Joint Communication and State Estimation over a State-Dependent Fading Gaussian Channel

Joint Communication and State Estimation Over a State-Dependent Fading Gaussian Channel

Viswanathan Ramachandran

***Abstract*—A joint message and state communication problem over a fading Gaussian channel with additive state is considered. The state process is assumed to be independent and identically distributed (i.i.d.) Gaussian, and known non-causally at the transmitter, while the instantaneous fading gain is assumed to be known at both the encoder and the decoder. The receiver not only has to decode the message from the transmitter, but also needs to estimate the state process to meet a prescribed distortion limit. A complete characterization of the optimal rate versus distortion performance is provided. Moreover, the converse proof presented here might be of independent interest, since it gives a simpler and more compact way of establishing the converse for settings with non-fading links studied in the literature – by virtue of relying only upon the differential entropy maximizing property of Gaussian random variables for a given variance.**

***Index Terms*—State estimation, MMSE distortion, fading channels, dirty paper coding, Gelfand-Pinsker, rate-distortion.**

I. Introduction

CHANNELS with state are used as an abstraction for communication settings in which the channel statistics are not fully known, and captured by an external *state* process. The communication rates achievable over a state-dependent channel depend upon the degree of state knowledge at the sender and the receiver. For instance, the channel state information might be available non-causally at the transmitter, but not at the receiver – this case was addressed in the seminal paper [1] for a discrete memoryless state-dependent channel. Reference [2] considered an additive white Gaussian noise counterpart of [1], with additive Gaussian state that is fully known to the sender but not the receiver. It was shown that by leveraging this known Gaussian noise rather than attempting to cancel it, one can make sure that it does not impact the channel capacity.

Reference [3] studied a state-controlled communication setting in which the dual goals of revealing the channel state to the receiver in addition to message communication must be met. The optimal trade-off region between message transmission rate and state estimation distortion was derived therein for the additive Gaussian case. The corresponding problem for discrete alphabets was investigated in [4], using a metric known as *state uncertainty reduction rate*.

The work of [3] has been extended in several directions. The dual problem where the sender attempts to conceal the state instead of conveying it, was addressed in [5]. Reference [6] considered a variant where only noisy state observations are available at the transmitter. More recently, [7] analyzed the benefits of feedback for simultaneous message and state communication over memoryless channels with memoryless states. State estimation over quantum channels was investigated in [8], while [9] analyzed it in a coordination framework. Very recently, [10] studied integrated sensing and communication from a more holistic viewpoint, focused on 6G applications. The framework has also been extended to multi-user channels – [11] studied joint state estimation and message communication over a state-dependent Gaussian broadcast channel (BC), [12] analyzed the same setup without additive state, while [13] investigated state estimation for a discrete memoryless BC with causal state information.

Notice that none of the aforementioned works consider fading links. In practical wireless communication channels, fading is an impairment that must be taken into account. This is the key gap in the literature which this letter addresses, by investigating joint communication and estimation over a fading channel. We note that fading channels (without any state process or state estimation) with channel gains known at both the sender and the receiver have been addressed in [14]. For a multiple sender, one-receiver (multiple access) extension of the model in [14], it is reasonable to assume that one of the (physically adjacent) senders could cooperate by making available its transmit codeword to its neighbour. This results in a fading channel (with channel gain known at both parties) with non-causal state information (where the state is nothing but the known interference from the neighboring sender's codeword) only at the transmitter. The receiver here is interested in recovering the intended message as well as obtaining an estimate of the other user's codeword. The main concern of the current paper is joint state estimation and message communication over such a channel – as such, the current work extends the framework of [3] to incorporate fading links.

A complete characterization of the optimal trade-off between message communication rate and state estimation distortion is the main contribution of this letter (see Theorem 1 in Section II). The key novel aspects here lie in the proof of converse – see Remark 1 and the ensuing explanation thereof in Section IV for the details. Moreover, the converse proof in this letter presents an alternative simpler way to establish the converse for [3, Th. 2].

Notations: Random variables/vectors are denoted by upper-case letters, while their realizations are denoted by the corresponding lower case letters. All logarithms are assumed to be with respect to base 2. A sequence $(X_1, X_2, \ldots, X_n)$ is denoted by X^n. $||\cdot||$ denotes the Euclidean norm of a vector.

Manuscript received October 25, 2021; accepted November 14, 2021. Date of publication November 17, 2021; date of current version February 17, 2022. The associate editor coordinating the review of this article and approving it for publication was L. Yang.

The author is with the EE Department, TU Eindhoven, 5612 AZ Eindhoven, The Netherlands (e-mail: v.ramachandran@tue.nl).

Digital Object Identifier 10.1109/LWC.2021.3128619

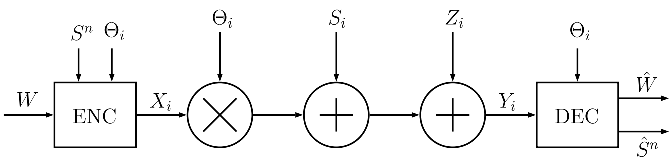

Fig. 1. Channel Model.

II. System Model and Results

Consider the system shown in Fig. 1. It consists of a single-user state-dependent Gaussian channel with fading as well as non-causal state information S^n at the transmitter, with *n* being the number of channel uses. The instantaneous fading gain $\Theta_i, i \in [1, 2, \ldots, n]$, is assumed to be known to both the transmitter and the receiver. A block-fading model is assumed, where the power constraint on the input is across blocks. The dual goals of message (*W* in Fig. 1) communication, and estimation of S^n to meet a distortion tolerance with respect to a squared error metric, must be met. The optimal trade-off between the average message communication rate (*R*) and the average distortion incurred in state estimation (*D*) over the fading process is sought. A single use of the channel is described by:

$$Y_i = \Theta_i X_i + S_i + Z_i, \;\; i \in [1, 2, \ldots, n], \tag{1}$$

where Θ^n is the fading process, S^n is the state process and Z^n is the noise process, which are independent of each other. The state process S^n is known only at the encoder, while the instantaneous fading gain Θ_i is known at both parties. The state and noise processes are i.i.d. Gaussian with zero mean and respective variances *Q* and σ_Z^2. Assuming the message *W* to be uniform over $\{1, \ldots, 2^{nR}\}$, the power constraint on the transmissions ($\forall\; w \in \{1, \ldots, 2^{nR}\}$) is:

$$\frac{1}{n}\mathbb{E}_\Theta\left[\mathbb{E}_{S^n}\left[\sum_{i=1}^{n} x_i^2(w, \theta, s^n)\right]\right] \leq \mathrm{P}. \tag{2}$$

Thus, for a given realization $\Theta = \theta$, $\mathbb{E}_{S^n}[||X^n||^2] = nP(\theta)$, and $\mathbb{E}_\Theta[\mathbb{E}_{S^n}[||X^n||^2]] \leq n\mathrm{P}$ averaging across the fading.

Definition 1: An (n, R, D, ϵ) state amplification scheme consists of an encoder map $\mathcal{E} : \{1, 2, \ldots, 2^{nR}\} \times \mathcal{S}^n \times \Theta \to \mathcal{X}$, a decoding map $\psi : \mathcal{Y}^n \times \Theta^n \to \{1, 2, \ldots, 2^{nR}\}$, and a receiver reconstruction map $\phi : \mathcal{Y}^n \times \Theta^n \to \mathbb{R}^n$ such that for *W* uniformly distributed over $\{1, 2, \ldots, 2^{nR}\}$, and $X_i = \mathcal{E}(W, S^n, \Theta_i)$, $i \in [1, 2, \ldots, n]$, the following holds:

$$\mathbb{E}_{\Theta^n}\left\{\frac{\mathbb{E}\left[\|S^n - \phi(Y^n, \Theta^n)\|^2\right]}{n}\right\} \leq D + \epsilon, \tag{3}$$

$$\mathbb{E}_{\Theta^n}\{\mathbb{P}(\psi(Y^n, \Theta^n) \neq W)\} \leq \epsilon, \tag{4}$$

under an average power constraint $\mathbb{E}_\Theta[\mathbb{E}_{S^n}[||X^n|]]^2] \leq n\mathrm{P}$.

A pair (*R*, *D*) is said to be achievable if an (n, R, D, ϵ) state amplification scheme exists for every $\epsilon > 0$, possibly by taking *n* large enough. Let $\mathcal{C}_{\text{est}}^{\text{fad}}(\mathrm{P})$ be the convex closure of the set of all achievable (*R*, *D*) tuples, with $0 \leq D \leq Q$. The main result of this letter is stated next.

Theorem 1: For the channel in (1) with state estimation constraints, the optimal trade-off region $\mathcal{C}_{est}^{fad}(\mathrm{P})$ is given by the convex closure of the set of all $(R, D) \in \mathbb{R}_+^2$ such that:

$$R \leq \mathbb{E}_\Theta\left[\frac{1}{2}\log\left(1 + \frac{\gamma\theta^2 P(\theta)}{\sigma_Z^2}\right)\right], \tag{5}$$

$$D \geq \mathbb{E}_\Theta\left[\frac{Q\left(\sigma_Z^2 + \gamma\theta^2 P(\theta)\right)}{\theta^2 P(\theta) + Q + \sigma_Z^2 + 2\theta\sqrt{(1-\gamma)P(\theta)Q}}\right], \tag{6}$$

for some $\gamma \in [0, 1]$, under the constraint $\mathbb{E}_\Theta[P(\theta)] \leq \mathrm{P}$.

Proof: The achievability is proved in Section III, while the converse proof is given in Section IV. ■

III. Achievability Proof of Theorem 1

The available power at the encoder $P(\theta)$ (for a given realization $\Theta_i = \theta$) is split into: $\gamma P(\theta)$ for message communication and $(1 - \gamma)P(\theta) \triangleq \bar{\gamma}P(\theta)$ for state amplification, for some $\gamma \in [0, 1]$. The state amplification signal is then generated as $X_S^n = \sqrt{\bar{\gamma}P(\theta)/Q}\,S^n$. Hence the channel model (for $\Theta_i = \theta$) in (1) becomes:

$$Y_i = \theta X_{M,i} + \left(1 + \theta\sqrt{\frac{\bar{\gamma}P(\theta)}{Q}}\right)S_i + Z_i, \tag{7}$$

where $\mathbb{E}[||X_M^n||^2] \leq n\gamma P(\theta)$ with X_M^n independent of S^n. Now in order to communicate the message across to the receiver, the result of [2] is applied. The average achievable rate using the result of [2] is the right-hand side of (5).

To meet the distortion bound, the receiver forms the (linear) minimum mean-squared error (MMSE) estimate based on the observation Y^n. The MMSE and thus the average distortion can be readily calculated to be the right-hand side of (6).

IV. Converse Proof of Theorem 1

The main idea of the converse is to first upper bound the sum $R + \frac{1}{2}\log(\frac{Q}{D})$, which leads to a lower bound on the distortion in terms of *R*. This combined with an upper bound on *R* characterizes the boundary of the optimal (*R*, *D*) trade-off region in Theorem 1. This method is in contrast to the converse proof in [3, Th. 2] – see Remark 1 and the ensuing explanation towards the end of the section. Now if (*R*, *D*) is achievable, for each $\epsilon > 0$ there is an (n, R, D, ϵ) state amplification scheme. For such a scheme, let K_i denote the covariance matrix of $(X_i, \Theta_i, S_i, Z_i, Y_i)$ for every $i \in [1, 2, \ldots, n]$. Let the power of X_i induced by the scheme be $P_i(\theta)$ for a realization $\Theta_i = \theta$, and $\frac{1}{n}\sum_{i=1}^{n} P_i(\theta) = P(\theta)$. Now define the correlation coefficient between X_i and S_i (for $\Theta_i = \theta$) to be $\rho_i = \mathbb{E}[X_i S_i|\Theta_i = \theta]/\sqrt{P_i(\theta)Q}$. Also, let $\gamma_i = 1 - \rho_i^2$. Notice that $\gamma_i \in [0, 1]$. This gives:

$$\mathbb{E}[X_i S_i|\Theta_i = \theta] = \sqrt{(1-\gamma_i)P_i(\theta)Q},$$

$$\sigma_{X_i|S_i,\Theta_i=\theta}^2 \triangleq \min_a \mathbb{E}\left[(X_i - aS_i)^2|\Theta_i = \theta\right] = \gamma_i P_i(\theta).$$

Now since $Y_i = \theta X_i + S_i + Z_i$ for a given realization $\Theta_i = \theta$, $h(Y_i)$ would be maximized when X_i is jointly Gaussian with S_i. Hence, without loss of generality, the following holds:

$$X_i|\{\Theta_i = \theta\} = V_i + \left(\sqrt{(1-\gamma_i)P_i(\theta)/Q}\right)S_i, \tag{8}$$

where V_i is Gaussian with zero mean, variance $\gamma_i P_i(\theta)$ and independent of S_i. On denoting $g(x) = (1/2)\log(2\pi e x)$ and using the differential entropy maximizing property of Gaussian random variables for a given variance, the following holds:

$$h(Y_i|S_i, \Theta_i = \theta) \leq g\Big(\sigma_Z^2 + \gamma_i \theta^2 P_i(\theta)\Big), \quad (9)$$

$$h(Y_i|\Theta_i = \theta) \leq g\Big(\theta^2 P_i(\theta) + Q + \sigma_Z^2 + 2\theta\sqrt{(1-\gamma_i)P_i(\theta)Q}\Big). \quad (10)$$

For the parameters mentioned above, choose $0 \leq \gamma \leq 1$ such that:

$$\gamma P(\theta) = \frac{1}{n}\sum_{i=1}^{n} \gamma_i P_i(\theta). \quad (11)$$

It will be shown that for any successful scheme, (*R*, *D*) satisfies (5) and (6) for the parameter γ defined above. The following bound holds for any scheme which achieves a distortion $D_n = \frac{1}{n}\mathbb{E}||S^n - \hat{S}^n||^2$ over block length *n*.

Lemma 1:

$$\frac{n}{2}\log\left(\frac{Q}{D_n}\right) \leq I(S^n; Y^n|\Theta^n). \quad (12)$$

Proof: It follows from [3] that $\frac{n}{2}\log(\frac{Q}{D_n}) \leq I(S^n; Y^n)$. Further, it follows that:

$$I(S^n; Y^n) \leq I(S^n; Y^n, \Theta^n) \overset{(a)}{=} I(S^n; Y^n|\Theta^n), \quad (13)$$

where (a) follows since S^n is independent of Θ^n. ■

The sum $R + \frac{1}{2}\log(\frac{Q}{D_n})$ is now bounded as follows:

$$\begin{aligned} nR + \frac{n}{2}\log\left(\frac{Q}{D_n}\right) &\overset{(a)}{\leq} H(W) + I(S^n; Y^n|\Theta^n) \\ &\overset{(b)}{=} H(W|S^n, \Theta^n) + I(S^n; Y^n|\Theta^n) \\ &\overset{(c)}{\leq} I(W; Y^n|S^n, \Theta^n) + I(S^n; Y^n|\Theta^n) + n\epsilon_n \\ &= I(W, S^n; Y^n|\Theta^n) + n\epsilon_n \\ &= h(Y^n|\Theta^n) - h(Y^n|W, S^n, \Theta^n) + n\epsilon_n \\ &\leq \sum_{i=1}^{n}\{h(Y_i|\Theta_i) - h(Z_i)\} + n\epsilon_n \\ &= \mathbb{E}_\Theta\left[\sum_{i=1}^{n} h(Y_i|\Theta_i = \theta) - h(Z_i)\right] + n\epsilon_n \\ &\overset{(d)}{\leq} \mathbb{E}_\Theta\left[\sum_{i=1}^{n}\frac{1}{2}\log\left(\frac{\theta^2 P_i(\theta) + Q + \sigma_Z^2 + 2\theta\sqrt{\bar{\gamma}_i P_i(\theta)Q}}{\sigma_Z^2}\right)\right] + n\epsilon_n \\ &\overset{(e)}{\leq} \mathbb{E}_\Theta\left[\frac{n}{2}\log\left(\frac{\theta^2 P(\theta) + Q + \sigma_Z^2 + 2\theta\sqrt{\bar{\gamma}P(\theta)Q}}{\sigma_Z^2}\right)\right] + n\epsilon_n, \end{aligned} \quad (14)$$

where (a) follows from Lemma 1, (b) follows since *W* is independent of (S^n, Θ^n), (c) follows by Fano's inequality with $\epsilon_n \to 0$ as $n \to \infty$, (d) follows from expression (10), while (e) follows from Jensen's Inequality. Since $\epsilon_n \to 0$ as $n \to \infty$, the desired bound on $R + \frac{1}{2}\log(\frac{Q}{D})$ results. Equation (14) can be recast as:

$$D \geq \mathbb{E}_\Theta\left[\frac{Q\sigma_Z^2 2^{2R}}{\theta^2 P(\theta) + Q + \sigma_Z^2 + 2\theta\sqrt{(1-\gamma)P(\theta)Q}}\right]. \quad (15)$$

Expression (15) specifies the coupling between the distortion *D* and the rate *R*.

An upper bound on the rate *R* is established next.

$$\begin{aligned} nR &= H(W) \overset{(a)}{=} H(W|S^n, \Theta^n) \\ &\overset{(b)}{\leq} I(W; Y^n|S^n, \Theta^n) + n\epsilon_n \\ &= h(Y^n|S^n, \Theta^n) - h(Y^n|W, S^n, \Theta^n) + n\epsilon_n \\ &\leq \sum_{i=1}^{n}\{h(Y_i|S_i, \Theta_i) - h(Z_i)\} + n\epsilon_n \\ &= \mathbb{E}_\Theta\left[\sum_{i=1}^{n} h(Y_i|S_i, \Theta_i = \theta) - h(Z_i)\right] + n\epsilon_n \\ &\overset{(c)}{\leq} \mathbb{E}_\Theta\left[\sum_{i=1}^{n}\frac{1}{2}\log\left(\frac{\sigma_Z^2 + \gamma_i\theta^2 P_i(\theta)}{\sigma_Z^2}\right)\right] + n\epsilon_n \\ &\overset{(d)}{\leq} \mathbb{E}_\Theta\left[\frac{n}{2}\log\left(1 + \frac{\gamma\theta^2 P(\theta)}{\sigma_Z^2}\right)\right] + n\epsilon_n, \end{aligned} \quad (16)$$

(a) follows since *W* is independent of (S^n, Θ^n), (b) follows by Fano's inequality, (c) follows from expression (9), while (d) follows from Jensen's Inequality. Since $\epsilon_n \to 0$ as $n \to \infty$, the following bound results:

$$R \leq \mathbb{E}_\Theta\left[\frac{1}{2}\log\left(1 + \frac{\gamma\theta^2 P(\theta)}{\sigma_Z^2}\right)\right]. \quad (17)$$

For a fixed rate *R*, the lower bound on *D* in (15) is monotonically increasing in γ. Hence, the minimal distortion *D*(*R*) for a given *R* will be obtained by searching for the minimal γ such that (17) is respected. This completes the proof of converse.

Remark 1: The technique proposed in the current paper of upper bounding $R + \frac{1}{2}\log(\frac{Q}{D})$ along with *R* bypasses the supporting hyperplane based approach and strict concavity arguments used in the proof of [3, Th. 2] to bound the weighted sum $R + \frac{\lambda}{2}\log(\frac{Q}{D}), \lambda \geq 0$. Thus, the current approach also presents a simpler alternative proof for the converse of [3, Th. 2] that only relies on the differential entropy maximizing property of Gaussian random variables for a given variance – which might be of independent interest.

Explanation for Remark 1 (Comparison to the proof of [3]): The proof of [3, Th. 2] uses a supporting hyperplane characterization of the convex (*R*, *D*) region based on a coordinate transformation from $D \to \log(\frac{Q}{D})$ (see [3, p. 1492, 4th equation from the top]). The strict concavity arguments therein (invoked for the functions $R(\gamma)$ and $D(\gamma)$ on pg. 1492 as well as expression (8) on pg. 1493 in [3]) were necessary to ensure that this coordinate transformation does not exclude any points within the original (*R*, *D*) region.

The current proof, in contrast, takes the fundamentally simpler view that the (*R*, *D*) trade-off region is completely specified by its boundary. This follows because the achievability of any pair (*R*, *D*) on the boundary of the region guarantees

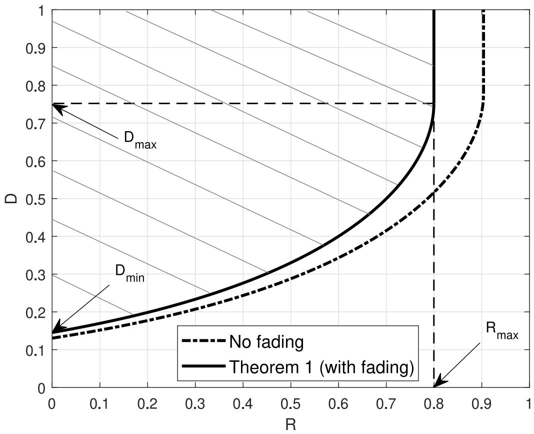

Fig. 2. Numerical illustration of the region in Theorem 1.

the achievability of the pair $(R-\delta_1, D+\delta_2)$ as well, for any $\delta_1, \delta_2 > 0$ (and hence, the interior of the region is covered). The other main insight in the current proof is that the boundary of the (*R*, *D*) trade-off region can be traced by simply collecting the distortion-rate functions $D(R) \triangleq \min\{D|(R,D) \in \mathcal{C}_{\text{est}}^{\text{fad}}(\mathrm{P})\}$ for all $0 \leq R \leq \mathbb{E}_\Theta[0.5\log(1+\theta^2 P(\theta)/\sigma_Z^2)]$. The steps leading to (15) and (17) precisely take care of this characterization of *D*(*R*) for various *R*.

Numerical Example: The rate versus distortion region is numerically depicted in Fig. 2, for an example where $\mathrm{P} = 2.5, Q = 1, \sigma_Z^2 = 1$. The fading process Θ^n is assumed to be i.i.d. with Rayleigh distributed components, with probability density function:

$$p_\Theta(\theta) = 2\theta e^{-\theta^2}, \theta \geq 0. \quad (18)$$

The optimal power allocation $P(\theta)$ for the plot is computed (numerically) as detailed in the Appendix.

For comparison, the rate-distortion trade-off region for a channel without fading but with the same average signal-to-noise-ratio is shown in dotted-dashed lines in Fig. 2. Naturally, the region in Theorem 1 is smaller compared to the case without fading impairments. In Fig. 2, the corner point corresponding to R_{max} is associated with pure message transmission and no state estimation, – i.e., $\gamma = 1$ in (5). The corresponding distortion obtained with $\gamma = 1$ in (6) is denoted by D_{max}. On the other hand, the corner point corresponding to zero rate is associated with pure state estimation and no message communication, – i.e., $\gamma = 0$ in (5). The corresponding distortion obtained with $\gamma = 0$ in (6) is denoted by D_{min}.

V. Conclusion

The discrete memoryless counterpart of the current model would be an interesting problem for further investigations. Moreover, multi-user versions of the fading state-dependent channel studied here present challenging research avenues.

Appendix
Solution for Optimal Power Allocation $P(\theta)$

Note that the characterization in Theorem 1 can be expressed as follows (where $\lambda \geq 0$):

$$R + \frac{\lambda}{2}\log\left(\frac{Q}{D}\right) \leq \max_{\substack{P(\theta):\\ \mathbb{E}_\Theta[P(\theta)] \leq \mathrm{P}}} \left[\mathbb{E}_\Theta\left[\frac{1}{2}\log\left(1+\frac{\gamma\theta^2 P(\theta)}{\sigma_Z^2}\right) + \frac{\lambda}{2}\log\left(\frac{\theta^2 P(\theta) + Q + \sigma_Z^2 + 2\theta\sqrt{\bar{\gamma} P(\theta) Q}}{\sigma_Z^2 + \gamma\theta^2 P(\theta)}\right)\right]\right].$$

By Lagrange multipliers method, the Lagrangian $J(\delta)$ is:

$$J(\delta) = \frac{1}{2}\log\left(1+\frac{\gamma\theta^2 P(\theta)}{\sigma_Z^2}\right) + \frac{\lambda}{2}\log\left(\frac{\theta^2 P(\theta) + Q + \sigma_Z^2 + 2\theta\sqrt{\bar{\gamma} P(\theta) Q}}{\sigma_Z^2 + \gamma\theta^2 P(\theta)}\right) - \delta P(\theta).$$

For simplicity, let $\sigma_Z^2 = 1$. Now the solution can be obtained by taking $\frac{\partial J}{\partial P(\theta)} = 0$. The partial derivative is given by:

$$\frac{\partial J}{\partial P(\theta)} = \frac{(1-\lambda)\gamma\theta^2}{2(1+\gamma\theta^2 P(\theta))} + \frac{\lambda\theta\left(\theta + \sqrt{\frac{\bar{\gamma} Q}{P(\theta)}}\right)}{2(\theta^2 P(\theta) + Q + 1 + 2\theta\sqrt{\bar{\gamma} P(\theta) Q})} - \delta.$$

Thus, the optimal allocation $P^*(\theta)$ is the solution to the above expression, with δ chosen such that $\mathbb{E}_\Theta[P^*(\theta)]_+ = \mathrm{P}$.

References

[1] S. Gelfand and M. Pinsker, "Coding for channels with random parameters," *Problems Control Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.
[2] M. H. Costa, "Writing on dirty paper (corresp.)," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
[3] A. Sutivong, M. Chiang, T. M. Cover, and Y.-H. Kim, "Channel capacity and state estimation for state-dependent Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1486–1495, Apr. 2005.
[4] Y.-H. Kim, A. Sutivong, and T. M. Cover, "State amplification," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1850–1859, May 2008.
[5] N. Merhav and S. Shamai, "Information rates subject to state masking," *IEEE Trans. Inf. Theory*, vol. 53, no. 6, pp. 2254–2261, Jun. 2007.
[6] C. Tian, B. Bandemer, and S. Shamai, "Gaussian state amplification with noisy observations," *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 4587–4597, Sep. 2015.
[7] S. I. Bross and A. Lapidoth, "The rate-and-state capacity with feedback," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1893–1918, Mar. 2018.
[8] U. Pereg, "Communication over quantum channels with parameter estimation," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2020, pp. 1818–1823.
[9] M. Le Treust and M. Bloch, "State leakage and coordination with causal state knowledge at the encoder," *IEEE Trans. Inf. Theory*, vol. 67, no. 2, pp. 805–823, Feb. 2021.
[10] F. Liu *et al.*, "Integrated sensing and communications: Towards dual-functional wireless networks for 6G and beyond," 2021, *arXiv:2108.07165*.
[11] W. Liu and B. Chen, "Message transmission and state estimation over Gaussian broadcast channels," in *Proc. 43rd Annu. Conf. Inf. Sci. Syst. (CISS)*, 2009, pp. 147–151.
[12] S. I. Bross and A. Lapidoth, "The Gaussian source-and-data-streams problem," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5618–5628, Aug. 2019.
[13] S. I. Bross, "Message and causal asymmetric state transmission over the state-dependent degraded broadcast channel," *IEEE Trans. Inf. Theory*, vol. 66, no. 6, pp. 3342–3365, Jun. 2020.
[14] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.