESQ para essas taxas podem ser observados nas Tabelas 3.5 e 3.6 respectivamente. A uma $TPQ = 28,6\%$ a qualidade da voz começa a ficar bastante comprometida em RSR mais baixa que $10dB$, nos diversos ambientes ruidosos testados. Com uma $TPQ = 38,5\%$, o desempenho em relação à qualidade da voz de todos os esquemas de QV fica bastante prejudicado. Esse resultado já era de se esperar, pois a análise utilizando a DE e percentual de *outliers* mostrou um desempenho muito ruim para todos os esquemas testados com a $TPQ = 38,5\%$. Entretanto, vale destacar que o desempenho do QVPC4 é ligeiramente melhor que o QVPCP2 e muito superior ao QVSM, mesmo em situações extremas de perda de quadros e RSR baixas.

Tabela 3.5: Resultado do teste de qualidade da voz em presença de ruído com $TPQ = 28,6\%$ utilizando o PESQ.

Ruído	$RSR(dB)$	-5	0	5	10	15	20	25	30
Falatório	QVPC4	1,764	1,877	2,010	2,153	2,311	2,392	2,482	2,532
	QVPCP2	1,737	1,868	1,992	2,140	2,279	2,360	2,447	2,530
	QVSM	1,681	1,765	1,880	2,021	2,143	2,199	2,264	2,307
Fábrica	QVPC4	1,911	2,009	2,180	2,308	2,411	2,480	2,529	2,557
	QVPCP2	1,900	1,999	2,168	2,296	2,398	2,474	2,513	2,535
	QVSM	1,766	1,844	1,995	2,132	2,209	2,270	2,316	2,332
Branco	QVPC4	2,007	2,091	2,236	2,367	2,474	2,536	2,577	2,587
	QVPCP2	2,003	2,091	2,227	2,365	2,457	2,521	2,553	2,566
	QVSM	1,989	2,008	2,090	2,192	2,260	2,311	2,336	2,339

Tabela 3.6: Resultado do teste de qualidade da voz em presença de ruído com $TPQ = 38,5\%$ utilizando o PESQ.

Ruído	$RSR(dB)$	-5	0	5	10	15	20	25	30
Falatório	QVPC4	1,706	1,807	1,958	2,085	2,209	2,295	2,303	2,340
	QVPCP2	1,690	1,796	1,948	2,076	2,185	2,296	2,300	2,328
	QVSM	1,641	1,704	1,841	1,968	2,068	2,131	2,203	2,238
Fábrica	QVPC4	1,866	1,970	2,105	2,205	2,224	2,287	2,320	2,361
	QVPCP2	1,854	1,947	2,101	2,201	2,216	2,254	2,314	2,321
	QVSM	1,718	1,823	1,952	2,056	2,143	2,196	2,237	2,240
Branco	QVPC4	1,972	2,055	2,176	2,299	2,301	2,340	2,360	2,363
	QVPCP2	1,972	2,042	2,171	2,289	2,295	2,326	2,341	2,346
	QVSM	1,970	1,988	2,059	2,144	2,204	2,230	2,256	2,258

Para melhor visualização e compreensão do impacto da perda de quadros na qualidade da voz, a Figura 3.32 mostra o desempenho em termos do PESQ para as vozes codificadas em presença de ruído de fábrica, utilizando o QVPC4 em diversas taxas de perda de quadros. Vale informar que os comentários a seguir, também são válidos para os outros esquemas de quantização vetorial testados nos diversos tipos de ruídos. Verifica-se que com uma $TPQ = 0,6\%$ a voz codificada, tem qualidade similar à obtida em uma rede sem perda ($TPQ = 0\%$). Já a uma $TPQ = 9\%$, a qualidade tem uma pequena queda, situação já observada através das tabelas apresentadas nessa seção. A partir da $TPQ = 28,6\%$ a qualidade da voz começa a sofrer uma queda significativa. Finalmente, considerando o pior cenário em relação à perda de quadros usado nessa dissertação, tem-se uma voz com qualidade bastante ruim. É importante destacar, que a queda de qualidade da voz codificada devido à perda de quadros de LSF, principalmente a uma $TPQ = 38,5\%$, é muito agravada em situações de RSR baixas.

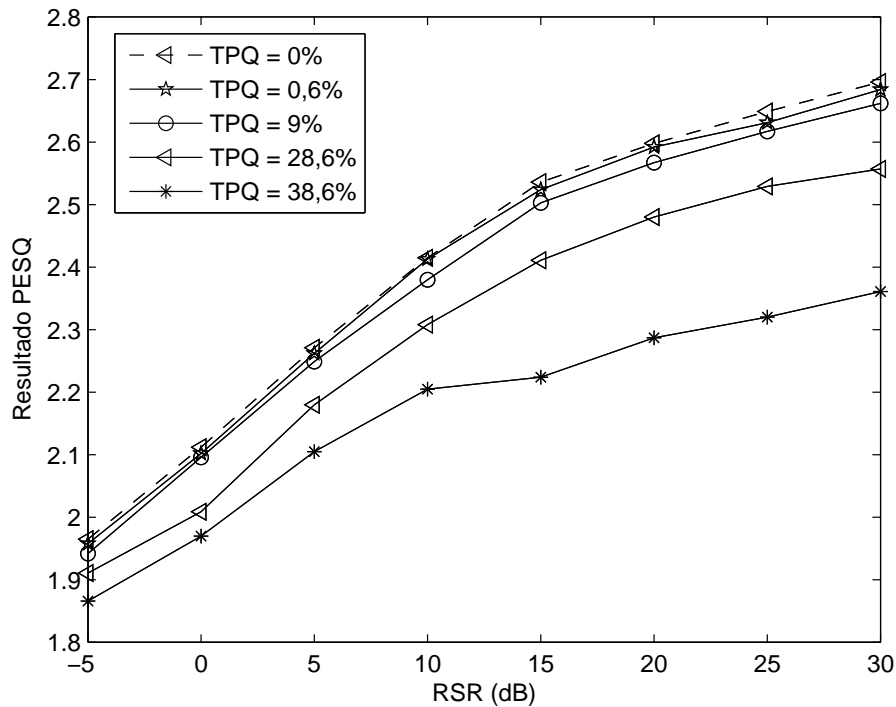


Figura 3.32: Resultado do PESQ para vozes codificadas em presença de ruído de fábrica utilizando QVPC4 e com perda de quadros.

3.4 Conclusões

Este capítulo apresentou a base de dados usada nos experimentos, as condições de simulação do sistema e uma breve descrição da recomendação ITU-T P.862 de avaliação perceptiva de qualidade da voz (PESQ). Em um primeiro momento, avaliou-se o desempenho dos quantizadores vetoriais QVPC4, QVPCP2 e QVSM em ambientes não ruidosos e com o codec operando em uma Rede IP com diversas taxas de perda de quadros. Verificou-se o melhor desempenho do esquema QVPC4 em todos os critérios adotados (DE, percentual de *outliers* e PESQ), ressaltando-se que em taxas de perda de quadros altas, $TPQ = 38,5\%$, o desempenho de todos os esquemas testados é baixo. Posteriormente, essa análise foi estendida para ambientes com diversos tipos de ruído (branco, falatório e fábrica). Foi observado que o ruído de falatório foi o mais prejudicial à qualidade da voz codificada e que RSR baixas (RSR entre $-5dB$ e $10dB$) trazem grande prejuízo ao desempenho dos esquemas de quantização vetorial. Verificou-se, também, que em todos os ambientes, o desempenho do QVPC4 é melhor que o dos outros esquemas testados.