School of Engineering and Applied Science (SEAS), Ahmedabad University

B.Tech(ICT) Semester V: Wireless Communication (CSE 311)

- **No** : BT S12
- **Group Members**
  Kesha Bagadia (AU1841011)
  Manal Shah (AU1840126)
  Yashvi Gandhi (AU1841033)
  Yashvi Pipaliya (AU1841092)
- **Base Article Title** :

  S. Ghavami, "Anomaly Detection in Molecular Communications With Applications to Health Monitoring Networks" [1]

1 Performance Analysis of Base Article

The typical network for anomaly detection in a practical bio-molecular set-up comprises a two-tier network architecture design. We detect the presence of the anomaly in the first tier using a nano network within the body and we report the aforementioned anomaly to the larger scale network on the outside in the second tier.

For this particular paper, the two of the tiers both rely on a diffusion based MC (molecular communication) paradigm model which operates over a cardiovascular network. In this model, we have considered an inhomogeneous Poisson point process to determine the number of biomarker molecules received by the ACs.

For tier one:
The detection set-up includes a number of artificial cells (ACs) disposed at random inside the body. ACs are the ones in charge of the measurement of biomarker concentration in the MC channel, where the diseased cells (DCs) act as transmitters and the ACs are considered receivers. These ACs send a signal to alarm the presence of the anomaly when the biomarker concentration is either too high or too low than the regular standard value.

For tier two:

This alarm signal is sent by the ACs, the transmitters in this case to the receiver, which is a bio-cyber interface (BC). This transmission is conducted by the release of molecules over another MC channel to communicate those decisions. By collectively considering all the received molecular signals, a decision is made by the BC, and it accordingly alarms the outer network if needed.

The effect of noise is considered in both DC-AC and AC-BC links.

Assumptions made:

• **A1**: Conditioned on hypothesis $Hj, j \in 0, 1$, the observations at distinct ACs are independent.

• **A2**: Conditioned on hypothesis $Hj, j \in 0, 1$, the observations at each AC, $y_{AC,m[n]}$ for $n_{ss} + 1 \leq n \leq n_{ss} + N$, are independent.

• **A3**: AC measures the number of received molecules after time of $t_s s$.

• **A4**: We assume $gm(t; \tau)$s are equal and almost constant over time.

• **A5**: We consider a symmetric topology for health monitoring network, hence $\overline{\eta}_m = \overline{\eta}$, probability of detection. and false alarm in ACs are equal, i.e., $P_D^{AC,m} = P_D^{AC}$ and $P_F^{AC,m} = P_F^{AC}$.

• **A6**: ACs are injected in the near location of cancer cells, and send their message to BC using the blood vessels medium.

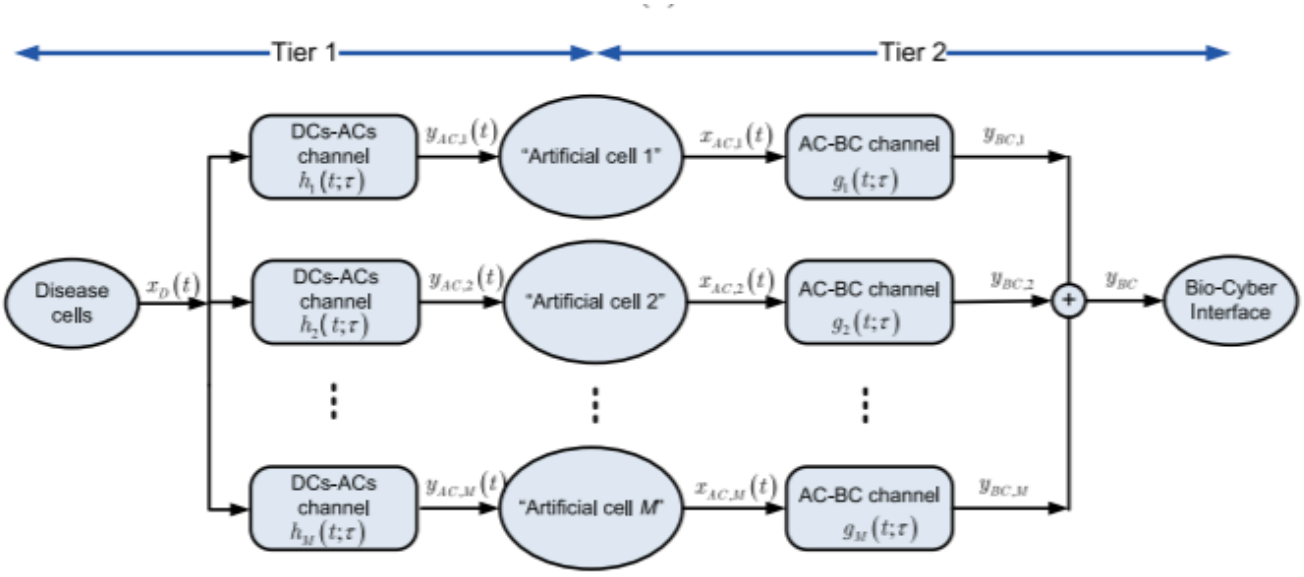


Figure 1: System model

1.1 DCs-AC System Model

We use Weibull function to model the release process i.e. the transmit waveform of biomarkers as a function of time. It is given as

$$x_D(t) = x_{D0}\left(1 - exp(-kt^b)\right) \tag{1}$$

x_{D0}: the released biomarker concentration at $t \to \infty$

b: the unitless biomarker power-law coefficient and $k > 0$: the biomarker release coefficient.

The received number of biomarkers in the nth time slot at the mth AC can be modeled as

$$y_{AC,m}[n] = \text{Poiss}(pr \times \lambda_{AC,m}[n] + pr \times \overline{\eta}_m) \tag{2}$$

p_r: the probability of reception

$\overline{\eta}_m$: the mean of Poisson noise. Poiss(.) denotes the Poisson distribution.

$\lambda_{AC,m}[n]$: average number of bio-marker molecules receive at the m^{th} AC, which is given as

$$\lambda_{AC,m}[n] = \int_{(n-1)T_s}^{nT_s} \int_{t-L^{(h_m)}T_s}^{t} h_m(t;\tau) x_D(\tau) d\tau dt \tag{3}$$

$h_m(t;\tau) > 0$: the time-varying channel between DCs and the m^{th} AC at transmission time τ and reception time t.

Consequentially, the propagation probability of receiving molecules on nth time slot from transmitted molecules in nl previous time slot, is given by

$$h_m[n,l] = \begin{cases} F_{W,m}(nT_s, lT_s) - F_{W,m}(nT_s, (l-1)T_s) & l \leq L^{h_m} \\ 0 & l \geq L^{h_m} \end{cases} \tag{4}$$

L^{h_m}: the memory length of the channel between the DCs and the m^{th} AC

Ts: the sampling period.

$F_{W,m}$: the cumulative distribution function (CDF) of the propagation time which is obtained using an additive inverse Gaussian (AIG) analysis and can be given as

$$F_{W,m} = \Phi\left(\sqrt{\frac{\lambda_m(t)}{w}}\left(\frac{w}{\mu_m(t)} - 1\right)\right) + e^{\frac{2\lambda_m(t)}{\mu_m(t)}}\left(-\sqrt{\frac{\lambda_m(t)}{w}}\left(\frac{w}{\mu_m(t)} + 1\right)\right. \tag{5}$$

$\mu_m(t) = l_m/v_m(t)$, where $vm(t)$: the drift velocity, l_m: the distance between DC and mth AC.

$\lambda_m(t) = (l_m^2)/\sigma_m^2(t)$, where $\sigma_m^2(t) = (\delta_m(t))/2$ is the the Brownian motion variance and $\delta_m(t)$ is the coefficient of diffusion between mth AC and BC at time t.

$\Phi(.)$ is the CDF of a standard Gaussian random variable.

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{-\frac{u^2}{2}\right\} du. \tag{6}$$

We assume that, $v_m(t)$ and $\delta_m(t)$ are constant over at least $L^{(hm)}Ts$ seconds.

1.2 ACs-BC System Model

Here, we use Weibull function again to model the release process of molecules

$$x_{AC,m}(t) = x_{AC_0} exp(-k^{'} t^{b'}) \tag{7}$$

x_{AC_0}: the transmitted molecules concentration at $t \to 0$

$b^{'}$: the unitless biomarker power-law coefficient and $k^{'} > 0$: the biomarker release coefficient.

The number of received molecules at the BC through the molecular ACs-BC channel is given by

$$y_{BC,m} = \sum_{m=1}^{M} \text{Poiss}(qr \times \lambda_{BC,m}) + qr \times \bar{\epsilon} \tag{8}$$

q_r : Reception probability

$\bar{\epsilon}$: mean of Poisson noise at the BC and the number of ACs

$$\lambda_{BC,m} = \int_0^{T_{BC}} \int_{t-L^{(g_m)}T_s}^{t} g_m(t;\tau) x_{AC,m}(\tau) d\tau dt \tag{9}$$

$g_m(t;\tau)$: the channel between mth AC and the BC,

$L^{(gm)}$: the memory length of the mth channel in the second tier

T_{BC} : the time duration of decision at BC

- **Hypothesis Test for Anomaly Detection:**

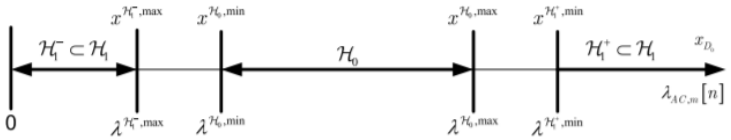


Figure 2: Illustration of hypothesis test

The ACs identify anomaly as proposed here,

$$\begin{cases} H_0 : x^{H_0,min} \leq x_{D_0} \leq x^{H_0,max} \\ H_1 : \begin{cases} H_1^- : x_{D_0} \leq x^{H_1^-,max} \\ H_1^+ : x_{D_0} \geq x^{H_1^+,min} \end{cases} \end{cases} \tag{10}$$

H_0 is considered a healthy setting where:
$x^{H_0,max}$ and $x^{H_0,min}$ denote the maximum and minimum values of bio-marker concentration.
H_1^- and H_1^+ is a non healthy setting where:
$x^{H_1{}^-,max}$ and $x^{H_0{}^+,min}$ denote the maximum and minimum values of bio-marker concentration.

Further, let us define, $n_{ss} = \lceil t_{ss}/T_s \rceil$ as the minimum sampling index for the steady state.
Due to the time varying nature and positivity of $h_m(t;\tau)$, and for the sake of tractability, consider $\lambda_{AC,m}^{H_0,min}$ and $\lambda_{AC,m}^{H_0,max}$ as a lower and upper bounds for the $\lambda_{AC,m}[.]$ under hypothesis test of H_0. Similarly, consider $\lambda_{AC,m}^{H_1^+,min}$ and $\lambda_{AC,m}^{H_1^-,max}$ as the lower and upper bounds for the $\lambda_{AC,m}[.]$ under hypothesis test of H_1^+ and H_1^- respectively, i.e. $\forall n \in n_{ss}+1,...,n_{ss}+N$ we have,

$$\begin{cases} H_0 : \lambda_{AC,m}^{H_0,min} \leq \lambda_{AC,m}[n] \leq \lambda_{AC,m}^{H_0,max} \\ H_1 : \begin{cases} H_1^- : \lambda_{AC,m}[n] \leq \lambda_{AC,m}^{H_1^-,max} \\ H_1^+ : \lambda_{AC,m}[n] \geq \lambda_{AC,m}^{H_1^+,min} \end{cases} \end{cases} \tag{11}$$

The BC uses the following hypothesis test to detect anomaly,

$$\begin{cases} W_0 : \sum_{m=1}^{M} \lambda_{BC,m} < k\overline{\lambda}_{BC} \\ W_1 : \sum_{m=1}^{M} \lambda_{BC,m} \geq k\overline{\lambda}_{BC} \end{cases} \tag{12}$$

where W_0 and W_1 are events corresponding to H_0 and H_1 in the BC. k is the minimum numbers of ACs, which detect anomaly in the DC-ACs link. We define $\overline{\lambda}_{BC}$ as the average number of received molecules at the BC using (9). It can be given for $\forall m \in 1,...,M$, i.e. as,

$$\overline{\lambda}_{BC} = \frac{1}{M} \sum_{m=1}^{M} \lambda_{BC,m} \tag{13}$$

- **Anomaly Detection Performance Of ACs**

A lower bound of detection probability and an upper bound of false alarm of the ACs can be derived by bounding the likelihood ratio (LR) of hypothesis test. This is where the Neyman Pearson (NP) framework is needed.

The decision rule of the hypothesis test based on generalized likelihood ratio test (GLRT) for the m^{th} AC with N independent Poisson observations can be given as:

$$\begin{aligned}\sum_{n=n_{ss}+1}^{n=n_{ss}+N} \hat{\lambda}_{AC,m}^{H_0}[n] - \hat{\lambda}_{AC,m}^{H_1}[n] \; + y_{AC,m}[n]\, log\, (\hat{\lambda}_{AC,m}^{H_1}[n] + p_r\overline{\eta}) \\ - \, y_{AC,m}[n]\, log\, (\hat{\lambda}_{AC,m}^{H_0}[n] + p_r\overline{\eta}) \; >^{H_1} \; log\gamma\end{aligned} \tag{14}$$

The ML estimator of $\lambda_{AC,m}[n]$ is given by

$$\hat{\lambda}_{AC,m}[n] = max(y_{AC,m}[n] - p_r\overline{\eta}) \tag{15}$$

We formulate a simplified decision rule from bounding the LR for the m^{th} AC with N independent Poisson observations and limited false alarm probability $P_F^{AC,m} < \xi_1$. It is given as

$$\begin{cases} H_0 : max(\gamma_l, 0) \leq \frac{1}{N}\sum_{n=n_{ss}+1}^{n_{ss}+N} y_{AC,m}[n] \leq \gamma_u \\ H_1 : \begin{cases} \frac{1}{N}\sum_{n=n_{ss}+1}^{n_{ss}+N} y_{AC,m}[n] > \gamma_u \qquad \text{or} \\ \frac{1}{N}\sum_{n=n_{ss}+1}^{n_{ss}+N} y_{AC,m}[n] < max(\gamma_l, 0) \end{cases} \end{cases} \tag{16}$$

where,

$$\gamma_l := \lambda_{AC,m}^{H_0,min} + p_r\overline{\eta} - \gamma^{'} \tag{17}$$

$$\gamma_u := \lambda_{AC,m}^{H_0,max} + p_r\overline{\eta} + \gamma^{'} \tag{18}$$

The lower bound of the detection probability is obtained as $P_D^{AC,m}$ and is given by

$$P_D^{AC,m} = min\left(\frac{\Gamma\Big(\lceil max(N\gamma_l - 1, 0) + 1\rceil,\ N\Big(\lambda_{AC,m}^{H_1^-,max} + p_r\overline{\eta}\Big)\Big)}{\lceil max\ (N\gamma_l - 1, 0)\rceil!},\ 1 - \frac{\Gamma\Big(\lfloor N\gamma_u + 1\rfloor\ N\Big(\lambda_{AC,m}^{H_1^+,min} + p_r\overline{\eta}\Big)\Big)}{\lfloor N\gamma_u\rfloor!}\right) \tag{19}$$

The decision rule for the BC with Poisson observation, where y_{BC}^{THR} is the decision threshold at the BC, is given as

$$\begin{cases} H_0 : y_{BC} < y_{BC}^{THR} \\ H_1 : y_{BC} \geq y_{BC}^{THR} \end{cases} \tag{20}$$

The lower bound of probability detection and upper bound for probability of false alarm at the BC are given as

$$P_D = \sum_{m=0}^{M} p_m^{'} \left(1 - \frac{\Gamma(\lceil y_{BC}^{THR}\rceil, m\overline{\lambda}_B C + q_r\overline{\epsilon})}{\lceil y_{BC}^{THR} - 1\rceil!}\right) \tag{21}$$

$$P_F = \sum_{m=0}^{M} p_m^{''} \left(1 - \frac{\Gamma(\lceil y_{BC}^{THR}\rceil, m\overline{\lambda}_B C + q_r\overline{\epsilon})}{\lceil y_{BC}^{THR} - 1\rceil!}\right) \tag{22}$$

Here, $p_m^{'}$ and $p_m^{''}$ denote the probability of detecting anomaly at m ACs, under H_1 and H_0, respectively, which are given by

$$\begin{aligned} p_m^{'} &= \binom{M}{m} \left(1 - P_D^{AC}\right)^{M-m} \left(P_D^{AC}\right)^m \\ p_m^{''} &= \binom{M}{m} \left(1 - P_F^{AC}\right)^{M-m} \left(P_F^{AC}\right)^m \end{aligned} \tag{23}$$

DESCRIPTION OF SYSTEM PARAMETERS

Symbol	Description
P_D^{AC}	Probability of detection at AC
P_F^{AC}	Probability of false alarm at AC
P_D	Probability of detection at BC
P_F	Probability of false alarm at BC
m	At a particular AC (m^{th} AC)
M	No. of ACs
$v_m(t)$	Drift velocity
Ts	Sampling Period
δ_m	The coefficient of diffusion
L^{h_m}	Memory length of channel between DCs
L^{g_m}	Memory length of channel between ACs
$\sigma_m^2(t)$	Brownian motion variance
$x_D(t)$	Released molecules by DCs at time t
$h_m(t;\tau)$	The MC channel between the DCs and m^{th} AC at reception time t and transmission time τ
$y_{AC,m}(t)$	No. of molecules received by the m^{th} AC
$\lambda_{AC,m}[n]$	Avg. no. of molecules received by m^{th} AC during n^{th} time slot
$x_{AC,m}(t)$	Released molecules by the m^{th} AC
$g_m(t;\tau)$	The MC channel between the m_{th} AC and BC at reception time t and transmission time τ
$y_{BC,m}(t)$	No. of received molecules by the BC from the m^th AC
$\lambda_{BC,m}$	Avg. no. of received molecules by the m^{th} AC
y_{BC}	Received molecules by the BC
N	No. of observations at AC
$\bar{\epsilon}_r$	Noise mean at BC
$\overline{\eta}$	Noise mean at ACs
p_r	Reception probability at ACs
q_r	Reception probability at BC
y_{BC}^{THR}	Decision threshold at BC
Δ_λ	$\min\Big(\Big\lvert\lambda_{AC,m}^{H_1^-,max} - \lambda_{AC,m}^{H_0,min}\Big\rvert, \Big\lvert\lambda_{AC,m}^{H_1^+,min} - \lambda_{AC,m}^{H_0,max}\Big\rvert\Big)$

γ_l	Lower bound satisfying the equation
γ_u	Upper bound satisfying the equation
$\gamma_{'}$	Minimum value satisfying the equation
$\lambda_{AC,m}^{H_0,min}$	Lower bound of $\lambda_{AC,m}[n]$ under H_0
$\lambda_{AC,m}^{H_0,max}$	Upper bound of $\lambda_{AC,m}[n]$ under H_0
$\lambda_{AC,m}^{H_1^-,max}$	Lower bound of $\lambda_{AC,m}[n]$ under H_1
$\lambda_{AC,m}^{H_1^+,min}$	Upper bound of $\lambda_{AC,m}[n]$ under H_1
$P_F^{AC,m}$	False alarm probability at m^{th} AC
n_{ss}	The minimum sampling index for the steady state
t_{ss}	The minimum time for the steady state

2 Numerical Results

2.1 Simulation Framework

Parameter used in Reproduced Figures

Symbol	**Value**
δ_m	0.5
l_m	1
l	[1 - 5]
$\lambda_m(t)$	4 (By calculation)
M	30
N	[5 10 20 40 80 160]
$\bar{\epsilon}_r$	100
q_r	0.05
$\overline{\eta}$	20
p_r	0.05
Δ_λ	1
$P_F^{AC,m}$	$[10^{-4} - 10^{-1}]$
$\lambda_{AC,m}^{H_0,min}$	10
$\lambda_{AC,m}^{H_0,max}$	20
y_{BC}^{THR}	[0 - 30]

2.2 Reproduced Figures

- **Reproduced Figure-1**

  This is the plot of the channel coefficient, $h_m[n,l]$, $l \in 1,...,5$ in terms of drift velocity (v_m) from (4) and (5). w.k.t.,

  $\mu_m(t) = l_m/v_m(t) = 1/v_m(t)$

  $\sigma_m^2(t) = (\delta_m(t))/2 = 0.5/2 = 1/4$

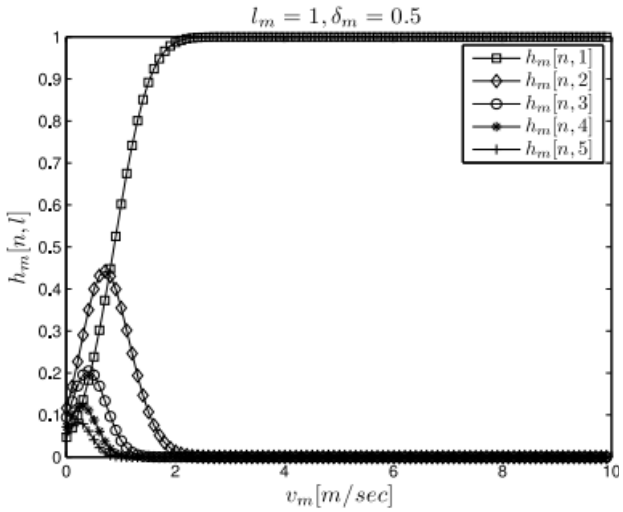
  $\lambda_m(t) = (l_m^2)/\sigma_m^2(t) = 4$

  $w = lTs$ where $Ts = 1$

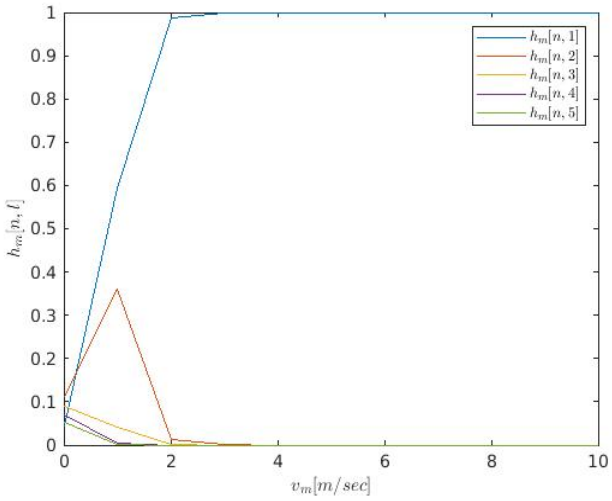
  Substituting above values in (5)

$$F_{W,m} = \Phi\left(\sqrt{\frac{4}{l}}\left(l \times v_m(t) - 1\right)\right) + e^{2\times 4\times v_m(t)}\left(-\sqrt{\frac{4}{l}}\left(l \times v_m(t) + 1\right)\right) \tag{24}$$

  This is substituted in (4) to give us the final plot.



(a) 1B



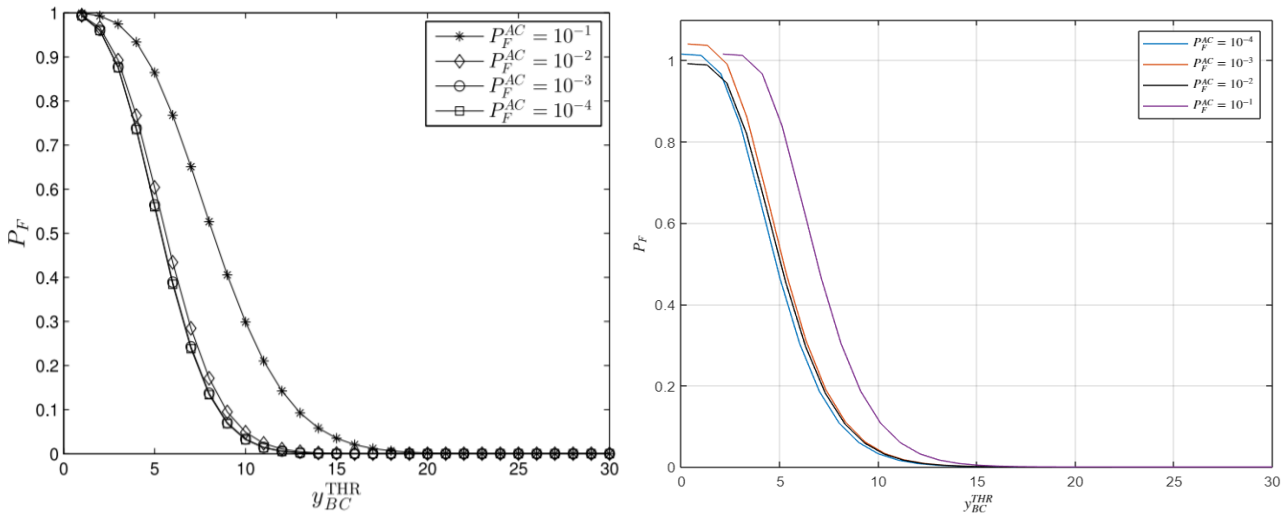
(b) 1R

Inference:

This plot shows the propagation probability of receiving molecules in the nth time slot from the transmitted molecules in $n-l$ prior ones, that is the channel coefficient $h_m[n,l]$. It's done for different values of l, against the different values of v_m. We can see that the probability of receiving molecules in same time slot, that is for $l = 1$ converges to 1 for values of $v_m > 2.5$. For values less that 2.5, we can see there are chances that it may take another slot to be received, but the probability that it takes more than 2 slots is negligible. For values $l > 1$, it reduces to 0 for the same value of $v_m = 2.5$, and are quite low for $l > 2$ to begin with. Most transmitted molecules are received within the same time slot for $vm > 2.5$.

- **Reproduced Figure-2**

This is the plot of P_F in terms of y_{BC}^{THR} for different values of P_F^{AC}, $\Delta\lambda = 1$ and $\bar{\epsilon} = 100$ using equations (22) and (23).



(a) 2B (b) 2R

Inference:

Here, we plot P_F in terms of y_{BC}^{THR} for different values of P_F^{AC}, and we see that P_F is monotonically decreasing function of y_{BC}^{THR} and also an increasing function of P_F^{AC}. Except for the greater value of P_F^{AC} that is 10^{-1}, the other lines appear more or less similar. This shows that with smaller values of P_F^{AC}, that is 10^{-2} and lower, the AC-BC channel has a dominant effect on P_F as in the occurrence of H_0 hypothesis, the DC-AC channel is error free. But for the the greater ones, like 10^{-1}, change is observed so which is the dominant effect of the DC-AC channel. This is why on changing the values of P_F^{AC} from 10^{-2} to 10^{-1}, a shift towards the right can be noticed on the P_F.

3 Problems faced in Reproduction.

1. **Lack of parameters.**

   To take one example, in (3), most of the parameter values, like L^{hm}, t, Ts, τ were undefined values. Further more, even for $x_D(\tau)$ from (1) which is a part of (3), the values of x_{D0}, k and b were unknown as well, and we were unable to assume correctly because there were a lot of factors affecting. This disabled us from attempting to reproduce Fig.3 (a) and (c) from the source paper.

2. **No equation or inequality to relate with the value of P_F^{AC}**

   There was a lack of any solid P_F^{AC} relation except for (23), that is with P_F. Hence, we were unable to successfully form an equation of P_D^{AC} with P_F^{AC} as a direct parameter or vice versa. This in turn also hindered us from forming a direct relation between P_D and P_F or P_D and P_F^{AC}. These relations are required for the majority of the figures in the source paper, that is Fig. 4-7, 9-12, so we had reached a dead end when it came to those.

3. **Lack of familiarity with plotting tools.**

   As this was a very fresh research paper, only from July this year, there was less to nothing guidance existing regarding the codes on the internet. Most existing plots relating to the concept of detection and false alarm are in terms of SNR. So it was a challenge to try and write the entire code on our own without base coding framework or modules to refer to or access tools like Mathematica. This was a slight handicap that we faced.

4 Contribution of team members

4.1 Technical contribution of all team members

Enlist the technical contribution of members in the table. Redefine the tasks (e.g Task-1 as simualtion of fig.1 and so on)

Tasks	**Kesha Bagadia**	**Yashvi Pipaliya**	**Yashvi Gandhi**	**Manal Shah**
Mathematical model	Yes	Yes	No	Yes
Plotting	Yes	Yes	Yes	No

4.2 Non-Technical contribution of all team members

Enlist the non-technical contribution of members in the table. Redine the tasks (e.g Task-1 as report writing etc.)

Tasks	**Kesha Bagadia**	**Yashvi Pipaliya**	**Yashvi Gandhi**	**Manal Shah**
Report Writing	Yes	Yes	Yes	Yes
Conceptual analysis (MIRO)	No	Yes	Yes	Yes
Research	Yes	Yes	Yes	Yes

References

[1] S. Ghavami, "Anomaly detection in molecular communications with applications to health monitoring networks," *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, vol. 6, no. 1, pp. 50–59, 2020.