Information Theory of Intercellular Signal Transduction

Andrew W. Eckford
Dept. of EECS
York University, Toronto, Canada
Email: aeckford@yorku.ca

Peter J. Thomas
Dept. of Mathematics
Case Western Reserve University, USA
Email: pjthomas@case.edu

***Abstract*—The cells in the human body use intercellular signal transduction to organize themselves and regulate their work; multicellular life could not exist without it. In this paper, we discuss why information theory provides a useful toolbox to approach this problem. We also discuss how to calculate the capacity of intercellular signal transduction in some specific examples. Finally, we describe the challenges and future prospects of this avenue of research.**

I. Introduction

Molecular communications, a rapidly growing subfield of information theory, addresses the fundamental limits of communication at micron and submicron length scales, and strategies for achieving those limits [1]. Driven in part by the anticipated need for nanoscale agents to communicate and coordinate their actions, for instance for nanomedicine applications [2], communications engineering can find inspiration in naturally evolved approaches for signaling systems in environments hostile to traditional wired and wireless solutions – such as the interior of the human body.

Living cells monitor their environments and signal to one another using chemical signals communicated via diffusion, and transduced by receptor proteins [3], [4]. Early information theoretic analyses of biochemical signaling focused on chemotaxis, the directed movement of migrating cells in response to chemical cues, in part because the input/output signal ensembles could be clearly defined [5]–[10]. Advances in high-throughput experimental techniques have facilitated gathering enough data to quantitatively measure the capacity of specific signaling pathways [11]–[13]. At the same time, interest has grown among information theorists in a constellation of microbiological signaling problems, for instance the problem of a population of agents collectively estimating local concentration of chemical species (the consensus problem [14], [15]), the problem of memory effects in diffusion-mediated communication [16], [17], bacterial quorum sensing [18], [19], and communication via bacterial cables [20]. In this paper we discuss recent work on the capacity of a simple intercellular signal transduction system based on the cAMP receptor of the social amoeba *Dictyostelium discoideum* [21].

AWE was supported in part by a grant from the Natural Sciences and Engineering Research Council (NSERC). PJT was supported in part by NSF grant DMS-1413770.

II. Markov models of signal transduction

As we show in this section, signal transduction can be modelled as a finite state Markov chain. Thus, we can take advantage of the rich toolbox of Markov chains [22] to analyze signal transduction.

A. Biological model

Throughout this paper, we use cyclic adenosine monophosphate (cAMP) as a motivating example to discuss signal transduction. This molecule is used in many biological processes, for example by the amoeba *Dictyostelium discoideum* in determining its social behaviour. Under normal circumstances, *Dictyostelium* act as unicellular individuals, but under stress, numerous individuals gather together (in response to cAMP) and form a multicellular "slug" in which the formerly independent cells take on specialized roles. *Dictyostelium* is well studied as a model organism for signal transduction.

cAMP has a simple (yet illustrative) model. The receptor can be in one of two states: unbound (U) or bound (B). In state U, the receptor awaits the arrival of a cAMP molecule; once one arrives, the receptor enters state B. In state B, the receptor cannot bind to other cAMP molecules (rendering it insensitive to the signal), while some processing time is required before returning to state U. We describe continuous-time and discrete-time mathematical models for this process below.

B. Signal transduction as a Poisson process

In continuous time, the binding process is represented as a Poisson process, transitioning between states U and B. The transition rate from U to B is $r_{\mathsf{UB}}c(t)$, proportional to ligand concentration $c(t)$; the transition rate from B to U is r_{BU}, independent of ligand concentration.

The transition rates are given as a time-varying matrix $R(t)$, as follows:

$$R(t) = \begin{bmatrix} -r_{\mathsf{UB}}c(t) & r_{\mathsf{UB}}c(t) \\ r_{\mathsf{BU}} & -r_{\mathsf{BU}} \end{bmatrix}. \tag{1}$$

(Note that the rows of R must sum to zero.) Let $p(t) = [p_{\mathsf{U}}(t), p_{\mathsf{B}}(t)]$ represent the time-varying row vector of state occupancy probabilities; these probabilities are related to the Poisson rates via the differential equation

$$\frac{d}{dt}p(t) = p(t)R(t). \tag{2}$$

Equation (2) forms the *master equation* of the system.

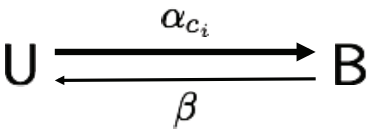

Fig. 1. State transition diagram for cAMP. Each transition is labelled with the transition probability. The $\mathsf{U} \to \mathsf{B}$ transition, depicted with a **bold** arrow, is sensitive to the input ligand concentration c_i; the $\mathsf{B} \to \mathsf{U}$ transition is insensitive.

C. Signal transduction as a Markov process

We are interested in representing the binding process as a discrete-time Markov chain. We can discretize time into steps of length Δt. Now, if I is the 2×2 identity matrix, the Poisson process becomes a discrete-time Markov chain with state transition probability matrix (at the ith step)

$$P_i = I - \Delta t R(i\Delta t) \tag{3}$$

$$= \begin{bmatrix} 1 - \Delta t\, r_{\mathsf{UB}} c(i\Delta t) & \Delta t\, r_{\mathsf{UB}} c(i\Delta t) \\ \Delta t\, r_{\mathsf{BU}} & 1 - \Delta t\, r_{\mathsf{BU}} \end{bmatrix}. \tag{4}$$

For consistency with notation in [23], we let

$$c_i = c(i\Delta t) \tag{5}$$

$$\alpha_{c_i} = \Delta t\, r_{\mathsf{UB}} c(i\Delta t) \tag{6}$$

$$\beta = \Delta t\, r_{\mathsf{BU}}, \tag{7}$$

and (4) becomes

$$P_i = \begin{bmatrix} 1 - \alpha_{c_i} & \alpha_{c_i} \\ \beta & 1 - \beta \end{bmatrix}. \tag{8}$$

The state transition diagram is given in Figure 1.

The ligand concentration c_i exists in a range, from the minimum allowed concentration $c_i = \mathsf{L}$ to the maximum $c_i = \mathsf{H}$. (Physically, there must exist a finite maximum concentration. The minimum may be zero.) For simplicity, we will use only the extreme concentrations $c_i = \mathsf{L}$ and $c_i = \mathsf{H}$; however, this restriction does not limit our analysis in the next section.

Thus, given the concentration, the signal transduction process is a time-inhomogeneous Markov chain with either

$$P_{\mathsf{H}} = \begin{bmatrix} 1 - \alpha_{\mathsf{H}} & \alpha_{\mathsf{H}} \\ \beta & 1 - \beta \end{bmatrix} \tag{9}$$

or

$$P_{\mathsf{L}} = \begin{bmatrix} 1 - \alpha_{\mathsf{L}} & \alpha_{\mathsf{L}} \\ \beta & 1 - \beta \end{bmatrix}, \tag{10}$$

selected by the ligand concentration.

III. Capacity of signal transduction

Here we review some of our main results on the capacity of signal transduction; more details can be found in [23], [24].

A. Signal transduction as a communication channel

To define a communication channel, we must define inputs, outputs, and channel input-output relationship:

- *Input.* The input X is the concentration of ligands in the environment. We have a binary channel input: $X \in \{\mathsf{L}, \mathsf{H}\}$.
- *Output.* The output Y is the state of the receptor. We have a binary channel output: $Y \in \{\mathsf{U}, \mathsf{B}\}$.
- *Input-output relationship.* As a Markov channel, the state of the output at time i, Y_i, depends on the current input X_i and the previous channel output Y_{i-1}. We can write $p_{Y_i|X_i,Y_{i-1}}(y_i \mid x_i, y_{i-1})$, where the correct probability is given by an entry in either (9) or (10). For example,

$$p_{Y_i|X_i,Y_{i-1}}(\mathsf{B} \mid \mathsf{H}, \mathsf{U}) = \alpha_{\mathsf{H}}, \tag{11}$$

selected from (9) since $X_i = \mathsf{H}$.

From now on, we will omit the subscripts for probability mass functions where unambiguous, e.g. $p_Y(y)$ becomes $p(y)$.

The conditional probability of a vector of outputs Y_1^n given a vector of inputs X_1^n is written

$$p(y_1^n \mid x_1^n) = \prod_{i=1}^{n} p(y_i \mid x_i, y_{i-1}), \tag{12}$$

where y_0 is null, and where each $p(y_i \mid x_i, y_{i-1})$ is selected from (9) or (10), as above.

B. Capacity and IID capacity

Logarithms are base 2 throughout. We will use the following special functions: let

$$\phi(p) = \begin{cases} 0, & p = 0 \\ -p \log p, & p \neq 0 \end{cases} \tag{13}$$

represent the partial entropy function, and let

$$\mathscr{H}(p) = \phi(p) + \phi(1 - p) \tag{14}$$

represent the binary entropy function.

For any communication system with inputs X_i and outputs Y_i, the mutual information rate is given by

$$\mathcal{I}(X;Y) = \frac{1}{n} \lim_{n\to\infty} I(X_1^n; Y_1^n), \tag{15}$$

where $I(X_1^n; Y_1^n)$ is the mutual information between vectors X_1^n and Y_1^n. In our specific case, it can be shown that

$$\mathcal{I}(X;Y) = \lim_{n\to\infty} H(Y_n \mid Y_1^{n-1}) - H(Y_n \mid X_n, Y_{n-1}). \tag{16}$$

The capacity of the system is given by

$$C = \max_{p(x_1^n)} \mathcal{I}(X;Y), \tag{17}$$

where the maximum is taken over all possible input distributions $p(x_1^n)$.

If we restrict the input distributions in (17) to be independent and identically distributed (IID), then we have the IID capacity, C_{IID}. Since the maximizing distribution in (17) is possibly (but not necessarily) IID, it should be clear that $C_{\mathrm{IID}} \leq C$.

To calculate C_{IID}, first note that the process Y_1^n is a *time-homogeneous* Markov chain if X_1^n is unknown but IID: (12) becomes

$$p(y_1^n) = \sum_{x_1} \cdots \sum_{x_n} \prod_{i=1}^{n} p(y_i \mid x_i) p(x_i) \tag{18}$$

$$= \prod_{i=1}^{n} \sum_{x_i} p(y_i \mid x_i) p(x_i). \tag{19}$$

Letting

$$p(y_i \mid y_{i-1}) = \prod_{i=1}^{n} \sum_{x_i} p(y_i \mid x_i) p(x_i), \tag{20}$$

and letting

$$\bar{\alpha} = p_X(\mathsf{L})\alpha_{\mathsf{L}} + p_X(\mathsf{H})\alpha_{\mathsf{H}}, \tag{21}$$

we have that Y_1^n is a Markov chain with transition probability matrix

$$P = p_X(\mathsf{L})P_{\mathsf{L}} + p_X(\mathsf{H})P_{\mathsf{H}} \tag{22}$$

$$= \begin{bmatrix} 1-\bar{\alpha} & \bar{\alpha} \\ \beta & 1-\beta \end{bmatrix} \tag{23}$$

Moreover, as a two-state Markov chain, the steady-state distribution of the process Y is given by

$$p_Y(y) = \begin{cases} \frac{\beta}{\bar{\alpha}+\beta}, & y = \mathsf{U} \\ \\ \frac{\bar{\alpha}}{\bar{\alpha}+\beta}, & y = \mathsf{B} \end{cases} \tag{24}$$

In (16), $H(Y_n \mid Y_1^{n-1})$ becomes $H(Y_n \mid Y_{n-1})$, since the process Y is a Markov chain. This quantity is given by

$$H(Y_n \mid Y_{n-1})$$

$$= -E[\log p(y_n \mid y_{n-1})] \tag{25}$$

$$= -\sum_{y_{n-1}} p(y_{n-1}) \sum_{y_n} p(y_n \mid y_{n-1}) \log p(y_n \mid y_{n-1}) \tag{26}$$

$$= \frac{\beta}{\bar{\alpha}+\beta}\mathscr{H}(\bar{\alpha}) + \frac{\bar{\alpha}}{\bar{\alpha}+\beta}\mathscr{H}(\beta). \tag{27}$$

Similarly,

$$H(Y_n \mid X_n, Y_{n-1})$$

$$= -E[\log p(y_n \mid x_n, y_{n-1})] \tag{28}$$

$$= \frac{\beta}{\bar{\alpha}+\beta}\Big(p_X(\mathsf{L})\mathscr{H}(\alpha_{\mathsf{L}}) + p_X(\mathsf{H})\mathscr{H}(\alpha_{\mathsf{H}})\Big) + \frac{\bar{\alpha}}{\bar{\alpha}+\beta}\mathscr{H}(\beta). \tag{29}$$

Substituting back into (16),

$$\mathcal{I}(X;Y)$$

$$= \frac{\beta}{\bar{\alpha}+\beta}\Big(\mathscr{H}(\bar{\alpha}) - p_X(\mathsf{L})\mathscr{H}(\alpha_{\mathsf{L}}) - p_X(\mathsf{H})\mathscr{H}(\alpha_{\mathsf{H}})\Big). \tag{30}$$

Finally, C_{IID} is found by maximizing (30) with respect to $p_X(\mathsf{L})$, recalling that $p_X(\mathsf{H}) = 1 - p_X(\mathsf{L})$. Since $C_{\mathrm{IID}} \leq C$, this provides us with a lower bound on capacity.

In [23], it is shown that even if inputs at intermediate concentrations other than L and H are allowed, the input distribution that maximizes C_{IID} has all probability mass on L and H, and none on any intermediate concentration.

C. Feedback capacity

Suppose the transmitter has causal feedback of all past outputs Y_1^{i-1}, and uses these outputs to determine its strategy in setting X_i; the best such strategy gives the *feedback capacity* C_{FB}.

In our model of the signal transduction channel, the output is the same as the Markov state of the channel. Such channels are sometimes called Previous Output is the STate (POST) channels [25], or unit output memory (UOM) channels [26]. For our specific case, C_{FB} is given by the maximum of the *directed information* [27] over all causal-conditional input distributions of the form $\prod_{i=1}^{n} p(x_i \,|\, y_1^{n-1}, x_1^{n-1})$. Since the transmitter has the option of disregarding the feedback (resulting in the regular capacity C), we now have the sequence of bounds

$$C_{\mathrm{IID}} \leq C \leq C_{\mathrm{FB}}. \tag{31}$$

To obtain C_{FB}, it seems a daunting task to specify the correct input distribution for all settings of y_1^{n-1} and x_1^n. However, from [26], we have a key simplification: in most UOM channels (including ours), the feedback-capacity-achieving input distribution is stationary, with the form

$$p(x_1^n \mid y_1^n) = \prod_{i=1}^{n} p(x_i \mid y_{i-1}). \tag{32}$$

That is, x_i is selected by considering only the two possibilities of $y_{i-1} \in \{\mathsf{U}, \mathsf{B}\}$. Moreover, in our specific case, it can be shown that the directed information rate has the same form as $\mathcal{I}(X;Y)$ in (16); we need only to maximize over the input distributions in (32) to obtain C_{FB}.

Consider $H(Y_i \mid Y_1^{i-1})$ from (16). From the above discussion, we need to specify only $p_{X_i|Y_{i-1}}(\mathsf{L} \mid \mathsf{U})$ and $p_{X_i|Y_{i-1}}(\mathsf{L} \mid \mathsf{B})$. Since the input x_i depends only on y_{i-1}, the process Y with feedback is still a time-homogeneous Markov chain: we can substitute

$$\bar{\alpha} = p_{X_i|Y_{i-1}}(\mathsf{L} \mid \mathsf{U})\alpha_{\mathsf{L}} + (1 - p_{X_i|Y_{i-1}}(\mathsf{L} \mid \mathsf{U}))\alpha_{\mathsf{H}} \tag{33}$$

into P from (23). Note that β (the transition rate in state B) is independent of the input x_i. Thus, **the transition probability matrix P is independent of** $p_{X_i|Y_{i-1}}(\mathsf{L}|\mathsf{B})$, as is $H(Y_i|Y_{i-1})$. By a similar argument, $H(Y_i \mid X_i, Y_{i-1})$ is also independent of $p_{X_i|Y_{i-1}}(\mathsf{L} \mid \mathsf{B})$.

Since the directed information rate is constant with respect to $p_{X_i|Y_{i-1}}(\mathsf{L} \mid \mathsf{B})$, it makes no difference to C_{FB} if we set

$$p_{X_i|Y_{i-1}}(\mathsf{L} \mid \mathsf{B}) = p_{X_i|Y_{i-1}}(\mathsf{L} \mid \mathsf{U}) = p_X(\mathsf{L}), \tag{34}$$

which is an IID input distribution.

Since C_{FB} is satisfied with an IID input distribution, we have $C_{\mathrm{IID}} = C_{\mathrm{FB}}$; thus, from (31),

$$C_{\mathrm{IID}} = C = C_{\mathrm{FB}}, \tag{35}$$

and the (non-feedback) capacity-achieving input distribution is IID.

Please see [23] for a more formal and complete description of this argument.

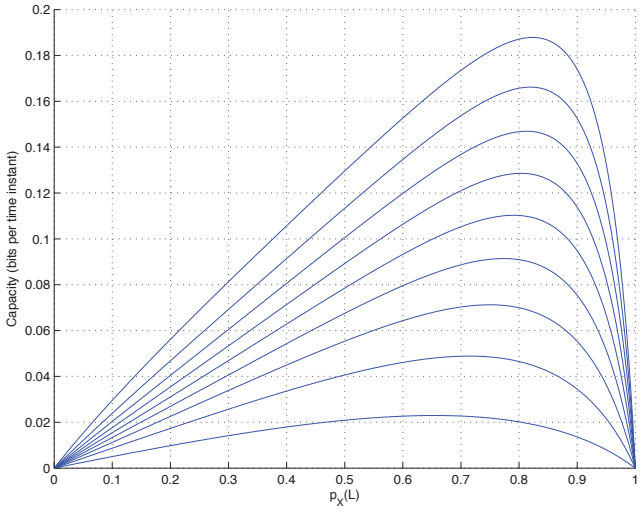

Fig. 2. Illustration of the information rate per time step Δt as a function of $p_X(\mathsf{L})$. In this figure, $\alpha_{\mathsf{L}} = 0.01$, $\beta = 0.1$, and α_{H} varies from 0.1 (bottom curve) to 0.9 (top curve) in increments of 0.1.

IV. Discussion

In Figure 2, we illustrate the information rates of our system for some sample parameter values. The capacity C may be determined by taking the maximum of each curve.

The techniques employed in this paper may be generalized to receptors with larger state spaces. In forthcoming work [28], we perform similar capacity calculations on the channelrhodopsin-2 (ChR2) receptor, with three states; and the acetylcholine (ACh) receptor, with five states. If there is exactly one state transition that is sensitive to the input signal, then the capacity-achieving input distribution is IID. This is the case for ChR2, but not for ACh. In the latter case, the IID capacity C_{IID} remains a lower bound on the true capacity C.

References

[1] I. F. Akyildiz, F. Fekri, R. Sivakumar, C. R. Forest, and B. K. Hammer, "Monaco: fundamentals of molecular nano-communication networks," *Wireless Communications, IEEE*, vol. 19, no. 5, pp. 12–18, 2012.

[2] S. K. Mallapragada, T. M. Brenza, J. M. McMillan, B. Narasimhan, D. S. Sakaguchi, A. D. Sharma, S. Zbarska, and H. E. Gendelman, "Enabling nanomaterial, nanofabrication and cellular technologies for nanoneuromedicines," *Nanomedicine: Nanotechnology, Biology and Medicine*, vol. 11, no. 3, pp. 715–729, 2015.

[3] H. C. Berg and E. M. Purcell, "Physics of chemoreception," *Biophysical Journal*, vol. 20, pp. 193–219, 1977.

[4] W. Bialek and S. Setayeshgar, "Physical limits to biochemical signaling," *PNAS*, vol. 102, pp. 10040–10045, July 19 2005.

[5] P. J. Thomas, D. J. Spencer, S. K. Hampton, P. Park, and J. Zurkus, "The diffusion-limited biochemical signal-relay channel," in *Advances in NIPS*, 2004.

[6] J. M. Kimmel, R. M. Salter, and P. J. Thomas, "An information theoretic framework for eukaryotic gradient sensing," in *Advances in Neural Information Processing Systems 19* (B. Schölkopf and J. Platt and T. Hoffman, ed.), pp. 705–712, Cambridge, MA: MIT Press, 2007.

[7] B. W. Andrews and P. A. Iglesias, "An information-theoretic characterization of the optimal gradient sensing response of cells.," *PLoS Comput Biol*, vol. 3, p. e153, Aug 2007.

[8] D. Fuller, W. Chen, M. Adler, A. Groisman, H. Levine, W.-J. Rappel, and W. F. Loomis, "External and internal constraints on eukaryotic chemotaxis," *Proc Natl Acad Sci U S A*, vol. 107, pp. 9656–9, May 2010.

[9] B. Hu, W. Chen, W.-J. Rappel, and H. Levine, "Physical limits on cellular sensing of spatial gradients," *Phys Rev Lett*, vol. 105, p. 048104, Jul 2010.

[10] B. Hu, W. Chen, H. Levine, and W.-J. Rappel, "Quantifying information transmission in eukaryotic gradient sensing and chemotactic response," *J Stat Phys*, vol. 142, pp. 1167–1186, Apr 2011.

[11] R. Cheong, A. Rhee, C. J. Wang, I. Nemenman, and A. Levchenko, "Information transduction capacity of noisy biochemical signaling networks," *Science*, Sep 2011.

[12] P. J. Thomas, "Cell signaling: Every bit counts," *Science*, vol. 334, pp. 321–2, Oct 2011.

[13] M. D. Brennan, R. Cheong, and A. Levchenko, "How information theory handles cell signaling and uncertainty," *Science*, vol. 338, no. 6105, pp. 334–335, 2012.

[14] A. Einolghozati, M. Sardari, A. Beirami, and F. Fekri, "Consensus problem under diffusion-based molecular communication," in *Information Sciences and Systems (CISS), 2011 45th Annual Conference on*, pp. 1–6, IEEE, 2011.

[15] A. Einolghozati, M. Sardari, A. Beirami, and F. Fekri, "Data gathering in networks of bacteria colonies: Collective sensing and relaying using molecular communication," in *Computer Communications Workshops (INFOCOM WKSHPS), 2012 IEEE Conference on*, pp. 256–261, IEEE, 2012.

[16] A. Einolghozati, M. Sardari, A. Beirami, and F. Fekri, "Capacity of discrete molecular diffusion channels," in *IEEE Intl. Symp. on Inform. Theory*, 2011.

[17] K. Srinivas, A. W. Eckford, and R. S. Adve, "Molecular communication in fluid media: The additive inverse gaussian noise channel," *Information Theory, IEEE Transactions on*, vol. 58, no. 7, pp. 4678–4692, 2012.

[18] A. Einolghozati, M. Sardari, and F. Fekri, "Design and analysis of wireless communication systems using diffusion-based molecular communication among bacteria," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 12, pp. 6096–6105, 2013.

[19] A. Einolghozati, M. Sardari, and F. Fekri, "Networks of bacteria colonies: A new framework for reliable molecular communication networking," *Nano Communication Networks*, 2015.

[20] N. Michelusi, S. Pirbadian, M. Y. El-Naggar, and U. Mitra, "A stochastic model for electron transfer in bacterial cables," *Selected Areas in Communications, IEEE Journal on*, vol. 32, no. 12, pp. 2402–2416, 2014.

[21] C. A. Parent and P. N. Devreotes, "A cell's sense of direction," *Science*, vol. 284, pp. 765–770, 30 April 1999.

[22] A. Goldsmith and P. Varaiya, "Capacity, mutual information, and coding for finite-state markov channels," *IEEE Trans. Inform. Theory*, vol. 42, pp. 868–886, May 1996.

[23] P. J. Thomas and A. W. Eckford, "Capacity of a simple intercellular signal transduction channel," 2015. http://arxiv.org/abs/1411.1650.

[24] A. W. Eckford and P. J. Thomas, "Capacity of a simple intercellular signal transduction channel," in *IEEE International Symposium on Information Theory*, 2013.

[25] H. Asnani, H. H. Permuter, and T. Weissman, "Capacity of a POST channel with and without feedback," in *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, pp. 2538–2542, IEEE, 2013.

[26] J. Chen and T. Berger, "The capacity of finite-state Markov channels with feedback," *IEEE Trans. Info. Theory*, vol. 51, pp. 780–798, Mar. 2005.

[27] J. L. Massey, "Causality, feedback and directed information," in *Proc. 1990 Intl. Symp. on Info. Th. and its Applications*, 1991.

[28] A. W. Eckford, K. Loparo, and P. J. Thomas, "Finite-state channel models for signal transduction in neural systems." To appear in Proc. IEEE ICASSP, 2016.