

Redes Neuronales Artificiales en la Selección Bayesiana de Modelos

Isidro Gómez-Vargas¹♣, JA Vázquez² y Ricardo García-Salcedo¹

¹ CICATA-Legaria, Instituto Politécnico Nacional, 11500 CDMX, México.

² Instituto de Ciencias Físicas, UNAM, 62210 Cuernavaca, Morelos, México.

♣ igomezv0701@alumno.ipn.mx

INTRODUCCIÓN

Teorema de Bayes:

$$P(\theta|D, H) = \frac{P(D|\theta, H)P(\theta|H)}{P(D|H)},$$

El denominador es conocido como evidencia Bayesiana:

$$P(D|H) = \int d^N \theta P(D|\theta, H)P(\theta|H),$$

Ésta sirve para la selección bayesiana de modelos, es decir, compararlos y determinar el mejor.

Por otro lado, el Teorema Universal de Aproximación, establece que cualquier Red Neuronal Artificial (RNA), con al menos tres capas, se puede aproximar a cualquier función bajo ciertos requerimientos simples. Por esto se propone calcular la integral de la evidencia Bayesiana mediante una Red Neuronal.

TRABAJO EN SEMESTRES PREVIOS

Aplicación de métodos MCMC, en particular, a observaciones y modelos cosmológicos. También se han analizado algoritmos (samplers) para muestrear el espacio de parámetros y hacer selección de modelos. Se han estudiado diversas arquitecturas de RNAs.

METODOLOGÍA Y PROPUESTA

Se eligen unas distribuciones de prueba:

- Función parabolode: $z = f(x, y) = x^2 + y^2$
- Distribución tipo *eggbox*: $z = [2,5 + \cos(\frac{x}{2})\cos(\frac{y}{2})]^5$

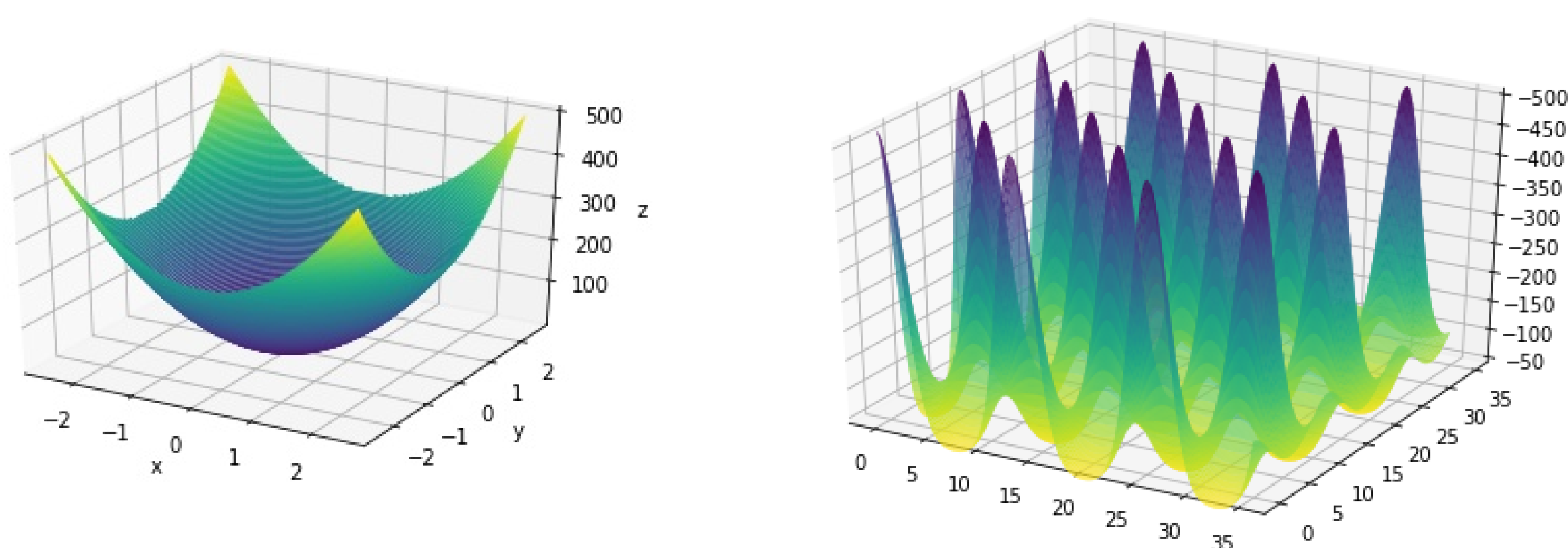


Fig. 1. Distribuciones de prueba: parabolode y caja de huevos (eggbox)

Para llevar a cabo el muestreo y cálculo de evidencia bayesiana se utiliza MULTINEST [2] (mejora del algoritmo Nested Sampling [1]). Éste se optimiza mediante una RNA sugerida en BAMBI (Ref. [3]) por medio de pyBAMBI^a.

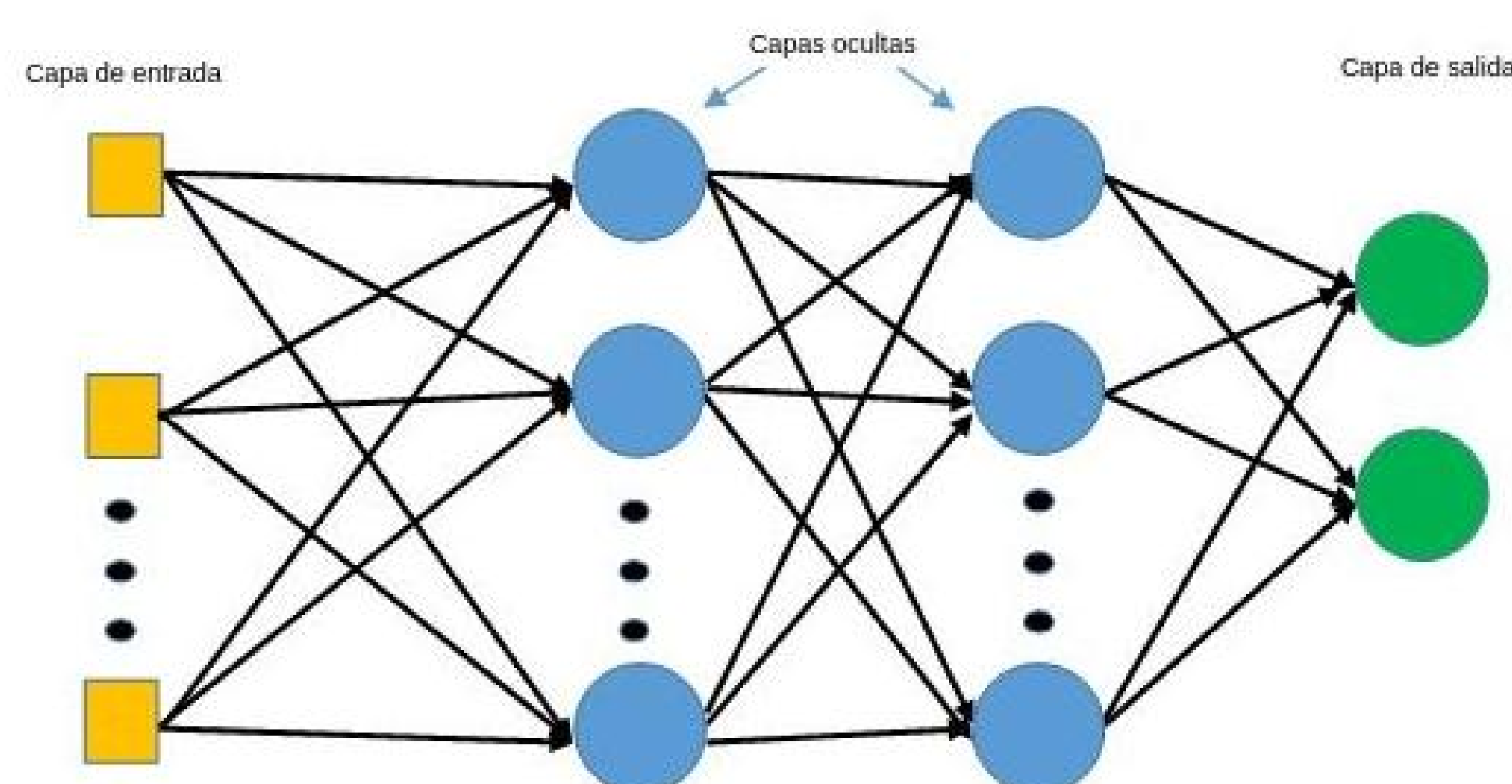


Fig. 2. Ejemplo de red neuronal artificial, perceptrón multicapa.

Además, hemos propuesto mejorar la RNA de pyBAMBI mediante un pre-entrenamiento *batch normalization*.

^a<https://pybambi.readthedocs.io/en/latest/>

RESULTADOS

No. Eval.	Tiempo MULTINEST (M)	Tiempo BAMBI (B)
1 000	0.0482 s	0.495 s
10 000	0.416 s	0.771 s
100 000	4.318 s	3.719 s
1 000 000	46.187 s	32.716 s

Tabla 1. Tiempos de ejecución para diferentes números de evaluaciones de la evidencia Bayesiana de la función parabolode.

Se obtuvo que:

$$\log(E)_M = -0,782 \pm -0,06 ; \log(E)_B = 0,783 \pm -0,06$$

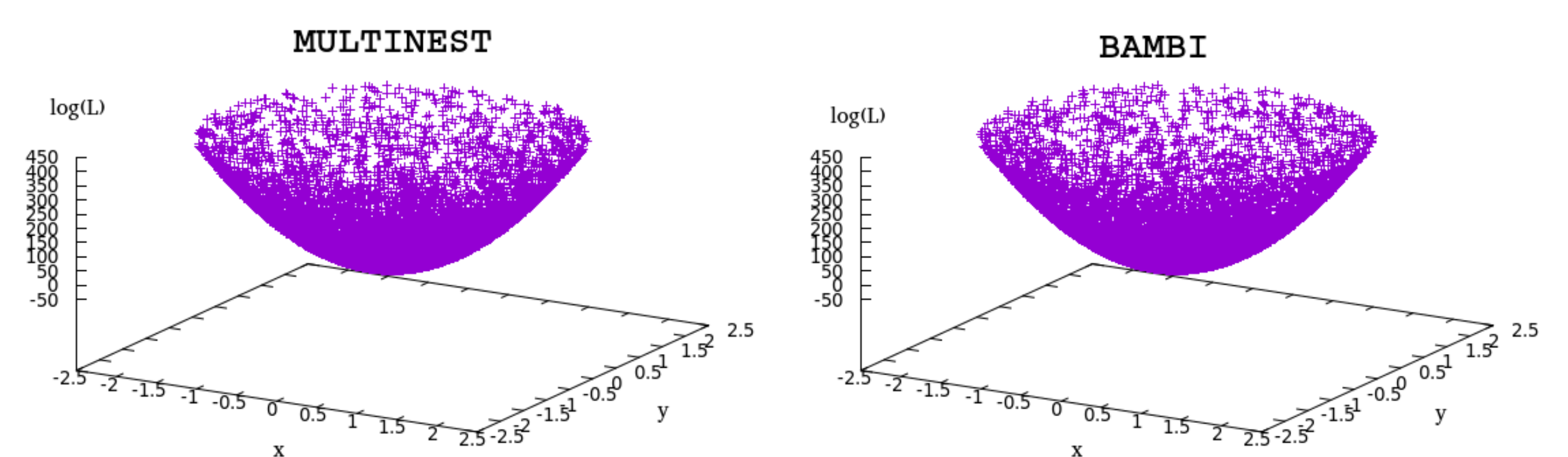


Fig. 3. Muestreo del parabolode.

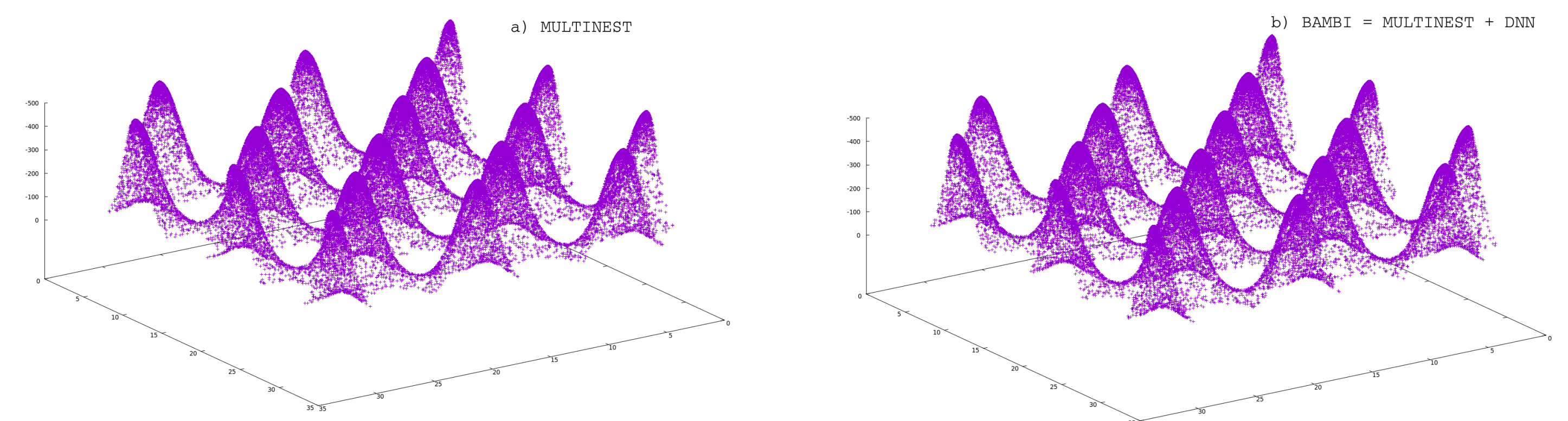


Fig. 4. Muestreo de la distribución *eggbox*

Sampler	steps	log(Evidence)	time (minutes)
MULTINEST	10 000	235,745 ± 0,025	4,48
+DNN	10 000	235,865 ± 0,025	6,84
MULTINEST	50 000	235,835 ± 0,011	132,70
+DNN	50 000	235,835 ± 0,011	166,28
MULTINEST	100 000	235,840 ± 0,008	47316,96
+DNN	100 000	235,854 ± 0,008	36743,78

Tabla 2. Comparación de evidencias y muestras de la distribución *eggbox*

También a la RNA le agregamos *batch normalization* y un regularizador *dropout*. Observamos que estos dos elementos pueden reducir aún más el tiempo, por ejemplo, en 50,000 pasos, el tiempo se reduce 6 minutos con relación a la RNA original.

CONCLUSIÓN Y TRABAJO FUTURO

Las RNAs pueden aprender, a partir de un *sampler* como MULTINEST, a calcular la evidencia y, una vez entrenadas, permiten ahorrar tiempo en su cálculo a través de las iteraciones del muestreo. A corto plazo, esperamos incluir este tipo de RNA a un código de estimación de parámetros cosmológicos y explorar nuevas técnicas de optimización para las RNAs.

REFERENCIAS

- [1] Skilling, J. (2006). Nested sampling for general Bayesian computation. *Bayesian analysis*. Vol. 1, No. 4.
- [2] Feroz, F., et al. (2009). MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Monthly Notices of the Royal Astronomical Society*. Vol. 398, No. 4.
- [3] Graff, P., et al. (2012). BAMBI: blind accelerated multimodal Bayesian inference. *Monthly Notices of the Royal Astronomical Society*. Vol. 421, No. 1.