SIGNAL MASKING IN GAUSSIAN CHANNELS

John A. Quinn

Faculty of Computing and I.T.
Makerere University
P.O. Box 7062, Kampala, Uganda

Christopher K.I. Williams

Institute for Adaptive and Neural Computation
University of Edinburgh
EH1 2QL, United Kingdom

ABSTRACT

We consider the problem of modifying the noise properties of a channel in order to make the source as indecipherable as possible given the output. Applications include jamming communications, maintaining confidentiality near spoken conversations and masking noise pollution. We present results as to how this can be done efficiently, assuming that we have a Gaussian channel and a constraint on the power of the noise. We go on to consider the case in which there is a positive signal which we want to remain coherent, as well as a negative signal which we wish to confound. We also discuss the application of the theory to acoustic signals, where we consider aspects of the human auditory system.

Index Terms— Gaussian channels, Signal masking, Acoustic signal processing.

1. INTRODUCTION

The degradation of a source signal by means of adding a noisy "counter-signal" is sometimes desirable. Consider, for example, a scenario in which two conspirators evade eavesdroppers by meeting in a bathroom with the taps running. In this paper, we consider signals sent over a set of parallel Gaussian channels, such that for each channel i, a source component $X_i \sim \mathcal{N}(0, P_i)$ is added to a noise component $S_i \sim \mathcal{N}(0, Q_i)$ to give the output Y_i. This is illustrated in Figure 1(a). In our problem, we want to minimise the mutual information between the source and the output as far as possible. Intuitively, the best results are obtained by adding noise which is close to the signal distribution, and we might naïvely set $Q_i \propto P_i$. However, we show that under a power constraint on S this is suboptimal. In section 2 we present results for generating an optimal counter-signal distribution $\{Q_i\}$ given such a constraint.

We then consider two extensions of this method. In section 3, we look at the case in which there is also a positive signal which we want to remain coherent, as well as a negative signal which we wish to confound. This arrangement is pictured in Figure 1(b), where we now have the positive signal $X_i^+ \sim \mathcal{N}(0, P_i^+)$ and the negative signal $X_i^- \sim \mathcal{N}(0, P_i^-)$. In section 4, we discuss an application of the theory which considers aspects of the human auditory system. We show how these can be incorporated into the suggested scheme in order to add noise to audio signals to reduce their subjective information content.

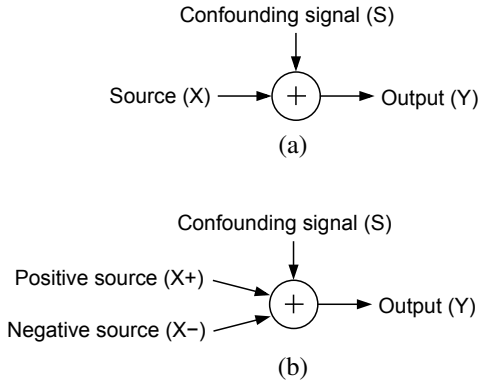

Fig. 1. Two cases considered in this paper. In panel (a), we attempt to confound the source with a noise signal S. In panel (b) there is both a positive source, which we want to preserve, and a negative source, which we want to degrade.

2. MINIMISING MUTUAL INFORMATION IN A GAUSSIAN CHANNEL

In most communication scenarios we want to manipulate the source distribution $P(X)$ in order to maximise the mutual information $I(X;Y)$ between the input X and output Y. Here we want to change $P(S)$ in order to minimise it. Where the input X_i is affected by additive noise to give the output $Y_i = X_i + S_i$, this mutual information is given by

$$I(X;Y) = \sum_i \frac{1}{2} \log \left(1 + \frac{P_i}{Q_i}\right) \quad (1)$$

(see e.g. [1]). If this is subject to the constraint that $\sum_i Q_i = Q$, where Q is some constant, then we can use a Lagrange

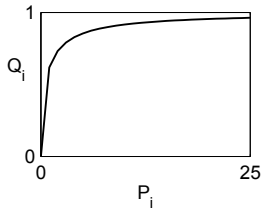

Fig. 2. Power of counter-signal Q_i for different signal strengths P_i, calculated using (5) with $\alpha = 1$.

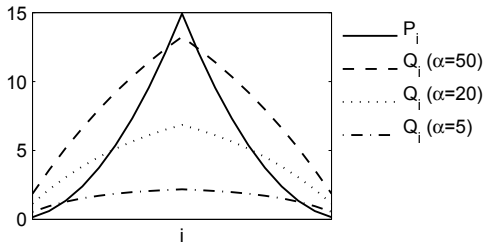

Fig. 3. An example source distribution, with corresponding counter-signal distributions for different settings of α.

multiplier to minimise the function

$$J = I(X;Y) + \lambda \left(\sum_i Q_i - Q \right) \tag{2}$$

where

$$\frac{\partial J}{\partial Q_i} = -\frac{P_i}{2Q_i(Q_i + P_i)} + \lambda \,. \tag{3}$$

Setting this partial derivative to zero and using the substitution $\lambda = \frac{1}{2\alpha}$ gives

$$Q_i^2 + Q_i P_i + \alpha P_i = 0 \,, \tag{4}$$

which has the following solution at the minimum:

$$Q_i = \frac{1}{2}\sqrt{P_i^2 + 4\alpha P_i} - \frac{1}{2} P_i \,. \tag{5}$$

We can therefore try different values of α until the power constraint is satisfied. Clearly the total noise power $\sum_i Q_i$ increases monotonically with α. The relationship is shown in Figure 2 for some sample values. As P_i becomes large, $Q_i \to \alpha$. In Figure 3 we show the effects of altering α for a given source distribution. At low α the counter-signal power is evenly distributed over all channels, whereas at high levels Q_i becomes approximately proportional to $\sqrt{P_i}$.

Note that the channel we examine here is similar to the "dirty-paper" channel [2] (in our case there is no additional noise component), for which Costa derived results about information rates when the masking signal S is known to the agent encoding the source X. We are interested in the inverse problem, where a masking agent has information about the source; however we only assume knowledge of P and not the value of X.

2.1. Limit in the case of infinite noise power

We can also examine the noise distribution as α tends to infinity, where we assume that the power of the noise signal Q is unbounded. In this case we are interested in the relative proportions of each Q_i,

$$\frac{Q_i}{Q_j} = \frac{-P_i + \sqrt{P_i^2 + 4\alpha P_i}}{-P_j + \sqrt{P_j^2 + 4\alpha P_j}} \,. \tag{6}$$

It can be seen that as α becomes large the expression approaches

$$\lim_{\alpha \to \infty} \left(\frac{Q_i}{Q_j} \right) = \sqrt{\frac{P_i}{P_j}} \,. \tag{7}$$

To work out the proportion of power for each channel in this limit, we can see that

$$\lim_{\alpha \to \infty} \left(\frac{Q_i}{\sum_j Q_j} \right) \propto \sqrt{P_i} \,. \tag{8}$$

We therefore have as a limiting case that the optimal strategy is to broadcast a counter-signal where the power at each frequency band is proportional to the square of the power in the signal in that frequency range.

3. DEALING WITH POSITIVE AND NEGATIVE SIGNALS

Given another signal X^+ which we wish to preserve, such that $Y_i = X_i^+ + X_i^- + S_i$, we now want to minimise the information about X^- while maintaining information about X^+. There is a trade-off between these two goals, and different solutions might be appropriate depending on the priority of each. The negative mutual information $-I(X^+;Y)$ can be used to constrain the solution, and the objective function

$$J' \;=\; I(X^-;Y) - \beta I(X^+,Y) + \lambda \left(\sum_i Q_i - Q \right) \tag{9}$$

associates an additional Lagrange multiplier β with this constraint in order to control the trade-off[1]. We can think of this as a specification that we want to maintain some fixed amount of information about X^+ in the output, while minimising $I(X^-;Y)$. In the Gaussian channel case we have

$$I(X^-;Y) = \frac{1}{2} \sum_i \log \left(1 + \frac{P_i^-}{P_i^+ + Q_i} \right) \tag{10}$$

and similarly

$$I(X^+;Y) = \frac{1}{2} \sum_i \log \left(1 + \frac{P_i^+}{P_i^- + Q_i} \right) \tag{11}$$

[1]Other objective functions are possible, though there are drawbacks to each. The ratio $I(X_i^-;Y)/I(X_i^+;Y)$ is differentiable, but it is difficult to find solutions for Q_i in this case. We can also consider removing the constraint associated with λ on the total power in (9).

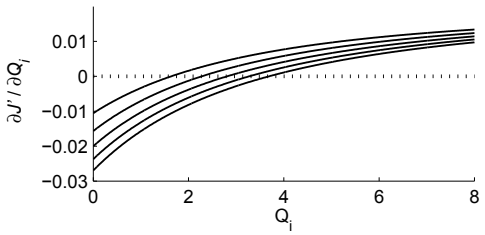

Fig. 4. Example solution for the two-source case. Here $P_i^+ = 5$ for each of the five channels, $P_{1:5}^- = 4, 4.5, 5, 5.5, 6$ (from top to bottom on the plot), $\beta = 0.2$, and $\lambda = 0.02$. Solutions for each of the Q_i are at the points $\frac{\partial J'}{\partial Q_i} = 0$.

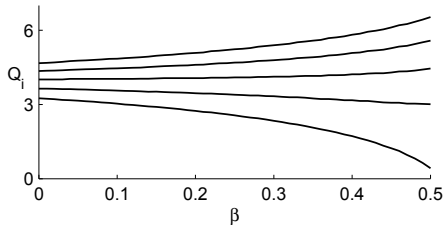

Fig. 5. The effect of varying β on solutions for Q_i, using λ to keep the total power constant ($Q = 20$). P_i^+ and P_i^- are the same as for Figure 4, with channels 1 to 5 appearing from bottom to top.

therefore

$$\frac{\partial J'}{\partial Q_i} = \frac{\beta P_i^+}{2a_i(P_i^- + Q_i)^2} - \frac{P_i^-}{2b_i(P_i^+ + Q_i)^2} + \lambda \qquad (12)$$

where $a_i = 1 + \frac{P_i^+}{P_i^- + Q_i}$ and $b_i = 1 + \frac{P_i^-}{P_i^+ + Q_i}$. In this case we have no analytical solution as with (5). However, solutions can be determined numerically by considering each Q_i independently. Note that when $P_i^+ = 0$, this formulation reduces to the single channel case in section 2.

In general, the solutions for the Q_i's as a function of β, λ and the P_i^+'s and P_i^-'s are quite complex. Solutions for an example case are shown graphically in Figures 4 and 5. In Figure 4 we can see what the consequences of reducing λ would be—effectively to move the function downwards, giving higher solutions for Q_i at $\frac{\partial J'}{\partial Q_i} = 0$. In Figure 5, we can see the effect of changing β, where λ is adapted to keep the total noise power constant. When $\beta = 0$, the situation is similar to the single-source masking case in section 2, P_i^+ being treated only as extra noise. As β is increased, the power is redistributed to channels in which P_i^- is high (i.e. the channels where there is less information about the positive source at the output, being more confounded by the negative source).

4. APPLICATION TO ACOUSTIC SIGNALS

Gaussian channels are often used to model audio signals, where each channel corresponds to a particular frequency bin. We therefore now turn our attention to using the methods developed in the previous sections to minimising the information in an acoustic signal from the perspective of a human listener.

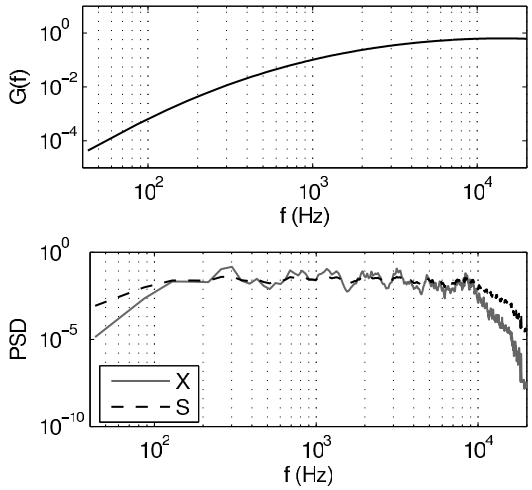

Fig. 6. Top: A-weighting transfer function. Bottom: power spectral densities of a weighted speech signal (solid) and corresponding counter-signal (dashed).

Because of the structure of the ear, not all frequencies are perceived equally. It is intuitively clear that where a signal has high power at a frequency which is undetectable to the ear, there is no need to broadcast a counter-signal at that frequency in order to block it. The A-weighting curve [3] can be used to approximate the subjective loudness of sounds at different frequencies, where the gain at each frequency is given by

$$G(f) = \frac{k_A f^4}{(f + 129.4)^2(f + 676.7)(f + 4636)(f + 76655)^2} \qquad (13)$$

for which the normalising factor is $k_A \approx 7.39 \times 10^9$. We assume that each channel i in the source is associated with a frequency f, and scale accordingly. We can then apply (5) using the mutual information with respect to the approximate relative magnitude of nerve impulses in the cochlea[2].

Figure 6 (bottom) shows the power spectral densities of a male speech signal and corresponding counter-signal. The speech power spectrum is weighted using (13). The counter-signal distribution was calculated using (5) by adjusting α such that the variance of the counter-signal is the same as for the source. It can be seen that the distribution of S is smoother than the source, where there is an averaging effect over different frequencies.

In cases such as the application of maintaining privacy near a spoken conversation, there might be difficulty in working out the relative magnitude of P_i and Q_i. For example, we

[2]Note that we are modelling the system here as $Y_f = G(f)X_f + S_f$, which is a convenient approximation. A more complete analysis would use the formulation of section 3, but with $Y_f = G(f)(S_f^+ + X_f^-) + T_f$, where T is background/cochlear noise.

may not know where a listener is situated, and therefore cannot calculate what mixture of source and counter-signal they will receive. If an estimate is impractical then a good strategy to mask a single source may be to use the limiting case (8).

Where each channel i is associated with a frequency bin, $\{Q_i\}$ defines a power spectral density. Standard methods can be used to generate noise from this power spectrum, for example the generation of sine waves with random phases.

Note that our method assumes a stationary signal, which in general (e.g. for speech masking) may not be the case. In practice it would likely be necessary to resample the distribution of the negative source X^- at intervals and adapt the confounding signal distribution accordingly.

4.1. Other psychoacoustic effects

There are many other pyschoacoustic phenomena which could also be incorporated into a sound masking scheme. There is a known masking effect in which loud constant tones mask quieter tones at different frequencies by exciting the same area of the cochlea. Temporal masking effects could also be exploited [4]. Another effect which we do not consider here is that the perception of the loudness of a sound is non-linear. Each of these effects could be incorporated to quantify a measure of subjective mutual information. See for example work on perceptual entropy in [5].

5. RELATION TO PREVIOUS WORK

The problem of reliably transmitting X^+ while masking X^- in the "dirty-paper" channel is considered in [6], though for the case in which the value of X^+ is known to the encoder which produces S (we only assume knowledge of the distribution P^+). A related case is that in which X and S are controlled by two players, where the 'watermarking game' emerges [7].

A number of proprietary systems exist, marketed as "white noise machines", which perform a similar masking function to that which we consider here. They typically broadcast modulated pink noise and are used for example to decrease distractions in open plan offices (a summary of noise masking systems for this purpose is given in [8]), outside doctors' consulting rooms to maintain privacy or to aid sleep. To the authors' knowledge, none of these systems adapt to the signals which are to be masked.

Another related application is radio jamming, which in practice often involves the use of noise that is white within a particular frequency range. The difference in that application to what we consider here is that it is often not possible to assume that the distribution of the source X is known.

6. DISCUSSION

In this paper we have presented novel results regarding minimising mutual information in a Gaussian channel by calculating the power distribution of an additive counter-signal. These results are intuitive and simple to implement. We have shown how to adapt the counter-signal when there is a positive signal whose coherency we wish to maintain, and suggested how the theory might be extended when we want to mask the signal from the perspective of a human listener by considering psychoacoustic effects.

There are several ways in which this work could be developed. First, we could extend away from the Gaussian channel case by investigating the use of parametric models for the counter-signal in order to make its characteristics closer to particular sources. For example, the best way to mask a speech signal is to generate a noise signal with characteristics close to speech (babble noise). Another extension would be to consider the case where a fixed competing signal is broadcast and one wants to decide whether to use power in degrading the other signal or in broadcasting one's own signal.

For acoustic signals in particular, rather than reducing the information of the source we could also try to minimise the *annoyance*, which has been quantitively defined in terms of spectral balance, fluctuation, roughness and tonality [9]. Such an effect would be useful for example to alleviate the symptoms of tinnitus. In this case we would try to broaden the spectrum of the source to make it similar to stationary pink noise.

Acknowledgements

We thank the anonymous reviewers, whose comments helped to improve this paper.

7. REFERENCES

[1] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley and Sons, 1991.

[2] M.H.M. Costa, "Writing on dirty paper," *IEEE Trans Inf Theory*, vol. 9, no. 3, pp. 439–441, 1983.

[3] M.T. Smith, *Audio Engineer's Reference Book, 2nd Ed*, Focal Press, 1999.

[4] B.C.J. Moore, *An Introduction to the Psychology of Hearing, 5th ed*, Elsevier Press, 2003.

[5] J.D. Johnston, "Estimation of perceptual entropy using noise masking criteria," *ICASSP*, 1988.

[6] N. Merhav and S. Shamai, "Information rates subjected to state masking," *IEEE Trans Inf Theory*, vol. 253, no. 6, pp. 2254–2261, 2007.

[7] P. Moulin and M. Kıvanç Mıhçak, "The Parallel-Gaussian Watermarking Game," *IEEE Trans Inf Theory*, vol. 9, no. 3, pp. 439–441, 2004.

[8] M. Navai and J.A. Veitch, "Acoustic Satisfaction in Open-Plan Offices: Review and Recommendations," Tech. Rep., Institute for Research in Construction, 2003.

[9] K. Waye and E. Öhrstrom, "Psycho-acoustic characters of relevance for annoyance of wind turbine noise," *J Sound and Vibration*, vol. 250, no. 1, pp. 65–73, 2002.