HOW MANY MODES CAN A CONSTRAINED GAUSSIAN MIXTURE HAVE?

NAVIN KASHYAP AND MANJUNATH KRISHNAPUR

ABSTRACT. We show, by an explicit construction, that a mixture of univariate Gaussians with variance 1 and means in $[-A, A]$ can have $\Omega(A^2)$ modes. This disproves a recent conjecture of Dytso, Yagli, Poor and Shamai [3] who showed that such a mixture can have at most $O(A^2)$ modes and surmised that the upper bound could be improved to $O(A)$. Our result holds even if an additional variance constraint is imposed on the mixing distribution. Extending the result to higher dimensions, we exhibit a mixture of Gaussians in $\mathbb{R}^d$, with identity covariances and means inside $[-A, A]^d$, that has $\Omega(A^{2d})$ modes.

$\delta_{a_1} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-a_1)^2}$

$Z = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$

1. INTRODUCTION

Let X be a random variable with distribution $\mu = p_1\delta_{a_1} + \ldots + p_N\delta_{a_N}$ where $-A \le a_1 < a_2 < \ldots < a_N \le A$ and $p_i > 0$ sum to 1. Throughout this note, Z denotes a standard Gaussian random variable that is independent of X. Then, $Y = X + Z$ has density $f_Y(t) = \sum_{k=1}^N p_k\varphi(t - a_k)$, where $\varphi(t) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}$. We want to know the maximum number of modes (local maxima) that f_Y, a mixture of Gaussians with centres (means) a_k constrained to be in $[-A, A]$, can have. Let this quantity be denoted as $\mathsf{m}(A)$. The main aim of this note is to give a proof of the following proposition.

Proposition 1. $\mathsf{m}(A) = \Omega(A^2)$, *i.e.*, $\mathsf{m}(A) \ge c_0A^2$ *for some constant* $c_0 > 0$ *and all* $A > 0$.

Remark 1. It was recently shown by Dytso, Yagli, Poor and Shamai [3, Theorem 6] that $\mathsf{m}(A) \le c_1A^2$, for some constant $0 < c_1 < \infty$. This, along with our Proposition 1 above, shows that $\mathsf{m}(A) = \Theta(A^2)$. In particular, this disproves the conjecture made by Dytso et al. [3, Remark 9] that $\mathsf{m}(A) = \Theta(A)$.[1]

The motivation for their conjecture was that, via [3, Eqs. (43) and (65)], $2\mathsf{m}(A)$ is an upper bound for $N^*(A)$, which is the number of points in the support of the optimal input distribution for an additive white Gaussian noise (AWGN) channel with amplitude constraint A. Thus, one consequence of their conjecture would have been that $N^*(A) = O(A)$. In fact, since they show that $N^*(A) = \Omega(A)$, their conjecture would have implied that $N^*(A) = \Theta(A)$. While our proposition

N.K. is partially supported by the SERB MATRICS grant MTR/2017/000368. M.K. is partially supported by UGC Centre for Advanced Study and the SERB MATRICS grant MTR/2017/000292.

[1] Independently of us, Polyanskiy and Wu [4] have also obtained a result that effectively disproves this conjecture. They give an example of a random variable X having a *density* π supported within $[-A, A]$ such that the density, $\pi * \varphi$, of $X + Z$ has $\Omega(A^2)$ modes.

shows that the route via their conjecture is blocked, numerical work does indeed suggest that $N^*(A) = \Theta(A)$. $\square$

The result of Proposition 1 does not change qualitatively if we further impose a variance constraint on the X in $Y = X+Z$. To be precise, consider now Gaussian mixtures $f_Y(t) = \sum_{k=1}^N p_k\varphi(t-a_k)$, with centres a_k again constrained to be in $[-A, A]$, but additionally requiring the random variable $X \sim \sum_{k=1}^N p_k\delta_{a_k}$ to have variance $\mathrm{var}(X) \leq 1$. (Of course, any constant bound on the variance will do; we take the bound to be 1 for simplicity.) Let $\mathsf{m}_\#(A)$ denote the maximum number of modes among such mixtures f_Y. We then have the following result.

Proposition 2. *$\mathsf{m}_\#(A) = \Omega(A^2)$, i.e., $\mathsf{m}_\#(A) \geq c_\# A^2$ for some constant $c_\# > 0$ and all $A > 0$.*

variance $\lesssim 1$

Our results extend to higher dimensions without substantial change. Let φ_d denote the standard Gaussian density (zero mean and identity covariance) in $\mathbb{R}^d$. Let $\mathsf{m}_d(A)$ denote the maximum number of modes that the Gaussian mixture density $f(t) = p_1\varphi_d(t-a_1) + \ldots + p_N\varphi_d(t-a_N)$ can have, subject to the constraints that $|a_i| \leq A$ for all i, and $p_i > 0$ sum to 1.

Proposition 3. *With the above notation, $\mathsf{m}_d(A) \geq cA^{2d}$ for a constant $c > 0$ that is independent of A.*

However, we are not aware of a corresponding upper bound. It is worth remarking here that there is considerable interest in counting modes of Gaussian mixtures. For instance, it was conjectured by Sturmfels (see [1, Conjecture 5]) that a Gaussian mixture (with identity covariances, as we have taken) with N components, has at most $\binom{N+d-1}{d}$ modes. In one dimension, this bound reduces to N, which is in fact proved in [5] — see also [2, Section 2.4]. These studies are without any constraint on the centers while the amplitude constraint is a key feature in this paper.

Sketch of the proofs. The main ingredients in our proofs of Propositions 1 and 2 are mixtures of the form

$$\gamma_{a,N}(x) := \frac{1}{2N+1}\sum_{n=-N}^{N}\varphi(x-an),$$

$$\boxed{f_Y(t) = \sum_{k=1}^N \underbrace{p_k}_{\frac{1}{2N+1}}\varphi(t-\underbrace{a_k}_{\frac{c\cdot k}{\sqrt{N}}}),}$$

with $a > 0$. This is an equally-weighted mixture of $2N+1$ Gaussians with centres (means) an, for integers n between $-N$ and N. Fig. 1 illustrates the shape of the unnormalized mixture

$$f_{a,N}(x) := \sum_{n=-N}^{N}\varphi(x-an).$$

We will show that by choosing $a = \frac{c}{\sqrt{N}}$ for a suitable constant $c > 0$, the resulting unnormalized mixture $f_{a,N}$ has centres in $[-c\sqrt{N}, c\sqrt{N}]$ and at least $N-1$ modes. Since scaling by a constant has no effect on the number of modes, the same holds for the mixture $\gamma_{a,N}$, which suffices to prove Proposition 1. The proof is elaborated in Section 2.

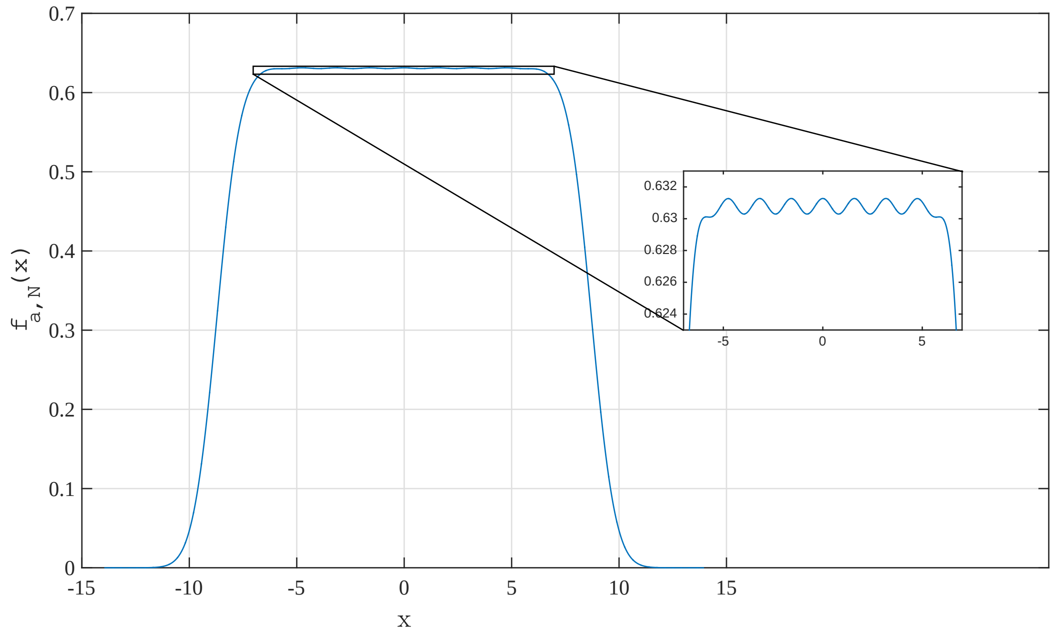

FIGURE 1. A plot of $f_{a,N}(x) = \sum_{n=-N}^{N} \varphi(x - an)$ for $N = 5$ and $a = 2\sqrt{\pi/N}$.

For Proposition 2, we work with the mixture

$$\begin{aligned}\Gamma_{\alpha;\,a,N}(x) &:= (1-2\alpha)\,\varphi(x) + \alpha\,\gamma_{a,N}(x+2aN) + \alpha\,\gamma_{a,N}(x-2aN)\\ &= (1-2\alpha)\,\varphi(x) + \frac{\alpha}{2N+1}\sum_{n=-3N}^{-N}\varphi(x-an) + \frac{\alpha}{2N+1}\sum_{n=N}^{3N}\varphi(x-an),\end{aligned} \tag{1}$$

where $a = \frac{c}{\sqrt{N}}$ is as above, and $\alpha \in (0, \frac{1}{2})$. This is a Gaussian mixture with centres at 0 and $\pm an$, $n = N, N+1, \ldots, 3N$, weighted by $1 - 2\alpha$ and $\frac{\alpha}{2N+1}$, respectively. It is easy to check that by taking $\alpha \sim \frac{1}{N}$, we can get the underlying random variable X to have variance at most 1. We will, moreover, show that for this choice of α, the mixture $\Gamma_{\alpha;\,a,N}$ has $\Omega(N)$ modes. Since $\Gamma_{\alpha;\,a,N}$ has all its centres within $[-3c\sqrt{N}, 3c\sqrt{N}]$, this will prove Proposition 2. The detailed proof is in Section 2.

The proof of Proposition 3 is entirely analogous to that of Proposition 1, and uses a mixture with equal weights and centers at ak, where $k = (k_1, \ldots, k_d) \in \mathbb{Z}^d$ with $-N \le k_i \le N$, for appropriately chosen a and N (the right choices turn out to be $a = 1/A$ and $N = A^2$). Details are in Section 3.

2. Proof of Proposition 1 and Proposition 2

Our analysis is based on the fact that, for any $a > 0$, the unnormalized mixture $f_{a,N}$ is a truncation of the infinite series

$$f_a(x) := \sum_{n\in\mathbb{Z}} \varphi(x - an).$$

$= \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-an)^2}{2}}$

$\ldots e^{-\frac{(x+1)^2}{2}} + e^{-\frac{x^2}{2}} + e^{-\frac{(x-1)^2}{2}} + \ldots$

Note that f_a is well-defined and periodic with period a. By standard real-analysis arguments, f_a is continuous on $\mathbb{R}$.

We first obtain an estimate for $h_a := f_a(0) - f_a(\frac{a}{2})$, which we will use in our proofs.

Lemma 4. *For any $a > 0$, we have*

$$\frac{4}{a}e^{-\frac{2\pi^2}{a^2}} \;\leq\; h_a \;\leq\; \frac{4}{a}e^{-\frac{2\pi^2}{a^2}}\left(1 - e^{-\frac{2\pi^2}{a^2}}\right)^{-1}.$$

Proof. We prove the lower bound first. By the Poisson summation formula[2], for any $x \in \mathbb{R}$,

$$f_a(x) = \sum_{n\in\mathbb{Z}} \varphi\left(a(\frac{x}{a} - n)\right) = \frac{1}{a}\sum_{n\in\mathbb{Z}} e^{-\frac{2\pi^2 n^2}{a^2}} e^{2\pi i n \frac{x}{a}}, \tag{2}$$

from which we get

$$f_a(0) \;=\; \frac{1}{a}\sum_{n\in\mathbb{Z}} e^{-\frac{2\pi^2 n^2}{a^2}} > \frac{1}{a} \;>\; \frac{1}{a}\sum_{n\in\mathbb{Z}} (-1)^n \, e^{-\frac{2\pi^2 n^2}{a^2}} \;=\; f_a(\tfrac{a}{2}).$$

In particular, we have

$$\begin{aligned} h_a \;&=\; \frac{1}{a}\sum_{n\in\mathbb{Z}} e^{-\frac{2\pi^2 n^2}{a^2}} - \frac{1}{a}\sum_{n\in\mathbb{Z}} (-1)^n \, e^{-\frac{2\pi^2 n^2}{a^2}} \\ &=\; \frac{2}{a}\sum_{\substack{n\in\mathbb{Z},\\ n \text{ odd}}} e^{-\frac{2\pi^2 n^2}{a^2}} \\ &=\; \frac{4}{a}\sum_{\substack{n>0,\\ n \text{ odd}}} e^{-\frac{2\pi^2 n^2}{a^2}} \\ &>\; \frac{4}{a}e^{-\frac{2\pi^2}{a^2}}. \end{aligned}$$

$|\Sigma a_i| \leq \Sigma |a_i|$

For the upper bound, consider

$$|f_a(x) - \frac{1}{a}| \;\leq\; \frac{1}{a}\sum_{n\neq 0} e^{-\frac{2\pi^2 n^2}{a^2}} \;\leq\; \frac{2e^{-\frac{2\pi^2}{a^2}}}{a\left(1 - e^{-\frac{2\pi^2}{a^2}}\right)},$$

the first inequality arising from (2), and the second inequality being obtained by replacing n^2 by n to get a geometric series. Thus,

$$h_a \;=\; |f_a(0) - \frac{1}{a}| + |f_a(\tfrac{a}{2}) - \frac{1}{a}| \;\leq\; \frac{4e^{-\frac{2\pi^2}{a^2}}}{a\left(1 - e^{-\frac{2\pi^2}{a^2}}\right)}, \tag{3}$$

$a > c > b$
$a - b = |a-c| + |b-c|$

which is the claimed upper bound. $\square$

[2]With the notation $\hat{f}(\lambda) = \int f(x)\, e^{-2\pi i \lambda x} dx$, we have $\sum_{n\in\mathbb{Z}} f(x+n) = \sum_{n\in\mathbb{Z}} \hat{f}(n) e^{2\pi i n x}$.

probably because the imaginary terms cancel out and the reals are less than 1.

Thus, for $a \ll 1$, we have $h_a \approx \frac{4}{a}\exp(-\frac{2\pi^2}{a^2})$. We actually need only the lower bound on h_a for our arguments.

Remark 2. A minor modification in the above proof shows that the bounds in Lemma 4 in fact apply to $\overline{h}_a = \max(f_a) - \min(f_a)$ as well. Indeed, the lower bound is obvious, since $\overline{h}_a \geq h_a$. For the upper bound, we observe that if x^* and x_* achieve the maximum and minimum, respectively, of f_a, then $\overline{h}_a = |f_a(x^*) - \frac{1}{a}| + |f_a(x_*) - \frac{1}{a}|$, so that the upper bound in (3) still holds.

It is clear from (2) that $f_a(0) > f_a(x)$ for all $x \in [-\frac{a}{2}, \frac{a}{2}]$, since there is non-trivial cancellation in the terms of the series unless x is an integer multiple of a. By the fact that f_a has period a, we see that na is a strict maximum of f_a in the interval $I_{a,n} := [na - \frac{a}{2}, na + \frac{a}{2}]$ for any $n \in \mathbb{Z}$. We wish argue that $f_{a,N}$ also has local maxima within those intervals $I_{a,n}$ that are contained in $[-\frac{1}{2}aN, \frac{1}{2}aN]$. For this, we will need the simple lemma stated next.

Lemma 5. *Let g be a continuous function such that $|f_a - g| < \frac{1}{2}h_a$ on a subset $S \subseteq \mathbb{R}$. Then, g has a local maximum in the interior of any interval $I_{a,n}$ that is contained within S.*

Proof. Recall that $I_{a,n} = [na - \frac{a}{2}, na + \frac{a}{2}]$, for $n \in \mathbb{Z}$. If $|f_a - g| < \frac{1}{2}h_a$ holds on $I_{a,n}$, then we have

$$\begin{aligned} g(na) - g(na - \tfrac{a}{2}) &= \big(g(na) - f_a(na)\big) + \big(f_a(na) - f_a(na - \tfrac{a}{2})\big) + \underbrace{\big(f_a(na - \tfrac{a}{2}) - g(na - \tfrac{a}{2})\big)}_{\text{min value is } \frac{h_a}{2}} \\ &> (-\frac{1}{2}h_a) + h_a + (-\frac{1}{2}h_a) \\ &= 0. \end{aligned}$$

Hence, $g(na) > g(na - \frac{a}{2})$. Analogously, $g(na) > g(na + \frac{a}{2})$. Therefore, the global maximum of g in $I_{n,a}$ is attained at and interior point. In particular, g has a local maximum strictly between $na - \frac{a}{2}$ and $na + \frac{a}{2}$. $\square$

We now have the facts necessary to furnish proofs of Propositions 1 and 2.

Proof of Proposition 1. We apply Lemma 5 with $g = f_{a,N}$. Note first that

$$\begin{aligned} |f_a(x) - f_{a,N}(x)| &= \frac{1}{\sqrt{2\pi}} \sum_{n:|n|>N} e^{-\frac{1}{2}(an-x)^2} \\ &\leq \frac{1}{\sqrt{2\pi}} \sum_{n:|n|>N} e^{-\frac{1}{2}(a|n|-|x|)^2} \quad \big(\text{since } |an - x| \geq \big|a|n| - |x|\big|\big) \\ &= \frac{2}{\sqrt{2\pi}} \sum_{n>N} e^{-\frac{1}{2}(an-|x|)^2} \\ &= \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}(aN-|x|)^2} \sum_{n>N} e^{-\frac{1}{2}a(n-N)(a(N+n)-2|x|)} \\ &\leq \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}(aN-|x|)^2} \sum_{n>N} e^{-a(n-N)(aN-|x|)} \end{aligned}$$

Now take $|x| \le \frac{1}{2}aN$ to get

$$
\begin{aligned}
|f_a(x) - f_{a,N}(x)| &\le \frac{2}{\sqrt{2\pi}}\, e^{-\frac{1}{8}a^2N^2} \sum_{n>N} e^{-\frac{1}{2}a^2N(n-N)} \\
&= \frac{2}{\sqrt{2\pi}}\, e^{-\frac{1}{8}a^2N^2}\, \frac{e^{-\frac{1}{2}a^2N}}{1-e^{-\frac{1}{2}a^2N}}. \qquad (4)
\end{aligned}
$$

If we take $a = \frac{2\sqrt{\pi}}{\sqrt{N}}$ and $S = [-\frac{1}{2}aN, \frac{1}{2}aN] = [-\sqrt{\pi N}, \sqrt{\pi N}]$, then (4) holds for all $x \in S$, so that

$$
|f_a(x) - f_{a,N}(x)| \;\le\; C_0 e^{-\frac{1}{2}\pi N} \qquad (5)
$$

with $C_0 = \frac{2}{\sqrt{2\pi}}\left(\frac{e^{-2\pi}}{1-e^{-2\pi}}\right)$. On the other hand, from the lower bound for h_a in Lemma 4, we have

$$
h_a \ge 2\sqrt{\frac{N}{\pi}}\, e^{-\frac{1}{2}\pi N}.
$$

As $C_0 < \frac{2}{\sqrt{\pi}}\left(\frac{e^{-2\pi}}{1-e^{-2\pi}}\right)$, we have for all $N \ge 1$, $C_0 e^{-\frac{1}{2}\pi N} < \left(\frac{e^{-2\pi}}{1-e^{-2\pi}}\right) h_a$, and consequently,

$$
|f_a(x) - f_{a,N}(x)| < \left(\frac{e^{-2\pi}}{1-e^{-2\pi}}\right) h_a \quad \text{for all } x \in S.
$$

Since $\frac{e^{-2\pi}}{1-e^{-2\pi}} \approx 0.0019$, the conclusion of Lemma 5 holds, i.e., $f_{a,N}$ has a local maximum in the interior of each of the intervals $I_{a,n}$ contained in $S = [-\frac{1}{2}aN, \frac{1}{2}aN]$. There are at least $N-1$ such intervals $I_{a,n}$, and hence, $f_{a,N}$ has at least $N-1$ local maxima within S. Thus, we conclude that the Gaussian mixture $\gamma_{a,N} = \frac{1}{2N+1} f_{a,N}$ (with $a = \frac{2\sqrt{\pi}}{\sqrt{N}}$), which has all its centres inside $[-2\sqrt{\pi N}, 2\sqrt{\pi N}]$, has at least $N-1$ modes (within $S = [-\sqrt{\pi N}, \sqrt{\pi N}]$). Choosing $N = A^2$ proves Proposition 1. □

Proof of Proposition 2. Consider $\Gamma_{\alpha;a,N}$ as defined in (1), with $a = \frac{2\sqrt{\pi}}{\sqrt{N}}$ as in the proof of Proposition 1. This is the density of $Y = X + Z$, where $Z \sim \mathcal{N}(0,1)$ is independent of $X \sim (1-2\alpha)\delta_0 + \frac{\alpha}{2N+1}\sum_{n=N}^{3N}(\delta_{-an} + \delta_{an})$. We then have

$$
\begin{aligned}
\mathrm{var}(X) &= \frac{\alpha}{2N+1}\sum_{n=N}^{3N} 2(an)^2 \\
&\le \frac{2\alpha a^2}{2N+1}\sum_{n=1}^{3N} n^2 \\
&= \frac{2\alpha a^2}{2N+1}\left(\frac{3N(3N+1)(6N+1)}{6}\right) \\
&\le \alpha a^2 (3N)(3N+1) \\
&= 12\pi(3N+1)\alpha \qquad (\text{using } a = \tfrac{2\sqrt{\pi}}{\sqrt{N}})
\end{aligned}
$$

Hence, setting $\alpha = \frac{1}{12\pi(3N+1)}$, we obtain $\mathrm{var}(X) \le 1$.

We will next show that, with a and α as above, $\Gamma_{\alpha;a,N}$ has $\Omega(N)$ modes. This suffices to prove the proposition, since $\Gamma_{\alpha;a,N}$ is a Gaussian mixture with all of its centres in $[-6\sqrt{\pi N}, 6\sqrt{\pi N}]$.

It is easy to check that $\Gamma_{\alpha;a,N}$ has a mode at 0. We will show that, when N is sufficiently large, $\Gamma_{\alpha;a,N}$ has at least $N-1$ modes in each of the intervals $[-5\sqrt{\pi N}, -3\sqrt{\pi N}]$ and $[3\sqrt{\pi N}, 5\sqrt{\pi N}]$. By symmetry, it is enough to show this for the interval $[3\sqrt{\pi N}, 5\sqrt{\pi N}]$. For this, we use Lemma 5 with $g = \left(\frac{2N+1}{\alpha}\right)\Gamma_{\alpha;a,N}$. For this choice of g, we have

$$
\begin{aligned}
|f_a(x) - g(x)| &= \left| \sum_{n:|n|<N \text{ or } |n|>3N} \varphi(x-an) - \left(\frac{1-2\alpha}{\alpha}\right)(2N+1)\varphi(x) \right| \\
&\leq \sum_{n:|n|<N \text{ or } |n|>3N} \varphi(x-an) + \left(\frac{1-2\alpha}{\alpha}\right)(2N+1)\varphi(x) \\
&\leq \sum_{n<N \text{ or } n>3N} \varphi(x-an) + \left(\frac{2N+1}{\alpha}\right)\varphi(x) \qquad (6)
\end{aligned}
$$

Consider the first term in (6) above. Writing $x' = x - 2aN$, we have

$$
\begin{aligned}
\sum_{n<N \text{ or } n>3N} \varphi(x-an) &= \sum_{n<N \text{ or } n>3N} \varphi(x' - a(n-2N)) \\
&= \sum_{n<-N \text{ or } n>N} \varphi(x'-an) \\
&= |f_a(x') - f_{a,N}(x')| \\
&\leq\ C_0 e^{-\frac{1}{2}\pi N}
\end{aligned}
$$

for $|x'| \leq \frac{1}{2}aN$ and $C_0 = \frac{2}{\sqrt{2\pi}}\left(\frac{e^{-2\pi}}{1-e^{-2\pi}}\right)$, by (5) in the proof of Proposition 1. Thus, for $|x - 2aN| \leq \frac{1}{2}aN$, i.e., for $x \in [3\sqrt{\pi N}, 5\sqrt{\pi N}]$, we see that the first term in (6) is bounded above by $C_0 e^{-\frac{1}{2}\pi N}$.

Turning our attention to the second term in (6), we first observe that $\frac{2N+1}{\alpha} \leq C_0' N^2$ for some constant C_0'. Thus,

$$
\left(\frac{2N+1}{\alpha}\right)\varphi(x) \leq C_0' N^2 \varphi(x) \leq \frac{1}{\sqrt{2\pi}} C_0' N^2 e^{-\frac{9}{2}\pi N},
$$

for $x \geq 3\sqrt{\pi N}$.

Combining these bounds, we obtain that for $x \in [3\sqrt{\pi N}, 5\sqrt{\pi N}]$,

$$
|f_a(x) - g(x)|\ \leq\ C_0 e^{-\frac{1}{2}\pi N} + \frac{1}{\sqrt{2\pi}} C_0' N^2 e^{-\frac{9}{2}\pi N}\ \leq\ 2C_0 e^{-\frac{1}{2}\pi N}
$$

when N is sufficiently large. As shown in the proof of Proposition 1, $C_0 e^{-\frac{1}{2}\pi N} < \left(\frac{e^{-2\pi}}{1-e^{-2\pi}}\right) h_a$. Consequently, when N is sufficiently large, for $x \in [3\sqrt{\pi N}, 5\sqrt{\pi N}]$, we have

$$
|f_a(x) - g(x)|\ <\ \left(\frac{2e^{-2\pi}}{1-e^{-2\pi}}\right) h_a\ <\ 0.004\, h_a.
$$

Then, applying Lemma 4, we obtain that, for all sufficiently large N, the function $g = \left(\frac{2N+1}{\alpha}\right)\Gamma_{\alpha;a,N}$ has at least $N-1$ modes within the interval $[3\sqrt{\pi N}, 5\sqrt{\pi N}]$. This naturally holds for $\Gamma_{\alpha;a,N}$ as well, thus proving the proposition. □

3. Proof of Proposition 3

Since the proof is entirely analogous to that of Proposition 1, we shall only sketch the modifications needed and omit the details. For $a > 0$ and integer $N \geq 1$ and define the functions

$$f_a(x) = \sum_{n \in \mathbb{Z}^d} \varphi_d(x - na),$$
$$f_{a,N}(x) = \sum_{n \in Q_N} \varphi_d(x - na),$$

where $Q_N = \{n \in \mathbb{Z}^d \,:\, -N \leq n_i \leq N \text{ for } 1 \leq i \leq d\}$. By the Poisson summation formula on $\mathbb{R}^d$ with respect to the lattice $\mathbb{Z}^d$, we get

$$\begin{aligned} f_a(x) &= \frac{1}{a^d} \sum_{p \in \mathbb{Z}^d} e^{-\frac{1}{2a^2}|p|^2 + \frac{2\pi i}{a}\langle p, x\rangle} \\ &= \frac{1}{a^d}\left(1 + 2e^{-\frac{1}{2a^2}} \sum_{j=1}^{d} \cos(2\pi x_j/a) + O(e^{-\frac{2}{a^2}})\right) \end{aligned}$$

where the big-O term includes the contribution of all p with $|p| \geq 2$. Since $\cos(2\pi t) \leq 1 - 8t^2$ for any $t \in \mathbb{R}$, we see that when $|x| = \frac{a}{2}$,

$$\begin{aligned} f_a(x) &\leq \frac{1}{a^d}\left(1 + 2e^{-\frac{1}{2a^2}} \sum_{j=1}^{d} (1 - \tfrac{8}{a^2}x_j^2) + O(e^{-\frac{2}{a^2}})\right) \\ &= \frac{1}{a^d}\big(1 + 2(d-2)e^{-\frac{1}{2a^2}} + O(e^{-\frac{2}{a^2}})\big). \end{aligned}$$

Since $f_a(0) = \frac{1}{a^d}(1 + 2de^{-1/(2a^2)} + O(e^{-2/a^2}))$, we see that

$$f_a(0) - \sup_{|x| = \frac{a}{2}} f_a(x) = \frac{1}{a^d}\big(4e^{-\frac{1}{2a^2}} + O(e^{-\frac{2}{a^2}})\big)$$

which is at least $h_a := \frac{3}{a^d}e^{-\frac{1}{2a^2}}$, for small enough a. By periodicity, in each cube of the form $na + [-\frac{1}{2}a, \frac{1}{2}a]^d$, the graph of f_a has a hill with peak at na and having height at least h_a. Further,

$$\begin{aligned} |f_a(x) - f_{a,N}(x)| &= \frac{1}{(2\pi)^{\frac{d}{2}}} \sum_{n \in \mathbb{Z}^d \setminus Q_N} e^{-\frac{1}{2a^2}|(x+na)|^2} \\ &= O(e^{-\frac{1}{8}a^2N^2}) \quad \text{for } |x| \leq \frac{1}{2}aN. \end{aligned}$$

Now take $a = \frac{c}{\sqrt{N}}$ to see that for suitable c, c',

$$\sup_{|x| \le c'\sqrt{N}} |f_a(x) - f_{a,N}(x)| < \frac{1}{2} h_a.$$

Therefore, the function $f_{a,N}$ has a local maximum in each cube of the form $na + [-\frac{1}{2}a, \frac{1}{2}a]^d$ that is contained inside the larger cube $[-c'\sqrt{N}, c'\sqrt{N}]^d$. This is because the perturbation is too small to wash away the local maximum of f_a located at na. The number of such cubes is about $(2c'\sqrt{N}/a)^d$, which is $\Theta(N^d)$.

Taking $N = \sqrt{A}$ gives us a function $f_{a,N}$ (with $a = c/A$) that is a mixture of Gaussians with centers in Q_A and having $\Theta(A^{2d})$ modes. This was the claim of Proposition 3.

Acknowledgement

We would like to thank Alex Dytso for asking us whether imposing a variance constraint on the mixing distribution would influence the number of modes, which led us to Proposition 2.

References

[1] C. Améndola, A. Engström, and C. Haase, "Maximum number of modes of Gaussian mixtures", *Information and Inference: A Journal of the IMA*, iaz013, June 2019. [Online] https://doi.org/10.1093/imaiai/iaz013

[2] M. Carreira-Perpinan and C. Williams, "On the number of modes of a Gaussian mixture", Informatics Research Report EDI-INF-RR-0159, School of Informatics, Univ. of Edinburgh, Feb. 2003. [Online] https://faculty.ucmerced.edu/mcarreira-perpinan/papers/EDI-INF-RR-0159.pdf

[3] A. Dytso, S. Yagli, H. V. Poor, and S. Shamai (Shitz), "The capacity achieving distribution for the amplitude constrained additive Gaussian channel: An upper bound on the number of mass points", *IEEE Transactions on Information Theory*, 66(4): 2006–2022, April 2020.

[4] Y. Polyanskiy and Y. Wu, "Self-regularizing property of nonparametric maximum likelihood estimator in mixture models," draft manuscript, May 2020.

[5] B.W. Silverman, "Using kernel density estimates to investigate multimodality", *Journal of the Royal Statistical Society, B*, 43(1): 97–99, 1981.

Navin Kashyap, Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, India.

E-mail address: nkashyap@iisc.ac.in

Manjunath Krishnapur, Department of Mathematics, Indian Institute of Science, Bangalore, India.

E-mail address: manju@iisc.ac.in