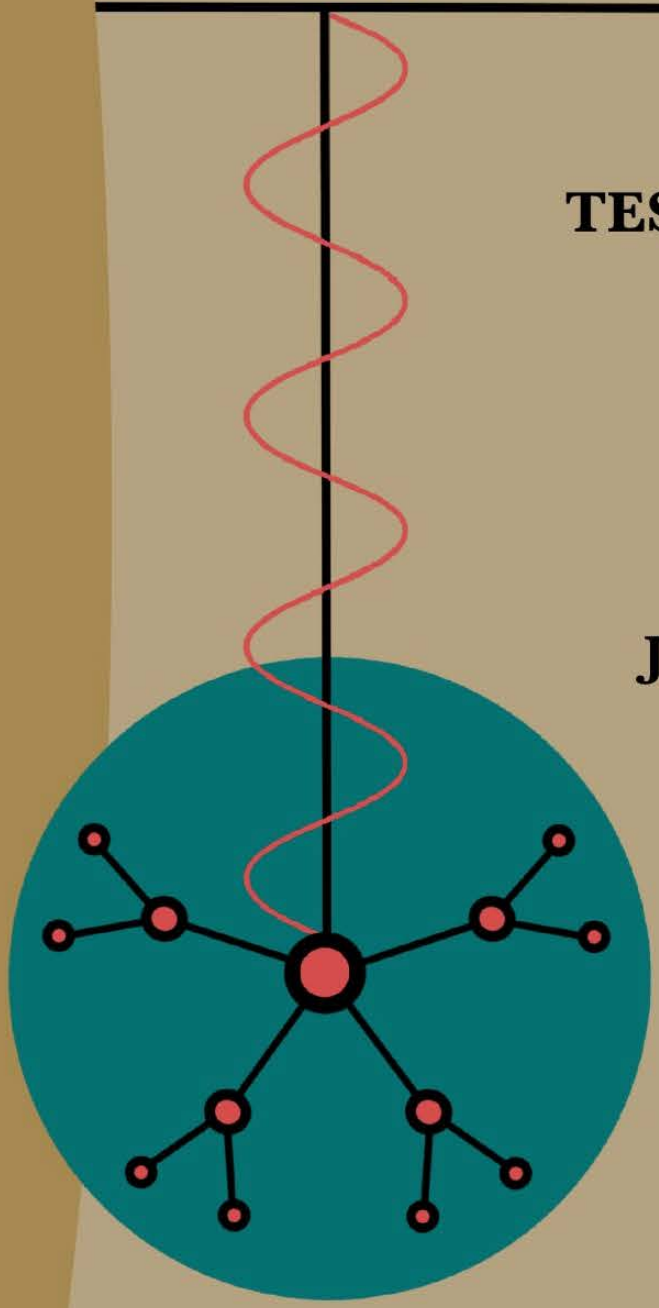
Programa de Doctorado en Bioquímica y Biología Molecular

Estudio del Impacto de la Dinámica Neuronal en el Procesamiento de Información en la Capa Granular del Cerebelo



TESIS DOCTORAL

Autora:

Milagros Marín Alejo

Directores de tesis:

María José Sáez Lara
Jesús A. Garrido Alcázar

Granada, 2021

UNIVERSIDAD DE GRANADA

Departamento de Bioquímica y Biología Molecular I

Centro de Investigación en Tecnologías de la Información y la Comunicaciones

Programa de Doctorado en Bioquímica y Biología Molecular

Universidad de Granada

Estudio del Impacto de la Dinámica Neuronal en el Procesamiento de Información en la Capa Granular del Cerebelo

Tesis Doctoral

Milagros Marín Alejo

2021

Editor: Universidad de Granada. Tesis Doctorales
Autor: Milagros Marín Alejo
ISBN: 978-84-1306-981-4
URI: http://hdl.handle.net/10481/70162

Estudio del Impacto de la Dinámica Neuronal en el Procesamiento de Información en la Capa Granular del Cerebelo

(Study of the Impact of the Neuronal Dynamics on Information Processing in the Granular Layer of the Cerebellum)

Memoria presentada por la Graduada en Bioquímica

(Con mención en Biomedicina Molecular)

Milagros Marín Alejo

para optar al

Título de Doctor

Fdo.: Milagros Marín Alejo

Bajo el VºBº de los Directores de tesis

Fdo.: María José Sáez Lara

Doctora en Bioquímica

Profesora Contratada Doctora del Departamento de Bioquímica y Biología Molecular I

Fdo.: Jesús A. Garrido Alcázar

Doctor en Ingeniería Informática

Profesor Contratado Doctor del Departamento de Arquitectura y Tecnología de Computadores

Universidad de Granada

2021

Esta Tesis Doctoral se ha realizado de forma interdisciplinar en el Departamento de Bioquímica y Biología Molecular I y en el Centro de Investigación en Tecnologías de la Información y de las Comunicaciones (CITIC) de la Universidad de Granada, y ha sido financiada por los siguientes proyectos:

- Proyecto Europeo "Human Brain Project" [HBP, Specific Grant Agreement 2 (SGA2 H2020-RIA 785907) and 3 (SGA3 H2020-RIA 945539)] financiado por EU (H2020), del 01-10-2013 hasta 31-03-2023.
- Proyecto Nacional "Integración sensorimotora para control adaptativo mediante aprendizaje en cerebelo y centros nerviosos relacionados. Aplicación en robótica" [INTSENSO (MICINN-FEDER-PID2019-109991GBI00)] financiado Ministerio de Ciencia e Innovación (MICINN), del 01-01-2020 hasta 31-12-2022.
- Proyecto Regional "Cerebelo y oliva inferior en tareas de adaptación sensorimotora" [CEREBIO (J.A. P18-FR-2378)] financiado por la Junta de Andalucía (Proyecto Excelencia), del 01-01-2020 hasta 31-12-2022.

Los resultados de esta Tesis Doctoral han sido publicados en:

1. **Marín, M.**, Esteban*, F. J., Ramírez-Rodrigo, H., Ros, E., & Sáez-Lara*, M. J. (2019). An integrative methodology based on protein-protein interaction networks for identification and functional annotation of disease-relevant genes applied to channelopathies. *BMC Bioinformatics*, *20* (1), 1-13.
2. **Marín*, M.**, Sáez-Lara, M. J., Ros, E., & Garrido*, J. A. (2020). Optimization of Efficient Neuron Models with Realistic Firing Dynamics. The Case of the Cerebellar Granule Cell. *Frontiers in Cellular Neuroscience*, *14*.
3. Cruz, N. C., **Marín*, M.**, Redondo, J. L., Ortigosa, E. M., & Ortigosa, P. M. (2021). A Comparative Study of Stochastic Optimizers for Fitting Neuron Models. Application to the Cerebellar Granule Cell. *Informatica*, 1-22.
4. **Marín*, M.**, Cruz, N. C., Ortigosa, E. M., Sáez-Lara, M. J., Garrido, J. A., & Carrillo, R. R. (2021). On the use of a multimodal optimizer for fitting neuron models. Application to the cerebellar granule cell. *Frontiers in Neuroinformatics*, *15*, 17.

** Autor de correspondencia*

A la memoria de mi padre, quien más ha creído en mí

A mi madre y mi hermana, pilares fundamentales

Agradecimientos

"El agradecimiento es la memoria del corazón"
Lao-tsé

A pesar de que es difícil encontrar las palabras para describir unos agradecimientos, no encuentro mejor forma de abrir esta sección que con esta frase. Y es que detrás del fin de una etapa tan importante para mí, y el comienzo de la siguiente, existe una red de apoyo incondicional que ha marcado los recuerdos que tendré de por vida.

En primer lugar, me gustaría mostrar mi gratitud a las personas que han dirigido y aportado a este trabajo.

A mis dos directores de tesis, María José y Jesús, y también a Eduardo. De vosotros no sólo me quedo con que me habéis ayudado a crecer profesionalmente, a adquirir las aptitudes, conocimientos, minuciosidad y perspectiva necesaria para poder llevar a cabo esta primera etapa del camino. Sino que también me llevo un crecimiento personal y un apoyo emocional que me ha enriquecido enormemente como persona, dándome el empujón necesario en mis momentos de mayor dificultad. Agradezco cada corrección a lo largo de todos estos años, vuestra paciencia, y todas las enseñanzas que he adquirido por vuestra parte, que son innumerables. Os admiro como profesionales, por vuestra entrega, vuestro compromiso y vuestra capacidad de ver más allá, privilegio que pocos pueden disfrutar, y también os admiro como personas.

Me siento orgullosa de trabajar en un laboratorio de investigación con compañeros y compañeras que, con nuestras diferentes formaciones y líneas de investigación, siempre están aportando nuevos conocimientos y perspectivas. A Eva, Richard, Paco, Niceto, Álvaro, Ignacio y Fran, que siempre me habéis ofrecido vuestra ayuda y simpatía sin dudarlo. Por todo ello, gracias.

También me gustaría agradecer a los colaboradores de estos años. Comenzando por Paco Esteban, quien me motivó y ayudó a dar el primer paso para realizar mi primer artículo. No sólo eres un referente como investigador, y como profesor (recuerdo las clases en el Máster tan sumamente interesantes y que me inspiraron a dar el paso a nuestra colaboración), sino que como persona siempre has tenido disposición y buenas palabras, muchas gracias. Agradezco también a los colaboradores de la Universidad de Almería: Nicolás, Pilar y Juani. Es sumamente cómodo trabajar con compañeros como vosotros: atentos, siempre dispuestos y con *feedback* continuo. Vuestros conocimientos son pasmosos, y sería un honor poder seguir produciendo científicamente en adelante, complementándonos de forma interdisciplinar. Nicolás, también te agradezco la ayuda fuera de lo meramente colaborativo, como echarme una mano en otro tipo de dudas académicas y el trabajo codo con codo durante las revisiones de revista.

Quiero hacer una mención especial a Hilario. De no ser por él, no estaría en este punto de mi carrera. Ha sido el mejor profesor que he tenido durante toda mi vida. Como él me dijo en su día: "Te he hecho que el *gusanillo* de la bioinformática te haya picado". Pues fíjate… ahora recordando sobre ello en los agradecimientos de mi tesis sobre neurociencia computacional.

Ha sido tutor, guía, compañero, coautor, pero sobre todo, lo siento como amigo. Hilario es sabiduría, referente, humildad y generosidad. Un polímata digno de toda admiración. Con llegar a ser la mitad que tú algún día, estaría más que satisfecha.

Gracias a las personas que he conocido todos estos años en el CITIC, en especial a Paco Illeras por haberme salvado en momentos puntuales muy críticos. También agradecer a mi tutor del programa de doctorado y coordinador Jesús Torres por su ayuda, y a todas las personas que han hecho posible este programa de doctorado y a la Escuela de Posgrado.

También me gustaría a agradecer a mis profesores y profesoras del grado y máster. Me habéis enseñado no solo las bases teóricas, sino las aptitudes necesarias para haber podido iniciar este camino. Sois muchísimos, pero os recuerdo a todos con cariño. Me gustaría hacer especial mención a Fernando, Marichu, Signe y Juan Antonio Marchal.

En segundo lugar, pero no menos importante, me gustaría dar las gracias a mi familia.

En especial a mi madre y hermana por todo su apoyo y haberme levantado, comprendido, y motivado cuando he caído. Nunca nos hemos rendido, y como decía papá: "la familia unida jamás será vencida". Espero que os sintáis orgullosas de este trabajo, porque en gran parte es por y para vosotras. Por mi parte me siento de sobra orgullosa de teneros en mi vida, y no os cambiaría por nadie ni nada.

La otra gran parte de este trabajo va dedicado a ti, papi. Es difícil escribir esto sabiendo que nunca lo vas a poder leer. Son demasiadas cosas las que te diría pero tengo que practicar lo que peor se me da, resumir. Una parte de ti vive en mí, así que te siento muy cerca. Fuiste el mejor padre que podía tener. Siempre creíste en mí, me inculcaste la capacidad de soñar y poder lograrlo, porque "tú puedes llegar a donde quieras, chica". Me enseñaste que a pesar de tener unas dificultades enormes en la vida, aun se puede salir de ellas siendo buena persona y dando mucho mucho mucho amor a tu familia. Me enseñaste el valor, la disciplina y el coraje. También me enseñaste a conservar a esa niña en tu regazo, a creer, a sentir, a volar. Eres el hombre más importante que jamás tendré en mi vida, y un héroe. Espero que desde el cielo estés orgulloso de mí, yo lo estoy de sobra de ti. Por ti siempre recordaré quién soy, de donde vengo, y a dónde puedo llegar. Sé que si hubieras estado para celebrar esta tesis, te habrías inundado a lágrimas de orgullo. Esas mismas lágrimas pero de nostalgia inundan ahora mis mejillas, aunque reconozco que también con trazas de tristeza. Aun así me siento en paz. Te quiero. Concluyo con que esto es el más alto grado académico al que puedo aspirar, y te lo dedico.

Especial mención a mis sobris, Carlos y Hugo, sois como mis hijos-hermanos. Quizás ahora ni lleguéis a leer esta parte, pero si alguna vez lo hacéis, quiero dejar aquí constancia de que creo en vosotros más que nadie en el mundo, os quiero muchísimo y llegaréis muy lejos. Este trabajo de "la tata" espero que os motive a nunca dejar de estudiar y aprender, y si hay miedo, hacedlo con miedo. También que, aunque vuestra madre y yo vengamos de una familia humilde con un mesón en las afueras de un pueblecito, hemos luchado y seguimos haciéndolo en la medida que podemos por nuestros sueños, entre éstos para mí el completar esta etapa.

También gracias a Alonso y sus parientes, de quienes he sentido apoyo y afecto todos estos años. Y a mi familia de Barcelona, es un placer teneros en mi vida y os tengo muchísimo cariño.

Gracias a mi padrino Antonio, y a Marina, espero que también estéis orgullosos de mí. Os tengo mucho aprecio.

En tercer lugar, me gustaría expresar mi gratitud a mis amigos y amigas.

A Antonio Felipe, AKA Tonicheeeen. Eres, más que amigo, hermano. Me haces sentir orgullosa por todo lo que logras y superas, y no sólo hablo de estudios. Cada recuerdo contigo no se me olvida. Y espero seguir creando muchísimos más. Tampoco suelo llorar de risa ya, pero cuando nos juntamos llevamos muchas papeletas. Eres MUY importante para mí.

A Josué, aunque nos conocemos de hace "relativamente" poco, ya te siento como otro hermano. Eres paz, estilo, risas, confianza e inteligencia. Eres esa persona que aparece en mi cabeza de las primeras posiciones cuando pasa algo, bueno o malo, porque sé que estarás ahí, sea como fuere, aportando una nueva visión o solución. Gracias por todo.

A Anita Medina, eres una mejor amiga muy coñazo jajaj Es broma. Mil gracias por todo lo que has hecho y haces por mí. Eres una brisa de aire fresco, metafórico y literal, porque contigo es cuando más me siento "desconectar", aunque no salgamos a ningún lado. En los momentos contigo me siento cómoda, me siento muy yo en toda mi esencia. Gracias por confiar en mí todos estos años, y por apoyarme cuando más lo he necesitado. Espero tenerte de amiga toda la vida. Y sí, la revisión de los ojos ya va tocando…pesada.

A "los de siempre", aquéllos que aunque pasen años y años seguimos y seguiremos estando los unos para los otros. Ana Biedma, Coral, Kake y Wendy, gracias por ser mis amigos de toda la vida, y lo que nos queda. Aunque la distancia nos separa, siempre os llevo conmigo. Os quiero.

Christian, AKA christianpwr. El campeonísimo de España. Qué decirte, nos conocemos de hace muy poco pero como buenos bros Marines eres muy importante para mí. No solo por escucharme siempre y estar ahí a diario apoyándome, sino que gran parte de la fase final de este trabajo ha sido gracias a tu apoyo. No solo das el consejo adecuado en el momento adecuado, sino que tu humanidad y principios son referencia. Gracias por aguantarme y por motivarme a creer "mejor" en mí. También gracias por las retransmisiones de nuestros entrenos en tiempo real por chat y por motivarme a luchar por mi hobby.

Gracias Fátima por servirme de guía y apoyo. Aunque nos conocemos de hace pocos años, y hay que sumar la distancia, te considero de mis más cercanas. Te tengo un cariño especial. Transmites paz, sensatez, risas y mucha motivación. Me has apoyado mucho todos estos años y estaré siempre agradecida. Ojalá poder vernos con más frecuencia, estoy segura de que sí.

Y hablando de hobby, gracias a todos y todas los que me apoyáis a desarrollarme en el entrenamiento. Realmente considero que ha sido un granito de aportación más en este trabajo de tesis, puesto que tiene una transferencia directa con mi salud mental y física y eso me ha ayudado a poder avanzar en este proceso. Especial mención a Miguel, por sus sabios consejos tanto de entreno como también de vida, y a Rewel, por nuestras charlas diarias y consejos sobre diversos temas.

Gracias a María del Mar (Granada), Raúl (Granada), Pablo (Alicante), Pablo y Alicia (Úbeda-Málaga), entre otros y otras. No sé si me he dejado a alguien, de ser así espero saber disculparme.

Y por último, gracias a Bubú, Simba y Puma, los individuos no humanos que me nutren el corazón a diario con su amor incondicional y trastadas. Son familia más que mascotas.

Contents

List of Figures

List of Abbreviations

AdEx: Adaptive exponential integrate-and-fire

DE: Differential Evolution (DE)

EA: Evolutionary algorithm

FF: Fitness function

GA: Genetic algorithm

GLIF: Generalized leaky integrate-and-fire

GoC: Golgi cell

GrC: Granule cell

GrL: Granular layer

HH: Hodgkin-Huxley

I-F: Intensity-frequency

MemeGA: Memetic algorithm derived from genetic algorithm

MSASS: Multi-start single-agent stochastic search

PPI: protein-protein interaction

STDP: Spike-timing-dependent plasticity

TLBO: Teaching-learning-based optimization

Prefacio

El cerebelo es un área cerebral crítica para funciones sensomotoras y no motoras como son los procesos cognitivos y emocionales. Las lesiones cerebelares contribuyen a síndromes patológicos como el autismo o la esquizofrenia. Sin embargo, aún se desconocen las primitivas bajo las que el cerebelo, y el cerebro en general, operan tanto a nivel funcional como disfuncional.

Para abordar la complejidad del sistema cerebral "enfermo" es necesario extraer los mecanismos moleculares relevantes que lo subyacen. La disponibilidad de grandes volúmenes de datos biomédicos a veces dificulta la extracción de esta información relevante y su interpretación completa. En esta tesis, hemos llevado a cabo una experimentación preliminar para analizar las correlaciones genéticas entre enfermedades con distintas sintomatologías clínicas y/o prognosis clínicas (y aún basadas en mecanismos moleculares similares). Para este fin, hemos desarrollado una metodología de identificación y anotación funcional de los genes más relevantes de enfermedad. Esta metodología integra métodos actuales de la biología de sistemas, como son las redes de interacción proteína-proteína (PPI), junto con conjuntos de datos multidimensionales de diferentes niveles biológicos. Los objetivos de esta primera parte de la tesis son: la identificación de biomarcadores diagnósticos potenciales (que corresponden a los nodos clave en los procesos biológicos y moleculares del interactoma); el análisis deductivo de datos multidimensionales como alternativa a otros sistemas de búsqueda; y la extracción de conexiones entre enfermedades (comorbilidades) que *a priori* no están relacionadas y que suele escapar a estos sistemas tradicionales. Aunque la metodología es de propósito general, la hemos aplicado a un conjunto de enfermedades denominado canalopatías, donde los canales iónicos se ven alterados y que generan una amplia variabilidad fenotípica. Concluimos que nuestra metodología es flexible, rápida y fácil de aplicar. Además, es capaz de encontrar más correlaciones entre los genes relevantes que otros dos métodos tradicionales.

Para entender la operación cerebelar en el procesamiento de la información se necesita decodificar las dinámicas funcionales intrínsecas de sus neuronas sanas. Las herramientas que nos brinda la neurociencia computacional permiten desarrollar modelos computacionales a gran escala para el estudio de estas primitivas de procesamiento de información. Las neuronas más abundantes y pequeñas no solo en la capa de entrada cerebelar, sino también en el cerebro completo, son las neuronas granulares cerebelares (GrCs). Estas neuronas juegan un papel determinante en la creación de representaciones de información somatosensorial. Sus características de disparo están relacionadas con la sincronización, ritmicidad y aprendizaje en el cerebelo. Una de estas características es la frecuencia de ráfagas de disparo mejorado (esto es, resonancia de disparo). Este patrón de disparo complejo se ha propuesto como clave para

facilitar la transmisión de señal de entrada en la banda de frecuencia theta (4-12Hz). Sin embargo, aún no está claro cuál es el rol funcional de esta característica en la operación de la capa granular (capa de entrada del córtex cerebelar). Además, estas dinámicas complejas inherentes, como es la resonancia, normalmente son ignoradas en la mayoría de modelos computacionales eficientes. El objetivo principal de esta tesis es la creación de diferentes modelos matemáticos de GrCs cerebelares que cumplan con dos requisitos: que sean suficientemente eficientes para poder simular redes neuronales a gran escala, y que sean lo suficientemente plausibles biológicamente para permitir la evaluación del impacto funcional de sus dinámicas no lineales en la transmisión de información. De hecho, un alto grado de realismo biológico en modelos eficientes permite investigar a niveles en los que la biología experimental *in vivo* o *in vitro* está limitada.

Metodológicamente, en esta tesis hemos elegido el modelo de tipo "adaptativo exponencial integrador-y-disparador"(AdEx) como el modelo de neurona simplificado (posee sólo dos ecuaciones diferenciales y pocos parámetros) que reúne tanto realismo como bajo coste computacional. Este modelo se ajusta bastante bien a las características de disparo de células reales, pero algunos de sus parámetros no pueden ajustarse de forma directa con los valores experimentales medibles. Por ello se necesita de un método de optimización para que ajuste mejor los parámetros a los datos biológicos. Nos hemos enfocado en abordar este problema de optimización complejo.

En primer lugar, hemos desarrollado una metodología de optimización paramétrica basada en algoritmos genéticos (GA) aplicado al caso de la GrC. Hemos presentado los modelos de neurona AdEx obtenidos y hemos demostrado su validez para reproducir no solo las propiedades de disparo principales de las GrCs reales (incluyendo la resonancia), sino también características emergentes no definidas en el GA (dentro dela función coste a optimizar).

En segundo lugar, nosotros evaluamos cuatro algoritmos alternativos, que son los más usados y exitosos en otros campos como la ingeniería.

Por último, en la última parte de este tesis hemos presentado una metodología de optimización avanzada basada en algoritmos multimodales. La ventaja de esta estrategia radica en que, tras un único proceso de optimización, en lugar de obtener un único candidato ganador numéricamente al resto de candidatos, como en los casos anteriores (solución única), obtenemos una población dispersa de diferentes modelos de neurona. Esto es, una población heterogénea de neuronas del mismo tipo con variaciones intrínsecas en sus propiedades. De entre este conjunto de modelos neuronales prometedores, el investigador puede elegir y filtrar en base a la plausibilidad biológica deseada (y configuración paramétrica neuronal). Así, hemos también estudiado cómo las propiedades diana de la neurona podrían obtenerse con estas diversas configuraciones de parámetros internos. Nosotros exploramos el espacio de

parámetros y su impacto en el subconjunto de propiedades de neurona que buscamos reproducir.

Estructura de la memoria de la tesis doctoral

Este prefacio termina resumiendo la estructura de esta tesis para orientar al lector:

- El primer capítulo describe el contexto en el que este trabajo ha sido realizado desde el principio del proyecto hasta su finalización.
- El segundo capítulo incluye la propuesta de investigación que dirigió esta tesis, describiendo la motivación, objetivos iniciales y explicando brevemente las contribuciones de este trabajo y el marco de proyecto en el que esta tesis está definida.
- El tercer capítulo contiene las publicaciones que conforman el núcleo de la investigación llevada a cabo. Cada uno de estos trabajos aborda una de las principales contribuciones indicada resumidamente en el capítulo anterior.
- El cuarto capítulo analiza los objetivos de la tesis y destaca las principales contribuciones de acuerdo a estos objetivos. Asimismo esboza el trabajo futuro a abordar y recoge las conclusiones que se derivan de esta tesis.

Preface

The cerebellum is a critical brain area for sensorimotor and also non-motor functions such as cognitive and emotional processes. Cerebellar lesions contribute to pathological syndromes such as autism or schizophrenia. However, the primitives under which the cerebellum, and the whole brain, operate at a functional and dysfunctional level are still unclear.

To address the complexity of the "diseased" brain system, it is necessary to extract the relevant underlying molecular mechanisms. The availability of large volumes of biomedical data usually makes it difficult to extract this relevant information and interpret it comprehensively. In this thesis, we have made a preliminary experimentation to analyze genetic correlations between diseases with different clinical symptomatologies and/or clinical prognosis (and still based on similar molecular mechanisms). For this purpose, we have developed a methodology for the identification and functional annotation of the most relevant genes in disease. This methodology integrates current systems biology methods, such as protein-protein interaction (PPI) networks, together with multidimensional data sets from different biological levels. The objectives of this first part of the thesis are: the identification of potential diagnostic biomarkers (corresponding to key nodes in the biological and molecular processes of the interactome); the deductive analysis of multidimensional data as an alternative to other search systems; and the extraction of connections between disorders (comorbidities) that are *a priori* unrelated and that usually escape these traditional systems. Although the methodology is of general purpose, we have applied it to a set of diseases called channelopathies, where ion channels are altered and which generate a wide phenotypic variability. We conclude that our methodology is flexible, fast and easy to apply. Furthermore, it is able to find more correlations between relevant genes than other two traditional methods.

Understanding the cerebellar operation in information processing requires decoding the intrinsic functional dynamics of healthy neurons. The tools provided by computational neuroscience allow developing large-scale computational models for the study of these information processing primitives. The most abundant and smallest neurons not only in the cerebellar input layer, but also in the whole brain, are the cerebellar granule cells (GrCs). These neurons play a crucial role in the creation of somatosensory information representations. Their firing characteristics are related to synchronization, rhythmicity and learning in the cerebellum. One of these features is the frequency of enhanced bursting (i.e., spiking resonance). This complex firing pattern has been proposed to facilitate input signal transmission in the theta-frequency band (4-12Hz). However, the functional role of this feature in the operation of the granular layer (the input layer of the cerebellar cortex) is still unclear. Moreover, inherent

complex dynamics such as resonance are usually ignored in most efficient computational models. The main goal of this thesis is the creation of different mathematical models of cerebellar GrCs that meet two requirements: to be efficient enough to allow the simulation of large-scale neuron networks, and to be biologically plausible enough to enable the evaluation of the functional impact of these nonlinear dynamics on the information transmission. Indeed, a high degree of biological realism in efficient models allows research at levels where *in vivo* or *in vitro* experimental biology is limited.

Methodologically, in this thesis we have chosen the "adaptive exponential integrate-and-fire" (AdEx) type of model as the simplified neuron model (it has only two differential equations and few parameters) that meets both realism and low computational cost. This model fits quite well the firing characteristics of real cells, but some of its parameters cannot be directly fitted with measurable experimental values. Therefore, an optimization method is necessary to best fit the parameters to the biological data. We have focused on addressing this challenging optimization problem.

First, we have developed a parametric optimization methodology based on genetic algorithms (GA) applied to the case of GrC. We have presented the obtained AdEx neuron models and demonstrated their suitability to reproduce not only the main firing properties of real GrCs (including resonance), but also emergent features not defined in the GA (within the cost function to be optimized).

Second, we evaluated four alternative algorithms, which are the most widely used and successful in other fields such as engineering.

Finally, in the last part of this thesis we have presented an advanced optimization methodology based on multimodal algorithms. The advantage of this approach is that, after a single optimization process, instead of obtaining an only one candidate numerically outperforming the other candidates, as in the previous cases (a single solution), we obtain a sparse population of different neuron models. That is, a heterogeneous population of neurons of the same type with intrinsic variations in their properties. From this set of promising neuron models, the researcher can choose and filter based on the desired biological plausibility (and neuronal parameter configuration). Thus, we also studied how the target properties of the neuron could be obtained with diverse internal parameter configurations. We explored the parameter space and its impact on the subset of neuronal properties that we aim to reproduce.

Structure of the Ph.D. thesis manuscript

This preface ends by summarizing the structure of this thesis to orientate the reader:

- The first chapter describes the context in which the work has been performed from the beginning of the project to its finalization.
- The second chapter includes the research proposal that led this thesis, describing the motivation and initial objectives and briefly explaining the contributions of this work and the project framework on which this thesis is defined.
- The third chapter contains the publications that form the core of the research carried out. Each of these works addresses one of the main contributions briefly indicated in the previous chapter.
- The fourth chapter analyses the thesis objectives and highlights the main contributions according to these objectives. It also outlines future work to be addressed and draws the conclusions that result from the thesis.

Chapter 1

Introduction

A neurological disease affects more than one-third of European citizens along their lifetime. Understanding the biological principles of brain and associated disorders have a high impact on the quality of life of a significant part of the population. It is associated to an overall cost of approximately 800 billion euros per year at the European level (estimated data for 2010). All this has motivated a strong research effort on the brain. One of the ultimate challenges for the current century is the decoding of the biological basis of consciousness and the mental processes by which we perceive, act, learn, and remember. Although this long-term challenge remains active, deep knowledge of the biological essence of the brain can lead us to several advances. Firstly, understanding the brain functions as a complex neural system can favor a deepening of its dysfunction in disease. This accumulative knowledge allows identifying the biological keys of neuropathophysiology for the design of innovative treatments and potential therapeutic interventions for disease. Secondly, one of the best and more complex dynamical systems of information processing is the brain, with low power consumption and reliability against gradual degradation of the system (with age or some neuronal diseases). An example of this situation would be the low-frequency oscillatory rhythms of certain brain areas that are fundamental for the inherent processes of learning, motor control or sleep (Buzsáki, 2006; Wang et al., 2019). Finally, emulating the way of representing somatosensory information by specific neuron types could develop new generations of processing architectures able to

replicate this massively-parallel computing performance. In this context, this thesis follows two of the aforementioned main open questions by the scientific community:

(1) Shedding light on the comprehension of relevant molecular mechanisms underlying related complex diseases, such as the extensive group of disorders that affect the ion channels.
(2) Understanding the brain operation with respect to the impact that potential intrinsic dynamics of neurons could take part in the information processing primitives of a specific brain area, the cerebellum

Computational approaches allow discovering the impact of neuronal behavior in the brain performance. However, other methods based on systems biology also allow addressing the complexity of the biological system of the brain (as detailed in the next subsection "**1.1. A preliminary study towards understanding the pathomechanisms of neurological disorders**". Molecular data-driven networks become powerful analytical tools in the identification of the most relevant components directly tied to biological functions. These structural aspects of gene relationships help to better understand the etiopathogenesis of disease. It should be emphasized that the complexity of disease comprises a combination of different symptoms, evolution, and it is the result of the effects of many genes. In this scenario, the identification of the most relevant molecular foundations of disease from integrating massive biomedical information available might become an arduous and complex task. Especially if the set of disorders to be analyzed has hundreds of genes involved. As mentioned above and in order to cope with the huge amount of current knowledge, it would be of utmost usefulness to explore valid association procedures beyond the current traditional search sources (typically based on exhaustive or systematic search systems). This thesis finds essential to develop a methodology for non-experts in systems biology approaches based on biologically data-driven networks (i.e., protein-protein interaction networks) as a preliminary experimentation towards understanding altered cellular dynamics of neurological disease.

The cerebellum (Latin for "little brain") is a structure located at the back of the brain, underlying the occipital and temporal lobes of the cerebral cortex. The cerebellum occupies 10% of the brain's volume and it contains approximately 50% of the overall amount of neurons in the brain (Welsch and Welsch, 2006). Historically, it has been known well for its role in motor coordination and cognitive functions, as was noted by Ramon y Cajal and Camillo Golgi in their histological analyses of brain tissues (Ramon Cajal, 1894; Golgi, 1906; De Carlos and Borrell, 2007). However, the cerebellum is gaining increasing attention for its involvement in other motor and non-motor skills of cognitive processing, such as language, and emotion (Schmahmann, 2019; Guevara et al., 2021). The cerebellum is considered as a time machine and it has expanded this concept to clinical disorders (Eccles, 1973; Bareš et al., 2019). Serious damage in this organ impairs its learning and results in motor disturbances. Recent work has evidenced its implication in motor disorders (e.g., ataxia, dystonia, epilepsy, Huntington,

Parkinson's disease, or Alzheimer's disease) and non-motor conditions (e.g., autism spectrum disorders, alcohol fetal syndrome, medulloblastoma, schizophrenia, visceral functions, or sleep apnea) (**figure 1**) (Reeber et al., 2013; Computational Models of Brain and Behavior, 2017). However, in spite of the increasing evidence of the operational primitives inside the cerebellum and their involvements in disease, its biological mechanisms and their possible dysfunctions are not yet being sufficiently explored.

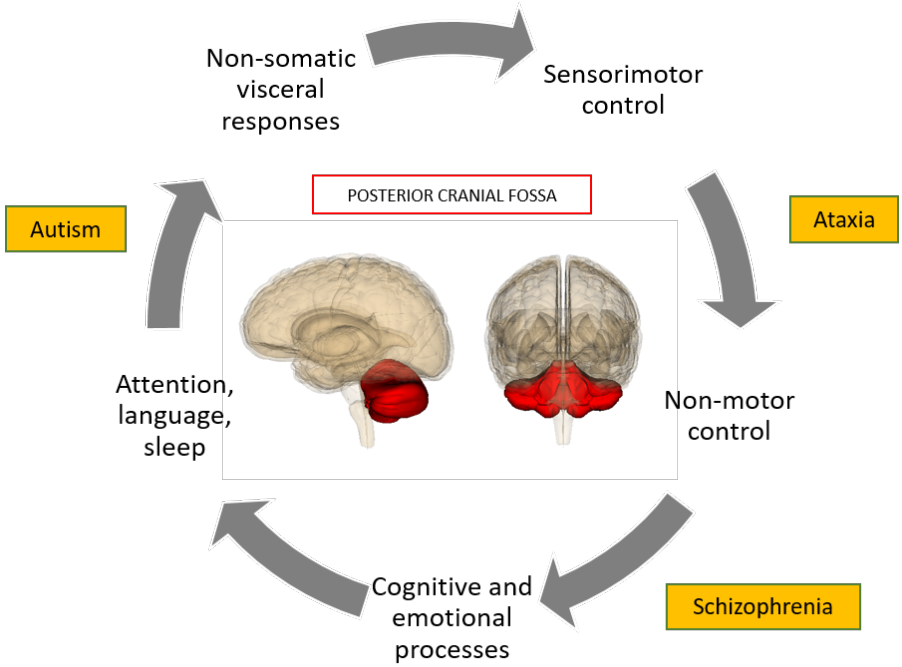


Figure 1. Representation of some of the cerebellar roles in motor and non-motor disturbances. The cerebellum is a key brain structure for the sensorimotor control and non-motor functions including cognitive and emotional processes such as attention, language, emotional behavior, sleep and even non-somatic visceral responses. Cerebellar lesions contribute to pathological syndromes such as autism, schizophrenia and ataxia. Images of the cerebellum are from Anatomography maintained by Life Science Databases (LSDB) (URL: https://commons.wikimedia.org/wiki/File:Cerebellum.png).

Regarding the understanding of the cerebellar operation, there is a significant number of studies on the cerebellar roles in motor learning and neuronal plasticity (D'Angelo et al., 2009). Current studies suggest that the cerebellum predict optimal movements by recognizing neuron patterns (Reeber et al., 2013). In addition, populations of neurons in several brain areas reflect complex synchronized temporal patterns typically modulated by coherent oscillations (Buzsáki, 2006). These oscillatory patterns are usually evidenced by the study of a complex neuronal behavior named resonance, i.e. a preferred neuronal frequency in response to oscillatory inputs (Hutcheon and Yarom, 2000). The cerebellum is one of the brain areas where the neuronal feature of resonance has received more attention (D'Angelo et al., 2009, 2011; Dugué et al., 2009; Gandolfi et al., 2013). The cerebellum is thought to generate low frequency

(5-30Hz) and higher-frequency activity rhythms, but its function for overall cerebellar information processing remains elusive. Nevertheless, new pieces of evidence suggested that somatosensory information representations might be determined by the cerebellar granule cells. The cerebellar granule cells are the most numerous cells in the entire mammalian brain (Herculano-Houzel, 2010). An intrinsic mechanism of resonance in these cells was evidenced at a concrete low-frequency band (i.e., the theta-frequency band, around 4-10Hz in rodents) and was proposed as facilitating the input signal transmission (which will be expanded in the later subsection "**1.2. The information processing in the granular layer of the cerebellum**"). However, the functional role of this cell feature in the operation of the cerebellar input layer remains largely unclear. The focus of this thesis concerns on how to facilitate the study of this fundamental research question by efficiently integrating this property (intrinsic resonance of burst frequencies) on efficient neuron models.

It is worthy to introduce the current *in silico* tools that are widely used in neuroscience (and more specifically in computational neuroscience). The field of computational neuroscience emerged in the mid-20th century with the purpose of studying the underlying mechanisms of brain function, behavior, and disease. This field strongly began to expand in the past few decades at the same time as the increasing capabilities of computational resources and computers. From this perspective, computational neuroscience developed into a promising discipline within system neuroscience through theoretical and computational modeling in the research of complex brain dynamics. That is, computational neuroscience challenges to link the structure, dynamics, and response properties of neural systems. However, this challenge is complicated since: (1) the human brain contains approximately 86 billion neurons each with an average of 7000 synapses per neuron, representing a complex biological system to untangle, and (2) the amount of data in experimental neuroscience, and overall in science, has enlarged. That implies, on the one hand, that the current computer power is limited in order to model the whole brain at this level of interconnectedness. On the other hand, a comprehensive interpretation of this current biological knowledge is needed in order to cope with these levels of biological complexity.

In this sense, a simpler approach has been adopted to face the biological complexity of the brain: the integration of different data sets into computer simulations as close approximations to biological data. Computational neuroscience provides a coherent framework involving many diverse methodological approaches, being all interconnected. As represented in **figure 2**, simulations take place at various separate organizational scales of experimental neuroscience, ranging from the molecular level, through the subcellular and to cellular levels, and up to the whole organ. With a constrained computing power, the level of detail decreases as the level rises towards the whole brain, i.e., from neuron models of the neurobiological substrate, network models of microcircuits, to high-level computation principles and models of behavior or even models of disease (D'Angelo et al., 2013a). In more detail, in the cellular level and below, neurophysiologists record the neuronal activity from

which they evidence possible neuronal characteristics that describe the cell dynamics. These characteristic response patterns are the result of a combination of membrane ion channels and specific interneuron synapses that are embedded in the neurons and microcircuits. These experimental findings are the starting point for the simulation of these neuronal dynamics through detailed biophysical models of neurons. Simplifying these complex neuron models but identifying and keeping functional characteristics is a sophisticated process on which this research has been focused (it is detailed in subsection "**1.3. Computational modeling of neurons**"). The simulations are multi-scale, ranging from a single-neuron with certain channels or specific synapses, to complex neuronal networks of a bunch of neurons. *In silico* simulators such as NEURON (Hines and Carnevale, 1997) or NEST (Peyser et al., 2017) allow dealing with neuron- and network- models at different levels of abstraction (molecular level, multi-compartmental level, point-neuron level, etc.). At this point, these simulations allow evaluating the behavioral impact of the neurobiological substrate, e.g. specific neuron characteristics, network topology or adaptation mechanisms. In consequence, studying the neuron behavior in a multiscale framework allows discovering functional characteristics of neurons behind basic brain functions (such as awareness, memory and stress) as well as addressing diagnosis and treatment of neurological disorders. Therefore, from detailed neuron models which involve small-scale simulations, to simplified neuron models that allow large-scale simulations, and from these neuron networks to even embodied simulations of behavior (behavioral experimental set ups in which neuronal simulations are done in the framework of an experimental task).

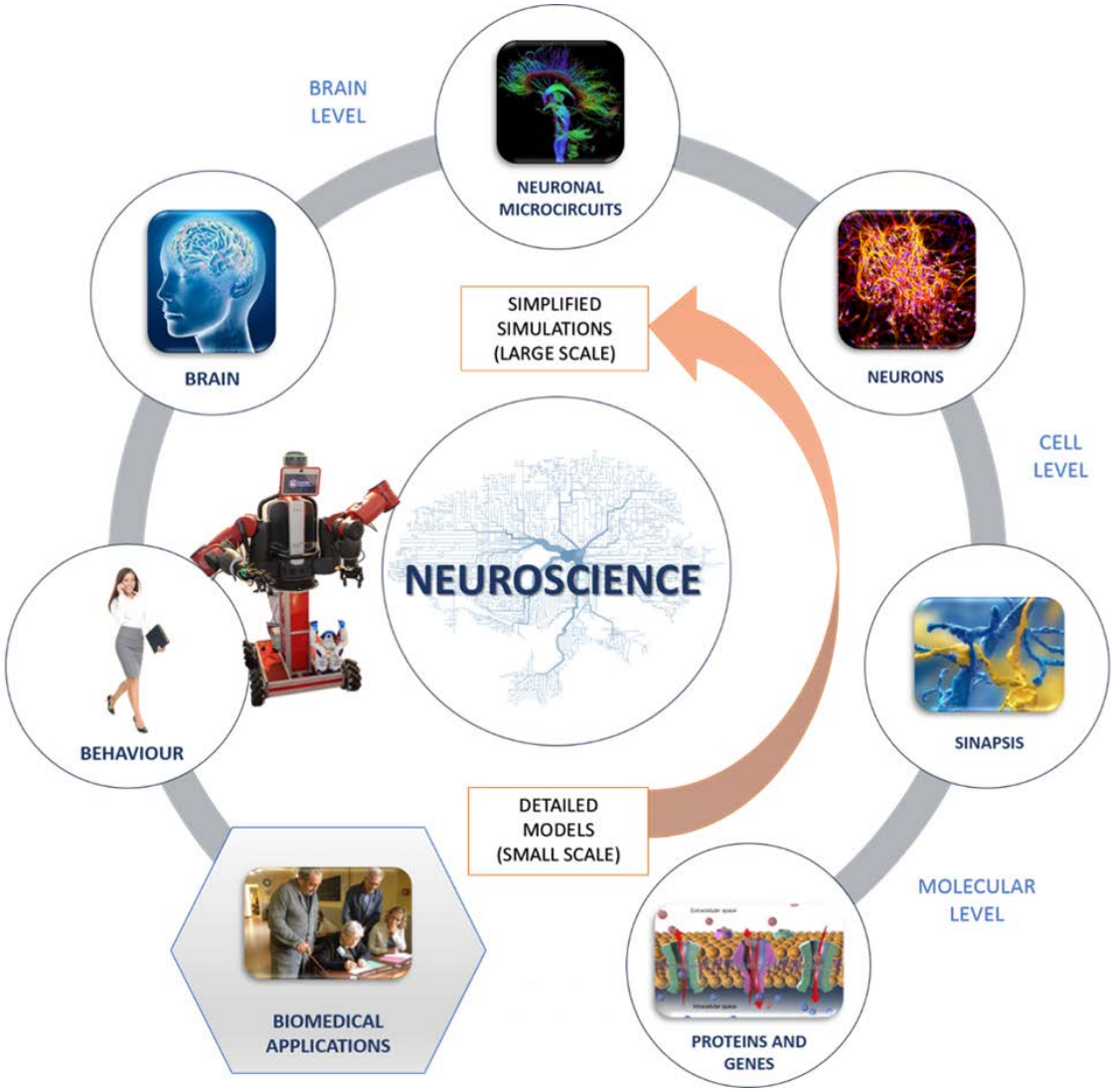


Figure 2. Conceptual framework for understanding the computational principles of the brain. At the molecular level, neurophysiology allows discovering characteristics that describe neuron dynamics, embedding a combination of channels and specific interneurons synapses in the neuron layer. Simulations are multiscale, from a single neuron with certain channels or specific synapses, to complex neuron networks of millions of neurons. This approach allows creating complex neuron models that can be simulated in detail. Simplifying these detailed models but identifying and keeping the most relevant spiking characteristics is a complex process. *In silico* simulators such as NEURON or NEST allow dealing with neuron models at different abstraction levels (molecular level, multi-compartmental models, point-neuron level, etc.). A higher abstraction level allows evaluating the behavioral impact of the neurological substrate (cell characteristics, network topologies and adaptation mechanisms). Finally, the study of the neuronal behavior in a multiscale framework allows discovering cell functional characteristics behind awareness, memory, stress, etc., and addressing diagnosis and treatment of neurobiological diseases.

This thesis is contextualized in the molecular and cellular scales, deepening in the study of the complex feature of resonance of cerebellar granule cells that could have a functional role in the information processing primitives. For this aim, part of the methodology addressed

in this research work is encompassed in the creation and simulation of mathematical models of these types of cells. Since the cerebellar granule cells show compact and simple morphology (D'Angelo et al., 2001; Delvendahl et al., 2015), it is appropriate to consider a mono-compartment model. In this thesis, we have deepened in the exploration of methodological approaches regarding the creation of neuron models that maintain biological realism and computational efficiency whilst capturing essential aspects of single-neuron processing. We have gone into detail about the development of refined methods of automatic parameter tuning of these neuron model configurations replicating from the neurophysiological data evidenced in the literature.

Hence, this thesis has interest in diverse research fields and topics of interest of international level by projects as the Human Brain Project (HBP in Europe, and where the Ph.D. candidate has been involved since 2019) or the Brain Research through Innovative Neurotechnologies (BRAIN Initiative in USA). These projects are among the largest initiatives that have caused greater expectation in recent years, and where huge economic resources to this research are dedicated. The HBP initiative is focused on an integrated effort and the development of various general platforms that facilitate research in different fields related to the human brain. The main goal of these platforms are based on giving scientists from all over the world a single point of access to neuroscience data, multi-omics clinical data and analysis tools. The American BRAIN Initiative is more focused on the development of new technologies that will enable disruptive advances in the research fields related to the brain. For further details on the early research and current frameworks, the reader is referred to the review by (Prieto et al., 2016). The interdisciplinary thematic, needed for both understanding the brain and treating neurological diseases, ranges from the alteration of the gene and its effect in the neuron behavior, in the network, through thousands of synapses, and even reproducing symptomatology and rehabilitation with an agent (that can be a robot avatar). Thus, the interdisciplinary nature of the field of this thesis ranges throughout various levels of biological organization (genetics, proteomics, micro- and macro-circuits and even behavioral and clinical level), integrating multidimensional data according to diverse up-to-date methodological approaches from different fields (such as neuroscience, biomedicine, physiology, molecular and systems biology, together with electrical engineering, computer science, mathematics and biophysics).

The general goal of this thesis is intended to provide some pieces of knowledge (contributions) to some of the questions that remain elusive about the brain and more specifically the cerebellum:

- ✓ Firstly, as a preliminary approach we look for a better understanding of the etiopathogenesis of neurological diseases. We propose to identify the most common polygenic influences and their associated manifestations (and still based on similar molecular mechanisms) throughout a set of diseases contributing to the overall

understanding of their pathomechanisms. This goal is intended to shed some light on the up-to-date knowledge of neurological disorders and related manifestations. Furthermore, we proposed a generic workflow that can be applied to other disorders.

- ✓ Secondly, we aim to develop a methodology for the simulation and simplification of mathematical neuron models. The intention is to be able to create computationally efficient neuron models that reproduce non-linear cellular dynamics. It is worthy to mention that we expect to maintain biological realism and computational efficiency in the neuron models whilst capturing essential aspects of single-neuron processing. We use the mentioned methodology to perform the study of a potential cellular mechanism in the somatosensory information representations, i.e. the intrinsic resonance of the cerebellar granule cells. The methodology to generate different computationally efficient models of granule cells with intrinsic resonance and other features considered functionally relevant is our contribution to this aim. Future simulations of microcircuits with these neuron models will allow evaluating how these complex neuronal dynamics support neuronal information processing at a system level.

1.1 A preliminary study towards understanding the pathomechanisms of neurological disorders

Many complex diseases, such as neurological disorders, are known to be an interplay of multifactorial conditions that are greatly influenced not only by genetic variants but also by environmental signals. Therefore, it is of utmost importance to identify their causal agents for the detection of suitable targets, the management of their diagnosis and the selection of the most adequate therapies.

Channelopathies are a group of genetically and phenotypically heterogeneous disorders that result from genetically determined defects in ion-channel function. Ion channels are transmembrane proteins that allow the passive flow of ions, both in and out of cells or cellular organelles, following their electrochemical gradients. This flux of ions across the membrane results in electrical currents, for which ion channels play an important role in generating membrane potential and function in diverse cellular activities. Ion channels are fundamental in neuronal signaling and thus, channelopathies are mainly considered as a group of neurologic disorders (e.g., generalized epilepsy with febrile seizures plus, episodic ataxia and familial hemiplegic migraine). However, its heterogeneity is due to mutations in the same gene that can cause different diseases and mutations in different genes can result in the same disease phenotype. Therefore, other manifestations of channelopathies also affect other systems such as the cardiovascular, the respiratory, the endocrine, the immune and the urinary systems (Knupp and Brooks-Kayal, 2017). The field of channelopathies is expanding rapidly, as

is the utility of molecular-genetic and electrophysiological studies. There are remarkable causal heterogeneity (especially genetic) and phenotypic variability in channelopathies, which make the diseases challenging to classify and thus to be effectively treated.

The increasing availability of large bibliographic data volumes lays the foundations for the identification of candidate genes as causal agents of a complex disease. However, the integration of all this knowledge requires understanding the diverse biomedical information sources available. The extraction of data performed by valid association procedures and the comprehensive interpretation of all this current knowledge is complex.

Traditionally, brain complexity has been addressed following a reductionist perspective, studying the anatomical regions of the brain and focusing on characterizing its cellular components and basic functions in isolation. Even though research endeavors are still focused on finding gene variants related with consciousness, memory and social capabilities, neuroscientists admit that the information processing skills and, in the last instance, brain behavior, are caused by the dynamic interaction of intricate synapse networks. However, this traditional perspective has not helped fully to clarify the mechanisms of interaction between components and has been unable to predict the effects of alterations in components upon the dynamics of the entire system (Diaz-Beltran et al., 2013).

Alternatively, the interdisciplinary research is developing new technologies and integrative computational methodologies in order to better understand pathogenesis. Systems biology emerged as a field of study aiming to understand biology at the system level, entailing the functional analysis of the structure and dynamics of cells and organisms and focusing on the interactions between isolated components. Most researchers agree that systems biology complements the classic reductionist approaches in biomedical research and represents one of the best strategies to understand the underlying complexity of biological systems (Diaz-Beltran et al., 2013). These new tools are able to manage deductive analyses by gaining insight into the connections among diseases, even between those *a priori* not related by the traditional bibliographic searches (such as exhaustive or systematic search systems), which usually tend to be subjective, time-consuming or not reproducible. However, the large range of diverse new tools created within different focuses hinders the existence of a unique approach to or a consensus on their usage. Thus, data extraction through ad-hoc approaches using specific tools may again be complex, not reproducible or subjective and would not allow discovering interconnections among genes involved in complex disease that would have utmost relevance and utility. In this way, network analyses and functional annotation tools from systems biology represent some of the best strategies for objective interpretation of biomedical data and cope with higher levels of biological complexity.

In this context, this thesis explores as a preliminary research a semi-automatic workflow based on the integration of biologically data-driven networks and databases for the

identification and functional annotation of the most relevant genes in disease. We planned to inspect the integration of different multi-dimensional data from various biological levels (genomics, transcriptomics and proteomics) in order to analyze genetic correlations among complex diseases with heterogeneous symptomatologies and/or clinical prognoses (and still based on similar molecular mechanisms). In order to illustrate the value of this integrative approach and demonstrate its usefulness, we select the case of channelopathies as proof-of-concept since this complex group of diseases is highly related to neurological pathologies and remains unexplored.

1.2 The information processing in the granular layer of the cerebellum

In order to better understand the role of the cerebellum as one of the key subsystems in the brain for the processing of temporal information, it is important to know its anatomy and structure. The cerebellum is organized in three main large-scale networks that involve a relatively modest number of cell types with highly-parallel and modular connectivity: the cerebellar cortex, deep cerebellar nuclei, and the inferior olive. This structure has inspired many theoretical models elaborating the combinatorial properties of this whole network (Eccles et al., 1967; Albus, 1971) and also cerebellar models of learning, movement prediction, and a timing machine (Eccles, 1973; Computational Models of Brain and Behavior, 2017). Cerebellar models help identifying basic research questions required to understand how the brain represents information and utilizes time (Bareš et al., 2019). The cerebellum is thought to have representations of internal models, transferring relevant information via their inputs and outputs (D'Angelo et al., 2013b). Its neuronal properties (at the intrinsic and synaptic levels) correlate with learning and memory processes, among others. In the last decade, our research group, the Applied Computational Neuroscience (ACN) Group (Applied Computational Neuroscience Research Group) at the University of Granada, has developed and applied computational tools to gain understanding on cerebellar information processing and motor control. The work developed in this thesis within this lab follows the research line in Computational Neuroscience about simulating biologically plausible neural systems at different levels of detail to contrast working hypotheses related to the cerebellar cortex and evaluate how the physiological substrate supports their computational features.

Introducing the cerebellar cortex, it has a large amount of neurons (estimated to up to 20^{10} neurons), with an approximately 1 mm-thick and a similar structure across the cerebellum (Computational Models of Brain and Behavior, 2017). Its neuronal circuits include seven types of neurons (namely, granule cells, Purkinje cells, basket cells, stellate cells, Golgi cells, Lugaro cells and unipolar brush cells). They also involve three types of afferents (i.e., mossy fiber

afferents, climbing fiber afferents, and the third type of afferent named beaded fibers through the white matter underlying the granular layer) (Rancz et al., 2007; Ito, 2014). The afferents lead the input of nerve impulses (spikes) to the cerebellar cortex (i.e., from which the cerebellum receives inputs), and the Purkinje cells provide the sole output of this layer (Rancz et al., 2007; Ito, 2014). From a histological point of view, the cerebellar cortex is divided into three layers (from outside to inside): the molecular layer, the Purkinje cell layer, and the granular layer (**figure 3**) (Welsch and Welsch, 2006).

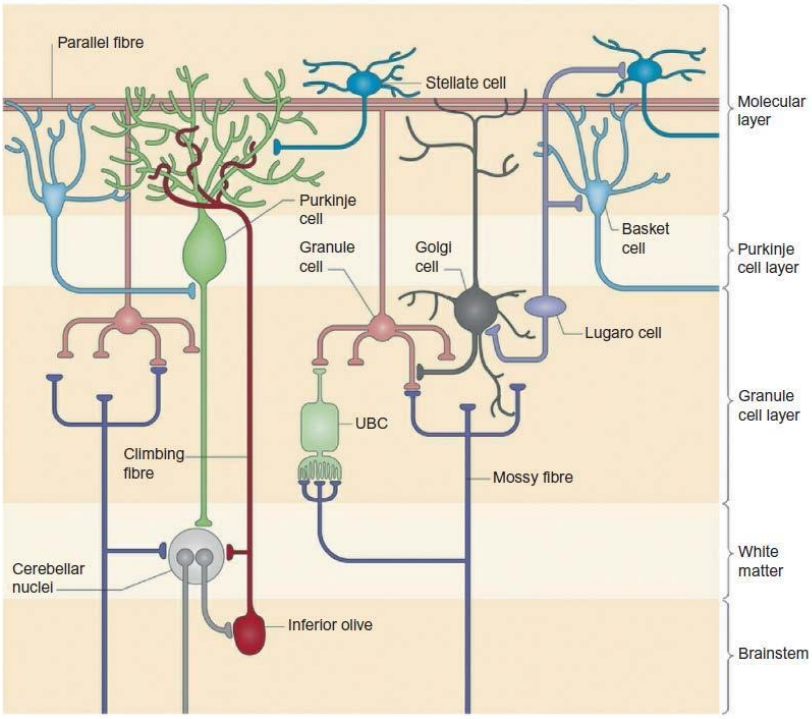


Figure 3. Schematic representation of the cerebellar cytoarchitecture. The cerebellar cortex is composed of three layers: the granular layer, the Purkinje cell layer and the molecular layer. The cellular types present in the cerebellar cortex are the Purkinje cells, granule cells, Golgi cells, Lugaro cells, unipolar brush cells (UBCs), stellate cells and basket cell interneurons. Climbing fibers, which synapse directly with Purkinje cells and a subset of neurons in the cerebellar nuclei, and the mossy fibers, which synapse with granule cells are the two main afferents to the cerebellum. Figure extracted from (Ashida et al., 2018).

Regarding the granular layer, it is the innermost layer and contains a high number of the smallest (5-8 μm in diameter) but the most numerous neurons (2-3 million/mm^{-3}, for a total of 10^{10}-10^{11} neurons in humans) in the brain, named granule cells (Ito, 2014). Each granule cell extends four to five short dendrites, each of which receives an excitatory synapse from a mossy fiber terminal and relays the input to Purkinje cells via their unmyelinated parallel fibers. Granule cell axons supply excitatory synapses to other neurons in the cerebellar cortex. The cerebellar granule cells (GrCs), Golgi cells (GoCs), Lugaro cells and unipolar brush cells (UBCs) are the interneurons present in this layer. Particularly, GoCs are the main inhibitory interneurons in the feedforward inhibition that contributes to the granular layer computations

of sensory and tactile inputs (Roggeri et al., 2008). Their specific ionic mechanisms (such as pacemaking activity and theta-frequency resonance) (Solinas et al., 2007) may allow the adaptation of firing frequency and spike precision (Solinas et al., 2007; Medini et al., 2012). In this line, GrCs also show a complex intrinsic mechanism of resonance in the theta range that may play a role in the phase-locking of the inhibitory circuit (Maex and De Schutter, 1998; D'Angelo et al., 2001). Theta-frequency resonance of GrCs also might be involved in regulating the induction of synaptic plasticity in the mossy fiber-granule cell pathway (Armano et al., 2000; D'Angelo et al., 2001). Given these pieces of evidences, the importance of studying this complex dynamics is of utmost importance in this thesis.

With respect to the neuronal feature of resonance, it is worthy to introduce some theoretical concepts. As mentioned before, resonance is a measurable property that describes the ability of neurons to respond selectively to oscillatory inputs at preferred frequencies (**figure 4**) (Hutcheon and Yarom, 2000). The characterization of resonance captures those essential properties of neurons that can serve as a substrate for coordinating network activity around a particular frequency in the brain. The mechanism of resonance assesses small-signal responses of neurons (thus ignoring or approximating their strongly nonlinear properties) in order to understand how neurons process oscillatory inputs at to neuron regimes: subthreshold potentials, and/or suprathreshold (spiking) potentials (Hutcheon and Yarom, 2000; Rotstein, 2017). As mentioned above, recent studies on the granular layer have added novel implications of its signals encoded for its function (D'Angelo et al., 2009). Specifically, most of the findings in this layer have focused on the subthreshold resonance, particularly *in vivo* studies evidenced theta-frequency resonance at 7Hz in rats (Hartmann and Bower, 1998) and 7-25Hz in monkeys (Pellerin and Lamarre, 1997; Courtemanche et al., 2009). However, the spiking resonance has been considered with less attention in spite of it being proposed to strengthen the input signal processing and transmission at the theta-frequency band in the granular layer (D'Angelo et al., 2009). More specifically, the spiking resonance has been claimed to be an intrinsic property of the cerebellar GrCs (D'Angelo et al., 2001).

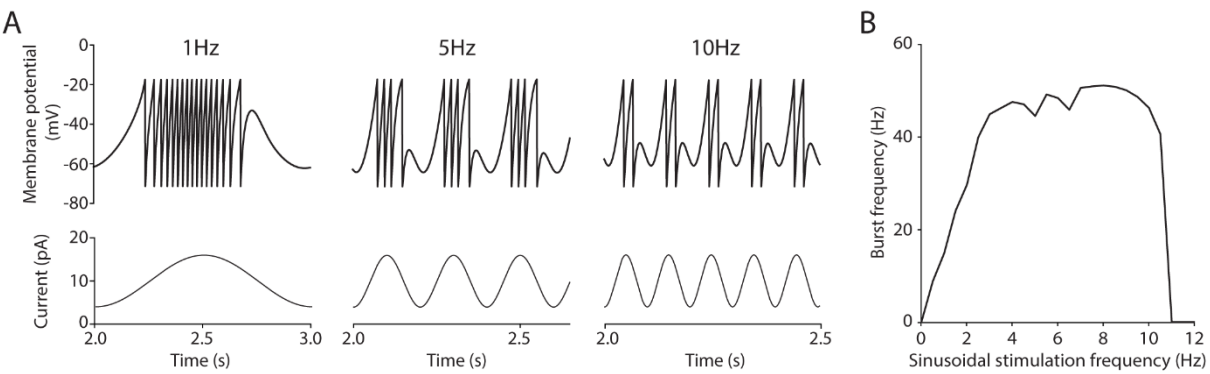


Figure 4. Intrinsic property of spiking resonance. (A) Model simulation in response to sinusoidal current injection of 10-pA offset and 6-pA amplitude. **(B)** Resonance curve (showing burst frequency) in response to the same stimulation protocol. These simulations are predicted by the cerebellar granule

cell (GrC) model obtained in (Marín et al., 2020) included in this thesis. This figure is extracted from figure 6 from (Marín et al., 2020).

Another general concept worthy to mention is the set of biochemical mechanisms that underlie the neuronal feature of resonance. This cellular dynamics emerge as a combination of low-pass and high-pass filter mechanisms, i.e., the interplay of the passive membrane properties and one or more ionic currents and their interaction with the oscillatory inputs (Hutcheon and Yarom, 2000; Magistretti et al., 2006; Das and Narayanan, 2017; Fox et al., 2017; Rotstein, 2017). A high-pass filter, as a "resonator" current, and a low-pass filter, caused by passive membrane properties, determine the resonance curve. In the case of the cerebellar GrCs, the pieces of evidence suggest that the core of the slow oscillatory mechanism are $I_{K\text{-}slow}$ and $_{INa\text{-}p}$. The high-pass filter is determined by $I_{K\text{-}slow}$ (i.e., a Ca^{2+}-independent TEA-insensitive K^+ current), which causes delayed repolarization terminating the positive phase of the oscillation promoted by $I_{Na\text{-}p}$. These interplay of filters is amplified by an "amplifier" current $I_{Na\text{-}p}$, which sustains theta-frequency oscillation and bursting. A resurgent current $I_{Na\text{-}r}$ facilitates bursting and resonance, intensifying the interplay of mechanisms by spike clustering, but not sufficient to induce them. The spiking resonance of GrCs selectively enhance low-frequency stimulation responses according to some studies (D'Angelo et al., 2001, 2009; Gandolfi et al., 2013).

This behavior of cerebellar GrCs is due to a high-pass filtering "resonator" current, generating the ascending branch of the resonance curve) in association with a low-pass filtering caused by passive membrane properties (generating the descending branch of the resonance curve). These mechanisms are amplified by an "amplifier" current ($I_{Na\text{-}p}$) and intensified by spike clustering (promoted by $I_{Na\text{-}r}$).

The cerebellar granule cells (GrCs) are the main components of the granular layer of the cerebellum since they are the smallest and most abundant in the entire brain (Herculano-Houzel, 2010). Following the above statements, it is suggested that GrCs determine the somatosensory information representation in the granular layer facilitating the input signal transmission. Nevertheless, the function of the spiking resonance in the theta band of GrCs for overall information processing remains elusive. For addressing this research question on the possible role of the oscillatory activity in the theta-frequency band of GrCs in the information processing capabilities of the granular layer, this thesis explores how to best fit this property in computationally efficient neuron models. Possibly, this cellular feature structures the information to be co-processed eventually by other components of the layer, such as the GoCs.

By means of computational neuroscience approaches, the aim of this thesis is to evaluate the computational capabilities of the spiking resonance in the granular layer, integrating non-linear cellular dynamics, such as the feature of spiking resonance, in efficient neuron models. Our contribution addresses the generation of efficient neuron models that reproduce accurately the essential properties of the cerebellar granule cells together with the

spiking resonant behavior in the theta range. This research leads to simulate large-scale networks of the granular layer for studying the functional impact of resonance since it is suggested to have a main role in the integration of signals in the cerebellar circuitry (Gandolfi et al., 2013).

1.3 Computational modeling of neurons

The technological advances in computational neuroscience have helped foster new insights into the processing of temporal information in the cerebellum. Regarding the understanding of neuronal information processing, theoretical approaches are yet be further developed. However, the theoretical principles are substantiated in: (1) experiment-dominated bottom-up (from biological data to computational models to functional concepts and structure-function relationships); and (2) theory-dominated top-down approaches (from functional working hypotheses to computational models to testable predictions). This thesis is based on this latter principle, where the functional working hypothesis is based on the study of unexplored nonlinear cellular dynamics in the information processing of granular layer and the methodology is based on the construction of computational models that allows evaluating this premise in large-scale networks.

Simulation approaches from computational neuroscience have proven an invaluable tool for turning these mathematical principles, theory, and data into new insights. For building a computational model based on biological data there are two different ways: (1) creating realistic neuron models based upon realistic information and with the lowest simplification possible, with the purpose of finding new structure-function relationships, or (2) creating very simplified models that are computationally efficient to predict and demonstrate a preconceived concept. It is important to mention that biophysical model predictions can simulate results that are not possible, or even replicate experiments on animal models, such as the mouse (Human Brain Project (HBP) - Brain Simulation; Computational Models of Brain and Behavior, 2017). Finally, it is also worth mentioning that computational neuroscience has become an auspicious technology for translational research, developing novel paradigms in both intelligent systems and clinical research. The computing environment used for simulation offers the possibility of studying disease processes electronically (*in silico*). Precisely, the study of tangled brain pathways would require the simulation of large-scale network models integrating the neurons and interconnections involved. However, there is a need to develop more comprehensive models and multi-level validations of these models through experimental studies to uncover the underlying molecular mechanisms leading to unknown dysfunctions.

Several simulation environments have been designed for building and using these mathematical classes of neuron- and network- models. On one hand, simulators as NEURON (Hines and Carnevale, 1997) are empirically based simulators of morphologically detailed neuron models (they allow describing neuron mechanisms, i.e. mathematical differential equations, mimicking the biological cell dynamics at a molecular level) widely used by theoreticians and electrophysiologists. On the other hand, NEST is focused on more specific neuronal simulations (i.e., dynamics and structure of spiking neuron network models) using simplified neuron models (defined by few mathematical equations) rather than including the exact morphology as the detailed single-neuron models. This approach is ideal for modeling information processing or network activity dynamics, and for models of learning and plasticity. Other current simulators at this level of abstraction are BRIAN (Stimberg et al., 2019) and EDLUT (Ros et al., 2006; Naveros et al., 2015, 2017). In this thesis, we have mainly used the NEST simulator as the modeling tool for our experimentation.

Detailed neuron models (i.e., integrating a high degree of biological plausibility) allowed fine-grained studies about intrinsic cellular dynamics, but the high computational cost associated with the simulation of these complex models makes them only suitable for small-scale models of networks or short simulations. Thus, simplified models appear to be a perfect alternative for exploring the functional role of resonant activity in information processing. Simplified models (such as point-neuron models) combine computational efficiency and realistic neuronal dynamics. Considering this, the adaptive exponential integrate-and-fire (AdEx) model (Brette and Gerstner, 2005) only includes two coupled differential equations which capture adaptation and resonance properties (Naud et al., 2008). In addition, the AdEx model enables large-scale implementations of neuronal circuits, which is key for testing predictions in the granular layer. Since the cerebellar GrCs show compact and simple morphology (D'Angelo et al., 2001; Delvendahl et al., 2015), it is appropriate to consider a mono-compartment model. For these reasons, this thesis models cerebellar GrCs based on AdEx neuron templates because of its computational efficiency and realistic firing modes.

The AdEx model can be seen as a two-dimensional reduction of the spike initiation of Hodgkin-Huxley models (the most common detailed models used by electrophysiologists). However, some of its parameters lack an experimental (measurable) counterpart (i.e., some parameter values to match with electrophysiological measurements cannot be experimentally determined). Finding an appropriate set of parameters becomes a challenging problem (Barranca et al., 2013; Venkadesh et al., 2018) which makes it being considered a suitable optimization problem that remains partially unsolved. Different optimization algorithms for tuning the parameters of computationally efficient neuron models and reproducing certain biological behaviors have been used (Van Geit et al., 2008). More generically, recently there are interesting tools being developed to optimize the parameters of these type of neuron models to fit different cell characteristics (Friedrich et al., 2014; Van Geit et al., 2016). But after exploring these tools four our parameter study, since, in our case, the burst frequency (spiking

resonance) was considered a key feature, we decided to use generic parameter optimization approaches. However, since no exact method is known to solve the nonlinear and simulation-based target problem, this thesis addresses the exploration of a methodology for building computationally efficient neuron models of cerebellar GrCs that replicate some inherent properties of the biological cell.

Moreover, the relevance of heterogeneity in neuron populations of the same type with variances in their properties, such as the cerebellar GrCs, is sound (Lengler et al., 2013; Migliore et al., 2018). The intrinsic variability of neurons in brain areas is thought to crucially change the microcircuit dynamics, and thus taking a relevant part in the information processing. However, the most used optimization strategies return a single optimal solution (a single neuron model). This thesis also aims to explore alternative optimizers that result in heterogeneous populations (exploring the parameter space) of promising neuron models whose properties are closely matched to the biological target data.

Chapter 2

Thesis Contextualization

2.1 Motivation

Currently, it is unknown how neuronal characteristics (and their biological basis at the molecular level) affect the representation of somatosensory information, their interaction with synaptic plasticity mechanisms and the development of neurological diseases. In this thesis, we develop neuron models with intrinsic dynamics that capture key features that we consider functionally relevant. In future work, we will study the impact of the characteristic firing patterns (spike bursting, resonance, delayed firing, among others) of a certain cerebellar neuron (the cerebellar granule cell) on the overall functioning of the cerebellum. By using cellular models (which reproduce the inherent dynamics obtained in electrophysiological studies), the functional implications on the neuronal system will be explored. This thesis will serve as a first step leading to advances in the basic knowledge of the brain, as well as in the prevention and treatment of neurological diseases.

This thesis is of interest in various fields of research and topics of international relevance:

- To understand the brain, we must cover all levels of abstraction: from the molecular, cellular and synaptic level to micro- and macro-circuitry and even behavioral impact. At the same time, dealing with neurological diseases requires a multidisciplinary approach: from a mutated gene to its effect on a neuron, on a network, to thousands of synapses (D'Angelo et al., 2013a), or even reproducing the symptomatology and rehabilitating with an agent (which can be a robot avatar) (Geminiani et al., 2018).

- This is an ambitious goal and it is already being tackled by international projects such as the Human Brain Project (HBP in Europe, in which our research group has been participating for more than six years) or the BRAIN initiative (in the USA). These international projects develop reverse engineering research in an interdisciplinary framework with the common aim of unraveling the brain, covering all the different possible levels of abstraction.

- This thesis aims at representing a differential value by addressing the optimization and modeling of simplified neurons from experimental data, capturing key functionally relevant features. This will help, in the future, to evaluate the processing capabilities of the granular layer based on specific characteristics of these neurons and their adaptation mechanisms. Thus, this work is approached from a thematic interdisciplinary perspective.

- Future implications of this thesis project may expand to the study of 1) the behavioral impact on biological tasks, understanding cerebellar neurological diseases and facilitating their treatments, 2) neuronal rehabilitation schemes, 3) reverse engineering, and 4) efficient computing (better understanding new forms of massively parallel computing).

Regarding the field of neuroscience, biological systems have traditionally been studied by using *in vitro* and *in vivo* experiments. Recently, computational neuroscience with its *in silico* (performed via computer simulation) experiments has emerged as a third approach to address the points mentioned above. It is important to also mention that biophysical model predictions can simulate results that are not possible experimentally, or even replicate experiments on animal models, such as the mouse (Computational Models of Brain and Behavior, 2017). This aims to also have an impact on the 3Rs ethics on animal experimentation (the principles of the 3Rs, abbreviations for Replacement, Reduction and Refinement): (1) Replacement of animals by simulation when possible; (2) Reduction on the need of the number of animals required for animal experimentation; and (3) Refinement of the specific experiments and pursued results that are needed (and thus reducing the number of experiments required). The impact on the 3Rs in animal experimentation can be seen in itself as an important motivation for the effort done of pushing the simulation capabilities of neuronal systems in this thesis.

Thus, the research in the field of computational neuroscience seeks (1) to reduce the need for animal experiments, (2) to exploit reverse engineering (theory-dominated top-down approaches) of smart computation primitives, (3) to study diseases in unprecedented *in silico* experiments, and (4) to improve the validation of data and experiments with computational validation. Long-term applications together with other interdisciplinary fields such as Neurorobotics can also lead to neuronal rehabilitation schemes, efficient computing, and acting principles of biologically relevant tasks. However, a deeper knowledge on how the biological systems process the information is yet of utmost importance. In this thesis we have used the results contained in different bibliographic source volumes, databases and also obtained from electrophysiological studies in order to develop novel methodologies (1) that ease the identification of relevant components in complex diseases and (2) that create computationally efficient neuron models of cerebellar cells. These neuron models will serve to simulate large-scale networks that will validate the hypothesis raised by *in vitro* and *in vivo* experiments and to propose new hypotheses that can be validated by these types of physiological experiments. Thus, computational neuroscience experiments are a perfect complement for cellular and animal experimentation.

With respect to the complexity of disease, there are multiple studies on distinct cellular dynamics and their associated pathologies (Necchi et al., 2008; Pappalardo et al., 2016; Spillane et al., 2016; Sathyanesan et al., 2019; Mitoma et al., 2020). Understanding how these properties relate to specific genes is of utmost importance. The large amount of biomedical literature sources available today often makes it difficult to integrate all this knowledge to identify the genetic causes underlying complex diseases. However, extracting these causal agents becomes essential to understand the pathomechanisms of complex diseases. What's more, it also allows a richer diagnostic management and exploring possible suitable targets as potential therapies. Complex diseases (those that combine a large number of genetic factors) are difficult to classify: some reviews deal with categorizations based on organ systems, others by the proteins involved or by specific mutations. Some genes involved may be involved in other systems that *a priori* are not related to the main organ system. Current traditional search systems (such as the exhaustive or systematic review) allow collecting this significant biomedical information (Ferreira González et al., 2011). However, these search systems used to be arduous, non-automatic and may not even allow the detection of interconnections between the genes involved or outcomes that are not retrievable. Therefore, in this thesis we have made a preliminary exploration of valid association procedures that allow coping with these levels of biological complexity. The aim of our methodological proposal is the comprehensive interpretation of current biological knowledge that acts as a bridge-builder between fields and extracts information that *a priori* may not seem important but which is extremely useful (Yang et al., 2015).

Focusing on the information transmission in the cerebellum, many *in vivo* studies have evidenced subthreshold resonance at theta frequency (Pellerin and Lamarre, 1997; Hartmann

and Bower, 1998; Courtemanche et al., 2009) and its possible role in the signal encoding from the granular layer (D'Angelo et al., 2009). However, according to previous *in vitro* and *in silico* studies (D'Angelo et al., 2001; Gandolfi et al., 2013), supra-threshold (spiking) resonance of single cerebellar granule cells unveils a possible role of this complex non-linear cellular dynamics in the processing and transmission of information at theta frequency rhythm. It is theoretically known that the spiking resonance feature depends on spiking mechanisms and intrinsic properties of single cells in most cases (Rotstein, 2017). Single-neuron responses in the granular layer have long been investigated in search of such theta-frequency activity patterns (D'Angelo et al., 2009; Gandolfi et al., 2013). However, how these neuronal dynamics interact with information processing primitives in terms of long-term plasticity and learning in the granular layer remains unclear (Masquelier et al., 2009; Solinas et al., 2010; Garrido et al., 2013b). In this thesis we have explored the cellular mechanisms underlying this behavior in order to generate efficient neuron models that allow us to simulate the granular layer on a large scale and thus to assess the impact of theta-frequency spiking resonance on information transmission.

Regarding the mathematical modeling of the information processing in the granular layer, it remains unclear how spiking resonance (demonstrated *in vitro* in GrCs and GoCs) interacts in a recurrent inhibitory loop with feed-forward excitation of the Golgi cell. Despite theoretical models of granular layer have addressed the information processing primitives (Solinas et al., 2010; Garrido et al., 2013a), these models either have not considered intrinsic resonance or they have neglected the role of the long-term plasticity in the GrC inputs. In addition, and from a functional perspective, mathematical models have demonstrated that an external oscillatory rhythm provides strong stability as an efficient learning scheme in excitatory synapses (Masquelier et al., 2009) and inhibitory recurrent networks (Garrido et al., 2016). With respect to the cerebellar GrCs, detailed models have suggested that the spiking resonance in the theta-frequency band may favor the phase-locking through the recurrent Golgi cell inhibitory circuit (Maex and De Schutter, 1998; D'Angelo et al., 2001). Moreover, these neuron models have suggested a regulating role in the induction of synaptic plasticity in the mossy fiber-granule cell pathway (Armano et al., 2000; D'Angelo et al., 2001). This represents a strong motivation towards integrating complex dynamics capturing the resonant behavior as a key functional feature of neurons (as has been done in this thesis).

To this end, in this thesis we have made use of the simulation tools offered by computational neuroscience together with optimization strategies. Mathematical models that capture "functional" neuron properties (i.e., relevant properties to the computational capacity of the neuron and the network) with sufficient biological plausibility while being computationally efficient. In order to create such models, parametric optimization methods are required to capture the properties that we select as relevant, and that are matched as closely as possible to the electrophysiological recordings of real cells.

2.2 Objectives

The general vision of this thesis is focused on providing some pieces of knowledge (our main contributions) to some of the questions that remain elusive about the brain and more specifically the cerebellum. The general goals of this thesis are summarized in **figure 5** and state as follows:

1. The first aim of this work is a preliminary exploration of complex diseases in a molecular level from the integration of a huge amount of data available, and extracting information and possible conclusions that *a priori* might not seem relevant, such as comorbidities.
2. The second aim is to create neuron models that capture particular properties of the of cerebellar granule cells which seem key for the processing of somatosensory information in the cerebellum's granular layer. The specific behavior under study, the spiking resonance under theta-frequency synchronization, is a cellular property usually ignored in most models.

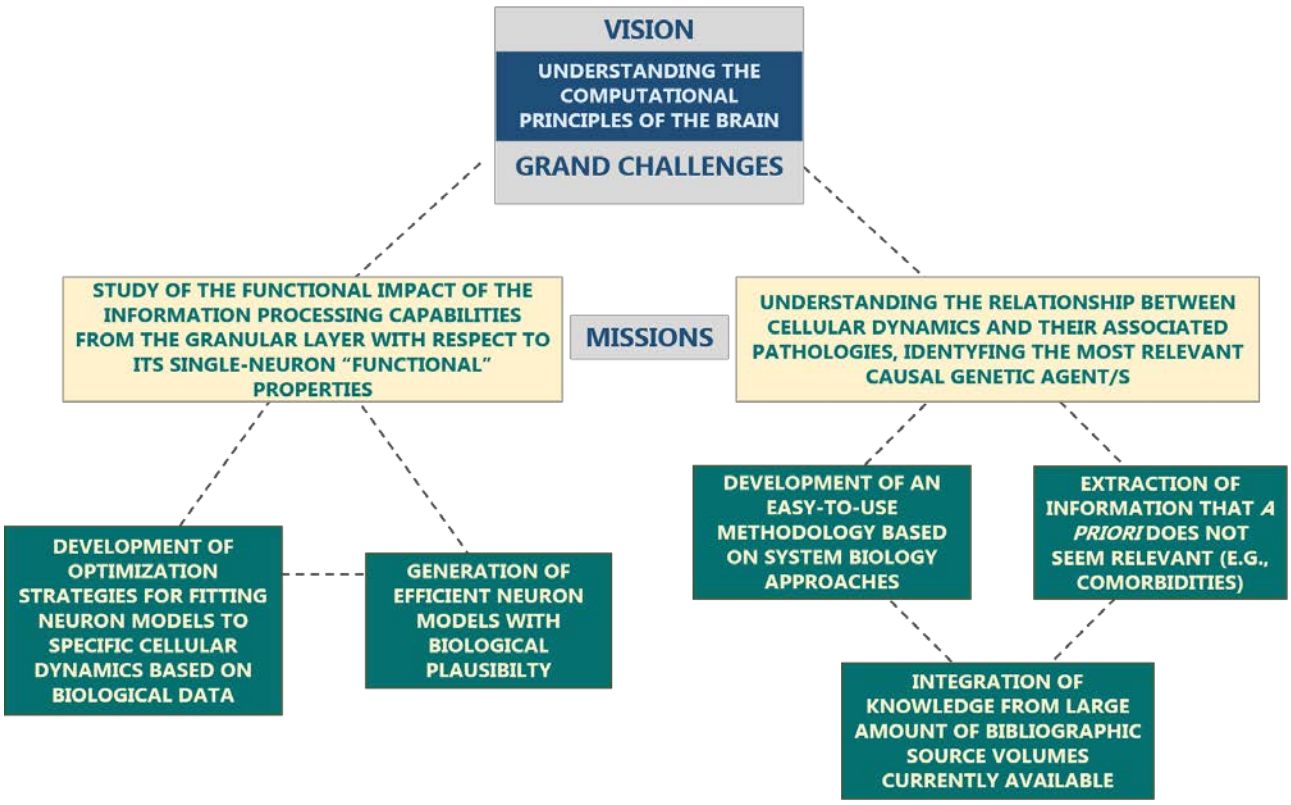


Figure 5. Vision and missions of this thesis.

In this thesis, we aim at modeling the cerebellum at different levels of detail, ranging from the computational primitives (simplified efficient behavioral neuron models) to their

neurobiological substrate (neuron models from biological measurements at molecular levels such as cellular dynamics of ionic channels). Particularly, this thesis addresses the identification of the specific properties of neurons from the granular layer of the cerebellum that are considered relevant for the somatosensorial information processing. Some of these features are the neuron resonance and the synchronization in the theta-frequency band that usually are ignored in most of the models.

That said, the specific objectives that this thesis addresses are:

1. Developing semi-automatic search methodology that allows integrating biomedical knowledge from different studies in the field. Deepening in the biochemical and physiological knowledge of the brain

2. Development of a methodology for simulation and simplification of neuron models

3. Development of computationally efficient neuron models that integrate nonlinear cellular dynamics

4. Study of the resonance at the cellular level in cerebellar granule cells and evaluation of its computational capabilities in the brain

5. Development of a model of cerebellar granule cell integrating nonlinear cellular dynamics (such as the spiking resonance)

6. Evaluation of the capabilities and efficiency in the information transmission of the granular layer with nonlinear cellular models

2.3 Our contribution

With respect to the first general goal of this thesis, our contributions are covered in the first (Marín et al., 2019) journal article included in this thesis. We have proposed a methodology for the objective interpretation of biomedical data and for coping with the higher level of biology complexity of complex diseases (Martin-Sanchez and Verspoor, 2014; Peek et al., 2014; Coveney et al., 2016). This methodology identifies and functionally annotates the relevant genes of complex diseases as molecular targets from the global analysis of multiple

interactions at different levels. We have integrated multidimensional data and used software tools of high impact based on protein-protein interaction (PPI) networks as representations of the biological complex interactions underlying these disorders. The central position of the nodes in the network are key in the main biological and molecular processes, i.e. they tend to be potential pharmacological targets (Goñi et al., 2008; Diaz-Beltran et al., 2013; Stoilova-McPhie et al., 2013; Di Silvestre et al., 2017). The proposed pipeline detects the central nodes that maintain the structure and information fluxes into the functional network. The final step consists on the exploration of the biological and molecular processes in which these candidate genes play a joint key role. In this step it is unveiled if these candidate genes could also be related with other diseases as comorbidity, which would provide a more comprehensive view of the importance of treatments (Stoilova-McPhie et al., 2013; Sánchez-Valle et al., 2017, 2020). We compared the usefulness of the proposal with other annotation and association systems such as the traditional bibliographic search systems (which are inefficient, subjective and time-consuming by hand) (Ferreira González et al., 2011), and with some of the highest impact tools for functional annotation (which are objective, quick and reproducible) (Yang et al., 2015). We applied this proposal to the case of channelopathies, whose remarkable causal heterogeneity is sound (Kim, 2014; Spillane et al., 2016). This application contributes to the overall understanding of pathomechanisms underlying these altered-channels diseases, in how mutations can modify disease severity (Musgaard et al., 2018), and to shed some light on effective treatments (Stoilova-McPhie et al., 2013; Spillane et al., 2016; Schorge, 2018). The proposed methodology proved to build as productive results as other non-automatic traditional search systems. At the same time, it was shown to work more flexibly, quicker, making it a convenient and easy-to-perform first-level approach compared to the methods analyzed in the article. This contribution serves as a bridge-builder among fields allowing the extraction of information that *a priori* might not seem relevant or related when the starting point is a very large group of genes in disease (i.e. comorbidities). It is worthy to mention that the proposed methodology and results obtained in this part of the thesis has helped in further studies by other research groups in related research areas such as Alzheimer (Meng et al., 2020), children obesity (Plaza-Florido et al., 2020), and neuroepithelial tumors in the posterior fossa (Wang et al., 2021).

Regarding the rest of the goals of the thesis, our main contributions are contained within the remaining journal articles included in this thesis (Marín et al., 2020, 2021; Cruz et al., 2021). In the following paragraphs, we will describe the specific contributions of each of these articles.

In (Marín et al., 2020), we developed an automatic parameter tuning method based on genetic algorithms (GAs) for adjusting neuron parameters to spiking features of real cells. Although it is of general purpose and flexible to be adapted to different neuronal optimization problems, we applied it to the case of cerebellar GrCs. Thus, we have contributed proposing different neuron models of cerebellar GrCs that are efficient enough to simulate large-scale

networks (as it is the granular layer). We showed that these GrC models suitably reproduce the spiking resonance at the theta-frequency band. This feature is usually ignored in most of the efficient models, and it is thought to have a key role in the information transmission. We also demonstrate the biological plausibility of these models not only reproducing the main electroresponsiveness typical of real GrCs (i.e., spike frequency of I-F curves and spike timing to the first spike) explicitly captured in the "cost function" that is minimized with the optimization engine, but also other emergent properties not defined in the "cost function" of the algorithm. The realism of these models is of utmost importance, and they are proposed as valid models in order to evaluate the functional impact of resonance on the information transmission.

In (Cruz et al., 2021), we explore four alternative optimization strategies (MemeGA, DE, TLBO, MSASS) in the same context of the previous work. These innovative algorithms are some of the most widely used and successful in other fields, such as engineering and technology. We compared them to the algorithm of reference from the previous published article, which is a standard GA. These alternative algorithms numerically improve the adjustment of the models to the neuronal features of interest. Particularly, TLBO obtained the best candidate solution (the best-fitted neuron model to the biological values) from all the methods compared. This neuron model improves the temporal accuracy to the first spike with respect to the reference candidate (the neuron model obtained with GA in the previous article). The contribution of this article lies on the proposal of more efficient and effective alternative optimization methods. In particular, with this article we propose a more sophisticated methodology for the optimization of neuron models. In addition, we propose a neuron model that is more closely tuned to the spiking characteristics of GrCs, with more accurate reproduction of the latency to the first spike.

The heterogeneity of neuron populations has been shown to be beneficial for improving information processing in network dynamics. In (Marín et al., 2021), we have contributed to this by presenting a methodology based on multimodal algorithms to the same case of the resonant GrCs. In this way, we would not propose only one best neuron model after the optimization process, but rather we provide a population of promising neuron models. From this population of models, the researcher can decide based on the filter or selection of the desired feature. This represents a powerful tool for studying the parameter space. With this work, we contribute with an unfamiliar approach in the field of neuron parameter fitting. More precisely, with these sparse population of GrC models (but fitted to the same spiking features of one real cell) we are now able to address the research question of the functional impact of resonance on the information processing of heterogeneous networks.

2.4 Project framework

The work described in this document has been developed in the framework of one European project, i.e. "Human Brain Project" [HBP, Specific Grant Agreement 2 (SGA2 H2020-RIA 785907) and 3 (SGA3 H2020-RIA 945539)], one national project, i.e. "Integración sensorimotora para control adaptativo mediante aprendizaje en cerebelo y centros nerviosos relacionados. Aplicación en robótica" [INTSENSO (MICINN-FEDER-PID2019-109991GBI00)], as well as one regional project, i.e. "Cerebelo y oliva inferior en tareas de adaptación sensori-motora" [CEREBIO (J.A. P18-FR-2378)].

This thesis has facilitated interdisciplinary collaborations with different groups close to systems biology and biomedical applications, such as the University of Jaén (Francisco J. Esteban) and the University of Granada (Hilario Ramírez). In addition, part of this thesis has been developed in collaboration with the research group of Computer Science and Technology led by Pilar Ortigosa at the University of Almería.

Our research group has been involved in the Human Brain Project (HBP) for more than six years, and this international project has partially funded the thesis of the doctoral candidate. This is a European Commission Future and Emerging Technologies Flagship that has involved in different stages more than 100 research groups. It was launched in October 2013, and it is scheduled to run for ten years. The HBP aims to produce a cutting-edge, ICT-based scientific Research Infrastructure for brain research, cognitive neuroscience and brain-inspired computing. This international project was originally divided in 12 subprojects. Our research group at the University of Granada has participated in previous stages of the HBP mainly in the framework of the Neurobotics sub-project, in which our role was to validate the different platforms that are being built for addressing an in-depth study of the brain. Thus, adapting our experiments to the research platforms (neuronal simulations such as NEST (Gewaltig et al., 2012)), and the Neurorobotics platform (NRP). One of the roles in this project by our research group at the University of Granada has been as an early adopter of such platforms, validating them for our studies on computational neuroscience. We firstly contributed to the second phase of this project (SGA2: H2020-RIA 785907) with a better understanding of the pathologies related to the nervous system throughout our proposed methodology for the identification of the most relevant genes in complex diseases. We also have explored some of the platforms proposed in this project for neuron fitting, i.e. Neurofitter (Van Geit et al., 2007), Optimizer (Friedrich et al., 2014) and BluePyOpt (Van Geit et al., 2016). Although these tools work appropriately for the optimization of neuron parameters, some of their limitations avoided us to create the simplified neuron models that we aim (these tools are more focused to detailed neuron models and the optimization of the whole electrophysiological traces). The optimization problem on which this thesis is based has required developing alternative

optimization methodologies. The methods proposed in this thesis address the optimization of simplified neuron models under different contexts for which each of them is more suitable. All of the proposed methodologies result in efficient neuron models of cerebellar GrCs that can be integrated in the cerebellar model to be developed at our research group. Thus, it could be connected to detailed simulations of robot bodies and environments.

The INTSENSO project (funded under the Ministry of Science and Innovation (MICINN-FEDER-PID2019-109991GB-I00) framework from the National Plan 2019) started in January 2020 and will finish in December 2022. The main goal of this project is to cerebellar neuron models and their impact on the processing capabilities of the cerebellum. This thesis has addressed the second main objective of this project namely "O2. Simulation of the cerebellum, inferior olive and basal ganglia). Specifically, we have contributed with the creation of cell models with internal dynamics (neurophysiological substrate) that could support specific computational capabilities.

The CEREBIO project (funded by Junta de Andalucía, J.A. FEDER P18-FR-2378) started in January 2020 and will finish in December 2022. The main goal of this project is the study of the computational primitives of the central nervous system during the body-environment interaction. With this approach, this project aims to stablish correlations between experimentation and simulation. The principal goal is to develop a methodology and a holistic vision of the cerebelar modeling. It is tackle ranging from the single-neuron features level, to network- and behavioral- levels that helps to understanding the brain computation. This thesis has contributed with a better understanding of the complex set of pathologies associated to "altered" ion-channels integrating genomics, transcriptomics and proteomics levels with the clinical data. In addition, this thesis has contributed to one of the main objectives of this project "Objective 2. Development and implementation of cerebellar model that allow testing working hypothesis related to neuronal mechanisms". Specifically, we created efficient cerebellar GrC models that capture the complex intrinsic feature of spiking resonance in the theta-frequency band. This work serves as a first step in the research of the functional impact of this feature in the granular layer.

Finally, the doctoral candidate has made a collaboration with the Pilar Ortigosa's group at the University of Almería. We have developed an interdisciplinary synergy. In particular, we have explored alternative optimization engines that they widely used in their research field (engineering and technology), but unfamiliar in our field of computational neuroscience. We are working together integrating their advanced algorithms within our biochemical context and neuron optimization (i.e. AdEx models of cerebellar GrCs that replicate the typical spiking features together with the spiking resonance in the theta-frequency band). With this collaborative work, we are seeking to contribute with sophisticated general optimization methodologies in the very specific context of optimizing point neuron models. We aim to

provide innovative methods that meet the requirements of emerging neuron optimization scenarios that arise in parallel with the advances in technology and neuroscience.

Chapter 3

Results

Contributions to scientific journals

Four journal articles support this thesis. All of them are ranked in the Journal Citation Reports (JCR): **three** of them in the first quartile (**Q1)** of the JCR categories, and **one** of them in the second quartile (**Q2**)**.** In agreement with the current normative of the University of Granada, doctoral theses presented as a compendium of scientific publications must be composed of at least three articles published in a journal indexed in the JCR. At least one publication must be located in the first tercile of the thematic category. In addition, the Ph.D. candidate must sign as first author or corresponding author.

The aforementioned journal articles have been enumerated, ranked and included in this chapter by following the normative for the compendium modality. Please note that each of the following articles contains its specific bibliography and methods section. Therefore, the references that the reader will find in each journal article do not correspond to the main bibliography section but to its particular one.

1. An integrative methodology based on protein-protein interaction networks for identification and functional annotation of disease-relevant genes applied to channelopathies

Authors **M. Marín**, F. J. Esteban*, H. Ramírez-Rodrigo, E. Ros & M. J. Sáez-Lara*

Journal *BMC bioinformatics*

Year 2019

Volume 20

Pages 565

DOI 10.1186/s12859-019-3162-1

Impact Factor (JCR 2019) 3.242

Categories

Mathematical & Computational Biology:	Ranking 9/59	**(Q1)**
Biochemical Research Methods:	Ranking 24/77	(Q2)
Biotechnology & Applied Microbiology:	Ranking 58/156	(Q2)

Marín et al. BMC Bioinformatics (2019) 20:565
https://doi.org/10.1186/s12859-019-3162-1

BMC Bioinformatics

METHODOLOGY ARTICLE Open Access

An integrative methodology based on protein-protein interaction networks for identification and functional annotation of disease-relevant genes applied to channelopathies

Milagros Marín[1,3], Francisco J. Esteban[2*], Hilario Ramírez-Rodrigo[3], Eduardo Ros[1] and María José Sáez-Lara[3*]

Abstract

Background: Biologically data-driven networks have become powerful analytical tools that handle massive, heterogeneous datasets generated from biomedical fields. Protein-protein interaction networks can identify the most relevant structures directly tied to biological functions. Functional enrichments can then be performed based on these structural aspects of gene relationships for the study of channelopathies. Channelopathies refer to a complex group of disorders resulting from dysfunctional ion channels with distinct polygenic manifestations. This study presents a semi-automatic workflow using protein-protein interaction networks that can identify the most relevant genes and their biological processes and pathways in channelopathies to better understand their etiopathogenesis. In addition, the clinical manifestations that are strongly associated with these genes are also identified as the most characteristic in this complex group of diseases.

Results: In particular, a set of nine representative disease-related genes was detected, these being the most significant genes in relation to their roles in channelopathies. In this way we attested the implication of some voltage-gated sodium (SCN1A, SCN2A, SCN4A, SCN4B, SCN5A, SCN9A) and potassium (KCNQ2, KCNH2) channels in cardiovascular diseases, epilepsies, febrile seizures, headache disorders, neuromuscular, neurodegenerative diseases or neurobehavioral manifestations. We also revealed the role of Ankyrin-G (ANK3) in the neurodegenerative and neurobehavioral disorders as well as the implication of these genes in other systems, such as the immunological or endocrine systems.

Conclusions: This research provides a systems biology approach to extract information from interaction networks of gene expression. We show how large-scale computational integration of heterogeneous datasets, PPI network analyses, functional databases and published literature may support the detection and assessment of possible potential therapeutic targets in the disease. Applying our workflow makes it feasible to spot the most relevant genes and unknown relationships in channelopathies and shows its potential as a first-step approach to identify both genes and functional interactions in clinical-knowledge scenarios of target diseases.

(Continued on next page)

* Correspondence: festeban@ujaen.es; mjsaez@ugr.es
[2]Systems Biology Unit, Department of Experimental Biology, University of Jaén, Jaén, Spain
[3]Department of Biochemistry and Molecular Biology I, University of Granada, Granada, Spain
Full list of author information is available at the end of the article

(Continued from previous page)
Methods: An initial gene pool is previously defined by searching general databases under a specific semantic framework. From the resulting interaction network, a subset of genes are identified as the most relevant through the workflow that includes centrality measures and other filtering and enrichment databases.

Keywords: Channelopathies, Protein-protein interaction networks, Genotype-phenotype relationships, Translational bioinformatics, Behavioural diagnosis, Genetic diseases, Systems medicine

Background

The genetic aetiology of many complex diseases comprises different specific clinical symptoms and evolution. The identification of their causal agents becomes essential for the detection of suitable targets, the management of their diagnosis and the selection of the most adequate therapies [1–3]. The increasing availability of large bibliographic data volumes lays the foundations for the identification of these candidate genes [2, 4]. However, the integration of all this knowledge requires understanding the diverse biomedical information sources available. The extraction of data performed by valid association procedures and the comprehensive interpretation of all this current knowledge is complex. This is in and of itself an issue of utmost importance for the purpose mentioned above [4–6].

Traditional reductionist strategies that deal with this diverse wealth of information focus on the study of particular molecules or signalling pathways that are useful for the identification of diagnostic biomarkers. Nevertheless, it does not seem enough to approach all the system complexity [2, 4]. Alternatively, interdisciplinary research is developing new technologies and integrative computational methodologies in order to better understand pathogeneses [7, 8]. Some studies that use these current integrative methodologies allow the discovery of comorbidities between Alzheimer's disease and some types of cancers [9] where genetic factors can play an important role along with other factors such as the environment, lifestyle, and drug treatments. They are also being used to perform a genome-wide search for Autism gene candidates [1]. These new tools are able to manage deductive analyses by gaining insight into the connections among diseases, even between those *a priori* not related by the traditional bibliographic searches, which usually tend to be subjective, time-consuming or not reproducible [10]. However, the large range of diverse new tools created within different focuses hinders the existence of a unique approach to or a consensus on their usage. Thus, data extraction through ad-hoc approaches using specific tools may again be complex, not reproducible or subjective. In this way, network analyses and functional annotation tools represent some of the best strategies for objective interpretation of biomedical data and cope with higher level of biological complexity [1–3, 11].

The identification of relevant genes is being addressed from the global analysis of multiple interactions at different levels, usually employing networks as representations of the biological complex interactions underlying clinical disorders [11–13]. A way to systematically decode the cellular signalling networks consists in the identification of interactome for the detection of the central nodes which maintain the structure and information fluxes into the functional network [11, 14]. Despite some limitations, protein-protein interaction (PPI) networks have been suitably applied to the definition of biological mechanisms by integrating PPI data with transcriptional changes [1, 13–15]. It is evidenced that in disease networks in which the alteration is produced by mutations, the node or nodes mutated play a primary role in the development of diseases and thus have a central position in the network [16]. In the case of multifactorial diseases, the nodes which seem to be the causal factor could be located in the periphery. However, the key nodes in the main biological and molecular processes affected, i.e. potential pharmacological targets, tend to have a central position in the network [1, 17–19]. Thus, for the identification of the most significant genes in a disease as molecular targets there are useful software tools of high impact [20–24]. One of them is STRING [22], a database used to build predicted and well-known PPI networks. The interactions in STRING are mainly derived from automated text-mining and databases of previous knowledge, among other resources. Other well-known tool is Cytoscape [23], an open-source software platform which has being designed for the purpose of visualizing, analysing and modelling complex biological networks and pathways.

Furthermore, in a system biology approach it is highly important to know the biological and molecular processes in which the complex set of genes involved play a joint key role. Though, if the aim were to identify pharmacological targets, it would also be mandatory to unveil if these candidate genes could also be related with other diseases as comorbidity [9, 25]. These annotations and associations can be performed through traditional bibliographic search systems, which are inefficient, subjective and time consuming by hand [10], or by using some of the highest impact tools from the large number of platforms developed for functional annotation in

objective, quick and reproducible ways [26]. This is the case of DAVID [24], which has been shown to provide an automatic comprehensive set of functional annotation tools for biological interpretation of large gene lists as pharmacological targets [7]. It is also very useful in unveiling other related diseases, providing a more comprehensive view of the importance of treatments [9, 19].

In this regard, the aim of this study is to present a semi-automatic workflow using PPI networks for the identification and functional annotation of the most relevant genes in diseases. This new contribution to the extant methods is based on the integration of a set of multidimensional data from different biological levels (genomics, transcriptomics and proteomics) in order to analyse genetic correlations among diseases with different clinical symptomatologies and/or clinical prognoses (and still based on similar molecular mechanisms). In order to illustrate the value of this integrative approach and demonstrate its usefulness, we applied this methodology to the case of channelopathies as proof-of-concept in order to understand their most common polygenic influences, which contributes to the overall understanding of pathomechanisms underlying these altered-channels diseases, in how mutations can modify disease severity [27] and to shed some light on effective treatments [19, 28, 29]. We showed that this proposed workflow is able to mine current available databases and platforms in the context of channelopathies.

Results

In this section, we illustrate the experimental application of the semi-automatic workflow (Fig. 1) to the case of channelopathies.

Semi-automatic workflow applied to channelopathies

Gene dataset of the disease under study

First, the gene dataset of channelopathies was created by introducing the term "channelopathies" in the first stage of the present workflow (Fig. 1 Stage 1), which generated a list of 42 genes involved in this complex group of disorders: *SCN5A, KCNH2, KCNQ1, HLA-B, RYR2, SCN2A, SCN4A, CACNA1C, KCNE1, KCNE2, CACNA1S, ATP8B4, DCHS1, SCN4B, SCN2B, SCN9A, SNTA1, CDKL5, STK11, STXBP1, TGFB1, TGFB2, TRPC4, SCN1A, SCN1B, HLA-DRB5, HSPB2, KCNQ2, LOXL2, CNGB3, SCN3B, PCDH19, KCNE3, AKAP9, PRRT2, CLCN1, ASB10, ARX, DMPK, SPESP1, ANK3, HLA-A.*

Identification of the most relevant genes

Then, the list of gene names was the input for Stage 2 (Fig. 1. Stage 2). Our target organism was *H. sapiens*, and a PPI network was generated through the STRING database (interactome network presented in Additional file 1) and then analysed by the Cytoscape platform (Fig. 2). We employed the main features used as centrality parameters, degree and betweenness (as described in methods) for the identification of the most important vertices within the

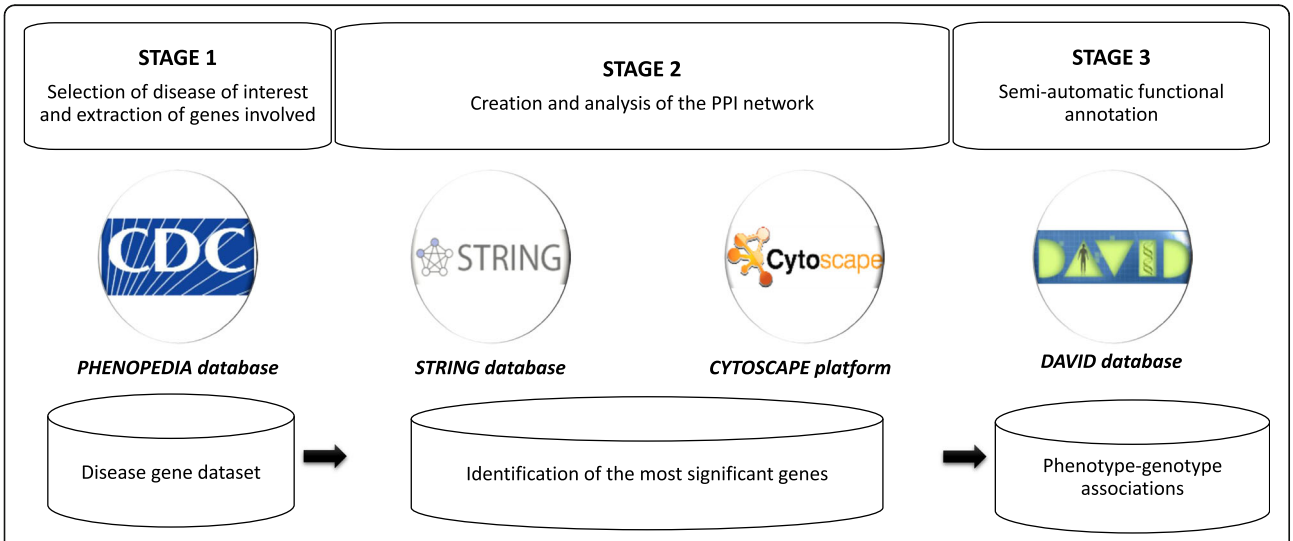


Fig. 1 Semi-automatic workflow for the identification and functional annotation of the most relevant genes in a pathology. Stage 1) *Phenopedia* [30] is a disease-centered view of genetic association studies summarized in the online *Human Genome Epidemiology (HuGE) encyclopedia*. It provides a list of genes involved in the disease of interest. It is in this stage that a complex disease, a set of diseases or a certain disease can be chosen to be studied. Stage 2) *STRING* [22] is a database of known and predicted protein-protein interactions that allows the discovery of relationships across disease genotype and thus the creation of the PPI network. *Cytoscape* [23] is an open source software platform for the visualization and analysis of complex networks that measures each gene and identifies network nodes. Stage 3) *DAVID* database [24] works as a semi-automatic functional annotation tool of the genes obtained after Stage 2

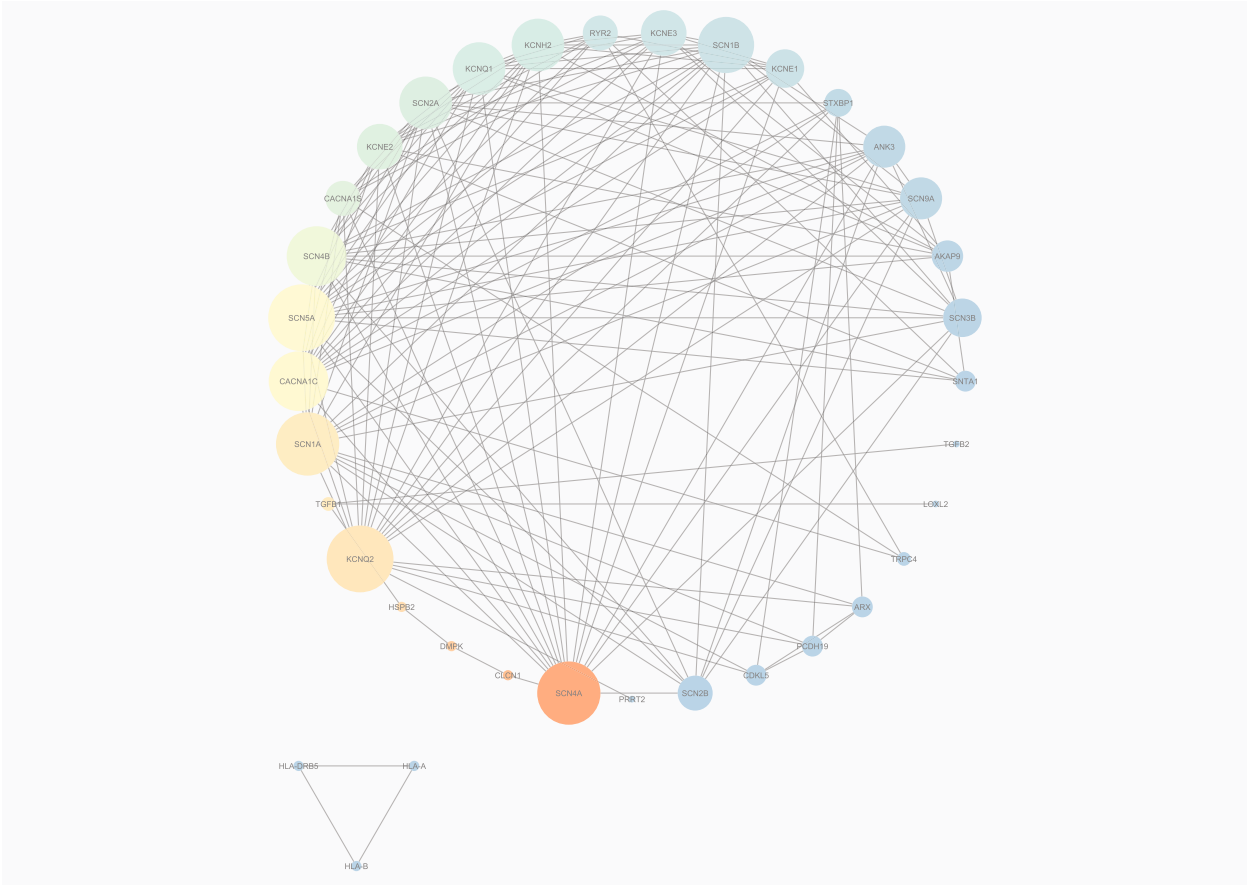


Fig. 2 Protein-protein interaction (PPI) network of channelopathies analysed in Stage 2. Each network node represents the protein produced by each single, protein-coding gene locus from the gene dataset of channelopathies. The representation is a circular layout based on the betweenness attribute with undirected edges (other layouts shown in Additional file 10). The node size marks the level of degree and therefore neighbourhood (the larger nodes represent proteins with a higher number of interactions). The node colour shows the level of betweenness and therefore the level of centrality (the warmer the colour of the protein, the shorter path between the two which indicates how important the node is within the wider context of the entire network) [31]. HLA proteins are discarded due to their disconnection from the principal component. The image was generated by Cytoscape [23]

graph. Thus, starting from 42 genes involved in channelopathies, nine genes with the highest degree of interactions and betweenness in their gene expressions were stemmed as the most relevant in channelopathies: SCN9A, ANK3, SCN5A, SCN2A, KCNQ2, SCN1A, KCNH2, SCN4B and SCN4A. This same set of nine relevant genes was also obtained using other connectivity features, such as closeness, EigenVector and radiality (Fig. 3). The result proves to be robust and concordant with that from Stage 2 of the workflow using only betweenness and centrality.

Gene functional annotation

Finally, the functional annotation of each gene was automatically generated in Stage 3 using DAVID search tool (Fig. 1. Stage 3). All the functional annotation results are detailed in section 4.1 in Additional file 4.

Validation of the workflow

To measure the quality of the results obtained, we carried out an alternative more conventional search with a view to comparing the workflow annotation results to the results offered by two other widespread family of bibliographic methods, such as systematic review and exhaustive review.

Comparison criterion

Using "MeSH" ontology [33], we selected four upper-level categories with their corresponding lower-level ones. Each of these lower-level categories refers to one or more diseases linked to these genes. We used "health disorder" as the specific comparator which contains up to four upper-level categories: 1) cardiovascular diseases, 2) nervous system diseases, 3) mental diseases, and 4) other diseases. This frame comprises all the phenotypes

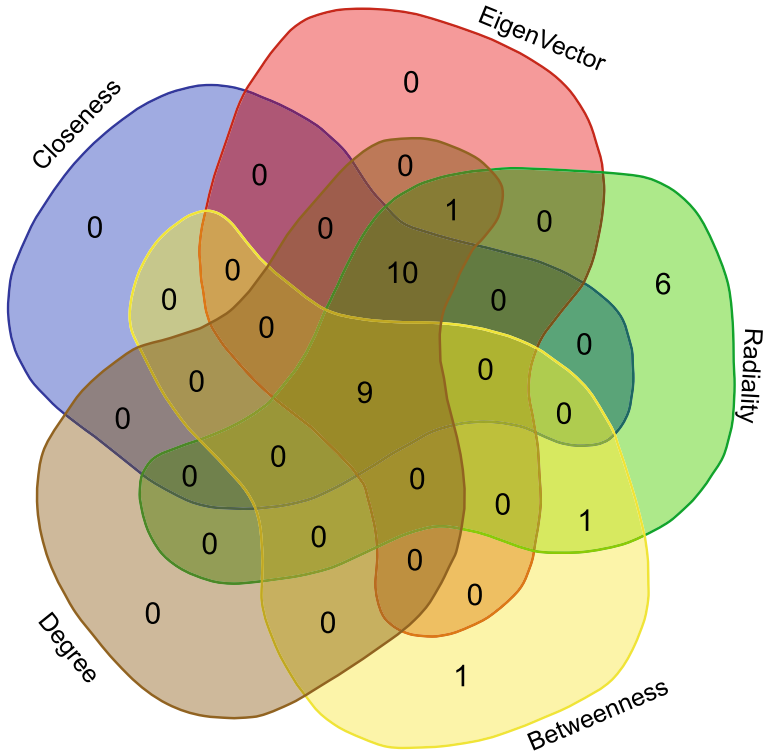


Fig. 3 Nine genes as the most relevant in Stage 2 using different centrality statistics. Venn diagram representing the intersections calculated through the use of other statistics for the proposed centrality measures such as closeness, Eigenvector and radiality. The same set of genes identified with degree and betweenness only still turned out to be the most relevant in channelopathies: SCN1A, SCN2A, SCN4A, SCN4B, SCN5A, SCN9A, KCNH2, KCNQ2 and ANK3. Venn diagram was obtained using a free available tool provided by Ghent University [32]. All the statistics information is included in Additional file 2

of each relevant gene in channelopathies to facilitate the visualization and comparison of functional annotation results (as specified in Table 1). In Additional file 3 we can find the "MeSH"-based terminological hierarchies of the selection of the lower-level categories.

From the functional annotation results through the last stage of the proposed workflow (using DAVID) (Table 4.1.8 in Additional file 4) and applying "health disorder" as the specific-domain category, we obtained the results (consigned in Table 4.2.1 in Additional file 4) that will be visually represented in the final results of this work (Figs. 5, 6 and 7).

Systematic review and exhaustive review as other traditional search systems

In the systematic review we searched by phenotype nomenclatures, filtered by *H. sapiens* as the target organism and removed duplicate entries (Fig. 4). Finally, we extracted nine gene entries from the OMIM and Gene databases and 32 evidences of diseases from the MedGene database (Table 2; all the diseases extracted through the systematic review can be found in Additional file 5). Following the same four upper-level categories, we created an equivalent table containing each disease or clinical manifestation related with its corresponding genes. We used "health disorder" as the specific-domain category and obtained the results shown in Additional file 6. We also compared DAVID against other phenotype-oriented databases of high impact, proving again the selection of this tool in Stage 3 (information included in Additional file 9).

As our third step, exhaustive review was performed by using the query words "gene product nomenclature" + "diseases" in the search box of PubMed and MEDLINE resources, the evidence filtering being the most time-consuming task. We took the same categories and created an equivalent table containing each disease or clinical manifestation related to the corresponding genes, its "health disorder" as the specific-domain category, and its bibliographic references (Additional file 7). While performing this traditional review, we could also expand the functional annotation of the most relevant genes with further information, detailed in Additional file 8.

Representation through genotype-phenotype association networks

From the genotype-phenotype relationships found by the three search systems used in this work – the last stage of the workflow (Table 4.2.1 in Additional file 4),

Table 1 "MeSH"-based categories selected. A total of four upper-level categories and their corresponding lower-level categories capture all the phenotypes manifested by more than one of these genes. We resorted to "MeSH" terminological-based hierarchical networks that include all the phenotypes as referred in the third column (included in Additional file 3)

Upper-level category	Lower-level category	Hierarchical network
Cardiovascular diseases	Vascular diseases	Figure 3.1 in Additional file 3
	Cardiac arrhythmias	
	Other diseases (heart arrest, cardiomyopathies, myocardial ischemia or cardiomegaly)	
Nervous system diseases	Neurobehavioral manifestations	Figure 3.2 in Additional file 3
	Febrile seizures	
	Epilepsy	
	Headache disorders	
	Neurodegenerative diseases	
	Neuromuscular diseases	
Mental disorders	Tobacco use disorder	Figure 3.3 in Additional file 3
	Other mental disorders (bipolar disorder, Alzheimer disease, autism, depression or schizophrenia)	
Other disorders	Sudden death	Figure 3.4 in Additional file 3
	Diabetes Mellitus type 2	
	Periodic paralyses	

Table 2 Gene accession numbers filtered through systematic review

GENE	OMIM ID	Gene ID
SCN1A	182389	6323
SCN2A	182390	6326
SCN4A	603967	6329
SCN4B	608256	6330
SCN5A	600163	6331
SCN9A	603415	6335
KCNQ2	602235	3785
KCNH2	152427	3757
ANK3	600465	288

the systematic review (Additional file 5), and the exhaustive review (Additional file 7) — and considering all the categories selected for every phenotype, we represented association networks for cardiovascular diseases (Fig. 5), nervous system diseases (Fig. 6), and mental diseases and other disorders (Fig. 7).

For cardiovascular diseases, DAVID search (Fig. 5a) found more diseases than the systematic review (Fig. 5b) and the exhaustive review (Fig. 5c), with the exception of a connection between the gene SCN4B and "other heart diseases" category retrieved by the exhaustive review but not by DAVID or systematic searches. This is due to the fact that the gene product of SCN4B is an

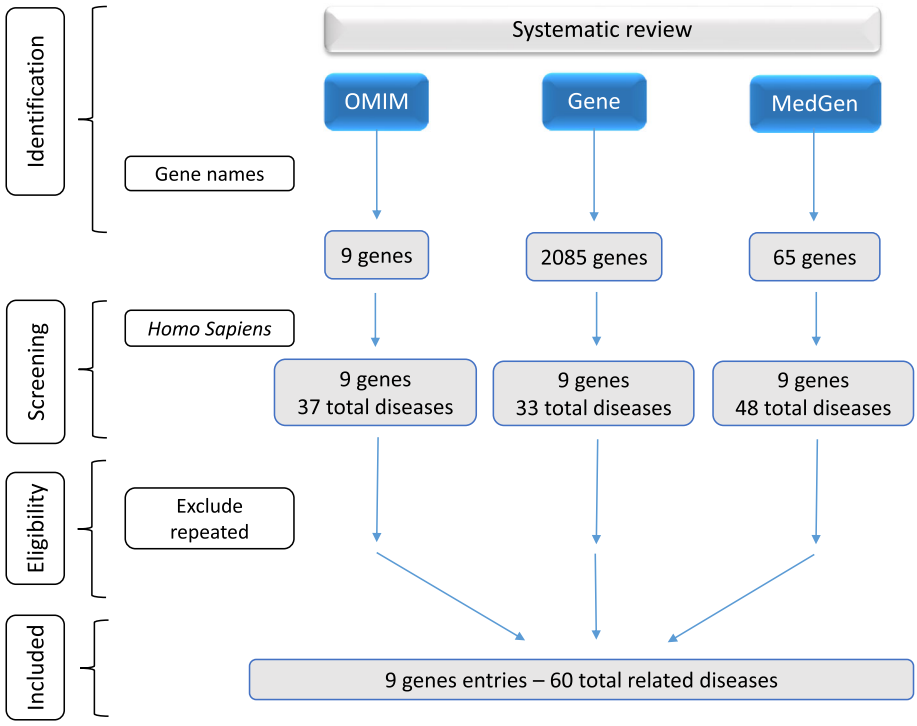


Fig. 4 Systematic review procedure for the genes of interest through OMIM, Gene and MedGen databases

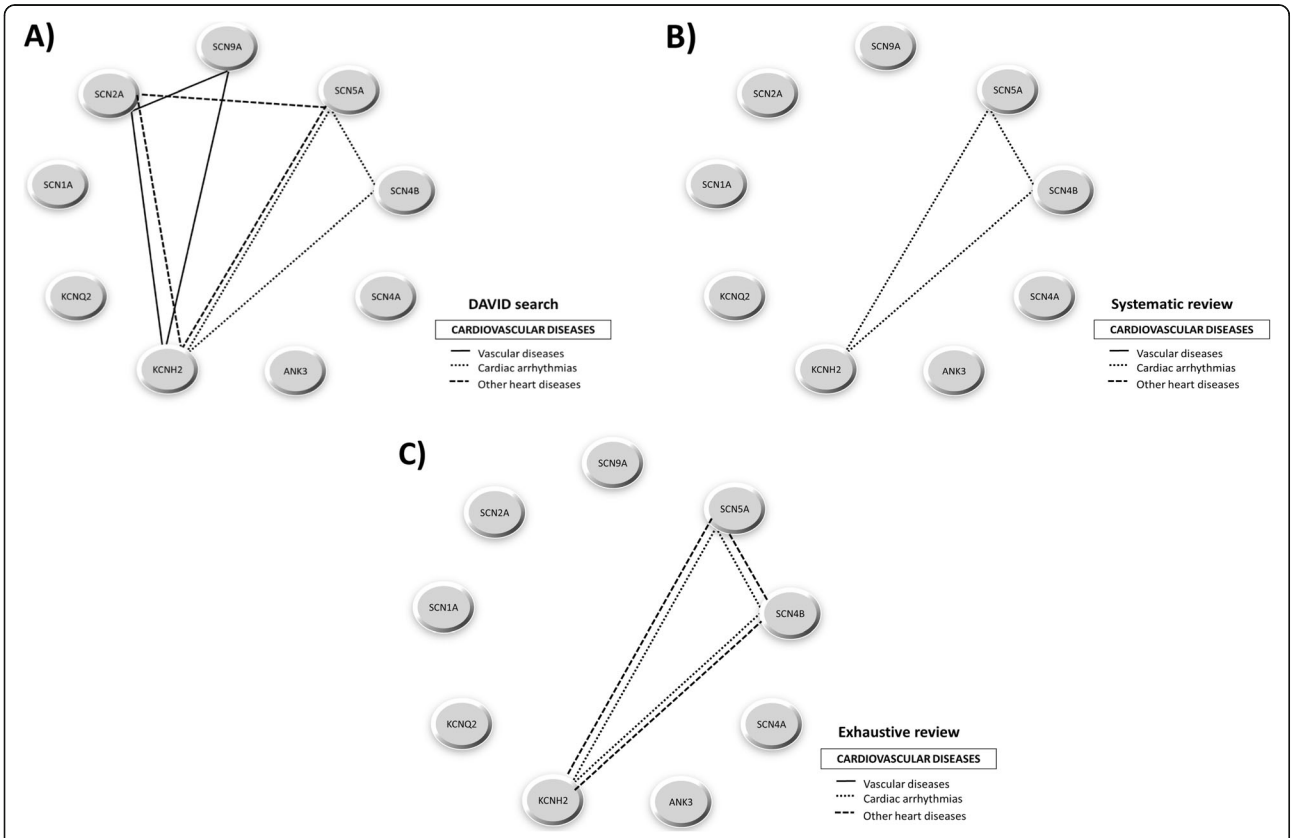


Fig. 5 Genotype-phenotype association network for cardiovascular diseases category**.** Association networks created for cardiovascular diseases category from the evidences obtained by **a**) last stage of the workflow, **b**) systematic review and **c**) exhaustive review. The "MeSH"-based categories include each pathophysiological evidence shared by two or more genes. Other heart diseases include heart arrest, cardiomyopathies, myocardial ischemia or cardiomegaly

auxiliary subunit, hence it influences but not directly causes the disease. In fact, it has been found to be associated with various inherited arrhythmia syndromes (Brugada syndrome, long-QT syndrome type 3, progressive cardiac conduction defect, sick sinus node syndrome, atrial fibrillation, and dilated cardiomyopathy) [34]. For nervous system diseases, DAVID search (Fig. 6a) provided many more phenotypic connections among genes than systematic review (Fig. 6b) or exhaustive review (Fig. 6c), which obtained the same amount of information. In fact, we could observe that the only gene with a lack of disease association is the SCN4B which, as mentioned above, is associated with cardiovascular diseases only. Finally, for mental and other disorders we only found phenotypic connections for the most relevant genes in DAVID (Fig. 7a), but not through systematic review (Fig. 7b) nor exhaustive review (Fig. 7c).

Discussion

In the present study, we addressed the prediction of the most relevant genes in the context of a group of pathologies not necessarily homogeneous but linked by a common term, as is the case of channelopathies. The identification of those genes may present several shortcomings: 1) finding key genes through scientific literature might be a burdensome task due to the fuzzy and textual nature of information, 2) completely objective criteria are hard to define, and 3) the comparison and validation of different search methodologies might not be objectively carried out. To tackle limitation 1), we developed an integrative methodology using a workflow which departs from genes linked to particular diseases. Then we built a protein-protein interaction network from which key genes are identified through the determination of the centrality measures. Finally, we proceeded to functionally annotate these key genes through the application of widely used data analysis tools in the bibliography.

Although the proposed methodology is of general purpose, in this study it was applied to the set of diseases termed channelopathies. In this clinical context, our method allowed the identification of the most relevant genes (with the highest degree of intermediation and centrality) related to channelopathies. The products of these genes are mostly channels of two different types, namely voltage-gated sodium channels — SCN1A,

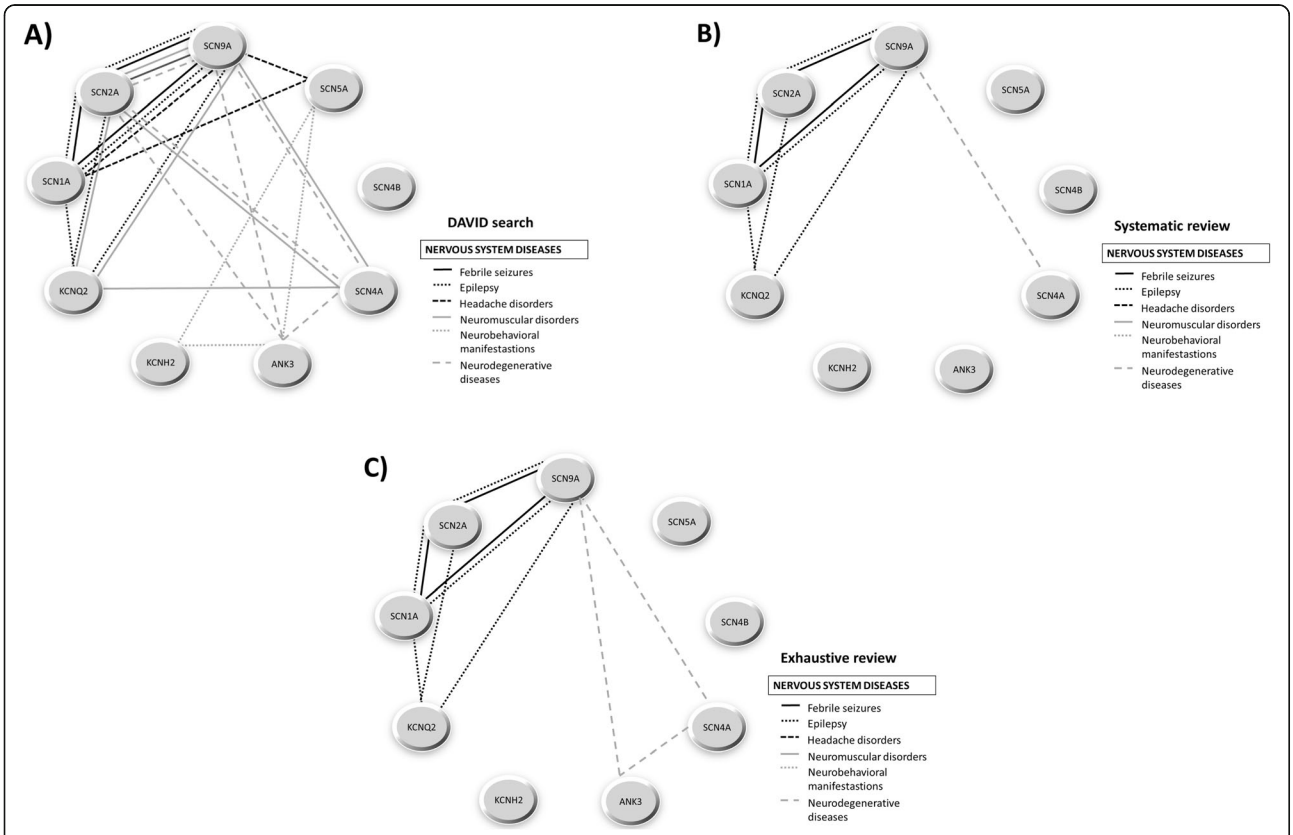


Fig. 6 Genotype-phenotype association network for nervous system diseases category**.** Association networks created for nervous system diseases category from the evidences obtained by **a**) last stage of the workflow, **b**) systematic review and **c**) exhaustive review. The "MeSH"-based categories comprise each pathophysiological evidence shared by two or more genes

SCN2A, SCN4A, SCN4B, SCN5A, and SCN9A — that are involved in the rapid depolarisation in the cardiac conduction (Reactome ID: R-HSA-5576892, Table 4.1.6 in Additional file 4), and voltage-gated potassium channels — KCNQ2 and KCNH2 — responsible for the activation of the voltage-gated potassium channels family in the neuronal system (Reactome ID: R-HSA-1296072, Table 4.1.6. in Additional file 4) [35–37]. KCNH2 is also involved in the rapid repolarisation of the cardiac conduction (Reactome ID: R-HSA-5576890, Table 4.1.6. in Additional file 4). On the other hand, Ankyrin-G (ANK3) is a protein which deals with the vesicle-mediated transport of the membrane trafficking (Reactome ID: R-HSA-374562, Table 4.1.6. in Additional file 4) and is also responsible for linking integral membrane proteins such as the voltage-gated sodium channel with the spectrin-based membrane skeleton [38]. Particularly, all the genes except KCNH2 contribute to the interaction between cytoskeleton adaptor ankyrins and a type of adhesion receptor (L1) which inhibits the nerve growth at the neural development pathway (Reactome ID: R-HSA-445095, Table 4.1.6. in Additional file 4) [35, 39].

Defects in the ion channels throughout the human body have been involved in a wide phenotypic variability in channelopathies. This remarkable causal heterogeneity makes the diseases hard to classify [40]. Some reviews deal with the categorization of channelopathies based on the organ system with which they are mainly associated in both clinical and pathophysiological aspects [28, 40–43]. Other reviews opt to classify channelopathies according to the ion channel proteins in order to improve the understanding of how their specific mutations can be linked to diseases [27, 44–46]. In current reviews the implication of voltage-gated sodium channels with cardiac pathologies (such as long-QT syndrome and fatal arrhythmias) and epilepsies is easily retrievable [27]. The role of some voltage-gated potassium channels with cardiac pathologies (heart arrhythmias, dilated cardiomyopathies), epilepsies and chronic pain is also well studied [27]. On the contrary, we do not know much about the clustering of Ankyrin-G at the axonal initial segments in the nervous system with voltage-gated sodium channels [47, 48] and some potassium channels [49]. In our work we found this implication of voltage-gated sodium and potassium channels in cardiovascular diseases (SCN2A-SCN9A-KCNH2 cluster for vascular diseases, SCN2A-SCN5A-KCNH2 cluster for cardiac arrhythmias and SCN5A-SCN4B-KCNH2

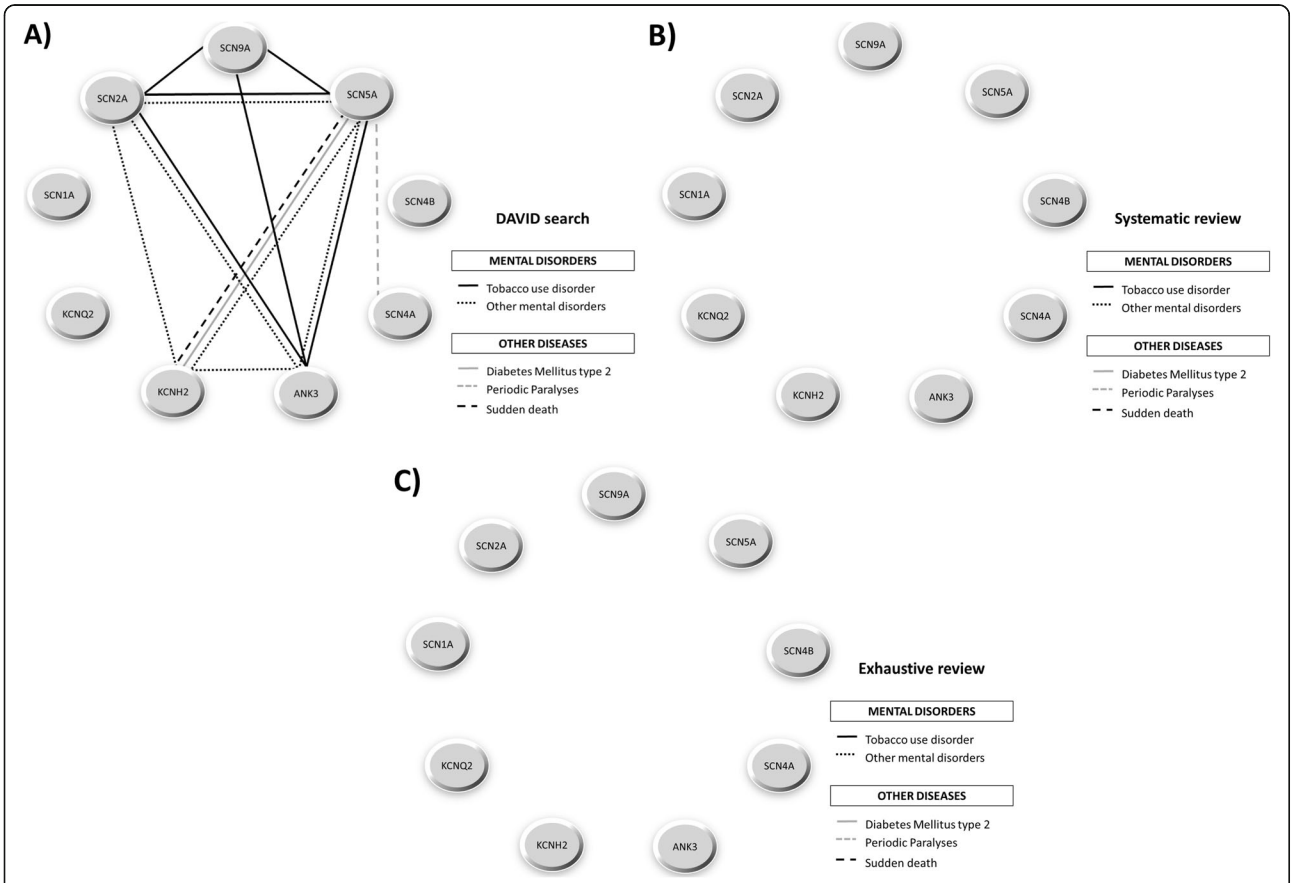


Fig. 7 Genotype-phenotype association network for mental and other diseases categories**.** Association networks created for mental and other diseases categories from the evidences obtained by **a**) last stage of the workflow, **b**) systematic review and **c**) exhaustive review. The "MeSH"-based categories comprise each pathophysiological evidence shared by two or more genes. Other mental disorders include bipolar disorder, Alzheimer disease, autism, depression or schizophrenia

cluster for other heart diseases) (Fig. 5). We also discovered a very high interconnection and participation of the genes selected not only in epilepsies, but also in febrile seizures, headache disorders, neuromuscular and neurodegenerative diseases and neurobehavioral manifestations (Fig. 6). It is interesting to highlight that in our results the above mentioned participation of Ankyrin-G in the nervous system (Fig. 6) is also reflected, specifically in neurobehavioral manifestations (ANK3-SCN5A-KCNH2 cluster) and neurodegenerative diseases (ANK3-SCN2A-SCN4A-SCN9A cluster). Finally, our results showed the implication of the genes obtained in other types of diseases, such as tobacco use disorder, diabetes mellitus type 2 or sudden death (Fig. 7), which consequently means the involvement of these genes in other systems, such as the immunological system [50] or the endocrine system [40]. As discussed above, we found that these results corroborate the conclusions collected by current literature about channelopathies, even outcomes which are not retrievable in comparative terms with respect to other traditional literature mining.

Approaching the above-mentioned validation of the proposed methodology by statistical comparison with other extant methods would be difficult due to their very different nature and properties. For that reason, we compared our proposal with two traditional and widespread family of methods, these being systematic review and exhaustive review. Among the three methods employed, our workflow and the systematic review proved to be the most objective approach when compared to the exhaustive review. Our results indicate that our methodology is actually able to find more correlations among the nine genes selected than any of the other two methods. Particularly, the present approach allows the detection of many more correlations than the systematic review (as seen in Figs. 5, 6 and 7).

Therefore, the proposed methodology is able to gather as much significant information as any other traditional literature search system mentioned in this work. At the same time, it was shown to work more flexibly, making it a convenient and easy-to-perform first-level approach compared to the above-mentioned methods.

Conclusion

We showed the usefulness of a semi-automatic integrative workflow with regard to successful, currently available mining databases and platforms based on protein-protein interaction networks applied to channelopathies. This workflow builds as productive results as a non-automatic research but in a quicker way, functioning as a bridge-builder among fields and allowing the extraction of information which *a priori* might not seem relevant when the starting point is a very large group of genes in disease. We encourage future line of research to focus on the full automatization of the workflow and the use of more specific statistical resources such as principal component analysis or machine learning classifiers.

Methods

In this section, we present the semi-automatic workflow (Fig. 1) and describe the current systems biology tools and processes used. Thus, the course of action runs as follows: first, a gene dataset of disease under study is extracted; second, a protein-protein interaction network is built and analysed and the most significant genes in disease are selected; third, the functional annotation for each relevant gene is performed.

Semi-automatic workflow

Gene dataset of the disease under study

In the first step of the workflow (Fig. 1, Stage 1), the "MeSH" term [33] of the disease at issue was obtained to know the unequivocal medical concept and introduced in Phenopedia [30]. Phenopedia is an online tool provided by the Center for Disease Control and Prevention (CDC) which allows linking genomic discoveries with health care and disease prevention. Through Phenopedia we extracted the list of genes which have been demonstrated to be involved in the disease so far.

Identification of the most relevant genes

The next step (Fig. 1, Stage 2) consisted in the generation of a protein-protein interaction (PPI) network from the list of genes through STRING [22]. We considered *Homo sapiens* as the target organism and extracted the PPI network. Then, the NetworkAnalizer available in Cytoscape [23] allows to compute and analyse a comprehensive set of topological parameters. The most highly connected proteins with a central role in the network are three times more likely to be essential than those with peripheral role, while at the same time being more associated with alterations that have a primary role in the development of diseases [51]. The identification of relevant genes in a disease has been addressed using two centrality parameters for the detection of the central nodes which maintain the structure and information fluxes into the functional network [17, 52, 53]. The network centrality features considered in the proposed workflow are degree and betweenness, two fundamental parameters in graph theory [17, 51–53]. Centrality degree is defined as the number of interactions in which a protein is involved. Betweenness is the number of shortest paths between all pairs of other proteins that pass through a certain protein [52, 53]. We set a threshold on both centrality parameters by their means and, after sorting them, those gene expressions exceeding this threshold were selected as the most relevant genes (Additional file 2).

Gene functional annotation

The last stage employed DAVID search (Fig. 1 Stage 3) for the functional annotation of genes, allowing the description of their main biological processes and the development of a functional enrichment analysis (providing information about Gene Ontology, protein interactions, functional protein domains, diseases associations, and signalling pathways, among others) from a list of genes (as official gene symbols) for the target organism *Homo sapiens.*

Validation of the workflow

Genotype-phenotype relationships of genes were obtained through the classification of the pathophysiological manifestations and diseases associated to the genes at issue. For the validation of the present workflow, we mapped those genotype-phenotype relationships of the genes obtained from the functional annotation onto phenotypic networks. We considered specific-domain category "health disorder" as our choice of interest from all the functional annotation results. This category was taken from the Medical Subject Headings ("MeSH"), a terminological database that captures biomedical information through ontological hierarchies [33]. MeSH offers a hierarchical organization of different pathological categories of every clinical manifestation that facilitates the representation of genotype-phenotype relationships. The pathophysiological implications shared by the most significant genes can thus be easily identified by means of their grade of intermediation and interaction.

Hence, we could compare the results of this workflow with the clinical manifestations associated to these genes through the use of two current traditional bibliographic search systems, systematic and exhaustive reviews (and other phenotype-oriented resources, Additional file 9). We followed the guidelines of the International Union of Pharmacology [54, 55] for the gene products nomenclature. Then we created genotype-phenotype association networks for each disease to clearly illustrate their relationships, helping visualize at a glance the different phenotypes found for every gene and thus to be able to validate the efficiency in the extraction of significant information by the presented methodology. Those diseases with no more than one gene associated were purposefully omitted in the network.

DAVID bases its disease annotation search on two human gene databases: Online Mendelian Inheritance in Man (OMIM, URL: /www.omim.org/) and Genetic Association Database from complex diseases and disorders (GAD DISEASE, URL: /geneticassociationdb.nih.gov/). The systematic review was performed using databases that focus on the relationships found between human genotypes and phenotypes of genetic alterations. Web resources for data presented herein are Online Mendelian Inheritance in Man (OMIM, URL: www.omim.org/); Gene, which integrates information about phenotypes and associated conditions (URL: www.ncbi.nlm.nih.gov/gene/); and MedGene, which offers search results about human medical genetics and conditions related to the genetic contribution (URL: www.ncbi.nlm.nih.gov/medgen/) (Fig. 4). The exhaustive review is sometimes an evidence-based review, more extensive and also takes much more time and significant effort than the systematic review, making it a tedious process in terms of filtering and selection of information. It is usually carried out by using a search equation with key words defining an unspecific question of interest [10].

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12859-019-3162-1.

Additional file 1. Protein-protein interaction (PPI) network obtained from the list of gene names involved in channelopathies. Each network node represents the protein produced by each single, protein-coding gene locus (Image generated by STRING). All the nodes are coloured to show that they are the query proteins used as input for the STRING platform. The nodes which are filled represent that some 3D structure is known or predicted; empty nodes do not present any 3D structure discovered as yet. The edges indicate protein-protein associations (full legend available in STRING).

Additional file 2. Connectivity statistics calculated from the PPI network. Raw statistics values of the two main centrality measures (degree and betweenness) are considered in Stage 2 of the workflow. Other connectivity features (closeness, Eigenvector and radiality) are included as evidence of the efficiency of the workflow and robustness of the results. The same nine genes identified as the most relevant are obtained from the average calculation of all these features. This intersection was represented in Fig. 3. HLA proteins were discarded due to their disconnection from the principal component, as shown in Additional file 1 and Fig. 2.

Additional file 3. Description of the upper-level and lower-level categories selected for the creation of "MeSH"-based terminological networks. Hierarchical trees of the upper-level categories (cardiovascular diseases, nervous system diseases, mental diseases, and other diseases) are described in detail. Lower-level categories are based on disease evidences obtained through DAVID search, systematic review and exhaustive review. The categories selected for the genotype-phenotype representations are highlighted in grey.

Additional file 4. Functional annotation results obtained through Stage 3 of the workflow. Section 1 refers to all the raw functional annotation results of the most relevant genes in channelopathies directly extracted from the last stage of the workflow (DAVID search). Section 2 refers to the extraction of the diseases from this source of information. The diseases then were classified by their lower-level categories according to the "MeSH" criterion described in methods. Some evidences could not be classified due to lack of enough information. These categories will allow the visual representation of genotype-phenotype associations obtained through DAVID search.

Additional file 5. Procedure and data extracted through the systematic review of the most relevant genes in channelopathies. Detailed procedure for the filtering and extraction of the relevant information of each gene and its diseases involved in channelopathies after the application of the systematic review.

Additional file 6. Dataset of genotype-phenotype relationships found through systematic review of the most relevant genes in channelopathies. Diseases related to the nine relevant genes through the systematic review based on three databases (Gene, OMIM and MedGen). Each phenotype is classified according to its MeSH category, as described in methods.

Additional file 7. Dataset of genotype-phenotype relationships found through exhaustive review of the most relevant genes in channelopathies. Diseases linked to the nine relevant genes through the exhaustive review. Each phenotype is classified according to its MeSH category, as described in methods. Some evidences cannot be classified due to lack of information.

Additional file 8. Further functional annotation results after the exhaustive review of the most relevant genes in channelopathies. Summary of the localization, distribution and functions of the genes after the exhaustive review, as well as a summary of further information found in this review. CNS: Central Nervous System; PNS: Peripheral Nervous System; DRG: dorsal root ganglion.

Additional file 9 Quantitative validation by significance analysis of DAVID search against other phenotype-oriented resources. We searched the nine relevant genes resulted from the workflow in PheGenI [56], ToppGene [57] and g:Profiler [58]. We quantitatively evaluated this search selecting those terms with a significance less than 0.05 using Benjamini-Hochberg FDR statistic. We obtained a minor result in DAVID search (OMIM search did not offer the phenotypes *p*-values, unlike GAP DISEASE database). Even so, results are useful to develop a quantitative comparison between semiautomatic platforms and bibliographic search systems (sheet 1). From these results we represented the genotype-phenotype association networks to compare easily each p-value phenotype obtained (sheet 2). It should be noted that p-values of clinical phenotypes could be only obtained from one of the two databases explored through DAVID (GAP DISEASE database), and so the genotype-phenotype association network is sparser than the network of the manuscript (section A in Figs. 5, 6, 7). Yet, it is demonstrated that the workflow results are statistically significant and are as valid as or even better than systematic or exhaustive reviews. Then, we created three Boolean tables (in sheets 3, 4, 5) comparing each phenotype obtained from each search; these tables were then converted to binary matrices and clustering multivariate statistical analyses and bootstrap validations were carried out. This approach demonstrated that the results provided in the manuscript, obtained from DAVID (DAVID_m) and systematic and exhaustive reviews, clustered together in a robust and significant way (sheets 3, 4, 5). Hence, this workflow builds as productive results as a non-automatic research but in a quicker way allowing the extraction of information which *a priori* might not seem relevant when the starting point is a very large group of genes in disease. Moreover, the results obtained using just significant FDR corrected *p*-values also cluster in particular branches.

Additional file 10. Protein-protein interaction (PPI) network of channelopathies analysed in Stage 2 represented by other layouts. The PPI network of channelopathies is represented as a circular layout with the betweenness attribute (Fig. 2), or with the degree attribute in the first page of this file. We also included this network with a hierarchical layout in the second page of this file as other type of representation of the same dense network. Both types of representations present each node with undirected edges. The node size marks the level of degree and therefore of neighbourhood (the larger nodes represent proteins with a higher number of interactions). The node colour shows the level of betweenness and therefore the level of centrality. HLA proteins are discarded due to their disconnection from the principal component. The images were generated by Cytoscape [23].

Abbreviations
BioGRID: Biological General Repository for Interaction Datasets;
DAVID: Database for Annotation, Visualization and Integrated Discovery;
HuGE: Human Genome Epidemiology; MeSH: Medical Subject Headings;
OMIM: Online Mendelian Inheritance in Man; PPI: Protein-protein interaction

Acknowledgments
We thank N.J. Fernández-Martínez for his linguistic assistance and the anonymous reviewers for their helpful comments and suggestions.

Authors' contributions
Study design: MM, MJS, ER; Literature and Database search: MM, MJS; Network analysis: MM, FJE; Analysis of results: MM, MJS, FJE; Wrote the paper: MM, HR. All authors have read and approved the final manuscript. This work is part of the results obtained in MM's PhD thesis.

Funding
This work was supported by funds from MINECO-FEDER (TIN2016–81041-R to E.R.), European Human Brain Project SGA2 (H2020-RIA 785907 to M.J.S.), Junta de Andalucía (BIO-302 to F.J.E.) and MEIC (Systems Medicine Excellence Network, SAF2015–70270-REDT to F.J.E.). The aforementioned bodies funded the research work described through scientific grants covering general and personnel costs. They did not play any direct role in the design of the study, collection, analysis nor interpretation of data in writing the manuscript.

Availability of data and materials
All data analysed during this study are included in this published article [and its supplementary information files].

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Computer Architecture and Technology – CITIC, University of Granada, Granada, Spain. [2]Systems Biology Unit, Department of Experimental Biology, University of Jaén, Jaén, Spain. [3]Department of Biochemistry and Molecular Biology I, University of Granada, Granada, Spain.

Received: 29 March 2019 Accepted: 15 October 2019
Published online: 12 November 2019

References
1. Diaz-Beltran L, Cano C, Wall D, Esteban FJ. Systems biology as a comparative approach to understand complex gene expression in neurological diseases. Behav Sci (Basel). 2013;3:253–72. https://doi.org/10.3390/bs3020253 .
2. Costa FF. Big data in biomedicine. Drug Discov Today. 2014;19:433–40. https://doi.org/10.1016/j.drudis.2013.10.012.
3. Diez D, Agustí A, Wheelock CE. Network analysis in the investigation of chronic respiratory diseases. From basics to application. Am J Respir Crit Care Med. 2014;190:981–8. https://doi.org/10.1164/rccm.201403-0421PP.
4. Coveney PV, Dougherty ER, Highfield RR. Big data need big theory too. In: Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences; 2016. p. 20160153. https://doi.org/10.1098/rsta.2016.0153.
5. Peek N, Holmes JH, Sun J. Technical Challenges for Big Data in Biomedicine and Health: Data Sources, Infrastructure, and Analytics. Yearb Med Inform. 2014;23:42–7. https://doi.org/10.15265/IY-2014-0018.
6. Martin-Sanchez F, Verspoor K. Big Data in Medicine Is Driving Big Changes. Yearb Med Inform. 2014;23:14–20. https://doi.org/10.15265/IY-2014-0020.
7. Gao E, Jiang Y, Li Z, Xue D, Zhang W. Association between high mobility group box-1 protein expression and cell death in acute pancreatitis. Mol Med Rep. 2017;15:4021–6. https://doi.org/10.3892/mmr.2017.6496.
8. Wu D, Rice CM, Wang X. Cancer bioinformatics: a new approach to systems clinical medicine. BMC Bioinformatics. 2012;13:71. https://doi.org/10.1186/1471-2105-13-71.
9. Sánchez-Valle J, Tejero H, Ibáñez K, Portero JL, Krallinger M, Al-Shahrour F, et al. A molecular hypothesis to explain direct and inverse co-morbidities between Alzheimer's disease, Glioblastoma and Lung cancer. Sci Rep. 2017;7:4474. https://doi.org/10.1038/s41598-017-04400-6.
10. Ferreira González I, Urrútia G, Alonso-Coello P. Systematic Reviews and Meta-Analysis: Scientific Rationale and Interpretation. Rev Española Cardiol (English Ed). 2011;64:688–96. https://doi.org/10.1016/j.rec.2011.03.027.
11. Mosca E, Bertoli G, Piscitelli E, Vilardo L, Reinbold RA, Zucchi I, et al. Identification of functionally related genes using data mining and data integration: a breast cancer case study. BMC Bioinformatics. 2009;10:S8. https://doi.org/10.1186/1471-2105-10-S12-S8.
12. Cervantes-Gracia K, Husi H. Integrative analysis of multiple sclerosis using a systems biology approach. Sci Rep. 2018;8:5633. https://doi.org/10.1038/s41598-018-24032-8.
13. Zou X-D, An K, Wu Y-D, Ye Z-Q. PPI network analyses of human WD40 protein family systematically reveal their tendency to assemble complexes and facilitate the complex predictions. BMC Syst Biol. 2018;12:41. https://doi.org/10.1186/s12918-018-0567-9.
14. Safari-Alighiarloo N, Taghizadeh M, Tabatabaei SM, Shahsavari S, Namaki S, Khodakarim S, et al. Identification of new key genes for type 1 diabetes through construction and analysis of protein-protein interaction networks based on blood and pancreatic islet transcriptomes. J Diabetes. 2017;9:764–77. https://doi.org/10.1111/1753-0407.12483.
15. Quan Z, Quan Y, Wei B, Fang D, Yu W, Jia H, et al. Protein-protein interaction network and mechanism analysis in ischemic stroke. Mol Med Rep. 2015;11:29–36. https://doi.org/10.3892/mmr.2014.2696.
16. Wuchty S, Almaas E. Peeling the yeast protein network. Proteomics. 2005;5:444–9. https://doi.org/10.1002/pmic.200400962.
17. Goñi J, Esteban FJ, de Mendizábal NV, Sepulcre J, Ardanza-Trevijano S, Agirrezabal I, et al. A computational analysis of protein-protein interaction networks in neurodegenerative diseases. BMC Syst Biol. 2008;2:52. https://doi.org/10.1186/1752-0509-2-52.
18. Di Silvestre D, Brambilla F, Scardoni G, Brunetti P, Motta S, Matteucci M, et al. Proteomics-based network analysis characterizes biological processes and pathways activated by preconditioned mesenchymal stem cells in cardiac repair mechanisms. Biochim Biophys Acta Gen Subj. 2017;1861:1190–9. https://doi.org/10.1016/j.bbagen.2017.02.006.
19. Stoilova-McPhie S, Ali S, Laezza F. Protein-protein interactions as new targets for Ion Channel drug discovery. Austin J Pharmacol Ther. 2013;1:1–6.
20. Oughtred R, Chatr-aryamontri A, Breitkreutz B-J, Chang CS, Rust JM, Theesfeld CL, et al. BioGRID | Database of Protein, Chemical, and Genetic Interactions. Cold Spring Harb Protoc. 2016, 2016:pdb.prot088880. https://doi.org/10.1101/pdb.prot088880.
21. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, et al. ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res. 2012;40:W597–603. https://doi.org/10.1093/nar/gks400.
22. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic Acids Res. 2017;45:D362–8. https://doi.org/10.1093/nar/gkw937.
23. Shannon P. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504. https://doi.org/10.1101/gr.1239303.
24. Huang DW, Sherman BT, Lempicki RA, Huang DW, Sherman BTLRA. DAVID Functional Annotation Bioinformatics Microarray Analysis. Nat Protoc. 2009;4:44–57 doi:4(1):44–57.
25. Gomollón F, Quintero R, Bastidas A, Ilzarbe D, Pintor L, Ilzarbe L, et al. Inflammatory bowel disease and eating disorders: a systematized review of comorbidity. J Psychosom Res. 2017;102:47–53. https://doi.org/10.1016/j.jpsychores.2017.09.006.
26. Yang C, Li C, Wang Q, Chung D, Zhao H. Implications of pleiotropy: challenges and opportunities for mining big data in biomedicine. Front Genet. 2015;6:229. https://doi.org/10.3389/fgene.2015.00229.
27. Musgaard M, Paramo T, Domicevica L, Andersen OJ, Biggin PC. Insights into channel dysfunction from modelling and molecular dynamics simulations. Neuropharmacology. 2018;132:20–30. https://doi.org/10.1016/j.neuropharm.2017.06.030.
28. Spillane J, Kullmann DM, Hanna MG. Genetic neurological channelopathies: molecular genetics and clinical phenotypes. J Neurol Neurosurg Psychiatry. 2016;87:37–48.

29. Schorge S. Channelopathies go above and beyond the channels. Neuropharmacology. 2018;132:1–2. https://doi.org/10.1016/j.neuropharm.2018.02.011.
30. Yu W, Clyne M, Khoury MJ, Gwinn M. Phenopedia and Genopedia: disease-centered and gene-centered views of the evolving knowledge of human genetic associations. Bioinformatics. 2010;26:145–6. https://doi.org/10.1093/bioinformatics/btp618.
31. Joy MP, Brock A, Ingber DE, Huang S. High-Betweenness proteins in the yeast protein interaction network. J Biomed Biotechnol. 2005;2005:96–103. https://doi.org/10.1155/JBB.2005.96.
32. VIB / UGent, Bioinformatics & Evolutionary Genomics. Draw Venn Diagram. http://bioinformatics.psb.ugent.be/webtools/Venn/. Accessed 18 Nov 2018.
33. Medical Subject Headings - MeSH. https://www.nlm.nih.gov/mesh/. Accessed 8 Jun 2017.
34. Zeng Z, Zhou J, Hou Y, Liang X, Zhang Z, Xu X, et al. Electrophysiological characteristics of a SCN5A voltage sensors mutation R1629Q associated with Brugada syndrome. PLoS One. 2013;8:e78382. https://doi.org/10.1371/journal.pone.0078382.
35. Reactome Pathway Database. http://www.reactome.org/. Accessed 5 Jun 2017.
36. McKeown L, Swanton L, Robinson P, Jones OT. Surface expression and distribution of voltage-gated potassium channels in neurons (review). Mol Membr Biol. 2008;25:332–43. https://doi.org/10.1080/09687680801992470.
37. Pongs O, Schwarz JR. Ancillary subunits associated with voltage-dependent K + channels. Physiol Rev. 2010;90:755–96. https://doi.org/10.1152/physrev.00020.2009.
38. Kapfhamer D, Miller DE, Lambert S, Bennett V, Glover TW, Burmeister M. Chromosomal localization of the AnkyrinG gene (ANK3/Ank3) to human 10q21 and mouse 10. Genomics. 1995;27:189–91. https://doi.org/10.1006/geno.1995.1023.
39. Whittard JD, Sakurai T, Cassella MR, Gazdoiu M, Felsenfeld DP. MAP kinase pathway–dependent phosphorylation of the L1-CAM Ankyrin binding site regulates neuronal growth. Mol Biol Cell. 2006;17:2696–706. https://doi.org/10.1091/mbc.e06-01-0090.
40. Kim J. Channelopathies. Korean J Pediatr. 2014;57:1. https://doi.org/10.3345/kjp.2014.57.1.1.
41. Hubner CA. Ion channel diseases. Hum Mol Genet. 2002;11:2435–45. https://doi.org/10.1093/hmg/11.20.2435.
42. Kullmann DM. Neurological Channelopathies. Annu Rev Neurosci. 2010;33:151–72. https://doi.org/10.1146/annurev-neuro-060909-153122.
43. Cannon SC. Pathomechanisms in channelopathies of skeletal muscle and brain. Annu Rev Neurosci. 2006;29:387–415. https://doi.org/10.1146/annurev.neuro.29.051605.112815.
44. Brouwer BA, Merkies ISJ, Gerrits MM, Waxman SG, Hoeijmakers JGJ, Faber CG. Painful neuropathies: the emerging role of sodium channelopathies. J Peripher Nerv Syst. 2014;19:53–65. https://doi.org/10.1111/jns5.12071.
45. Baruscotti M, Bottelli G, Milanesi R, DiFrancesco JC, DiFrancesco D. HCN-related channelopathies. Pflügers Arch Eur J Physiol. 2010;460:405–15. https://doi.org/10.1007/s00424-010-0810-8.
46. Bidaud I, Lory P. Hallmarks of the channelopathies associated with L-type calcium channels : A focus on the Timothy mutations in Cav1.2 channels. Biochimie. 2011;93:2080–6. https://doi.org/10.1016/j.biochi.2011.05.015.
47. Mohler PJ, Rivolta I, Napolitano C, LeMaillet G, Lambert S, Priori SG, et al. Nav1.5 E1053K mutation causing Brugada syndrome blocks binding to ankyrin-G and expression of Nav1.5 on the surface of cardiomyocytes. Proc Natl Acad Sci. 2004;101:17533–8. https://doi.org/10.1073/pnas.0403711101.
48. Kretschmer T, England JD, Happel LT, Liu ZP, Thouron CL, Nguyen DH, et al. Ankyrin G and voltage gated sodium channels colocalize in human neuroma – key proteins of membrane remodeling after axonal injury. Neurosci Lett. 2002;323:151–5. https://doi.org/10.1016/S0304-3940(02)00021-6.
49. Xu M, Cooper EC. An Ankyrin-G N-terminal gate and protein kinase CK2 dually regulate binding of voltage-gated sodium and KCNQ2/3 potassium channels. J Biol Chem. 2015;290:16619–32. https://doi.org/10.1074/jbc.M115.638932.
50. RamaKrishnan AM, Sankaranarayanan K. Understanding autoimmunity: the ion channel perspective. Autoimmun Rev. 2016;15:585–620. https://doi.org/10.1016/j.autrev.2016.02.004.
51. Jeong H, Mason SP, Barabási A-L, Oltvai ZN. Lethality and centrality in protein networks. Nature. 2001;411:41–2. https://doi.org/10.1038/35075138.
52. Newman MEJ. A measure of betweenness centrality based on random walks. Soc Networks. 2005;27:39–54. https://doi.org/10.1016/j.socnet.2004.11.009.
53. Alvarez-Ponce D, Feyertag F, Chakraborty S. Position matters: network centrality considerably impacts rates of protein evolution in the human protein–protein interaction network. Genome Biol Evol. 2017;9:1742–56. https://doi.org/10.1093/gbe/evx117.
54. Catterall WA. International Union of Pharmacology. XLVII. Nomenclature and structure-function relationships of voltage-gated sodium channels. Pharmacol Rev. 2005;57:397–409. https://doi.org/10.1124/pr.57.4.4.
55. Gutman GA. International Union of Pharmacology. LIII. Nomenclature and molecular relationships of voltage-gated potassium channels. Pharmacol Rev. 2005;57:473–508. https://doi.org/10.1124/pr.57.4.10.
56. Ramos EM, Hoffman D, Junkins HA, Maglott D, Phan L, Sherry ST, et al. Phenotype-genotype integrator (PheGenI): synthesizing genome-wide association study (GWAS) data with existing genomic resources. Eur J Hum Genet 2014. Accessed 18 Nov 2018.
57. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res. 2009;37(Web Server issue):W305–11.
58. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. 2019;47(W1):W191–8.

Publisher's Note

2. Optimization of efficient neuron models with realistic firing dynamics. The case of the cerebellar granule cell

Authors	**M. Marín***, M. J. Sáez-Lara, E. Ros & J. A. Garrido*		
Journal	*Frontiers in Cellular Neuroscience*		
Year	2020		
Volume	14		
Pages	161		
DOI	10.3389/fncel.2020.00161		
Impact Factor (JCR 2019)	3.921		
Category	Neurosciences:	Ranking: 93/272	**(Q2)**

frontiers in Cellular Neuroscience

ORIGINAL RESEARCH
published: 14 July 2020
doi: 10.3389/fncel.2020.00161

Optimization of Efficient Neuron Models With Realistic Firing Dynamics. The Case of the Cerebellar Granule Cell

Milagros Marín[1,2], María José Sáez-Lara[2], Eduardo Ros[1] and Jesús A. Garrido[1]**

[1]Department of Computer Architecture and Technology—CITIC, University of Granada, Granada, Spain, [2]Department of Biochemistry and Molecular Biology I, University of Granada, Granada, Spain

OPEN ACCESS

Edited by:
Egidio D'Angelo,
University of Pavia, Italy

Reviewed by:
Laurens Bosman,
Erasmus Medical Center,
Netherlands
William Martin Connelly,
University of Tasmania, Australia

***Correspondence:**
Milagros Marín
mmarin@ugr.es
Jesús A. Garrido
jesusgarrido@ugr.es

Specialty section:
This article was submitted to Cellular Neurophysiology, a section of the journal Frontiers in Cellular Neuroscience

Received: 10 December 2019
Accepted: 13 May 2020
Published: 14 July 2020

Citation:
Marín M, Sáez-Lara MJ, Ros E and Garrido JA (2020) Optimization of Efficient Neuron Models With Realistic Firing Dynamics. The Case of the Cerebellar Granule Cell. Front. Cell. Neurosci. 14:161. doi: 10.3389/fncel.2020.00161

Biologically relevant large-scale computational models currently represent one of the main methods in neuroscience for studying information processing primitives of brain areas. However, biologically realistic neuron models tend to be computationally heavy and thus prevent these models from being part of brain-area models including thousands or even millions of neurons. The cerebellar input layer represents a canonical example of large scale networks. In particular, the cerebellar granule cells, the most numerous cells in the whole mammalian brain, have been proposed as playing a pivotal role in the creation of somato-sensorial information representations. Enhanced burst frequency (spiking resonance) in the granule cells has been proposed as facilitating the input signal transmission at the theta-frequency band (4–12 Hz), but the functional role of this cell feature in the operation of the granular layer remains largely unclear. This study aims to develop a methodological pipeline for creating neuron models that maintain biological realism and computational efficiency whilst capturing essential aspects of single-neuron processing. Therefore, we selected a light computational neuron model template (the adaptive-exponential integrate-and-fire model), whose parameters were progressively refined using an automatic parameter tuning with evolutionary algorithms (EAs). The resulting point-neuron models are suitable for reproducing the main firing properties of a realistic granule cell from electrophysiological measurements, including the spiking resonance at the theta-frequency band, repetitive firing according to a specified intensity-frequency (I-F) curve and delayed firing under current-pulse stimulation. Interestingly, the proposed model also reproduced some other emergent properties (namely, silent at rest, rheobase and negligible adaptation under depolarizing currents) even though these properties were not set in the EA as a target in the fitness function (FF), proving that these features are compatible even in computationally simple models. The proposed

Abbreviations: AdEx, Adaptive exponential integrate-and-fire; AP, Action potential; EA, Evolutionary algorithm; FF, Fitness function; GrC, Granule cell; GrL, Granular layer; HH, Hodgkin-and-Huxley; I-F, Intensity-frequency; I-V, Intensity-voltage.

methodology represents a valuable tool for adjusting AdEx models according to a FF defined in the spiking regime and based on biological data. These models are appropriate for future research of the functional implication of bursting resonance at the theta band in large-scale granular layer network models.

Keywords: neuron model, granule cell, cerebellum, model simplification, spiking resonance, point neuron, adaptive exponential integrate-and-fire

INTRODUCTION

Neuronal populations in the brain reflect complex synchronized temporal patterns typically modulated by coherent oscillations (Buzsáki, 2006). This oscillatory behavior is usually evidenced by the study of resonance as the preferred frequency in response to oscillatory inputs (Hutcheon and Yarom, 2000). In particular, one of the brain centers where resonance has received more attention is the cerebellum (Dugué et al., 2009; D'Angelo et al., 2009, 2011; Gandolfi et al., 2013). The cerebellum is thought to generate low-frequency (5–30 Hz) and higher-frequency activity rhythms, depending on the circuit sections or the neurons involved (D'Angelo et al., 2009; Dugué et al., 2009). Previous findings suggest that theta-frequency activity (around 4–10 Hz in rodents) contributes to signal integration in the cerebellum (Gandolfi et al., 2013), but its function for overall cerebellar information processing remains elusive.

The cerebellar granular layer (GrL) represents one of the main inputs to the cerebellar cortex and low-frequency rhythms at this layer is fundamental for motor control, learning, and sleep (Buzsáki, 2006; D'Angelo et al., 2009; Wang et al., 2019). Most studies have focused on subthreshold (membrane potential oscillations) resonance. In particular, *in vivo* studies of cerebellar GrL evidenced theta-frequency resonance at 7 Hz in rats (Hartmann and Bower, 1998) and 7–25 Hz in monkeys (Pellerin and Lamarre, 1997; Courtemanche et al., 2009). However, much less attention has been paid to the suprathreshold (spiking) resonance (Rotstein, 2017). The spiking resonance has been proposed to strengthen input signal processing and data transmission at the theta-frequency band in the GrL (D'Angelo et al., 2001, 2009). In most cases, this feature depends on the spiking mechanisms and the intrinsic properties of single cells (Rotstein, 2017).

Single-neuron responses in the GrL have long been investigated in search of theta-frequency activity patterns (Ros et al., 2009; Gandolfi et al., 2013). Spiking resonance has been claimed to be an intrinsic property of the cerebellar granule cells (GrCs), the most abundant cells not only in the cerebellum but also in the whole mammalian brain (Herculano-Houzel, 2010). Although many experimental studies have registered the electrophysiological activity of single GrCs from rat cerebellar recordings, both from slices *in vitro* (Brickley et al., 2001; Diwakar et al., 2009; Osorio et al., 2010; Delvendahl et al., 2015; Masoli et al., 2017) and *in vivo* (Chadderton et al., 2004; Jörntell and Ekerot, 2006), they have traditionally neglected the presence of spiking resonance. However, only *in vitro* recordings have reported spiking resonance (as enhanced bursting activity) at theta-frequency band of single cerebellar GrCs in response to low-frequency sinusoidal stimulation (D'Angelo et al., 2001; Gandolfi et al., 2013). According to these studies, the spiking resonance could emerge from an intrinsic property of the neurons that selectively enhance low-frequency stimulation responses due to a combination of passive and active membrane properties (Hutcheon and Yarom, 2000; Magistretti et al., 2006; Das and Narayanan, 2017). However, the functional role of resonance at the theta band in the processing of the cerebellar GrCs remains largely unclear.

Computational modeling has demonstrated to be an effective strategy in exploring the origin of resonant behavior in the GrCs. Detailed models (i.e., integrating a high degree of biological plausibility) allowed fine-grained studies about the intrinsic mechanisms involved at isolated GrCs (D'Angelo et al., 2001). Additionally, a conductance-based Hodgkin-and-Huxley (HH) mono-compartmental GrC model evidenced that the subthreshold voltage-dependent potassium current (I_{KSlow}) is at the core of the intrinsic resonance during sinusoidal stimulation (Nieus et al., 2006; Solinas et al., 2010; Gandolfi et al., 2013; Rössert et al., 2014; Masoli et al., 2017). However, the high computational cost associated to the simulation of this type of detailed model makes them only suitable for small scale models of the GrL network or short simulations (Nieus et al., 2006; Diwakar et al., 2009; Solinas et al., 2010; Gandolfi et al., 2013).

Thus, simplified models appear to be an exceptional alternative for exploring the functional role of resonant activity in information processing. Simplified models combine computational efficiency and realistic neuronal dynamics. Considering this, the adaptive exponential integrate-and-fire (AdEx) model (Brette and Gerstner, 2005) only includes two coupled differential equations that capture adaptation and resonance properties (Naud et al., 2008), while enabling large scale implementations of neuronal circuits. Although the AdEx model can be seen as a two-dimensional reduction of the spike initiation in HH models, the specific parameter values of the model configuration to match with electrophysiological measurements (Jolivet et al., 2008; Hanuschkin et al., 2010; Barranca et al., 2013; Venkadesh et al., 2018) cannot be experimentally determined as they require an automatic parameter tuning algorithm.

In this article, we present a methodology for the development of simplified neuron models based on the AdEx generic model template that consider both biological relevance and computational efficiency. Evolutionary algorithms (EAs) have been used to find suitable sets of parameters to capture specific

firing dynamics. The application to the use case of cerebellar GrC models allows the replication of the most essential properties of the biological cell that are key for the frequency and timing of firing patterns in the neural code. We particularly focus on the spiking resonance of bursts in the theta-frequency band that has been experimentally evidenced in previous studies in the literature. We also address how the inclusion of different spiking properties in the fitness function (FF) affects the behavior of the optimized neuron configuration.

MATERIALS AND METHODS

Neuron Model

The proposed mathematical model of the cerebellar GrC aims to maintain biological realism (to capture important aspects of single-neuron processing) as well as a low computational cost. We have selected the AdEx neuron model (Brette and Gerstner, 2005) as the generic template model. Since GrCs have a compact and simple morphology (D'Angelo et al., 1995, 2001; Delvendahl et al., 2015), a mono-compartment model, such as an AdEx point neuron model, represents a reasonable approach. Previous studies have addressed how this model can be tuned to capture biological realism and compared to more detailed models (Brette and Gerstner, 2005; Nair et al., 2015) as well as recordings in pyramidal neurons, in which this model has been demonstrated to fit, at least qualitatively, a rich set of observed firing patterns (Brette and Gerstner, 2005; Jolivet et al., 2008; Naud et al., 2008).

The AdEx model accounts for only two coupled differential equations and a reset condition regulating two state variables, the membrane potential (V) and the adaptation current (w), according to the following equations:

$$C_m \frac{dV}{dt} = -g_L(V - E_L) + g_L \Delta_T \exp\left(\frac{V - V_{\mathrm{T}}}{\Delta_{\mathrm{T}}}\right) + I(t) - w \quad (1)$$

$$\tau_w \frac{dw}{dt} = a(V - E_L) - w \quad (2)$$

Equation (1) describes the evolution of the membrane potential (V) during the injection of the current [$I(t)$]. When the membrane potential is driven beyond the threshold potential (V_T), then the exponential term of the slope factor (Δ_T) models the action potential (AP). This depolarization ends when the membrane potential reaches the reset threshold potential (V_{peak}). Then, the membrane potential (V) is instantaneously reset to V_r and the adaptation current (w) is increased a fixed amount (b).

The first term in equation (1) models the passive membrane mechanisms dependent on the total leak conductance (g_L), the leak reversal potential (E_L), and the membrane capacitance (C_m), all regulating the integrative properties of the neuron. The second (exponential) term represents the activation of the sodium channel in a Hodgkin-Huxley type neuron model (Naud et al., 2008), whose dynamics are determined by the parameters Δ_T and V_T. Equation (2) describes the evolution of the recovery variable (w). It depends on the adaptation time constant parameter (τ_w) and the subthreshold adaptation (a), while (b) defines the spike-triggered adaptation. In our simulations, the refractory period (τ_{ref}) was set to 1 ms. The membrane potential was initially set to the same value as the leak reversal potential ($V_{init} = E_L$).

To sum up, 10 parameters define the dynamics of the AdEx neuron model that need to be tuned to reproduce the firing properties of the cerebellar GrCs.

Model Optimization With Evolutionary Algorithms

Our optimization method is based on an EA that allows multiple parameter exploration to fit the experimentally recorded firing behavior (Jolivet et al., 2008; Hanuschkin et al., 2010; Barranca et al., 2013; Venkadesh et al., 2018). After the execution of the EA, it provides sets of parameters that minimize the FF, i.e., the function which associates each parameter set with a single value quantifying the goodness of such a neuron configuration. Our FF (score) includes a weighted sum of specific features related to spike firing that we consider biologically relevant, according to equation (3).

$$score = \sum_{\mathrm{i}=1}^{n} [abs(feat_i - \exp_i) \cdot w_i] \quad (3)$$

The score is defined as the sum of every firing pattern feature (i) in response to the corresponding experimental stimulation protocols. The score of each feature is calculated as the absolute value (abs) of the difference between the feature value extracted from the simulated neuron trace with the parameter configuration of the individual ($feat_i$) and the feature value extracted from the experimental recordings ($\exp_i$). This is multiplied by the weight associated with each feature (w_i; see "Simulations" section below). The EA will perform progressive parameter optimization to select the individual (set of parameters) that minimizes the fitness value (score). Thus, each individual represents a set of GrC model parameters and the FF quantifies the similarity between the firing pattern in the simulated neuron model and the experimental recording of the neuron in response to the same stimulation protocols.

We also explore an alternative method which aims to evaluate the variability of the burst frequency over successive oscillatory cycles. The score of the burst frequency has been complemented with an additional multiplicative term related to the standard deviation of the burst frequency over consecutive oscillatory cycles, according to equation (4). By using this method as the FF of the burst frequency feature, the EA will prefer neuron model configurations whose burst frequency not only keep close to the target (experimentally recorded) value but are also stable over oscillatory cycles.

$$score_{\mathrm{BF}} = \sum_{j=1}^{N} [abs(\overline{BF_{sim_{\mathrm{j}}}} - \overline{BF_{\exp_{\mathrm{j}}}}) \cdot w_{\mathrm{BF}} \cdot (std(BF_{sim_{\mathrm{j}}}) + 1)] \quad (4)$$

According to this formula, the score of the burst frequency feature ($score_{\mathrm{BF}}$) is defined as the sum of each score for all the sinusoidal stimulation frequencies (N; 14 sinusoidal frequencies, see **Table 3**). The individual score for each stimulation frequency is calculated as the absolute difference (abs) between the burst frequency (averaged over 10 oscillatory cycles) of the simulated

TABLE 1 | Parameters boundaries used for the neuron optimization procedures.

Parameter name (unit)	Fixed boundaries	Parameter name (unit)	Fixed boundaries
C_m (pF)	[0.1, 5.0]	V_T (mV)	[−60, −20]
Δ_T (mV)	[1, 1000]	a (nS)	[−1, 1]
E_L (mV)	[−80, −40]	b (nA)	[−1, 1]
V_r (mV)	[−80, −40]	g_L (nS)	[0.001, 10.0]
V_{peak} (ms)	[−20, 20]	τ_w (ms)	[1, 1000]

The minimum and maximum values of the parameters are indicated in square brackets.

neuron ($\overline{BF_{sim_j}}$) and the experimental value at that stimulation frequency ($\overline{BF_{\exp_j}}$). This term is multiplied by the weight of the burst frequency feature (w_{BF}) and the standard deviation of the simulated neuron burst frequency [$std(BF_{sim_j})$] plus one.

For every execution, the EA runs for 50 generations and 1,000 individuals in the population. The initial generation is set with 1,000 simulated neurons with parameters created according to a uniform distribution ranging between the boundaries indicated in **Table 1**. During each generation of the EA, each model configuration (individual) is simulated and ranked according to the FF (equation 3). The next generation is created using basic genetic operators, such as crossover and mutation. The one-point crossover operator was used with a 60% probability and the uniform mutation operator was used with a 10% probability. In those individuals randomly selected to mutate, each parameter was mutated with a probability of 15%. To select those individuals to be included in the population in the next generation, the selection was carried out by three-individual tournaments. Therefore, the new population is composed of the winners (minimum score) resulting from 1,000 tournaments with randomly-chosen individuals. Finally, the individual with the minimum score obtained during each complete EA execution is selected as the best neuron model (the most suitable parameter configuration to the target behavior).

Fitness Functions and Feature Quantification

Biological Data Used as Reference

The experimental data used as a reference for EA optimization are taken from two different sources (D'Angelo et al., 2001; Masoli et al., 2017). In particular, the burst frequency in response to sinusoidal current stimulation is obtained from D'Angelo et al. (2001). The authors recorded cerebellar GrCs in acute cerebellar slices obtained from 20 ± 2-day-old rats. The slice preparation and whole-cell patch-clamp were performed as reported previously (see their references). The presence of bicuculline prevented GrC rhythmic inhibition by Golgi cells and that spontaneous EPSPs were too rare to affect spike generation. Injection of sinusoidal currents at various frequencies (0.5–40 Hz) revealed resonance in burst spike frequency in correspondence with the positive phase of the stimulus. Spike frequency within bursts increased and then decreased according to the injected current frequency showing spiking resonance. The preferred frequency was 6 Hz with sinusoidal currents of 6-pA amplitude, and 8 Hz with 8-pA amplitude (reference dots in **Figures 1B**, **2B**, **4A**, **6**). It is important to highlight that in these *in vitro* recordings, the burst frequencies with stimulation frequencies beyond 10.19 Hz in 6-pA amplitude and 14.23 Hz in 8-pA amplitude fell to zero as one or no spikes were obtained.

On the other hand, the average firing rate and first-spike latency in response to current pulses were obtained from Masoli et al. (2017). In this case, the authors performed *in vitro* patch-clamp recordings under step current injections. They recorded cerebellar GrCs in acute cerebellar slices from 21-day-old rats.

EA Fitness Functions (FFs)

The FF described in equation (3) weights the similarity of different quantified features with the experimental recordings (i.e., the value of the feature extracted from the traces). Since our models aim to reproduce the spiking resonance of bursting, it is required to estimate this resonance as a set of values. Thus, the burst frequency under sinusoidal current injection was calculated as the inverse of the average inter-stimulus interval (ISI) of the output neuron (the cerebellar GrC) during each stimulation cycle. Then, the average burst frequency was measured throughout 10 consecutive cycles of sinusoidal stimulation. The total simulation time was set to 22.5 s. Sinusoidal amplitude values of 6 pA and 8 pA (in addition to 12-pA offset) were used according to the available experimental data (specified in the subsection above), generating spike bursts in correspondence with the positive phase of the stimulus (sinusoidal phase of 270°). To reproduce the differential effect of the oscillatory stimulation frequency, the burst frequency in response to its stimulation frequency was included in the EA as an individual feature (all of them equally weighted). As it occurred in the *in vitro* recordings [see "Biological Data Used as Reference" section above], we have set the burst frequency to zero when the same firing pattern (one or no spike per cycle) has been obtained in the simulated neurons.

Although the firing dynamics of cerebellar GrCs are complex, these cells implement a mechanism of linear frequency encoding through repetitive firing discharge under current stimulation which might help to sustain the spiking resonance of burst frequency at the theta-frequency band (D'Angelo et al., 2001, 2009). Recent literature has characterized the fast repetitive discharge in the GrCs based on the mean frequency (the number of spikes divided by the stimulation time) and the latency to first spike (time of the first spike firing) in response to three different step-current injections (10 pA, 16 pA, and 22 pA) of 1-s stimulation (Masoli et al., 2017).

The optimizations were carried out with FFs that considered different combinations of the minimal number of features that

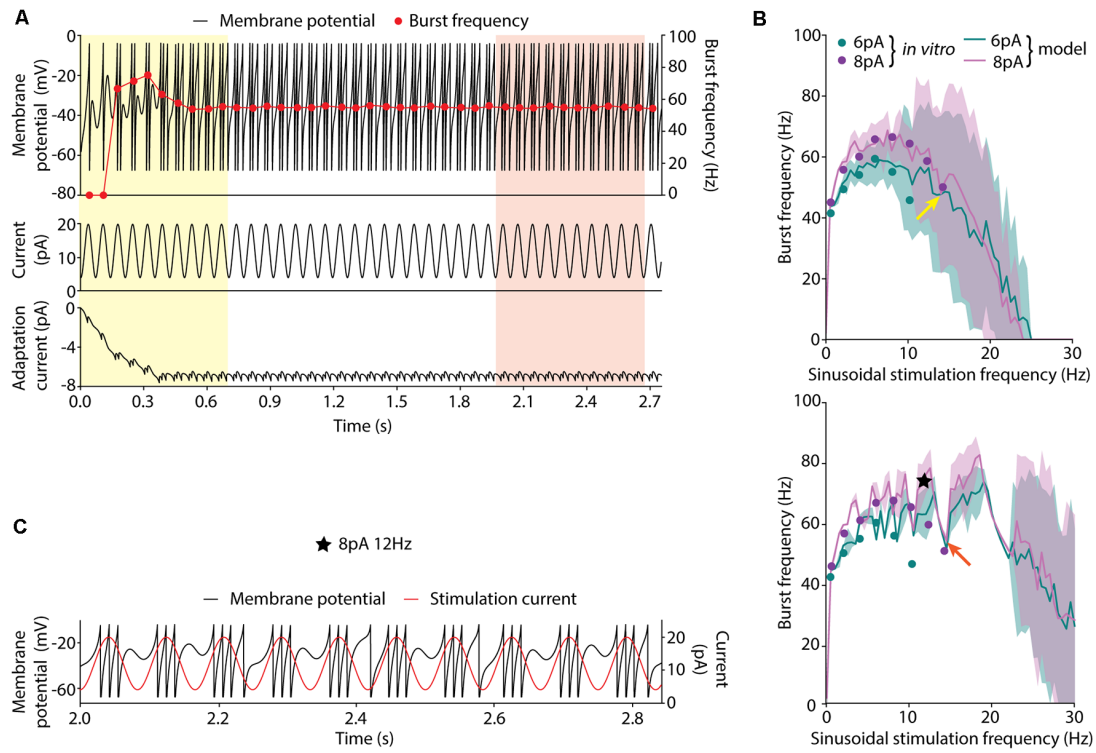


FIGURE 1 | Burst frequency analysis. **(A)** Membrane potential (*V*; top, black line, left axis), burst frequencies (top, red line, right axis), stimulation current (middle) and adaptation current (*w*; bottom) of a neuron model instance stimulated with the sinusoidal current of 8-pA amplitude, 12-pA offset current and 14.23-Hz frequency. Two strategies for measuring the average burst frequency are shown: (i) during the initial 10 cycles (yellow shadow); and (ii) 10 cycles after 2 s of initial stabilization (red shadow). **(B)** Burst frequency in response to sinusoidal stimulation with different stimulation frequencies (in steps of 0.5 Hz) up to 30 Hz, 12-pA offset current and 6-pA (green), or 8-pA (purple) amplitude measured during the initial cycles (top) and after 2-s stabilization period (bottom). The solid lines represent the average burst frequency from the simulated neuron model configuration and the shaded area shows the standard deviation. The arrows indicate the average burst frequency corresponding to the stimulation frequency shown in **(A)** during the initial cycles (yellow arrow, 49.74 ± 25.81 Hz) and after 2 s of stabilization (red arrow, 55.04 ± 0.38 Hz). The dots correspond to the burst frequency data obtained from *in vitro* recordings of cerebellar GrCs in D'Angelo et al. (2001) with the same stimulation protocols (**Table 2**). The black star indicates the point further explored in plot C. **(C)** Simulated membrane potential (*V*; black line, left axis) during 10 oscillatory cycles selected after 2-s initial stabilization in response to 12-Hz sinusoidal stimulation. The red line (right axis) represents the stimulation current signal (8-pA amplitude and 12-pA offset).

characterize the typical firing of cerebellar GrCs: (1) burst frequency in response to different sinusoidal current stimulations (stimulation at different frequencies of the sinusoidal current); (2) burst frequency feature (as in point 1) in addition to the average mean frequency in response to step-currents; (3) the burst frequency (as in point 1) and the latency to the first spike under step-current stimulations; and (4) all the previous features (burst frequency under sinusoidal stimulation, mean frequency and latency to the first spike under current stimulations; **Table 3**). Later on, we will refer to these combinations of features as FFs from 1 to 4.

Simulations

The score of each individual approaches zero as the measured firing features approximate target values. We ran each optimization protocol (EA algorithm) with five different seeds and selected the individual with a minimal score from those executions. The weight of the burst frequency and the mean frequency features were set to 1 (as they both were measured in Hz and present values in comparable scales). The latency to the first spike feature was weighted to 1,000 as it was measured in seconds. Thus, our algorithm equally weights 1 Hz-error in the average mean frequency feature and 1 ms-lag in the latency to the first spike feature.

The EA algorithm was implemented using the DEAP library (Fortin et al., 2012) for Python (version 2.7.12). The GrC model was simulated using NEST (version 2.14.0; Peyser et al., 2017). The model uses the embedded 4th order Runge-Kutta-Fehlberg solver with adaptive step-size to integrate the differential equations. The simulations were run in parallel with SCOOP on a 6-cores 3.30 GHz CPU (32 GB RAM) PC allowing each optimization protocol to run (five simulations with different seeds) in around 7 h.

RESULTS

Bursting Frequency Optimization

We conducted preliminary experimentation to determine the best strategy to measure the burst frequency in response

TABLE 2 | Feature and scores obtained with simulated neurons after EA optimization with different fitness functions.

	Burst frequency 6 pA (Hz)					Burst frequency 8 pA (Hz)				
		Simulation					Simulation			
Sinusoidal Stim. Freq. (Hz)	**Experimental**	***FF1***	***FF2***	***FF3***	***FF4***	**Experimental**	***FF1***	***FF2***	***FF3***	***FF4***
0.58	41.43	36.77	**37.66**	35.73	35.19	45.00	**45.62**	42.63	45.78	42.68
2.12	49.29	47.36	46.29	**47.52**	46.15	55.71	55.42	**55.75**	56.70	53.97
4.04	54.00	51.78	**52.82**	51.81	50.74	60.00	59.03	61.01	**59.84**	60.39
5.96	59.29	**55.22**	54.32	52.36	53.28	65.71	63.16	**65.57**	63.83	63.07
8.08	55.00	**55.04**	53.93	55.25	54.74	66.43	64.43	**66.23**	64.94	64.52
10.19	45.71	50.51	57.97	**50.48**	55.25	64.29	69.44	68.94	69.93	**67.57**
12.31	-	-	-	-	-	58.57	66.23	**58.62**	66.94	66.01
14.23	-	-	-	-	-	50.00	49.30	71.43	**50.02**	51.74
Score		**20.71**	26.26	**21.61**	28.45		**19.94**	29.90	**19.33**	21.45

	Mean frequency (Hz)					First spike latency (ms)				
		Simulation					Simulation			
Step-current amp. (pA)	**Experimental**	***FF1***	***FF2***	***FF3***	***FF4***	**Experimental**	***FF1***	***FF2***	***FF3***	***FF4***
10	30	(1)	**30**	(2)	19	31.90	(45.8)	(9.9)	**36.10**	14.90
16	45	(35)	49	(35)	**45**	19.00	(12.8)	(6.4)	**12.40**	9.00
22	60	(72)	67	(73)	**66**	14.65	(8.5)	(5.0)	**8.40**	6.70
Score		(51)	**11**	(51)	17		(26.25)	(44.25)	**17.05**	34.95

The score corresponding to different features (burst frequency, IF curve, and latency to the first spike) of the best-performing individuals are shown against the target experimental values. Mean frequency and latency to the first spike experimental values were extracted from Masoli et al. (2017), and experimental recordings of burst frequency from D'Angelo et al. (2001). Note that, although the standard deviation of the burst frequency is considered by the EAs, they are not included in the burst frequency scores to obtain an overall view of the distance between the simulated and experimental values of every feature. The values in brackets correspond to the features not included in the FFs (and are just evaluated for comparison). In bold are the closest values and minimal scores.

TABLE 3 | Parameter values of the best-performing neuron models.

Parameter name (unit)	***FF1***	***FF2***	***FF3***	***FF4***
C_m (pF)	3.10	4.21	3.36	2.80
Δ_T (mV)	5.42	1.09	7.01	22.07
E_L (mV)	−64.06	−51.42	−59.92	−58.00
V_{peak} (mV)	−13.49	6.80	−12.24	−17.56
V_r (mV)	−70.28	−73.66	−64.86	−71.31
V_T (mV)	−40.59	−38.00	−40.31	−24.01
a (nS)	0.26	0.36	0.36	0.23
b (nA)	0.19	0.65	0.15	0.37
g_L (nS)	0.49	0.17	0.67	0.25
τ_w (ms)	327.25	338.75	365.41	619.07

Parameter values of the individuals resulting from the optimization process with different FFs. Membrane capacitance (C_m), slope factor (Δ_T), leak reversal potential (E_L), reset value for membrane potential after a spike (V_r, map to resting potential V_{init}), spike detection threshold (V_{peak}), spike initiation threshold (V_T), subthreshold adaptation (a), spike-triggered adaptation (b), leak conductance (g_L) and adaptation time constant (τ_w).

to sinusoidal stimulation. As a first approach, we calculated the average burst frequency during the 10 initial cycles of the simulation (as explained in the Methods section). The EA was set to minimize only the error of the average burst frequency in response to all the available data. The resulting neuron model configuration (individual) showed high instability (i.e., highly variable burst frequency) during the initial cycles (**Figure 1A**, yellow shaded area in the top plot). The same optimization was carried out with five different random seeds and all the individual winners showed similar behavior. Particularly, it can be observed that in response to high stimulation frequencies (namely, 10–14 Hz), the burst frequency remained unsteady for eight oscillatory cycles (**Figure 1A**, red dots in the top plot). This observation can be explained based on: (i) the model configuration emerging from the EA combines high membrane capacity (C_m ranging between 2.63 and 4.87 pF), low leakage conductance (g_L ranging between 0.49 and 4.31 nS) and low initial membrane voltage (V_{init} equal to E_L as specified in the "Materials and Methods" section; both ranging between −58.18 and −49.88 mV), so that the membrane potential required between cycles 1 and 4 until stable values were reached for several consecutive cycles (**Figure 1A**, yellow shaded area in the top plot); and (ii) the neuron configuration included long adaptation time constants (τ_w) so that the adaptation current (w) required six cycles to reach steady-state (**Figure 1A**, yellow shaded area in the bottom plot). Although the average burst frequency stays close to the experimental measures for every stimulation frequency, the standard deviation of the burst frequency over the measuring cycles is higher than desired, especially for high stimulation frequencies (**Figure 1B**, top plot).

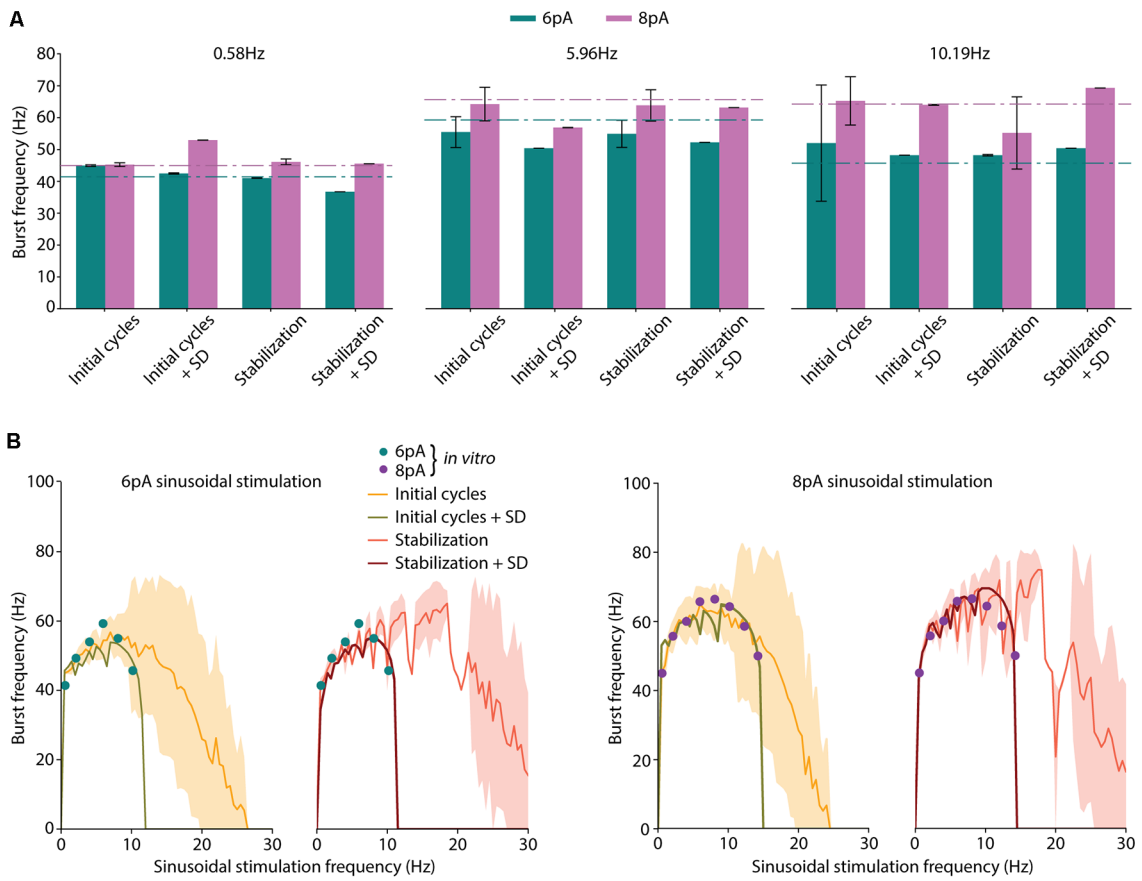


FIGURE 2 | Neuron model optimization with different strategies. **(A)** Average burst frequencies in response to low (0.58 Hz), medium (5.96 Hz), or high (10.19 Hz) sinusoidal stimulation frequencies. The mean value over 10 oscillatory cycles and the standard deviation from the best performing individual are represented for each considered fitness function (FF). Corresponding *in vitro* values from D'Angelo et al. (2001) for each stimulation frequency and amplitude (6 pA, green and 8 pA, purple) are shown as dashed-dotted lines. **(B)** Spiking resonance curves in response to 6 pA and 8 pA of sinusoidal stimulation. The stimulation frequency of the individuals ranges from 0.5 Hz to 30 Hz with 0.5-Hz steps. Standard deviation is shown shaded in yellow for individuals optimized with *initial cycles* FF, and in red for individuals optimized with the *stabilization* FF. Those individuals optimized including the standard deviation in their score (namely, *initial cycles + SD* and *stabilization + SD*) obtained near-zero standard deviation (this is indeed represented in the green and dark red shaded areas but are hardly visible). The available *in vitro* data extracted from D'Angelo et al. (2001) and specified in **Table 3** are shown with colored dots.

Aiming to overcome the instability produced during the initial period of simulation, we tested whether averaging over 10 bursts (oscillatory cycles) after 2 s of initial stabilization produced different results. This period was chosen as it corresponds to twice the maximum allowed adaptation time constant (τ_w; see **Table 1**). In this way, the neuron membrane potential reached steady state (**Figure 1A**, red shaded area at top plot) before measuring and averaging the burst frequency. Not unexpectedly, the EA set with this second estimation method resulted in neuron configurations whose average burst frequencies closely matched the experimental measures (**Figure 1B**, green and purple lines in the bottom plot). However, the standard deviation of the burst frequencies remained higher than desired (although remarkably lower than using the first estimation method) in response to high stimulation frequencies (**Figure 1B**, green, and purple shaded areas in the bottom plot). E.g., when stimulated with 14.23 Hz (**Figure 1B**, red arrow in the bottom plot) the average burst frequency is stable with almost no standard deviation (55.04 ± 0.38 Hz; **Figure 1A**). On the contrary, when stimulated with 12 Hz (**Figure 1B**, a black star at bottom plot) the simulated neuron showed an increased standard deviation of the average burst frequency (74.96 ± 5.99 Hz). It occurred because the neuron did not fully recover from one oscillatory cycle to the next one (**Figure 1C**). This situation produces enhanced variability in the burst frequency values for some neuron configurations. To prevent this issue, we set the EA with a third method for calculating the score of each individual based on the 2-s-stabilization method (described in the "Materials and Methods" section).

Four implementations of FFs were considered once defined the period considered for burst frequency calculation and the inclusion of penalization for instability: (1) average burst frequency calculation over initial cycles (shortly, *initial cycles*); (2) average burst frequency calculation over initial cycles with the penalization of the standard deviation (shortly, *initial cycles + SD*); (3) average burst frequency calculation after

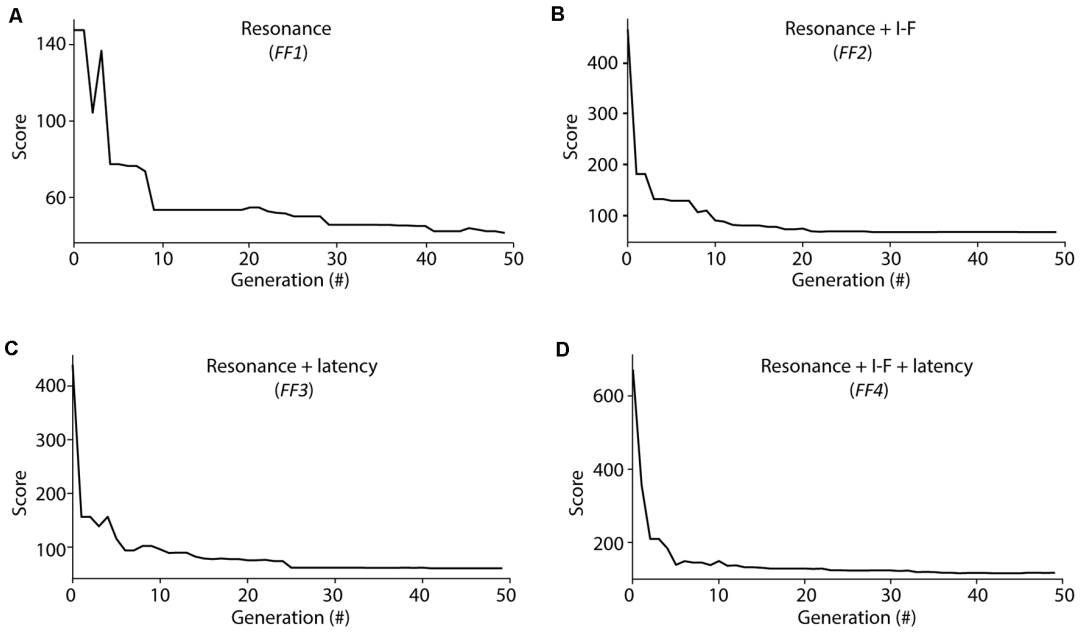


FIGURE 3 | Score evolution during Evolutionary algorithm (EA) optimization. Evolution of the minimal score from the individuals considered at each generation during the optimization processes of EAs configured with different fitness functions (FFs; as described in Methods section): **(A)** *FF1* (Resonance), **(B)** *FF2* (Resonance + I-F curves), **(C)** *FF3* (Resonance + Latency to the first spike), **(D)** *FF4* (Resonance + I-F curves + Latency to the first spike). The score was calculated as the weighted sum of the individual feature scores considered for each FF (equations 3 and 4).

2-s simulation (shortly, *stabilization*); and (4) average burst frequency calculation after 2-s stabilization period with the penalization of the standard deviation (shortly, *stabilization + SD*). The score was calculated according to equation (3) in the cases of *initial cycles* and *stabilization* (cases 1 and 3), and according to equation (4) in the cases of *initial cycles + SD* and *stabilization + SD* (cases 2 and 4).

The EA was executed five times for each case considering different random seeds to obtain the neuron model configurations that best matched the experimental values of average burst frequencies (**Table 3**). The individuals with the lowest score were selected. Not unexpectedly, all the individuals presented similar values to the experimental data (**Figure 2A**), validating the operation of the EA. In response to low (0.58 Hz) stimulation frequency, negligible standard deviations were obtained with all the considered FFs (left plot in **Figure 2A**). However, higher stimulation frequencies (i.e., 5.96 Hz and 10.19 Hz) resulted in increased standard deviations for those functions which did not include SD penalization (namely, *initial cycles* and *stabilization*; **Figure 2A**, middle and right plots, respectively).

We then compared the spiking resonance curves of the individuals obtained using each FF (**Figure 2B**). When the penalization of standard deviation is not included in the FF (namely, *initial cycles* and *stabilization*), the average burst frequencies (**Figure 2B**, yellow and orange lines) are near the experimental values (the sum of the distances between simulated and experimental burst frequency features are 19.44 Hz and 43.15 Hz in *initial cycles* and *stabilization*, respectively; **Figure 2B**, colored dots), but with large standard deviation (the sum of the SDs of burst frequency features are 104.62 Hz and 43.44 Hz in *initial cycles* and *stabilization*, respectively; **Figure 2B**, yellow and orange shadow). Additionally, resonance curves fall to zero (indicating one or zero spikes per oscillatory cycle) with remarkably higher stimulation frequencies (beyond 25 Hz), especially with the *stabilization* function. Thus, the model configuration resulting from the usage of the *initial cycles* FF appropriately reproduced spiking resonance at theta-frequency band but with considerable variability. Differently, the *stabilization* FF drove to the neuron models whose resonance peaks were beyond the theta-frequency band (around 20 Hz). This situation makes these two neuron models unsuitable for our aim.

When the penalization of standard deviation is included in the FF (namely, *initial cycles + SD* and *stabilization + SD*), average burst frequencies are also close to the experimental data (the sum of the distances between simulated and experimental burst frequency features are 50.79 Hz and 41.50 Hz in *initial cycles + SD* and *stabilization + SD*, respectively; **Figure 2B**, green and dark red lines) and they are stable, with almost negligible standard deviations (the sum of the SDs of burst frequency features are 0.69 Hz and 1.05 Hz in *initial cycles + SD* and *stabilization + SD*, respectively; **Figure 2B**, green and dark red shadow areas representing the standard deviations that are almost negligible and hardly visible in the plots). Interestingly, the neuron models resulting from these individuals show resonance curves falling to zero just above the last stimulation frequency points (10.19 Hz

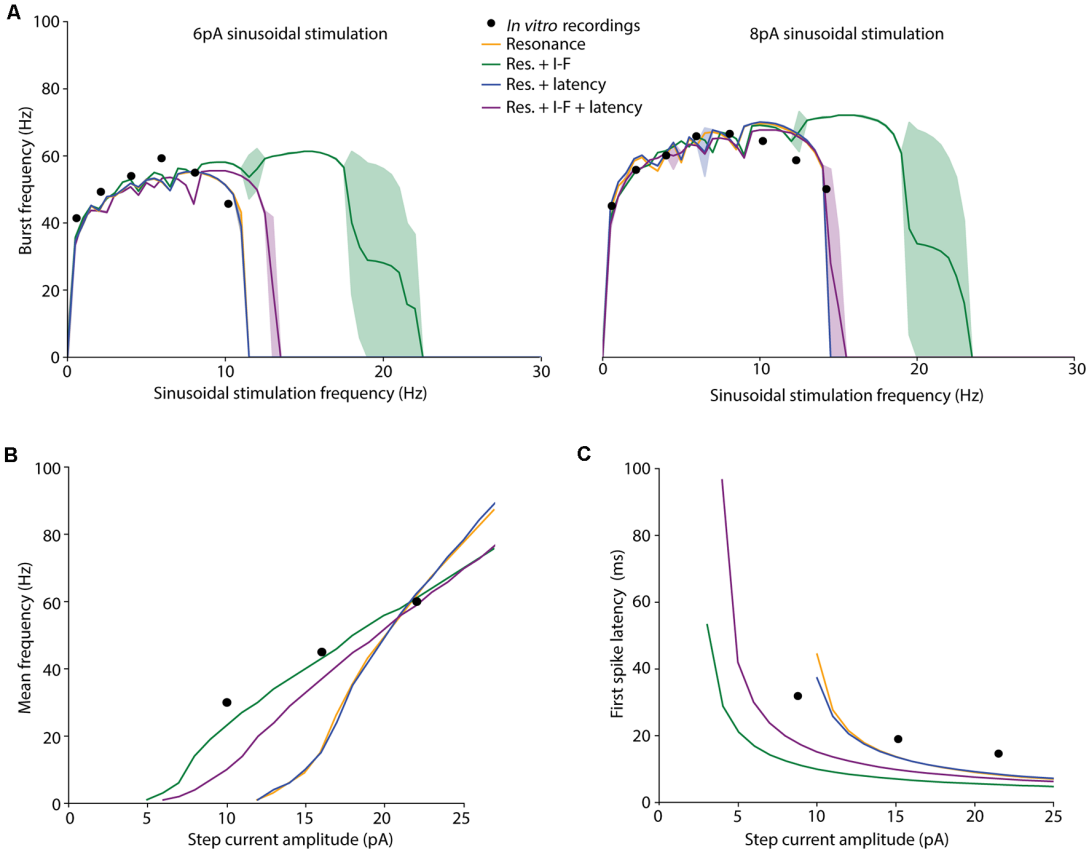


FIGURE 4 | Spiking dynamics of the individuals resulting from different EA implementations. Simulated features of the selected (best fitted) neuron models obtained from the EAs configured with different FFs: only burst frequency (yellow), burst frequency and intensity-frequency (I-F) curve (green), burst frequency and first-spike latency (blue), and burst frequency, I-F curve and first-spike latency (purple). The target *in vitro* data from D'Angelo et al. (2001) and Masoli et al. (2017); specified in **Table 3**) are represented as black dots. **(A)** Burst frequency in response to 6-pA (left) or 8-pA (right) sinusoidal stimulation with varying frequencies (in steps of 0.5 Hz). The lines represent the simulated average burst frequency over 10 cycles after 2 s of initial stabilization, and the standard deviations are shown as shaded areas. **(B)** I-F curves of the neuron models obtained from the EAs configured with different FFs. The Y-axis represents the average firing frequency in response to 1 s-step-currents. **(C)** Latency to the first spike in response to 1-s-step-currents.

at 6-pA and 14.23 Hz at 8-pA sinusoidal stimulations) as was experimentally tested in real GrCs, generating one or no spikes per cycle at higher frequency sinusoidal stimulations (**Figure 2B**, green and dark red lines; D'Angelo et al., 2001). Thus, these model configurations are considered to better reproduce the spiking resonance at the theta-frequency band.

According to these preliminary results, it is preferable to calculate the burst frequency after the 2-s-stabilization period and including the standard deviation as part of the FF (*stabilization* + *SD*; named *FF1* in the next subsection). This FF drives our EA to penalize unstable configurations, resulting in neuron model configurations that match the spiking resonance at the theta-frequency band of biological cerebellar GrCs and maintain stable neuronal behavior during the oscillatory cycles.

Parameter Fitting With Other Suprathreshold Dynamics

Once, we had explored the most convenient definition of FF for burst frequency feature optimization, we aimed to demonstrate whether additional electrophysiological properties could also be optimized and reproduced by the neuron model. Thus, we considered other representative firing properties of GrCs that are seemingly relevant in neural transmissions such as the intensity-frequency (I-F) curve and the latency to the first spike in response to different stimulation currents. We carried out additional optimization experiments with different combinations of features in the FF: burst frequency under sinusoidal stimulation (namely, *FF1*), burst frequency under sinusoidal stimulation and mean frequency (I-F) under step current injection (namely, *FF2*), burst frequency under sinusoidal stimulation and latency to the first spike in response to step current injection (namely, *FF3*), and all the three mentioned features together (namely, *FF4*).

The evolution of the minimum score of the individuals in the explored population from these EAs showed fast convergence during the optimization processes among generations (**Figure 3**). We aim to determine whether the usage of different FFs

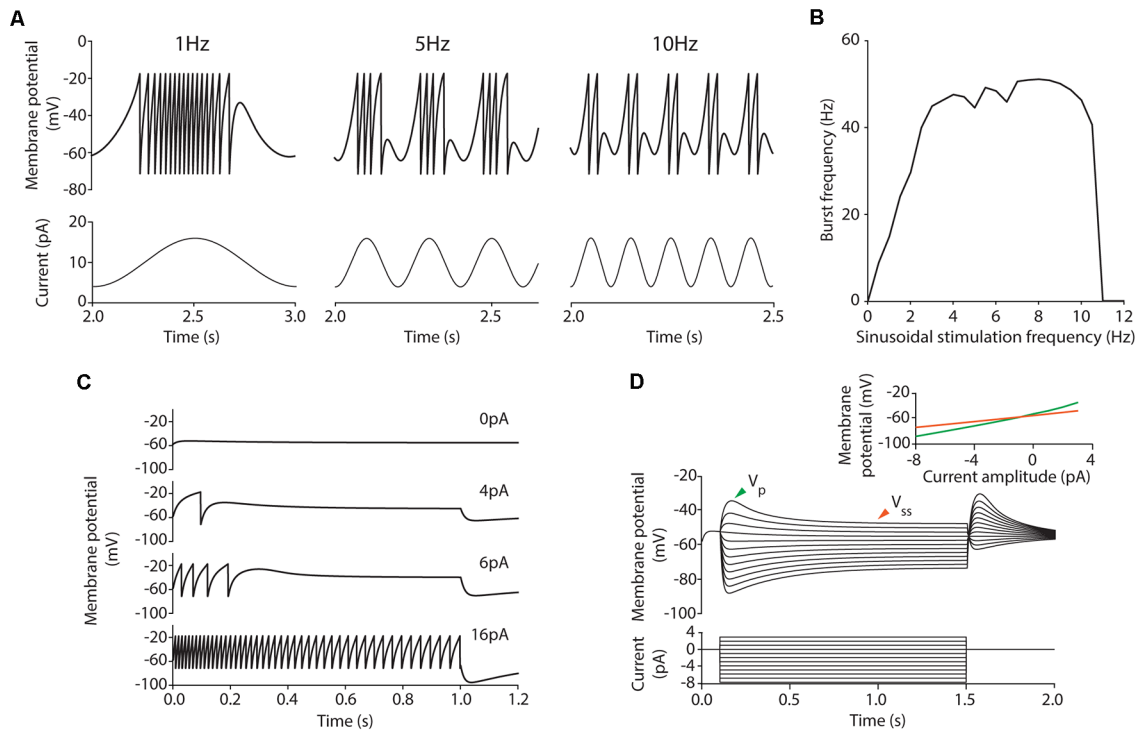


FIGURE 5 | Intrinsic properties predicted by the cerebellar granule cell (GrC) model. **(A)** Neuron model simulation in response to sinusoidal current injection of 10-pA offset and 6-pA amplitude. Bursts are shown after 2 s of stimulation (stabilization). **(B)** Resonance curve (showing burst frequency) in response to the same stimulation protocol. **(C)** Membrane potential evolution in response to 1-s step-current injections with variable amplitude. **(D)** Step pulse hyperpolarizing and depolarizing currents in the subthreshold regimen (from −8 pA to 3 pA in steps of 1 pA from 100 ms to 1500 ms) cause the membrane potential to reach an early peak followed by a decayed (sag response) at a stable level. Current-voltage relationship plots (I–V plots) from the voltage peak (sag; green colored) and from the steady-state (considered at 1,000 ms; orange-colored) demonstrate the absence of inward rectification in the model during the polarization with both hyperpolarizing and depolarizing currents.

affect the capability of the resulting neuron models to resonate in the theta-frequency band as well as determining whether the proposed AdEx model can reproduce all these different firing features in a single parameter configuration. The scores obtained from the evaluation of all the features (included or not in its EA implementation) simulated by the individuals are shown in **Table 2**, and the corresponding parameters of the best performing individuals with each EA configuration are in **Table 3**.

Spiking Resonance in the Theta-Frequency Band

Cerebellar GrCs have been demonstrated to resonate in a rather broad theta-frequency band. The spiking resonance peak has been described around 6–12 Hz (D'Angelo et al., 2001) in experimental measurements, and around 4–10 Hz in previous detailed GrC models (D'Angelo et al., 2001, 2009; Magistretti et al., 2006; Gandolfi et al., 2013; Masoli et al., 2017). The proposed EAs selected neuron models matching the burst frequency of the experimental curves when the configured FFs included only the burst frequency (*FF1*; preferred resonance frequency within 7–11 Hz), the burst frequency and the latency to the first spike (*FF3*; preferred resonance frequency within 7–11 Hz) and all the three features considered in this work (*FF4*; preferred resonance frequency within 8–12 Hz; **Figure 4A**). The simulation of the selected individuals closely fitted the experimental data with stable burst frequencies (between 0.5–1.5 Hz SD) and burst frequency falling to zero (one or zero spikes per cycle) with stimulation frequencies beyond 10.19 Hz (6-pA amplitude) and 14.23 Hz (8-pA amplitude), respectively.

The best-fitted individuals for the experimental spiking resonance were the neuron model resulting from the EA with *FF1* (the sum of the distances between simulated and experimental burst frequency features is 40.65 Hz) and the neuron model from the EA with *FF3* (the sum of the distances between simulated and experimental burst frequency features is 40.94 Hz), closely followed by the neuron model from the EA with *FF4* (the sum of the distances between simulated and experimental burst frequency features is 49.90 Hz) and, finally, the neuron model from the EA with *FF2* (the sum of the distances between simulated and experimental burst frequency features is 56.16 Hz). Not unexpectedly, the EAs with FFs which included only burst frequency features resulted in neuron models with the best fitting of the resonance to the experimental data. On the contrary, the individuals resulting from EAs with the FF that included all the features (*FF4*) showed a shifted resonance curve only with 6-pA-amplitude sinusoidal stimulation (**Table 2** and

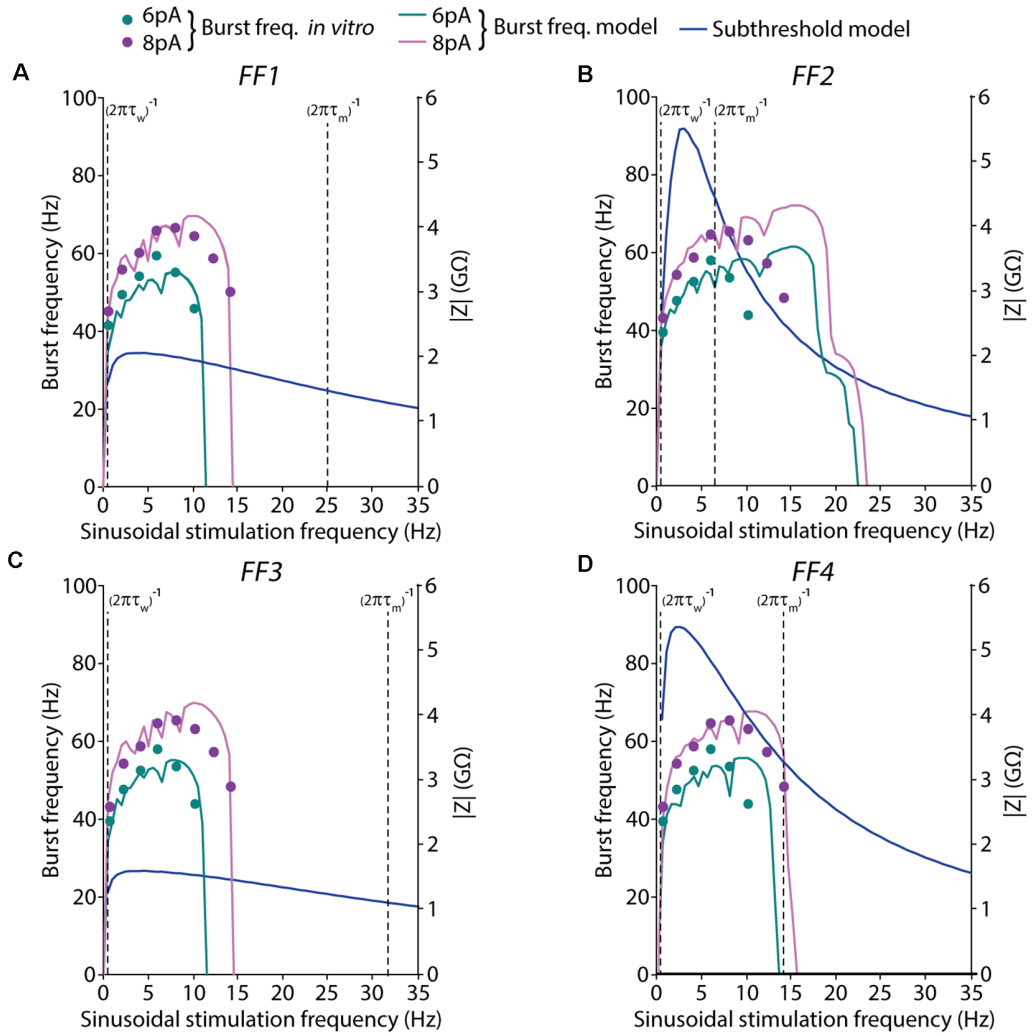


FIGURE 6 | Spiking and subthreshold resonance properties obtained with the models resulting from different EA implementations. Simulated resonance properties under suprathreshold and subthreshold regimes of the selected (best fitted) neuron models obtained from the EAs configured with different FFs: **(A)** only burst frequency, **(B)** burst frequency and I-F curve, **(C)** burst frequency and first-spike latency, and **(D)** burst frequency, I-F curve, and first-spike latency. Burst frequencies are shown in response to 6-pA (green) or 8-pA (pink) sinusoidal stimulation with varying frequencies (in steps of 0.5 Hz). The pink and green lines represent the simulated average burst frequency over 10 cycles after 2 s of initial stabilization (left axis). The target *in vitro* data, from D'Angelo et al. (2001), of the burst frequencies, are represented as dots. Subthreshold resonance properties are represented as the module of the impedances that were calculated with the FFT algorithm in response to 2-pA (blue lines) sinusoidal current stimulation with varying frequencies (in steps of 0.5 Hz; right axis). The cut-off frequency of the high-pass filter is represented as a vertical dashed line in the form of $1/(2\pi\tau_w)$, while the cut-off frequency corresponding to the low-pass filter is represented as a vertical dashed line in the form of $1/(2\pi\tau_m)$.

Figure 4A, purple line). The individuals resulting from the EAs with the FF including burst frequency and I-F curve (but not first spike latency; *FF2*) showed resonance beyond the theta range (**Figure 4A**, green-shaded lines) with unstable behavior (large standard deviations).

I-F Curve

We also evaluated the membrane voltage response while injecting step currents of increasing amplitude. Beyond specific thresholds of injected current fast repetitive firing was reproduced (**Figure 4B**). The individual resulting with *FF1* and *FF3* showed rheobases (understood as the minimum current injected required to fire a single AP) at 10 pA, the individual resulting with *FF2* at 3 pA, and the individual resulting with *FF4* at 4 pA. This is in agreement with the experimental rheobases obtained for GrCs (ranging between 2 pA and 10 pA; D'Angelo et al., 2001; Bezzi et al., 2004; Gandolfi et al., 2013; Masoli et al., 2017). The best fitting to the experimental frequency values were obtained, as expected, by the neuron models resulted from those FFs that included the I-F curve in their features to optimize (*FF2* and *FF4*; **Table 2** and in **Figure 4B**, green and purple lines).

Linear coding of stimulus intensity (I-F curve) is usually used as a measure of the intrinsic excitability of GrCs. I-F plots were constructed (using 1-s current stimulation with amplitude

ranging between the rheobase and 25 pA; **Figure 4B**) and fitted to a linear function ($r > 0.9$). The slope of such a linear function is usually representative of the intrinsic excitability of the neurons. The output frequency values of the neuron model with the FF containing all the features (*FF4*) were slightly further away from the experimental values (slightly higher score; see mean frequency in **Table 2**) than those resulting from the EA with the FF including burst frequency and mean frequency (*FF2*). However, their I-F slopes were very similar (3.83 Hz/pA and 3.39 Hz/pA, respectively) and near the slope of the experimental points used in the EAs (2.5 Hz/pA, which was calculated from the experimental frequency values in **Table 2**). The neuron models that were not optimized for this frequency (using *FF1* and *FF3*) resulted in higher scores (their firing frequencies fell far from the experimental points) and higher I-F slopes (6.27 Hz/pA and 6.36 Hz/pA, respectively; **Figure 4B**, yellow and blue lines). Despite this, their I-F slopes were coherent to those slopes reported in previous GrC models (7 Hz/pA in D'Angelo et al., 2001; Bezzi et al., 2004; Masoli et al., 2017). Both the I-F slope ranges obtained (around 3.5 Hz/pA and around 6.3 Hz/pA) are then considered biologically plausible since they fall within the experimentally recorded values (6.5 ± 3.2 Hz/pA in D'Angelo et al., 1995).

Latency to the First Spike

Another central behavior of biological GrCs is that the latency to the first spike decreases and spike frequency increases when the injected current intensity is increased (D'Angelo et al., 2001). Similar behavior is observed in the neuron models resulting when using all the proposed FFs (**Figure 4C**). Experimental *in vitro* recordings evidenced that the latency to the first spike decreased from 31.9 ± 16.2 ms with 10-pA step current to 14.65 ± 9.4 ms with 22-pA current (Masoli et al., 2017). The closest latencies to these data were obtained by the neuron model resulted from the EA with the FF that included first-spike latency in its definition (*FF3*; **Figure 4C**, blue line). The neuron model resulting from the EA with *FF1* (**Table 2** and **Figure 4C**, yellow line) obtained fitted results too. Differently, those individuals from the EA with FFs that included the step-current firing rate (*FF2* and *FF4*) generated higher latencies than those reported, mainly with low stimulation currents (**Figure 4C**, green and purple lines). The individual from the EA with *FF2* reproduced the closest fitting to the I-F curve but the least suitable to fit either the theta-frequency band or latency to the first spike. However, according to other *in vitro* recordings (D'Angelo et al., 1995, 1998; Brickley et al., 1996; Cathala et al., 2003) and computational GrC models (D'Angelo et al., 2001; Diwakar et al., 2009; Masoli et al., 2017), latencies to the first spike decreased from around 100 ms at the rheobase to around 1 ms, similar to that from the individual resulting from the EA with the *FF4* (see first spike latency in **Table 2**).

Selecting a Biologically Plausible GrC Model

Overall, the most accurate neuron model according to all the features (with the lowest sum of the distances between experimental and simulated features of burst frequency, mean frequency and first spike latency) corresponded to the individual obtained including all the features in the EA (the sum of the distances is 101.85 using *FF4*), followed by the neuron model obtained from the inclusion of the average burst frequency and latency to the first spike (the sum of the distances is 108.99 using *FF3*). The individual resulting from the FF only defined by the average burst frequency had, unexpectedly, a higher total score (the sum of the distances is 117.9 using *FF1*) than the individual from the FF of average burst frequency and I-F curve (the sum of the distances is 111.41 using *FF2*). Therefore, the simplified model configuration with the best fitting to the spiking dynamics of a real GrC is the individual resulted from the EA implementation that contains all the spiking properties (namely, *FF4*).

The behavior of this model is presented in **Figure 5**. When stimulated by just-threshold sinusoidal stimulation, the model generated spikes clustered in doublets-triplets or longer bursts (as in D'Angelo et al., 1998, 2001; Gandolfi et al., 2013; **Figure 5A**) with specific tuning in the theta frequency band (7–10 Hz; **Figure 5B**). In response to step-current stimulations, the model resulted in regular spike discharge (**Figure 5C**) with latency compatible with the experimental data in real cells. Additionally, the model exhibited other emergent properties (i.e., not selected during the EA optimization). First, the neuron is silent at rest (**Figure 5C**). When stimulated by depolarizing step-current injections, the neuron model elicited a single spike with 4 pA as in D'Angelo et al. (2001). The firing rate showed no adaptation with 0, 4, and 6 pA and little adaptation with 16 pA which is similar to the experimental recordings (Masoli et al., 2017; **Figure 5C**). However, we evaluated some other emergent properties from the subthreshold regime typical of a cerebellar GrC, such as the inward rectification (D'Angelo et al., 1995). The model did not reproduce the inward rectification during the application of current steps in the hyperpolarizing direction neither its I-V relationships (**Figure 5D**). Simulations using detailed neuron models based on *in vitro* recordings suggested that some well-demonstrated features of the intrinsic excitability of cerebellar GrCs—namely fast repetitive firing, oscillations, bursting and resonance in theta-range—had in common the dependence upon the same mechanism (a slow K^+ current component; D'Angelo et al., 2001; Gandolfi et al., 2013). However, the inward rectification of a cerebellar GrC was fully explained by another type of current (a fast K-dependent inward rectifier; D'Angelo et al., 2001). Despite there is evidence that an exponential integrate and fire model can fit and reproduce deflective I-V curves in the near-threshold range (Badel et al., 2008), it seems complicated to obtain an AdEx model configuration able to fully reproduce all these different behaviors in different regimes (suprathreshold and subthreshold, respectively) with a single set of parameters configuration (and especially considering that the optimization algorithm only fitted the spiking dynamics).

Bursting Resonance vs. Subthreshold Resonance in AdEx Neuron Models

The formal analysis of the resonance in integrate-and-fire neuron models represents a well-studied field. However, how this resonance extends to the suprathreshold regime is still under

exploration. The subthreshold intrinsic resonance in a biological neuron is shaped by the dynamics of voltage-gated ionic currents, which can be expressed in variable levels but may have a stable resonant frequency (Fox et al., 2017). The resonant frequencies result from a combination of low-pass and high-pass filter mechanisms produced by the interplay of the passive membrane properties and one or more ionic currents and their interaction with the oscillatory inputs (Hutcheon and Yarom, 2000; Fox et al., 2017). The slow resonant currents (or currents having resonant gating variables) oppose voltage changes and act as high-pass filters. Finally, fast amplifying currents (or currents having amplifying gating variables) favor voltage changes and can make resonance more pronounced (Hutcheon and Yarom, 2000; Fox et al., 2017).

One of the main advantages of the AdEx model is the low computational requirements derived from accounting only two differential equations (and state variables; equations 1 and 2). The AdEx model describes a capacitive current (CdV/dt) balanced by membrane currents compressed in three terms: (1) the leak current describes the passive membrane properties and determines an equivalent low-pass filter according to the membrane time constant (Hutcheon and Yarom, 2000); (2) the exponential term describes the activation-dependent on the Na^+ voltage; and (3) the adaptation current, which has proven effective in reproducing more complex subthreshold dynamics such as resonance (Richardson et al., 2003; Brette and Gerstner, 2005; Badel et al., 2008; Naud et al., 2008). The adaptation current could implement a high-pass filter (representative of slow voltage-gated current). This high-pass filter needs to have slow activation according to the adaptation time constant (τ_w), which drives it to turn on or off with a relative delay with respect to the passive membrane charge.

Many types of neurons show membrane potential resonance through a peak in the impedance in contrast with the frequency curve (Z-profile; Fox et al., 2017). The resonance frequency peak can be estimated depending on the adaptation time constant from the high-pass filter [cut-off frequency defined as $1/(2\pi\tau_w)$] and the membrane time constant from the low-pass filter [cut-off frequency defined as $1/(2\pi\tau_m)$, where $\tau_m = C_m/g_L$].

To better understand the oscillatory behavior of our resulting AdEx models, we have explored how resonance frequencies relate in both the subthreshold and suprathreshold regimes. To analyze the subthreshold resonance, we used the impedance profile measured as the amplitude of the membrane voltage response to sinusoidal current stimulation with different frequencies (applying the Fast Fourier Transform algorithm). We have compared both (subthreshold and suprathreshold) resonant peaks and evaluate if both falls within the cut-off frequencies range of the high-pass and low-pass filters that control subthreshold resonance.

In all the models under study, the resonance peaks resulting from the subthreshold regime fall remarkably far from the preferred frequency in the suprathreshold regime. In the case of individuals optimized using *FF1* and *FF3* (**Figures 6A,C**), spiking resonance peaks around theta-band (10 Hz) while subthreshold resonance peaks are under 3 Hz. Both types of resonance fall into the wide window between their low-pass and high-pass filters. Even further, the neuron models obtained from the optimizations using *FF2* and *FF4* showed subthreshold resonance profiles notably sharper. However, the neuron model *FF2* showed a spiking resonance peak markedly shifted to higher frequencies (as it was highlighted in **Figure 4A** and addressed in the "Discussion" section), which falls out of the band between low-pass and high-pass cut-off frequencies (**Figure 6B**). On the other hand, the neuron model *FF4* showed spiking resonance peak around theta-band that falls into the interval between the low-pass and high-filter cut-off frequencies.

DISCUSSION

Computational models represent an essential strategy in neuroscience for researching the function of certain neuronal properties which remain insufficiently explored, as is the case of cerebellar resonance in the theta frequency band (4–12 Hz; Buzsáki, 2006). The convenience of having single-compartment GrC models (point neuron models) reconstructing this behavior with both biological realism and computational efficiency represents an initial step towards understanding these firing dynamics and their involvement in the cerebellar synchronization and learning. This study develops a methodological workflow and explores the best alternatives (in terms of FFs and biological features defined in them) for creating simplified models through optimization of their parameters using EAs. As a result, a set of efficient cerebellar GrC models that closely reflect realistic spiking dynamics are proposed.

The suggested methodology has shown to be successful in generating efficient neuron models capturing the fundamental properties of firing in real cells (e.g., cerebellar GrCs). Interestingly, just the inclusion of the burst frequency as an optimization criterion resulted in neuron models essentially reproducing the main properties of a biological GrC neuron. This seems to suggest that this property is dependent on the main parameters of the cell model and can thus be considered a pivotal property that integrates the main features of the GrC. In addition to this, the optimized neuron models proved suitable against a set of properties that could be relevant in neural information transmission and can be used as features for neuron model optimization. They include linear frequency coding (implemented as repetitive firing under step-current stimulation; D'Angelo et al., 2001, 2009) and the latency to the first spike upon current injection (D'Angelo et al., 1995, 2001; Masoli et al., 2017).

The resulting simplified model evidenced electrical properties characteristic in a biological GrC that were not explicitly integrated into the FF. These are rare spontaneous activity (Chadderton et al., 2004; Jörntell and Ekerot, 2006; Rössert et al., 2014) with high attainable spike frequency (low current needed for the spike generation; D'Angelo et al., 1995, 2001) and non-adapting spike discharge with high firing frequencies (D'Angelo et al., 1995, 1998; Brickley et al., 1996; Chadderton et al., 2004). The proposed model also reproduced properties not so closely related to the firing pattern, such as a strong inward rectification [as in the detailed models of

D'Angelo et al., 2001; Masoli et al., 2017, *in vivo* (Chadderton et al., 2004) and *in vitro* (D'Angelo et al., 1995) experiments]. These emergent properties were predicted uniquely from the suitability of the whole set of AdEx parameter values. This reinforces the biophysical plausibility (in terms of realistic firing dynamics) of the model with very low computational costs. These results make this neuron model of cerebellar GrC a good candidate for large-scale simulations of realistic networks and analysis of these spiking properties.

According to our results, a single set of parameters (specific configuration) of the AdEx model can reproduce a variety of spiking features (wrapped in the FF), but also some emergent behaviors (not explicitly integrated into the FF) since they are governed by compatible suprathreshold dynamics. However, the resulting model failed to reproduce other subthreshold properties like inward rectification (as observed in the I-V relationships). The oscillatory behavior of the cerebellar GrCs is governed by a slow K^+ current component (D'Angelo et al., 2001; Gandolfi et al., 2013), while the inward rectification of the subthreshold regime strongly depends on a fast K^+-dependent component (D'Angelo et al., 2001). Thus, since the AdEx neuron model only includes an additional current component (the adaptation current), we do not have to expect a single set of AdEx parameters fitted to certain spiking properties to also describe both regimes appropriately. Given the computational efficiency but complex adjustment (following a formal analysis) of bursting behaviors of the AdEx model (Brette and Gerstner, 2005), the proposed methodology is presented as a valuable tool to generate a single combination of these few but highly-interrelated parameters for the spiking resonance. The application of this methodology further extending the FF with additional properties from the subthreshold regime would be of interest in helping us to understand how intrinsic properties could affect at the neuron- and also network- level.

The proposed model parameters selected by the EAs (**Table 3**) are consistent with those equivalent values of biological cerebellar GrCs reported both through the literature and the electrophysiological database (Tripathy et al., 2015; NeuroElectro database, 2019). The resting membrane potential (E_L) in our models are within the experimental range from the electrophysiological database (-73.91 ± 9.46 mV from Storm et al., 1998; Brickley et al., 2001; Cathala et al., 2003; Gall et al., 2003; Goldfarb et al., 2007; Prestori et al., 2008; Osorio et al., 2010; Usowicz and Garden, 2012; NeuroElectro database, 2019) and further bibliography (from -60 to -85 mV in D'Angelo et al., 1995, 2001; Brickley et al., 1996; Armano et al., 2000), and they are closer to the mono-compartmental detailed model's values (-65 mV in D'Angelo et al., 2001; Masoli et al., 2017). The spike emission (V_T) values are triggered close to the mean value from the database [at around -41.50 ± 6.43 mV (Brickley et al., 2001; Cathala et al., 2003; Goldfarb et al., 2007; Prestori et al., 2008; Usowicz and Garden, 2012; NeuroElectro database, 2019)] and the computational model of D'Angelo et al. (2001), with a spike peak (V_{peak}) near the experimental evidence [around 20.23 ± 7.04 mV (NeuroElectro database, 2019) from D'Angelo et al., 1998; Osorio et al., 2010; Usowicz and Garden, 2012]. The membrane capacitance (C_m) values appear low as it is notably characterized in a typical GrC (D'Angelo et al., 1995, 2001; Gandolfi et al., 2013) within the range of experimental evidence [3.46 ± 0.82pF (NeuroElectro database, 2019) from D'Angelo et al., 2001; Cathala et al., 2003; Gall et al., 2003; Goldfarb et al., 2007; Prestori et al., 2008; Osorio et al., 2010; Usowicz and Garden, 2012; Gandolfi et al., 2013; Masoli et al., 2017].

It should be noted that the proposed models resulting from EAs with the I-F curve featured in their FFs (*FF2* and *FF4*) show resonance curves in response to sinusoidal current shifted out of the theta band (higher preferred frequencies). Also, the latencies to the first spike remain longer than those experimentally reported, mainly with low stimulation currents. These differences are more severe to the case of the individual from the *FF2*. This fact may indicate an incompatibility of both firing properties (mean frequency under step-current pulses vs. burst frequency resonance under sinusoidal currents) within the simplified AdEx model. Thus, the GrC behavior complexity being beyond the capabilities of these AdEx models with a single parameter configuration (GrCs have different functioning modes).

Based on the analysis of resonance in subthreshold and suprathreshold (spiking) resonance, it seems clear that the preferred frequencies in these two regimes fall in notably different ranges (while spiking resonance tends to fall between 8 and 10 Hz, as driven through the FF in the EA processes, subthreshold preferred frequency peaks about 2 Hz because this regime was not explicitly selected in the FF that drove the parameter tuning). These results may reassert the possibility that the complexity of the spiking resonance in the AdEx model cannot be directly addressed through the analytical adjustment of the parameters. It has to be noted that the subthreshold resonance analysis considers the neuron as the composite of a capacitive current, a passive current, and an adaptive current, neglecting in this way the influence of exponential current of spike firing and the effect of the dynamics of the refractory period. For this reason, the proposed optimization methodology represents a valuable tool to obtain neuron models fitted to complex features. The EA allowed us to tune the two differential equations of the AdEx model according to a complex set of spiking patterns (spiking resonance, regular firing, and delayed firing) under different stimulation protocols (sinusoidal and step current injections).

Realistic modeling based on recent experimental data has provided novel insights on how intrinsic and extrinsic mechanisms interact in other neural systems as the inferior olive (Negrello et al., 2019). According to these results, strong synaptic activity in the awake brain of mice could vanish the functional impact of subthreshold oscillations. Our methodology provides an initial but fundamental tool for the construction of computationally tractable but realistic computational models for future large-scale studies of the functional impact of neuronal resonance in information processing in the GrL. In the particular case of the cerebellar input layer, it remains unclear how spiking resonance (demonstrated *in vitro* in the granule cells and the Golgi cells) interacts in a recurrent inhibitory loop with feed-forward excitation of the Golgi cell. In this sense, theoretical models have addressed information processing in the GrL (Solinas et al., 2010; Garrido et al., 2013), but these models

either have not considered neuronal intrinsic resonance or they have neglected the role of the long-term plasticity in the GrC inputs. In addition to this, theoretical models have demonstrated that external oscillatory activity strongly facilitates learning in excitatory synapses (Masquelier et al., 2009) and inhibitory recurrent networks (Garrido et al., 2016). In any case, further experimental data will be required to fit future computational models to address the functional impact of oscillations in GrL operation.

Based on our results, the AdEx model has shown to be a computationally light approach for the close reproduction of the firing patterns reported from cerebellar GrCs. Recent articles in the literature have proposed modified GLIF point-neuron equations (the so-called Extended Generalized Leaky Integrate-and-Fire model) for the reproduction of experimental traces recorded in different cerebellar cells (Geminiani et al., 2018; Casali et al., 2019). That model allowed the direct application of some experimentally testable parameters together with other optimized ones. We, however, propose a methodology based on automatic parameter tuning through an EA-based exploration (all the behavioral target is integrated through FF definition). It has shown to be effective in fitting the model parameters to diverse spiking responses. Therefore, the optimization process is fast, versatile, and able to capture relevant firing features. Contrary to the methodology proposed in Geminiani et al. (2018) where the optimization algorithm fitted the recorded voltage traces, our approach aims to reproduce the firing characteristics (namely, the burst frequency, the firing rate, and the first-spike latency) of the biological neuron.

To sum up, in this study we present an automatic optimization strategy for the development of computationally efficient neuron models that reproduce realistic firing properties under different stimulation protocols. This methodology was applied to the case of the cerebellar GrC. As a result, a simplified GrC model is proposed, suitable for predicting the main suprathreshold dynamics, such as the spiking resonance at the theta range and the linear frequency coding. This contribution serves as an initial step towards a better understanding of the functional implication of the theta-frequency-band resonance for information processing at the cerebellar cortex. This model provides both efficiency and biological plausibility which will facilitate further computational work in the reconstruction of large-scale models of microcircuits to better understand the computational role of the suprathreshold dynamics of the cell on a large scale.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

MM, JG, and ER: study design. MM and MS-L: literature and database search. JG and MM: EAs methodology. MM, JG, and ER: analysis and interpretation of results. MM and JG: writing of the article. All the authors have read and approved the final manuscript. All the results included in this article are part of MM's Ph.D. thesis.

FUNDING

JG was supported by FEDER/Junta de Andalucía-Consejería de Economía y Conocimiento under the EmbBrain (A-TIC-276-UGR18) project and by the University of Granada under the Young Researchers Fellowship. This work was also supported by funds from Ministerio de Economía y Competitividad (MINECO)-FEDER (TIN2016-81041-R), the European Human Brain Project SGA2 (H2020-RIA 785907), and SGA3 (European Commission; H2020-RIA 945539), and CEREBIO (J. A. P18-FR-2378) to ER. The aforementioned bodies funded the research work described through scientific grants covering general and personnel costs. They did not play any direct role in the design of the study, collection, analysis or the interpretation of data in the manuscript.

REFERENCES

Armano, S., Rossi, P., Taglietti, V., and D'Angelo, E. (2000). Long-term potentiation of intrinsic excitability at the mossy fiber-granule cell synapse of rat cerebellum. *J. Neurosci.* 20, 5208–5216. doi: 10.1523/JNEUROSCI.20-14-05208.2000

Badel, L., Lefort, S., Brette, R., Petersen, C. C. H., Gerstner, W., and Richardson, M. J. E. (2008). Dynamic I-V curves are reliable predictors of naturalistic pyramidal-neuron voltage traces. *J. Neurophysiol.* 99, 656–666. doi: 10.1152/jn.01107.2007

Barranca, V. J., Johnson, D. C., Moyher, J. L., Sauppe, J. P., Shkarayev, M. S., Kovačič, G., et al. (2013). Dynamics of the exponential integrate-and-fire model with slow currents and adaptation. *J. Comput. Neurosci.* 37, 161–180. doi: 10.1007/s10827-013-0494-0

Bezzi, M., Nieus, T., Coenen, J.-M., and D'Angelo, E. (2004). An integrate-and-fire model of a cerebellar granule cell. *Neurocomputing* 58, 593–598. doi: 10.1016/j.neucom.2004.01.100

Brette, R., and Gerstner, W. (2005). Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J. Neurophysiol.* 94, 3637–3642. doi: 10.1152/jn.00686.2005

Brickley, S. G., Cull-Candy, S. G., and Farrant, M. (1996). Development of a tonic form of synaptic inhibition in rat cerebellar granule cells resulting from persistent activation of GABAA receptors. *J. Physiol.* 497, 753–759. doi: 10.1113/jphysiol.1996.sp021806

Brickley, S. G., Revilla, V., Cull-Candy, S. G., Wisden, W., and Farrant, M. (2001). Adaptive regulation of neuronal excitability by a voltage-independent potassium conductance. *Nature* 409, 88–92. doi: 10.1038/35051086

Buzsáki, G. (2006). *Rhythms of the Brain.* New York: Oxford University Press. doi: 10.1093/acprof:oso/9780195301069.001.0001

Casali, S., Marenzi, E., Medini, C., Casellato, C., and D'Angelo, E. (2019). Reconstruction and simulation of a scaffold model of the cerebellar network. *Front. Neuroinformatics* 13:37. doi: 10.3389/fninf.2019.00037

Cathala, L., Brickley, S., Cull-Candy, S., Farrant, M., Necchi, D., Diwakar, S., et al. (2003). Maturation of EPSCs and intrinsic membrane properties enhances precision at a cerebellar synapse. *J. Neurosci.* 23, 6074–6085. doi: 10.1523/JNEUROSCI.23-14-06074.2003

Chadderton, P., Margrie, T. W., and Häusser, M. (2004). Integration of quanta in cerebellar granule cells during sensory processing. *Nature* 428, 856–860. doi: 10.1038/nature02442

Courtemanche, R., Chabaud, P., and Lamarre, Y. (2009). Synchronization in primate cerebellar granule cell layer local field potentials: basic anisotropy and dynamic changes during active expectancy. *Front. Cell. Neurosci.* 3:6. doi: 10.3389/neuro.03.006.2009

D'Angelo, E., De Filippi, G., Rossi, P., and Taglietti, V. (1995). Synaptic excitation of individual rat cerebellar granule cells *in situ*: evidence for the role of NMDA receptors. *J. Physiol.* 484, 397–413. doi: 10.1113/jphysiol.1995.sp020673

D'Angelo, E., De Filippi, G., Rossi, P., and Taglietti, V. (1998). Ionic mechanism of electroresponsiveness in cerebellar granule cells implicates the action of a persistent sodium current. *J. Neurophysiol.* 80, 493–503. doi: 10.1152/jn.1998.80.2.493

D'Angelo, E., Koekkoek, S. K., Lombardo, P., Solinas, S., Ros, E., Garrido, J. A., et al. (2009). Timing in the cerebellum: oscillations and resonance in the granular layer. *Neuroscience* 162, 805–815. doi: 10.1016/j.neuroscience.2009.01.048

D'Angelo, E., Mazzarello, P., Prestori, F., Mapelli, J., Solinas, S., Lombardo, P., et al. (2011). The cerebellar network: from structure to function and dynamics. *Brain Res. Rev.* 66, 5–15. doi: 10.1016/j.brainresrev.2010.10.002

D'Angelo, E., Nieus, T., Maffei, A., Armano, S., Rossi, P., Taglietti, V., et al. (2001). Theta-frequency bursting and resonance in cerebellar granule cells: experimental evidence and modeling of a slow K^+-dependent mechanism. *J. Neurosci.* 21, 759–770. doi: 10.1523/JNEUROSCI.21-03-00759.2001

Das, A., and Narayanan, R. (2017). Theta-frequency selectivity in the somatic spike-triggered average of rat hippocampal pyramidal neurons is dependent on HCN channels. *J. Neurophysiol.* 118, 2251–2266. doi: 10.1152/jn.00356.2017

Delvendahl, I., Straub, I., and Hallermann, S. (2015). Dendritic patch-clamp recordings from cerebellar granule cells demonstrate electrotonic compactness. *Front. Cell. Neurosci.* 9:93. doi: 10.3389/fncel.2015.00093

Diwakar, S., Magistretti, J., Goldfarb, M., Naldi, G., and D'Angelo, E. (2009). Axonal Na^+ channels ensure fast spike activation and back-propagation in cerebellar granule cells. *J. Neurophysiol.* 101, 519–532. doi: 10.1152/jn.90382.2008

Dugué, G. P., Brunel, N., Hakim, V., Schwartz, E., Chat, M., Lévesque, M., et al. (2009). Electrical coupling mediates tunable low-frequency oscillations and resonance in the cerebellar golgi cell network. *Neuron* 61, 126–139. doi: 10.1016/j.neuron.2008.11.028

Fortin, F.-A., De Rainville, F.-M., Gardner, M. A., Parizeau, M., and Gagné, C. (2012). DEAP: evolutionary algorithms made easy. *J. Mach. Learn. Res.* 13, 2171–2175. doi: 10.5555/2503308.2503311

Fox, D. M., Tseng, H. A., Smolinski, T. G., Rotstein, H. G., and Nadim, F. (2017). Mechanisms of generation of membrane potential resonance in a neuron with multiple resonant ionic currents. *PLoS Comput. Biol.* 13:e1005565. doi: 10.1371/journal.pcbi.1005565

Gall, D., Roussel, C., Susa, I., D'Angelo, E., Rossi, P., Bearzatto, B., et al. (2003). Altered neuronal excitability in cerebellar granule cells of mice lacking calretinin. *J. Neurosci.* 23, 9320–9327. doi: 10.1523/JNEUROSCI.23-28-09320.2003

Gandolfi, D., Lombardo, P., Mapelli, J., Solinas, S., and D'Angelo, E. (2013). Theta-frequency resonance at the cerebellum input stage improves spike timing on the millisecond time-scale. *Front. Neural Circuits* 7:64. doi: 10.3389/fncir.2013.00064

Garrido, J. A., Ros, E., and D'Angelo, E. (2013). Spike timing regulation on the millisecond scale by distributed synaptic plasticity at the cerebellum input stage: a simulation study. *Front. Comput. Neurosci.* 7:64. doi: 10.3389/fncom.2013.00064

Garrido, J. A., Luque, N. R., Tolu, S., and D'Angelo, E. (2016). Oscillation-driven spike-timing dependent plasticity allows multiple overlapping pattern recognition in inhibitory interneuron networks. *Int. J. Neural Syst.* 26:1650020. doi: 10.1142/s0129065716500209

Geminiani, A., Casellato, C., Locatelli, F., Prestori, F., Pedrocchi, A., and D'Angelo, E. (2018). Complex dynamics in simplified neuronal models: reproducing golgi cell electroresponsiveness. *Front. Neuroinform.* 12:88. doi: 10.3389/fninf.2018.00088

Goldfarb, M., Schoorlemmer, J., Williams, A., Diwakar, S., Wang, Q., Huang, X., et al. (2007). Fibroblast growth factor homologous factors control neuronal excitability through modulation of voltage-gated sodium channels. *Neuron* 55, 449–463. doi: 10.1016/j.neuron.2007.07.006

Hanuschkin, A., Kunkel, S., Helias, M., Morrison, A., and Diesmann, M. (2010). A general and efficient method for incorporating precise spike times in globally time-driven simulations. *Front. Neuroinform.* 4:113. doi: 10.3389/fninf.2010.00113

Hartmann, M. J., and Bower, J. M. (1998). Oscillatory activity in the cerebellar hemispheres of unrestrained rats. *J. Neurophysiol.* 80, 1598–1604. doi: 10.1152/jn.1998.80.3.1598

Herculano-Houzel, . (2010). Coordinated scaling of cortical and cerebellar numbers of neurons. *Front. Neuroanat.* 4:12. doi: 10.3389/fnana.2010.00012

Hutcheon, B., and Yarom, Y. (2000). Resonance, oscillation and the intrinsic frequency preferences of neurons. *Trends Neurosci.* 23, 216–222. doi: 10.1016/s0166-2236(00)01547-2

Jolivet, R., Kobayashi, R., Rauch, A., Naud, R., Shinomoto, S., and Gerstner, W. (2008). A benchmark test for a quantitative assessment of simple neuron models. *J. Neurosci. Methods* 169, 417–424. doi: 10.1016/j.jneumeth.2007.11.006

Jörntell, H., and Ekerot, C.-F. (2006). Properties of somatosensory synaptic integration in cerebellar granule cells *in vivo*. *J. Neurosci.* 26, 11786–11797. doi: 10.1523/JNEUROSCI.2939-06.2006

Magistretti, J., Castelli, L., Forti, L., and D'Angelo, E. (2006). Kinetic and functional analysis of transient, persistent and resurgent sodium currents in rat cerebellar granule cells *in situ*: an electrophysiological and modelling study. *J. Physiol.* 573, 83–106. doi: 10.1113/jphysiol.2006.106682

Masoli, S., Rizza, M. F., Sgritta, M., Van Geit, W., Schürmann, F., and D'Angelo, E. (2017). Single neuron optimization as a basis for accurate biophysical modeling: the case of cerebellar granule cells. *Front. Cell. Neurosci.* 11:71. doi: 10.3389/fncel.2017.00071

Masquelier, T., Hugues, E., Deco, G., and Thorpe, S. J. (2009). Oscillations, phase-of-firing coding, and spike timing-dependent plasticity: an efficient learning scheme. *J. Neurosci.* 29, 13484–13493. doi: 10.1523/JNEUROSCI.2207-09.2009

Nair, M., Subramanyan, K., Nair, B., and Diwakar, S. (2015). "Parameter optimization and nonlinear fitting for computational models in neuroscience on GPGPUs," in *2014 International Conference on High Performance Computing and Applications (ICHPCA)*, (Bhubaneswar, India: IEEE), 1–5. doi: 10.1109/ICHPCA.2014.7045324

Naud, R., Marcille, N., Clopath, C., and Gerstner, W. (2008). Firing patterns in the adaptive exponential integrate-and-fire model. *Biol. Cybern.* 99, 335–347. doi: 10.1007/s00422-008-0264-7

Negrello, M., Warnaar, P., Romano, V., Owens, C. B., Lindeman, S., Iavarone, E., et al. (2019). Quasiperiodic rhythms of the inferior olive. *PLoS Comput. Biol.* 15:e1006475. doi: 10.1371/journal.pcbi.1006475

NeuroElectro database. (2019). NeuroElectro: organizing information on cellular neurophysiology. Available online at: https://www.neuroelectro.org/publications/. Accessed October 4, 2019.

Nieus, T., Sola, E., Mapelli, J., Saftenku, E., Rossi, P., and D'Angelo, E. (2006). LTP regulates burst initiation and frequency at mossy fiber-granule cell synapses of rat cerebellum: experimental observations and theoretical predictions. *J. Neurophysiol.* 95, 686–699. doi: 10.1152/jn.00696.2005

Osorio, N., Cathala, L., Meisler, M. H., Crest, M., Magistretti, J., and Delmas, P. (2010). Persistent Nav1.6 current at axon initial segments tunes spike timing of cerebellar granule cells. *J. Physiol.* 588, 651–670. doi: 10.1113/jphysiol.2010.183798

Pellerin, J.-P., and Lamarre, Y. (1997). Local field potential oscillations in primate cerebellar cortex during voluntary movement. *J. Neurophysiol.* 78, 3502–3507. doi: 10.1152/jn.1997.78.6.3502

Peyser, A., Sinha, A., Vennemo, S. B., Ippen, T., Jordan, J., Graber, S., et al. (2017). NEST 2.14.0. Zenodo. doi: 10.5281/zenodo.882971

Prestori, F., Rossi, P., Bearzatto, B., Lainé, J., Necchi, D., Diwakar, S., et al. (2008). Altered neuron excitability and synaptic plasticity in the cerebellar granular layer of juvenile prion protein knock-out mice with impaired motor control. *J. Neurosci.* 28, 7091–7103. doi: 10.1523/jneurosci.0409-08.2008

Richardson, M. J. E., Brunel, N., and Hakim, V. (2003). From subthreshold to firing-rate resonance. *J. Neurophysiol.* 89, 2538–2554. doi: 10.1152/jn.00955.2002

Ros, H., Sachdev, R. N. S., Yu, Y., Sestan, N., and McCormick, D. A. (2009). Neocortical networks entrain neuronal circuits in cerebellar cortex. *J. Neurosci.* 29, 10309–10320. doi: 10.1523/JNEUROSCI.2327-09.2009

Rössert, C., Solinas, S., D'Angelo, E., Dean, P., and Porrill, J. (2014). Model cerebellar granule cells can faithfully transmit modulated firing rate signals. *Front. Cell. Neurosci.* 8:304. doi: 10.3389/fncel.2014.00304

Rotstein, H. G. (2017). Spiking resonances in models with the same slow resonant and fast amplifying currents but different subthreshold dynamic properties. *J. Comput. Neurosci.* 43, 243–271. doi: 10.1007/s10827-017-0661-9

Solinas, S., Nieus, T., and D'Angelo, E. (2010). A realistic large-scale model of the cerebellum granular layer predicts circuit spatio-temporal filtering properties. *Front. Cell. Neurosci.* 4:12. doi: 10.3389/fncel.2010.00012

Storm, D. R., Hansel, C., Hacker, B., Parent, A., and Linden, D. J. (1998). Impaired cerebellar long-term potentiation in type I adenylyl cyclase mutant mice. *Neuron* 20, 1199–1210. doi: 10.1016/s0896-6273(00)80500-0

Tripathy, S. J., Burton, S. D., Geramita, M., Gerkin, R. C., and Urban, N. N. (2015). Brain-wide analysis of electrophysiological diversity yields novel categorization of mammalian neuron types. *J. Neurophysiol.* 113, 3474–3489. doi: 10.1152/jn.00237.2015

Usowicz, M. M., and Garden, C. L. P. (2012). Increased excitability and altered action potential waveform in cerebellar granule neurons of the Ts65Dn mouse model of Down syndrome. *Brain Res.* 1465, 10–17. doi: 10.1016/j.brainres.2012.05.027

Venkadesh, S., Komendantov, A. O., Listopad, S., Scott, E. O., De Jong, K., Krichmar, J. L., oetet al. (2018). Evolving simple models of diverse intrinsic dynamics in hippocampal neuron types. *Front. Neuroinform.* 12:8. doi: 10.3389/fninf.2018.00008

Wang, H., Sun, M. J., Chen, H., Zhang, J., Zhang, L. B., Zhang, W. W., et al. (2019). Spontaneous recovery of conditioned eyeblink responses is associated with transiently decreased cerebellar theta activity in guinea pigs. *Behav. Brain Res.* 359, 457–466. doi: 10.1016/j.bbr.2018.11.030

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

3. A comparative study of stochastic optimizers for fitting neuron models. Application to the cerebellar granule cell

Authors	N.C. Cruz, **M. Marín***, J. L. Redondo, E. M. Ortigosa & P. M. Ortigosa		
Journal	*Informatica*		
Year	2021		
Pages	22		
DOI	10.15388/21-infor450		
Impact Factor (JCR 2019)	3.312		
Categories	Applied Mathematics:	Ranking: 9/261	**(Q1)**
	Computer Science, Information Systems:	Ranking: 46/156	(Q2)

INFORMATICA, 2021, Vol. 0, No. 0, 1–22

DOI: https://doi.org/10.15388/21-INFOR450

A Comparative Study of Stochastic Optimizers for Fitting Neuron Models. Application to the Cerebellar Granule Cell

Nicolás C. CRUZ[1], Milagros MARÍN[2,*], Juana L. REDONDO[1], Eva M. ORTIGOSA[3], Pilar M. ORTIGOSA[1]

[1] *Department of Informatics, University of Almería, ceiA3 campus, Spain*
[2] *Department of Biochemistry and Molecular Biology I, University of Granada, Spain*
[3] *Department of Computer Architecture and Technology, University of Granada, Spain*
e-mail: ncalvocruz@ual.es, mmarin@ugr.es, jlredondo@ual.es, ortigosa@ugr.es, ortigosa@ual.es

Received: November 2020; accepted: April 2021

Abstract. This work compares different algorithms to replace the genetic optimizer used in a recent methodology for creating realistic and computationally efficient neuron models. That method focuses on single-neuron processing and has been applied to cerebellar granule cells. It relies on the adaptive-exponential integrate-and-fire (AdEx) model, which must be adjusted with experimental data. The alternatives considered are: i) a memetic extension of the original genetic method, ii) Differential Evolution, iii) Teaching-Learning-Based Optimization, and iv) a local optimizer within a multi-start procedure. All of them ultimately outperform the original method, and the last two do it in all the scenarios considered.
Key words: granule cell, neuron model, model tuning, optimization, meta-heuristics.

1. Introduction

One of the main pillars in Computational Neuroscience is understanding the brain operation by studying information processing primitives of brain areas. For this purpose, it is necessary to simulate brain microcircuits using large-scale neural networks with thousands or millions of neurons. Neuron computational models aim to reproduce neuronal firing patterns as well as the information contained in electrophysiological recordings. However, biological realism frequently requires high computational resources. Thus, neuron models for large-scale simulations need to be computationally efficient.

The adaptive-exponential integrate-and-fire (AdEx) model (Brette and Gerstner, 2005) is a simplified neuron model that meets both requirements of realism and efficiency. It consists of only two differential equations and performs reasonably well in fitting real electrophysiological recordings with a few parameters and low computational cost (Naud

*Corresponding author.

et al., 2008). Nevertheless, some of the parameters in the AdEx model lack an experimental (measurable) counterpart, and finding an appropriate set of parameters becomes a challenging problem (Barranca *et al.*, 2014; Hanuschkin *et al.*, 2010; Venkadesh *et al.*, 2018). Fitting mathematical neuron models to real electrophysiological behaviour can be considered a suitable optimization problem that remains partially unsolved.

The cerebellum is a centre of the nervous system involved in fine motor control, somatosensory processing, and non-motor control (emotional, cognitive, and autonomic processes such as attention and language) (Schmahmann, 2019). In its anatomical structure, there exists one input layer named granular layer (GrL), and it is compounded of cerebellar granule cells (GrCs). The GrCs are the smallest and the most numerous neurons in the human brain (Lange, 1975; Williams and Herrup, 1988). The cerebellar GrCs are thought to regulate the information processing through the main afferent system of the cerebellum (Jörntell and Ekerot, 2006). These neurons show regular repetitive spike discharge in response to a continuous direct stimulus, and their first-spike latency under direct stimulation is well characterized (D'Angelo *et al.*, 2009; Masoli *et al.*, 2017). Besides, previous findings suggest that the theta-frequency oscillatory activity (around 4–10 Hz in rodents) contributes to signal integration in the GrL. Indeed, the spiking resonance (as enhanced bursting activity) at the theta-frequency band of single cerebellar GrCs in response to low-frequency sinusoidal stimulation has been proposed to strengthen information transmission in the GrL (D'Angelo *et al.*, 2009, 2001; Gandolfi *et al.*, 2013). However, the functional role of resonance at the theta band in the information processing of cerebellar GrCs remains elusive.

Previous work by Marín *et al.* (2020) proposed a methodology for building computationally efficient neuron models of cerebellar GrCs that replicate some inherent properties of the biological cell. Since the cerebellar GrCs show compact and simple morphology (D'Angelo *et al.*, 2001; Delvendahl *et al.*, 2015), it is appropriate to consider a monocompartment model. Thus, the cerebellar GrC was modelled with the AdEx neuron model because of its computational efficiency and realistic firing modes. This fact has been supported by several comparisons with detailed models and experimental recordings (Brette and Gerstner, 2005; Nair *et al.*, 2014; Naud *et al.*, 2008). As part of their method, Marín *et al.* (2020) tuned the parameters of the AdEx model to fit the neuronal spiking dynamics of real recordings. In this context, the authors model the tuning procedure as an optimization problem and study different objective functions to conduct the optimization. These functions combine the accumulated difference between the set of *in vitro* measurements and the spiking output of the neuron model under tuning. However, their evaluation involves launching computer simulations with uncertainty and nonlinear equations. For this reason, the authors proposed a derivative-free, black-box or direct optimization approach (Price *et al.*, 2006; Storn and Price, 1997). Namely, they successfully implemented a standard genetic algorithm (Boussaïd *et al.*, 2013; Cruz *et al.*, 2018; Lindfield and Penny, 2017) to face the parametric optimization problem.

This work focuses on the optimization component of the methodology proposed by Marín et al. and evaluates alternative algorithms. Since no exact method is known to solve the nonlinear and simulation-based target problem, the choice of Marín *et al.* is sound.

Genetic algorithms are widely-used and effective meta-heuristics, i.e. generic problem resolution strategies (Boussaïd *et al.*, 2013; Lindfield and Penny, 2017). Nevertheless, the optimization part of the referred work has been tangentially addressed, and this paper aims to assess the suitability of some alternative meta-heuristics that perform well in other fields. More precisely, this paper compares four more meta-heuristics in the context of the reference work. One of them is Differential Evolution (DE) (Storn and Price, 1997), which is arguably one of the most used methods for parametric optimization in Engineering. Another option is the Teaching-Learning-Based optimizer (TLBO) (Rao *et al.*, 2012), which features configuration simplicity, low computational cost and high convergence rates. A third option is the combination of a simple yet effective local optimizer, the Single-Agent Stochastic Search (SASS) method (Cruz *et al.*, 2018) by Solis and Wets (1981), with a generic multi-start component (Redondo *et al.*, 2013; Salhi, 2017). This compound optimizer will be referred to as MSASS. The last method, which will be called MemeGA, is an ad-hoc memetic algorithm (Cruz *et al.*, 2018; Marić *et al.*, 2014) that results from replacing the mutation part of the genetic method by Marín *et al.* (2020) by the referred local search procedure, SASS.

The rest of the paper is structured as follows: Section 2 describes the neuron model with the parameters to tune and the corresponding optimization problem. Section 3 explains the optimizers considered as a potentially more effective replacement of the genetic method. Section 4 presents the experimentation carried out and the results achieved. Finally, Section 5 contains the conclusions and states some possible future work lines.

2. Neuron Model

This section starts by describing the neuron model, whose parameters must be adjusted, and by defining the tuning process as an optimization problem to solve. After that, the section explains both the neuron simulation environment and how the experimental pieces of data have been recorded according to the reference paper.

2.1. *Model Structure and Problem Definition*

The adaptive exponential integrate-and-fire (AdEx) model consists of two coupled differential equations and a reset condition that regulate two state variables, i.e. the membrane potential (V) and the adaptation current (w):

$$C_m \frac{dV}{dt} = -g_L(V - E_L) + g_L \Delta_T \exp\left(\frac{V - V_T}{\Delta_T}\right) + I(t) - w, \tag{1a}$$

$$\tau_w \frac{dw}{dt} = a(V - E_L) - w. \tag{1b}$$

Equation (1a) describes how V varies with the injection of current, $I(t)$. When V exceeds the threshold potential (V_T), the slope factor (Δ_T) models the action potential. This depolarization persists until V reaches the reset threshold potential (V_{peak}), which defines the reset condition aforementioned. At that point, V resets to V_r, and w increases the

Table 1
Parameters to tune for configuring the AdEx model including the boundaries and units.

Parameter	Boundaries	Parameter	Boundaries
C_m (pF)	[0.1, 5.0]	V_T (mV)	[−60, −20]
Δ_T (mV)	[1, 1000]	a (nS)	[−1, 1]
E_L (mV)	[−80, −40]	b (pA)	[−1, 1]
V_r (mV)	[−80, −40]	g_L (nS)	[0.001, 10]
V_{peak} (mV)	[−20, 20]	τ_w (ms)	[1, 1000]

fixed amount b. The first term models the passive mechanisms of the membrane potential that depend on the total leak conductance (g_L), the leak reversal potential (E_L) and the membrane capacitance (C_m), which regulate the integrative properties of the neuron. The second exponential term models the spike generation and shape, whose dynamics depend on Δ_T and V_T (Naud *et al.*, 2008).

Equation (1b) models the evolution of w. It depends on the adaptation time constant parameter (τ_w), the sub-threshold adaptation (a), and the spike-triggered adaptation parameter (b).

There are ten parameters to tune for reproducing the firing properties or features of the cerebellar GrCs with the AdEx neuron model. Table 1 contains the list of model parameters to tune including their units and the range in which they must be adjusted.

Roughly speaking, the tuning process is equivalent to minimizing the accumulated difference between each firing feature for the studied parameter set and the corresponding experimental recordings. Let f be the function that models this computation. It is defined in (2) as an abstract function that depends on the ten model parameters included in Table 1. This configuration is that tagged as *FF*4 in Marín *et al.* (2020), and it is further explained below.

$$f(C_m, \ldots, \tau_w) = \sum_{i \in [MF, LF]} \left(|feat_i - \exp_i| \cdot w_i \right) + \sum_{j=1}^{N} \left(|\overline{BF_{sim_j}} - \overline{BF_{\exp_j}}| \cdot W_{BF} \cdot \left(std(BF_{sim_j}) + 1 \right) \right). \quad (2)$$

In practical terms, f implies a neuron simulation procedure gathering the output of the behaviour of interest. In order to obtain a neuron model that reproduces the properties of the cerebellar GrCs, the following features were integrated into f: i) repetitive spike discharge in response to continuous direct stimulation (measured as the mean frequency of spike traces) (*MF*), ii) first-spike latency under direct current stimulation (measured as the time to the first spike) (*LF*), iii) spiking resonance in the theta range under sinusoidal current stimulation (measured as the average burst frequency with different oscillation frequencies) (*BF*).

MF and *LF* are included in the first term of f, where the score of each feature is the accumulative absolute difference between the model output of the i feature ($feat_i$)

under different step-current amplitudes, and its experimental equivalent ($\exp_i$), for $i \in \{MF, LF\}$. In order to integrate several components in this function, this term is multiplied by a weighting factor (w_i) as described in the next section.

Regarding BF, it is considered in the second part of f. This term accumulates the summation of each separate absolute difference for $j = 1, \ldots, N$ sinusoidal stimulation frequencies. The score for each one results from the absolute difference between: i) the average burst frequency of the simulated neuron after a certain number of oscillatory cycles, $\overline{BF_{sim_j}}$, and ii) the experimental value, $\overline{BF_{\exp_j}}$, at that stimulation frequency, j. This term is multiplied by the weight of this particular feature, w_{BF} (as detailed in the section below). It is also multiplied by the standard deviation of the simulated neuron burst frequency ($std(BF_{sim_j})$) plus one. This additional multiplicative term promotes configurations both close to the target and stable.

The lower the value of f is for a given set of parameters, the more appropriate it is for replicating the desired neuronal behaviour. It is hence possible to formally define the target optimization problem according to (3). The constraints correspond to the valid range of every parameter, which is generally defined by the max and min superscripts linked to the parameter symbol referring to the upper and lower bounds, respectively. The numerical values considered are those shown in Table 1.

$$\begin{aligned} \underset{C_m,\ldots,\tau_w}{\text{minimize}} \quad & f(C_m, \ldots, \tau_w) \\ \text{subject to} \quad & C_m^{\min} \leqslant C_m \leqslant C_m^{\max} \\ & \ldots \\ & \tau_w^{\min} \leqslant \tau_w \leqslant \tau_w^{\max}. \end{aligned} \tag{3}$$

2.2. *Model Context and Feature Measurement*

According to Marín *et al.* (2020), the initial membrane potential starts with the same value as the leak reversal potential, i.e. $V_{init} = E_L$. The burst and mean frequency have been weighted by 1 because they are measured in hertz and are in comparable ranges. In contrast to them, since it is measured in seconds, the weight of the first spike feature was set to 1000. Consequently, for instance, an error of 1 Hz at burst frequency is weighted as an error of 1 Hz at repetitive spike discharge and a lag of 1 ms at latency to the first spike.

For the experimental measurements of the spiking resonance, the burst frequency is computed as the inverse of the average inter-spike interval (ISI) of the output neuron (the cerebellar GrC) during each stimulation cycle. Then, the average burst frequency is measured from 10 consecutive cycles during a total of 22.5 seconds of stimulation. The sinusoidal amplitude is of 6 and 8 pA (including a 12 pA offset), and the spike bursts are generated corresponding with the positive phase of the stimulus (sinusoidal phase of 270°). The burst frequencies with stimulation frequencies beyond 10.19 Hz in amplitude of 6 pA and 14.23 Hz in amplitude of 8 pA are set to zero because either one or no spikes were observed in the experimental measurements (D'Angelo *et al.*, 2001).

The typical behaviour of the cerebellar GrCs implements a mechanism of repetitive spike discharge in response to step-current stimulation. It could help to support the spiking

resonance of burst frequency at the theta-frequency band (D'Angelo *et al.*, 2001). According to recent literature (Masoli *et al.*, 2017), the fast repetitive discharge in the GrCs has been characterized based on the mean frequency and the latency to the first spike in response to three different step-current injections (10, 16 and 22 pA) of stimulation of 1 s.

3. Optimization Methods

As introduced, five numerical optimization algorithms have been considered for solving the problem stated in Section 2.1: i) the genetic algorithm used in the reference paper (GA) (Marín *et al.*, 2020), ii) a memetic optimizer based on it (MemeGA), iii) Differential Evolution (DE), iv) Teaching-Learning-Based Optimization (TLBO) (Rao *et al.*, 2012), and v) Multi-Start Single-Agent Stochastic Search (MSASS). The first four cover the two main groups of population-based meta-heuristics (Boussaïd *et al.*, 2013): Evolutionary Computation and Swarm Intelligence. The optimizers in the first group are inspired by the Darwinian theory, while those in the second rely on simulating social interaction. Namely, GA, MemeGA, and DE belong to the first class, and TLBO can be classified into the second one. Regarding MSASS, it is a single-solution-based meta-heuristic (Boussaïd *et al.*, 2013) that iteratively applies a local search process to independent random starts.

Since all the algorithms rely on randomness, they can be classified as stochastic methods. The following subsections describe these optimizers for the sake of completeness. However, the interested reader is referred to their corresponding references for further information.

3.1. *Genetic Algorithm (GA)*

Genetic algorithms, popularized by Holland (1975), were developed as an application of artificial intelligence to face hard optimization problems that cannot be rigorously solved (Boussaïd *et al.*, 2013; Salhi, 2017). Roughly speaking, genetic algorithms work with a pool of candidate solutions. Although they are randomly generated at first, the solutions are ultimately treated as a population of biological individuals that evolve through sexual reproduction (involving crossover and mutation). Based on these ideas, genetic algorithms define a general framework in which there are different options for every step and are used in a wide range of problems (Boussaïd *et al.*, 2013; Salhi, 2017; Shopova and Vaklieva-Bancheva, 2006). The problem addressed in this work has also been previously faced with a genetic algorithm in Marín *et al.* (2020). It will be referred to as GA, and it is described next as one of the methods compared.

GA starts by creating as many random candidate solutions as required by the parameter that defines the population size. Every one of these so-called 'individuals' is represented by a vector in which the i component corresponds to the i optimization variable. Random generation in each dimension follows a uniform distribution between the boundaries indicated in Table 1. After generating the initial individuals, they must be evaluated according to the objective function and linked to their resulting fitness. As the problem at hand is

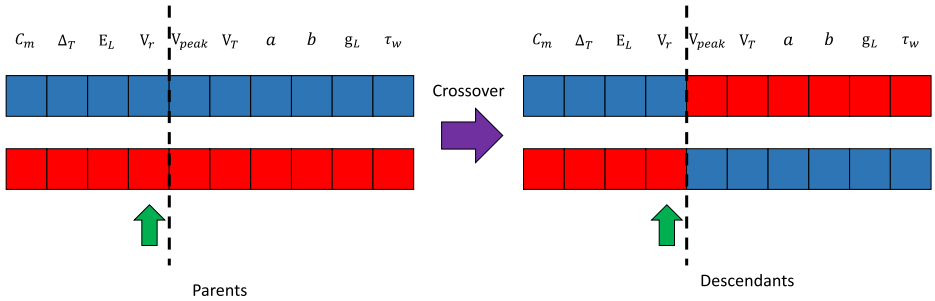


Fig. 1. Single-point crossover between two individuals.

a minimization one, the lower the fitness is, the better it is. The population size remains constant during the search even though individuals change due to evolution.

After creating the initial population, GA simulates as many generations as required by the parameter that defines the number of cycles. Each of them consists of the ordered execution of the following steps or genetic operators:

Selection: This step selects those individuals from the current population that might become the progenitors and members of the next one. It chooses as many individuals as defined by the population size according to a tournament selection process (Salhi, 2017; Shopova and Vaklieva-Bancheva, 2006). This strategy consists in choosing every individual as the best out of a random sample whose size is another user-given parameter. In contrast to other selection approaches, this one tends to attenuate strong drifts in the population because the scope of selection is limited by the tournament size (Salhi, 2017). Thus, not only the top best are selected, but also more regular ones, which increases variability and avoids premature convergence to local optima.

Crossover: This step simulates sexual reproduction among those individuals previously selected to explore new regions of the search space based on the information provided by the progenitors. The process iterates through the selected individuals and pairs those at even positions with those at odd ones. For each pair, there is a user-given probability of performing single-point crossover (Salhi, 2017; Shopova and Vaklieva-Bancheva, 2006). It consists in randomly selecting a dimension of the individuals and swapping the rest of the vector between the two involved. Figure 1 depicts the single-point crossover procedure. As shown, the two progenitors (left) result in the two descendants (right) after having selected the fourth as the splitting dimension.

Mutation: This step tries to randomly alter the offspring that comes from the two previous ones. While crossover explores in depth the area of the search space covered by progenitors, mutation aims to allow reaching new regions of it, which avoids stagnation at local optima. The mutation procedure has a user-given probability to be launched for every individual of the offspring. Every time that it happens to apply, the individual is traversed, and each of its components has another user-given probability to be randomly reset.

After the three evolutionary steps described, the individuals involved are evaluated according to the objective function (only those that changed after being selected) and replace the current population for the next cycle. It is also relevant to highlight that the algorithm keeps a record with the best individuals found through the process. The solution to the problem is taken as the best one among them.

3.2. *Memetic Algorithm Derived from GA (MemeGA)*

Memetic algorithms (Cruz *et al.*, 2018; Marić *et al.*, 2014) are an extension of standard genetic methods formalized by Moscato (1989). Their name comes from the term 'meme', defined by Dawkins (1976) as an autonomous and cultural entity similar to biological genes. However, while the individuals of genetic algorithms are mainly passive, those of memetic methods can be seen as active agents that improve their aptitude autonomously. In practical terms, this is achieved by adding the use of complementary local search techniques to the underlying process of Darwinian evolution. This approach, which is useful to avoid premature convergence while improving the exploitation of the search space, is popular in many different fields. For instance, in Marić *et al.* (2014), the authors design an effective and efficient parallel memetic algorithm for facility location, an NP-hard combinatorial optimization problem. A memetic algorithm is also the best-performing method for parametric heliostat field design in the study published in Cruz *et al.* (2018). Similarly, the optimization engine of the recent framework for drug discovery extended in Puertas-Martín *et al.* (2020) is a memetic method.

Considering the theoretical capabilities and the previous success cases of this kind of algorithms, a memetic method has been specifically designed and included in the present study. As introduced, the referred method is MemeGA, and it is based on the genetic method designed by Marín *et al.* (2020) for the problem at hand. The algorithm maintains the same structure as GA with the only exception of the mutation stage, which is replaced by the use of a local search algorithm. Namely, the original mutation probability is treated as the percentage of individuals to be randomly selected for the local search to improve them at each cycle. This approach is aligned with the proposal in Marić *et al.* (2014) since the local search is applied to a random subset of the population rather than to all of them, which increases diversity. As introduced, the local optimizer used is SASS, by Solis and Wets (1981), which is also the local method included in the memetic algorithms applied in Cruz *et al.* (2018), Puertas-Martín *et al.* (2020). Thus, every individual is a potential starting point for an independent execution of SASS.

As a method, SASS is a stochastic hill-climber with an adaptive step size that starts at a certain point of the search space, x. At the beginning of every iteration, SASS generates a new point, x', according to (4). The term ξ is a random perturbation vector in which every component i (there is one for each decision variable) follows a Gaussian distribution with component-specific mean b_i and common standard deviation σ (assuming a normalized search space), i.e. $\xi_i = \mathcal{N}(b_i, \sigma)$. Both, the b vector (also known as the bias) and σ, will be varied during the search. However, b starts as a zero vector, and σ is a user-given parameter.

$$x' = x + \xi. \tag{4}$$

Having generated x', SASS computes the value of the objective function at it. If x' represents a better solution, the algorithm moves its focus from x to x', and the iteration is considered successful. The b vector is then recomputed as $b = 0.2b + 0.4\xi$ for the next iteration. Otherwise, SASS explores the opposite direction by computing an alternative new point, $x'' = x - \xi$. Again, if the evaluation of x'' returns a better value for the objective function, SASS moves from x to x'', and the iteration is also considered successful. Under this circumstance, b is updated as $b = b - 0.4\xi$. However, if neither x' nor x'' were better than x, the iteration is supposed to be failed, and b is recomputed as $b = 0.5b$.

In contrast to the bias vector, the standard deviation of perturbation is not modified after every iteration but considering consecutive failures or successes for stability. Namely, if the number of consecutive successful iterations reaches a user-given parameter, *Scnt*, σ is expanded by a factor ex, which is also user-defined and supposed to be greater than 1, i.e. $\sigma = ex \cdot \sigma$. Analogously, if the number of consecutive failed iterations reaches a user-given parameter, *Fcnt*, σ is contracted by a factor c, which is also user-defined in $(0, 1) \in \mathbb{R}$, i.e. $\sigma = c \cdot \sigma$. However, notice that σ is also bounded by the user, and if it goes out from the valid range, σ is automatically reset to its upper bound.

SASS executes as many iterations, i.e. attempts to modify its current solution, as it can perform according to the number of objective function evaluations allowed by the user. After this process, the current solution of the method is finally returned. Under no circumstance will it move to a worse solution in the search space. Hence, the result of SASS will be the same initial point or a better one in the worst and best scenarios, respectively.

3.3. *Differential Evolution (DE)*

DE is a simple yet powerful genetic-like numerical optimizer that was proposed by Storn and Price (1997) and has become widely used (Dugonik *et al.*, 2019; Price *et al.*, 2006). It maintains a user-defined number (*NP*) of randomly-generated candidate solutions (individuals) and progressively alters them to find better ones. Like GA, every individual is a vector with a valid value for each optimization variable. The workflow of this method does not try to imitate aspects such as the selection of progenitors and sexual reproduction as closely as standard genetic methods like GA. However, the terminology of DE also comes from traditional Genetic Algorithms, so each iteration applies mutation, crossover, and selection stages to every individual.

The step of mutation follows (5) to compute for each individual S^j ($j = 1, \ldots, NP$), a mutant vector v_{S^j}. Both r_2 and r_3 are different and random integer indexes in the range $[1, \ldots, NP]$. Regarding r_1, it can be either another random population index or that of the best candidate solution in the population. The former is known as 'rand' strategy, and the latter is called 'best'. Regarding F, which controls the amplification of the differential variation, it is a user-given real and constant factor in the range $[0, 2]$ in the traditional method. However, it can be randomly redefined in the range $[0.5, 1]$ either for each iteration or for every mutant vector during the search. This approach is known as 'dither' and improves the rate of convergence with noisy objective functions. Finally, notice that the term $S^{r_2} - S^{r_3}$ defines a single difference vector between candidate solutions, but it is possible to use more than one. The most popular alternative uses two instead, which results in

$S^{r_2} - S^{r_3} + S^{r_4} - S^{r_5}$ (assuming that r_4 and r_5 are two random population indices more).

$$v_{S^j} = S^{r_1} + F\left(S^{r_2} - S^{r_3}\right). \tag{5}$$

The step of crossover merges each candidate solution, S^j, with its mutant vector, v_{S^j}. The result is a mutated version of S^j, $S^{j'}$, which is known as 'trial vector'. The computation follows (6), which is in terms of the i component of the trial vector. According to it, $S_i^{j'}$ can come from either the current individual S^j or its mutant vector. The component selection is controlled by the user-given crossover probability, $CR \in [0, 1]$, which is linked to uniform random number generation between 0 and 1. Regarding k, it is the index of one of the components defining the individuals, i.e. one of the optimization variables. This index is randomly selected for each computation of (6) to ensure that at least one of the components of the trial vector comes from the mutant one.

$$S_i^{j'} = \begin{cases} v_{S_i^j} & \text{if rand()} \leqslant CR \text{ or } i = k, \\ S_i^j & \text{otherwise.} \end{cases} \tag{6}$$

Additionally, there exists an additional stage that can be linked to crossover. It is a probabilistic random-alteration process proposed by Cabrera *et al.* (2011) to define a variant of DE aimed at mechanism synthesis. However, the addition is not limited to that field but is of general-purpose in reality. It entails an in-breadth search component that increases variability and is similar to the traditional mutation phase of genetic algorithms such as GA. This addition aims to help the algorithm to escape from local optima, especially when working with small populations. In practical terms, it follows (7) to make more changes to each trial vector. According to it, each i component of the input trial vector can be randomly redefined around its current value with a user-given probability. The origin of the components, i.e. the current individual or its associated mutant vector, is not relevant. Therefore, this stage only adds two expected parameters: the per-component mutation probability, MP, and the modification range, $range$.

$$S_i^{j'} = \begin{cases} rand() \in [S_i^{j'} \pm range] & \text{if rand()} \leqslant MP, \\ S_i^{j'} & \text{otherwise.} \end{cases} \tag{7}$$

After the previous steps, DE evaluates each trial vector as a solution according to the objective function. Only those trial vectors that outperform the candidate solutions from which they come persist and replace the original individuals. This process defines the population for the next iteration. DE concludes after having executed the user-defined number of iterations. At that point, the algorithm returns the best individual in the population as the solution found.

3.4. *Teaching-Learning-Based Optimization (TLBO)*

TLBO is a recent numerical optimization method proposed by Rao *et al.* (2012). As a population-based meta-heuristic, this algorithm also works with a set of randomly-generated candidate solutions. However, instead of representing a group of individuals

that evolve through sexual reproduction like the previous method, TLBO treats the set of solutions as a group of students. The algorithm simulates their academic interaction, considering a teacher and plain pupils, to find the global optimum of the problem at hand. This optimizer has attracted the attention of many researchers in different fields, from Engineering to Economics, because of its simplicity, effectiveness and minimalist set of parameters (Cruz *et al.*, 2017; Rao, 2016).

The algorithm only requires two parameters to work, namely, the population size and the number of iterations to execute. Based on this information, TLBO first creates and evaluates as many random candidate solutions or individuals as indicated by the population size. As it occurs with the previous method, every individual is a plain vector containing a valid value for each optimization variable of the objective function. After this initial step, TLBO executes as many iterations as required, and every one consists of the teacher and learner phases.

The teacher phase models the way in which a professor improves the skills of his/her students. In practical terms, this step tries to shift each candidate solution towards the best one in the population, which becomes the teacher, T. After having identified that reference solution, TLBO computes the vector M. Every i component in M results from averaging those of the individuals in the current population. This information serves to create an altered or shifted solution, S', from every existing one, S, according to (8). The computation is defined in terms of every component or optimized variable, i. r_i is a random real number in range [0, 1] and linked to component i. Similarly, T_F, known as 'teaching factor', is a random integer that can be either 1 or 2. Both, r_i and T_F, are globally computed for the current step. Finally, every S' is evaluated and replaces S if it obtains a better value from the objective function (S' is discarded otherwise).

$$S'_i = S_i + r_i(T_i - T_F M_i). \tag{8}$$

The learner phase simulates the interaction between students to improve their skills. At this step, TLBO pairs every student, S, with any other different one, W. The goal is to generate a modified individual, S', which will replace S if it is a better solution according to the objective function. Every i component of S' results from (9), where r_i is a random number in range [0, 1] linked to component i and globally computed for the current step. According to it, the movement direction in every pair goes from the worst solution to the best.

$$S'_i = \begin{cases} S_i + r_i(S_i - W_i) & \text{if } error(S) < error(W), \\ S_i + r_i(W_i - S_i) & \text{otherwise.} \end{cases} \tag{9}$$

Additionally, although it is usually omitted, there is also an internal auxiliary stage to remove duplicate solutions in the population (Waghmare, 2013). At the end of every iteration, duplicity is avoided by randomly altering and re-evaluating any of the involved candidate solutions. Finally, after the last iteration, TLBO returns the best solution in the population as the result of the problem at hand.

3.5. *Multi-Start Single-Agent Stochastic Search (MSASS)*

The Multi-Start Single-Agent Stochastic Search, as introduced, consists in the inclusion of the SASS algorithm (already described as part of MemeGA in Section 3.2) within a standard multi-start component (Redondo *et al.*, 2013; Salhi, 2017). The former is a simple yet effective local search method that will always try to move from a starting point to a nearby better solution. The latter is in charge of randomly generating different initial points, provided that the computational effort remains acceptable.

Considering the local scope of SASS, the multi-start component serves to escape from local optima by focusing the local exploration on different regions of the search space. Namely, at each iteration, the multi-start module randomly defines a feasible point. It then independently starts SASS from there. Once the local search has finished, so does the iteration, and the process is repeated from a new initial point while keeping track of the best solution found so far, which will be finally returned as the result. As a compound method, MSASS works until consuming a user-given budget of objective function evaluations.

The multi-start component expects from the user the total number of function evaluations allowed. It will always launch SASS with its maximum number of evaluations set to the remainder to consume all the budget, and the rest of its traditional parameters as defined by the user (see Section 3.2). However, in this context, SASS is enhanced with an extra parameter, i.e. the maximum number of consecutive failures that force the method to finish without consuming all the remaining function evaluations. By proceeding this way, once SASS has converged to a particular solution, the method will not waste the rest of the evaluations. Instead, it will be able to return the control to the multi-start component to look for a new start. This aspect is of great importance to increase the probabilities of finding an optimal solution.

4. Experimentation and Results

4.1. *Environment and Configuration*

The present study inherits the technical framework defined by the reference work in Marín *et al.* (2020). Consequently, the GrC model is simulated with the neural simulation tool NEST (Plesser *et al.*, 2015), in its version 2.14.0, and using the 4th-order Runge-Kutta-Fehlberg solver with adaptive step-size to integrate the differential equations. The cost function, known as FF4 in the reference work as introduced, is implemented in Python (version 2.7.12) and linked to the NEST simulator. The simulation environment uses a common and fixed random-generation seed to define a stable framework and homogenize the evaluations. In this context, the GA method by Marín *et al.* (2020) was also implemented in Python 2.7.12 using the standard components provided by the DEAP library (Kim and Yoo, 2019).

Marín et al. configured their GA method to work with a population of 1000 individuals for 50 generations and selection tournaments of size 3. They adjusted the crossover, mutation, and per-component mutation probabilities to 60%, 10%, and 15%, respectively.

Table 2
Varied parameters for each population-based method and computational cost.

		Computational cost (f.e.)			
Method	Parameters	15k	22.5k	30k	60k
GA	Population	1500	1500	1000	1000
	Cycles	16	25	50	100
MemeGA	Population	500	500	300	300
	Local f.e.	10	12	20	35
	Cycles	20	26	40	50
DE	Population	250	250	200	150
	Cycles	60	90	150	400
TLBO	Population	250	250	200	150
	Cycles	30	45	75	200

Approximately, this configuration results in 30 000 evaluations of the objective function (f.e.) and also defines a general reference in computational effort. This aspect is especially relevant since the development and execution platforms differ from the original work. Namely, the four new optimizers compared (MemeGA, DE, TLBO, and MSASS) run in MATLAB 2018b after having wrapped the objective function of Marín et al. to be callable from it. The experimentation machine features an Intel Core i7-4790 processor with 4 cores and 32 GB of RAM.

The comparison takes into account four different computational costs in terms of the reference results: 50%, 75%, 100%, and 200% of the function evaluations consumed by GA, i.e. 15 000, 22 500, 30 000, and 60 000, respectively. Nevertheless, before launching the final experiments, the population-based optimizers have been tested under different preliminary configurations to find their most robust set of parameters. After having adjusted them, the focus was moved to those parameters directly affecting the overall computational cost, i.e. function evaluations. In practical terms, those are limited to the population sizes and the number of cycles for DE and TLBO. Similarly, they are the only parameters of the reference method, GA, that have been ultimately re-defined to encompass the new execution cases. However, for its memetic version, MemeGA, the number of objective function evaluations for each independent local search (local f.e.) has also been varied with the computational cost. In contrast to them, MSASS is directly configured in terms of function evaluations. Its local search component has kept the constant configuration explained below.

Table 2 shows the previous information for each population-based optimizer and computational cost. Concerning the population sizes, the general paradigm followed opted for spawning larger populations with the lower computational costs to increase and speed up the exploration. More precisely, when the computational cost is below 100% of the reference method, the population is larger to accelerate the identification of promising regions in the search space and to compensate for the impossibility of allowing the optimizer to run more cycles. Numerically, the population sizes of GA remain in the range of the reference paper for the most demanding cases and get an increment of 50% for the others. Those of MemeGA come from GA but are approximately divided by 3 to allow for the

function evaluations of its independent instances of SASS, which grow from 10 to 35 for the 60k configuration. The population sizes of DE are in the range of multiplying by ten the number of variables, as suggested by the authors, and doubled to maximize diversity. TLBO successfully shares the same strategy.

Regarding the number of cycles, it is adapted from the corresponding population size to achieve the desired computational effort. GA has statistical components, but its approximate number of function calls is 60% of the population size each cycle. MemeGA keeps the crossover and mutation probabilities of GA and redefines the latter as the percentage of individuals to be locally improved. Hence, it approximately executes 60% of the population size plus 10% of the referred value multiplied by the number of local evaluations each cycle. DE consumes as many f.e. as individuals per cycle, and TLBO takes twice. The cost of evaluating the initial populations is neglected.

Concerning the rest of the parameters that have been fixed, the internal SASS component of MemeGA considers $Scnt = 5$, $Fcnt = 3$, $ex = 2.0$, and $c = 0.5$ with $\sigma \in \{1e{-}5, 0.25\}$. This is mainly the configuration recommended by Solis and Wets (1981) with the only exception of the upper bound of sigma being 0.25 instead of 1. It is because preliminary experimentation revealed that when allowing few function evaluations, most individuals were directly moved to the bounds by big perturbations without the possibility to improve. Thus, the smaller upper bound of σ allows for a better local exploration within MemeGA. However, for MSASS, all the parameters of SASS coincide with the recommended values because its execution approach does not have to share resources. That said, every instance of SASS will be stopped after 50 non-improving or failed iterations to save function evaluations for later independent starts. Regarding DE, after preliminary experimentation, it has been configured to use 'rand' selection, a single difference vector, and per-generation dither. Notice that the latter aspect avoids adjusting the F parameter of the optimizer and takes into account its potential convergence advantages. Finally, the crossover rate has been fixed to 0.8, which is in the general-purpose range and between the values used in the extension proposed in Cabrera *et al.* (2011).

4.2. *Numerical Results*

Table 3 contains the results achieved by each method. There is a column for every optimizer, and each one consists of groups of cells that cover the different scenarios of computational cost. The algorithms have been independently launched 20 times for each case considering their stochastic nature, i.e. their results might vary even for the same problem instance and configuration. The referred table includes the average (ave.) and standard deviation (STD) of each case. Since the problem addressed is of minimization, the lower the values are, the better result they represent. The best average of each computational cost is in bold font. Notice that the results of GA and 30 000 f.e. (i.e. 100%) combine the five ones obtained by Marín *et al.* (2020) plus the remaining executions to account for 20 cases in total. Finally, to orientate the readers about the real computational cost, it is interesting to mention that completing each cell of 15 000 f.e. approximately takes 9 hours in the experimentation environment and scales accordingly.

Table 3
Objective function values achieved for each compared optimizer.

Effort (f.e.)	GA	MemeGA	DE	TLBO	MSASS
15 000 (50%)	Ave.: 115.38 STD: 4.19	Ave.: 113.51 STD: 10.83	Ave.: 115.78 STD: 7.38	Ave.: **105.76** STD: 5.90	Ave.: 106.88 STD: 4.50
22 500 (75%)	Ave.: 110.32 STD: 5.98	Ave.: 111.46 STD: 8.03	Ave.: 112.18 STD: 5.96	Ave.: **104.61** STD: 7.09	Ave.: 104.71 STD: 5.29
30 000 (100%)	Ave.: 108.70 STD: 6.59	Ave.: 103.45 STD: 4.47	Ave.: 107.98 STD: 4.74	Ave.: **98.63** STD: 5.97	Ave.: 101.20 STD: 3.52
60 000 (200%)	Ave.: 104.17 STD: 6.70	Ave.: 100.06 STD: 5.60	Ave.: 99.44 STD: 4.08	Ave.: **97.67** STD: 3.37	Ave.: 100.52 STD: 3.04

4.3. *Discussion*

From the results in Table 3, it is possible to make several overall appreciations. Firstly, none of the optimizers stably converges to a single global solution even after doubling the computational effort allowed in the reference paper. This aspect indicates that the search space is hard to explore and features multiple sub-optimal points. For this reason, ensuring stable and global convergence might not be feasible in a reasonable amount of time. Fortunately, it is not a requirement in this context as long as the found configurations make the simulated neurons behave as expected. Secondly, all the optimizers tend to improve the average quality of their results when the computational cost increases, but TLBO always achieves the best average quality. It is also the method that has found the best individual solution known so far, with a value of 87.45. Thirdly and last, the reduction in standard deviations is not as regular as that of the averages, but for most methods, the standard deviation of the results after 60 000 f.e. is approximately half of that observed for 15 000 f.e. The only exception is the reference optimizer, GA, whose STD was better for the lightest configuration than for the heaviest. Hence, in general, the probability of obtaining particularly divergent results in quality is reduced when the optimizers are provided with enough computational budget.

As previously commented, all the methods provide better average results after increasing the computational budget, but they effectively evolve at different rates. Certainly, the average quality achieved by a certain optimizer with a particular computational cost turns out to be nearly equivalent to that of another one. Nevertheless, some of the algorithms require more computational effort to be at the same level as others. For instance, the average of GA for 60 000 f.e. is very similar to those of MSASS and TLBO for 22 500 f.e. In fact, between 15 000 and 30 000 f.e., MSASS and TLBO are a step beyond the rest. Moreover, not all the methods ultimately achieve the same degree of quality. More specifically, for 60 000 f.e., MemeGA, DE and MSASS perform similarly, but GA is worse than them, and TLBO remains numerically ahead. Thus, according to the computational cost, some of the methods stand out from the rest. TLBO and MSASS do it positively with the best averages, and GA does it negatively considering the two heaviest configurations. Regarding MemeGA and DE, the former starts to outperform the reference method with 30 000 f.e., and the latter does it after doubling this value.

After the preliminary analysis, it is necessary to test whether the methods exhibit statistically significant differences considering the impact of uncertainty. For this reason, the individual results have been studied according to the Kruskal-Wallis test (Kruskal and Wallis, 1952; Mathworks, 2021), which is a non-parametric method for testing whether the samples come from the same distribution. By proceeding this way, it is possible to assess if the results registered for each optimizer and cost seem to significantly differ from each other with reduced datasets and without making distribution assumptions. The overall significance of the tests is 0.05, i.e. the corresponding confidence level is 95%.

Observing in depth the results in Table 3 for the lightest computational cost, it seems that there are two groups. Specifically, TLBO and MSASS exhibit the best performance with close values between them, while GA, MemeGA, and DE show worse averages (and also numerically similar between them). According to the Kruskal-Wallis test, there is no significant statistical difference within each group either. Thus, for 15 000 f.e., TLBO and MSASS return indistinguishable results within a significance of 0.05. In other words, for half of the computational budget of the reference work, TLBO and MSASS are equivalent according to the results achieved. Moreover, both perform better than the rest with a statistically significant difference, including the reference method, GA. Analogously, the performances of GA, MemeGA, and DE are equivalent in this context, so there is no more expected difference than the effect of uncertainty between the three. Additionally, the fact that the memetic variant of GA, MemeGA, does not significantly outperform it can be attributed to the lack of computational budget to let local search be effective.

For 22 500 f.e., the previous situation persists: TLBO and MSASS define the group of the best-performing optimizers, without significant difference between using one or the other. GA, MemeGA, and DE remain the worst-performing methods without practical difference between them. However, for 30 000 f.e., the situation changes: MemeGA moves from the group of GA and DE to that of TLBO and MSASS, which feature the best average. At this point, the methods achieving the worst results are GA and DE, which keep being statistically indistinguishable. Those with better performance are now TLBO, MSASS, and MemeGA. Among them, TLBO and MSASS keep being statistically indistinguishable within a significance of 0.05, but the same occurs between MemeGA and MSASS. Finally, the previous trend is confirmed with 60 000 f.e: another method, DE, separates from the reference one, GA, and outperforms it. Therefore, GA persists as the only member of worse-performing methods in the end, while MemeGA, DE, TLBO, and MSASS become statistically indistinguishable between them and achieve better results than the reference method.

Based on the previous analysis, TLBO and MSASS are the best choices for 50 and 75% of the function evaluations allowed in the reference paper. For the same computational budget considered by Marín *et al.* (2020), the performance of MemeGA reaches the level of those two referred methods, which remain ahead. Finally, with the highest computational cost, MemeGA, DE, TLBO, and MSASS significantly outperform the reference method, and there is no meaningful difference between the four. Thus, the reference method, GA, is always outperformed by at least TLBO and MSASS. MemeGA and DE also tend to separate from it to join the other two at 30 000 and 60 000 f.e., respectively.

Additionally, for the sake of completeness, the significance of the difference between the results of each method has also been systematically assessed with the Kruskal-Wallis test under the previous significance level. According to the study, there is no statistically significant difference between the results of GA after 22 500 and 30 000 f.e. Hence, it seems possible to reduce the computational effort in the reference work by 25% without expecting worse results for any cause apart from stochasticity. TLBO, MemeGA and DE experience the same phenomenon, but they do between 15 000 and 22 500 f.e, while the performance of the last two is still similar to that of GA. Finally, MSASS stagnates between 30 000 and 60 000 f.e., which is well aligned with its conceptual simplicity and lack of sophisticated components to globally converge.

4.4. *Insight into the Best Solution*

To conclude this section, the best result found will be analysed with further details in Fig. 2. The best-fitted neuron model according to all the features (with the lowest score) is an individual obtained by TLBO featuring a quality of 87.45 (blue lines). The neuron model from the reference work (GA optimizer) (Marín *et al.*, 2020) with a score of 104.24 is also compared (orange lines). Finally, the *in vitro* data used to define the fitness function are also represented (black dots) (D'Angelo *et al.*, 2001; Masoli *et al.*, 2017).

The selected neuron model has successfully captured the well-demonstrated features of the intrinsic excitability of cerebellar GrCs, i.e. repetitive spike discharge in response to injected currents (implemented as the mean frequency), latency to the first spike upon current injection (implemented as the time to the first spike), and spiking resonance in the theta-range (implemented as the average burst frequencies in response to sinusoidal stimulations). The neuron model resonates in the theta frequency band as expected, i.e. 8–12 Hz (Fig. 2(A)). The model practically reproduces identical resonance curves as the model of reference (GA model) (Marín *et al.*, 2020). These resonance curves are the graphical representation of doublets, triplets, or longer bursts of spikes generated when stimulated by just-threshold sinusoidal stimulation at different frequencies (Fig. 2(B)). The main behaviour of biological GrCs is the increase of spike frequency when the latency to the first spike decreases as current injections increase. A sample of the neuron behaviour from which these features are calculated is shown in Fig. 2(E). The repetitive spike discharge of the TLBO model is similar to that of the model of reference and in accordance with the experimental measurements in real cells (Fig. 2(C)). The real improvement obtained by the neuron model of the proposed optimizer lies in the first-spike latency feature. The model from the reference work exhibited longer latencies than those experimentally reported, mainly with low stimulation currents (Fig. 2(D)). However, the TLBO model achieves an adjustment in its score of 7.95 ms compared to the 34.95 ms-error obtained by the GA model, both with respect to the *in vitro* data. Thus, the TLBO optimizer proves not only to be effective in fitting the model parameters to diverse spiking features, but also to improve both the quantitative and qualitative predictions of these supra-threshold characteristics against the methodology of reference (GA).

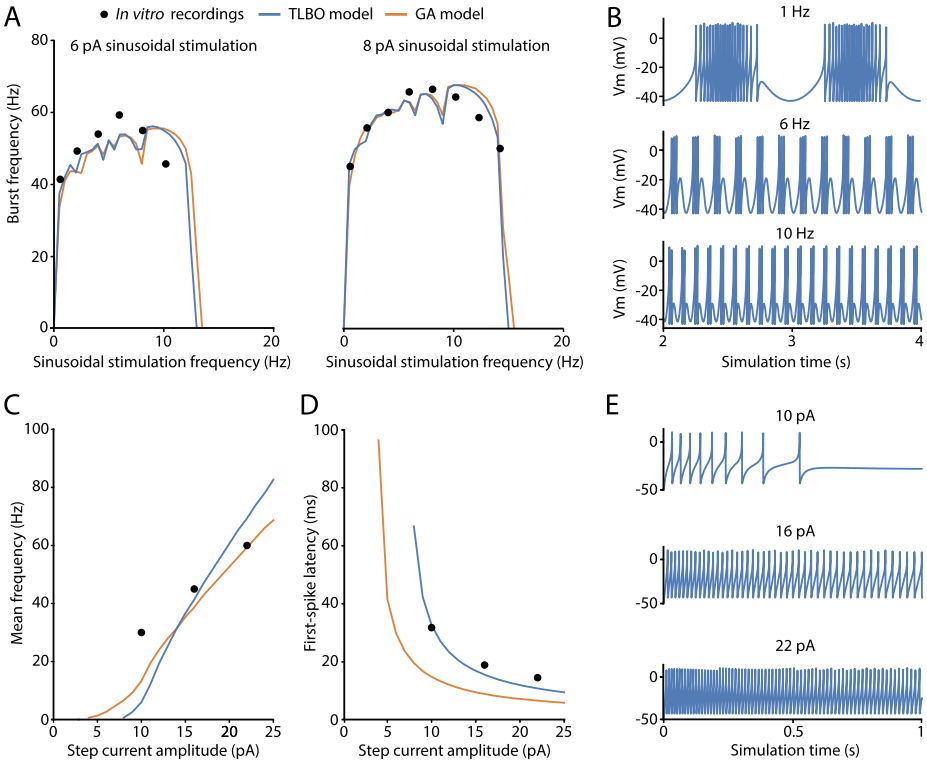


Fig. 2. Spiking properties predicted by the best-fitted cerebellar granule cell (GrC) model obtained by TLBO optimizer. Simulated features and electrophysiological traces of the best solution from TLBO (blue) and compared to the neuron model from the reference work (orange) and the target *in vitro* recordings used in the fitness function (black dots). A) Spiking resonance curves of the models computed as average burst frequencies in response to sinusoidal stimulation of 6 pA (left) and 8 pA (right) with increasing frequencies (in steps of 0.5 Hz). B) Membrane potential evolution of the TLBO model generates spike bursts in response to sinusoidal current injections with offset of 10 pA and amplitude of 6 pA. This is shown after 2 s of stimulation (stabilization). C) Repetitive spike discharge (intensity-frequency curves) of the models computed as the mean frequency in response to step-currents of 1 s. D) Latency to the first spike in response to step-currents of 1 s. E) Membrane potential traces of the TLBO model in response to step-current injections of 1 s with various amplitudes.

5. Conclusions and Future Work

This article has studied the optimization component of a recent methodology for realistic and efficient neuron modelling. The referred method focuses on single-neuron processing, relies on the use of the adaptive-exponential integrate-and-fire (AdEx) model, and has been applied to cerebellar granule cells. It requires fitting its parameters to mimic experimental recordings, which can be defined as an optimization problem of ten variables. The original paper implemented a traditional genetic algorithm (GA) to address the resulting problem. This work has compared that optimizer to four alternatives: an ad-hoc memetic version of the referred genetic algorithm (MemeGA), Differential Evolution (DE), Teaching-Learning-Based Optimization (TLBO), and a multi-start method built on the same local optimizer used for MemeGA (MSASS).

All of the methods are stochastic heuristic algorithms based on solid principles that exhibit high success rates in different problems, which is why they have been selected. The comparison has been performed in terms of the context defined by the reference paper. It has considered four computational budgets for the problem resolution: 50, 75, 100, and 200% of that invested in the referred work. The effect of stochasticity at optimization has been considered by making 20 independent executions for each method and computational budget. Moreover, the assessment has not been limited to comparing their average quality. It has been complemented by studying whether the different sets of results significantly differ according to the Kruskal-Wallis test.

The original hypothesis for this work was that the optimization engine defined for the referred method could be improved, and it has been confirmed. The reference optimizer is outperformed by some of the candidates in all the scenarios, while the others equal it at least. Moreover, they tend to positively differentiate from GA as the computational budget increases. Namely, either TLBO or MSASS is equally the best choice for 15 000 and 22 500 function evaluations. Once the computational budget is the same as in the reference work, i.e. 30 000 function evaluations, MemeGA joins TLBO and MSASS. Before that point, there were no enough function evaluations to fully exploit the theoretical advantage of using a dedicated local search component over its original version. Then, no unambiguous evidence support one of them over the other two, but the three outperform GA and DE in that scenario. Ultimately, with 60 000 f.e., DE also achieves better results than GA and merges itself with TLBO, MSASS and MemeGA.

The real benefit of increasing the computational cost has also been separately studied. According to the analysis performed, the results achieved by the reference method do not significantly vary between 22 500 and 30 000 function evaluations, so the computational cost of just applying GA can be reduced by 25% without expecting a reduction in its performance. Regarding MemeGA, DE, and TLBO, they experience the same situation between 15 000 and 22 500 function evaluations. Similarly, the computational effort if using any of them could be reduced by 25% without executing more than 15 000 evaluations (logically, unless the budget can reach 30 000). Concerning MSASS, this phenomenon occurs between 30 000 and 60 000 evaluations. In this situation, 50% of the computing time could be saved by not doubling the budget for this method.

Additionally, the best configuration known so far has been found at experimentation. TLBO is the method that achieved it, and the resulting model features higher temporal accuracy of the first spike than that of the reference paper. This aspect is key for the reproduction of the relevant properties that could play a role in neuronal information transmission. This finding supports the relevance of using an effective and efficient optimization engine in the referred methodology. The gain in biological realism in simple neuron models is expected to allow the future simulation of networks compounded of thousands of these neurons to better mimic the biology. Obtaining more realistic yet efficient neuron models also allows research at levels in which *in vitro* or *in vivo* experimental biology is limited. Thus, simulations of sufficiently realistic neuronal network models can become valid to shed light on the functional roles of certain neuronal characteristics or on the interactions that may have various mechanisms among each other.

In future work, the existence of multiple sub-optimal solutions will be further studied. For that purpose, the aim is to use a state-of-the-art multi-modal optimization algorithm that can keep track of the different regions throughout its execution. That study might identify patterns that allow reducing the search space proposed in the reference work.

Acknowledgements

The authors would like to thank R. Ferri-García, from the University of Granada, for his advice on Statistics. The results included in this article are part of Milagros Marín's PhD thesis.

Funding

This research has been funded by the Human Brain Project Specific Grant Agreement 3 (H2020-RIA. 945539), the Spanish Ministry of Economy and Competitiveness (RTI2018-095993-B-I00), the National Grant INTSENSO (MICINN-FEDER-PID2019-109991GB-I00), the Junta de Andalucía (FEDER-JA P18-FR-2378, P18-RT-1193), and the University of Almería (UAL18-TIC-A020-B).

References

Barranca, V.J., Johnson, D.C., Moyher, J.L., Sauppe, J.P., Shkarayev, M.S., Kovačič, G., Cai, D. (2014). Dynamics of the exponential integrate-and-fire model with slow currents and adaptation. *Journal of Computational Neuroscience*, 37(1), 161–180.

Boussaïd, I., Lepagnot, J., Siarry, P. (2013). A survey on optimization metaheuristics. *Information Sciences*, 237, 82–117.

Brette, R., Gerstner, W. (2005). Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *Journal of Neurophysiology*, 94(5), 3637–3642.

Cabrera, J.A., Ortiz, A., Nadal, F., Castillo, J.J. (2011). An evolutionary algorithm for path synthesis of mechanisms. *Mechanism and Machine Theory*, 46(2), 127–141.

Cruz, N.C., Redondo, J.L., Álvarez, J.D., Berenguel, M., Ortigosa, P.M. (2017). A parallel teaching–learning-based optimization procedure for automatic heliostat aiming. *Journal of Supercomputing*, 73(1), 591–606.

Cruz, N.C., Redondo, J.L., Álvarez, J.D., Berenguel, M., Ortigosa, P.M. (2018). Optimizing the heliostat field layout by applying stochastic population-based algorithms. *Informatica*, 29(1), 21–39.

D'Angelo, E., Nieus, T., Maffei, A., Armano, S., Rossi, P., Taglietti, V., Fontana, A., Naldi, G. (2001). Theta-frequency bursting and resonance in cerebellar granule cells: experimental evidence and modeling of a slow K+-dependent mechanism. *Journal of Neuroscience*, 21(3), 759–770.

D'Angelo, E., Koekkoek, S.K.E., Lombardo, P., Solinas, S., Ros, E., Garrido, J., Schonewille, M., De Zeeuw, C.I. (2009). Timing in the cerebellum: oscillations and resonance in the granular layer. *Neuroscience*, 162(3), 805–815.

Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press, London.

Delvendahl, I., Straub, I., Hallermann, S. (2015). Dendritic patch-clamp recordings from cerebellar granule cells demonstrate electrotonic compactness. *Frontiers in Cellular Neuroscience*, 9, 93.

Dugonik, J., Bošković, B., Brest, J., Sepesy Maučec, M. (2019). Improving statistical machine translation quality using differential evolution. *Informatica*, 30(4), 629–645.

Gandolfi, D., Lombardo, P., Mapelli, J., Solinas, S., D'Angelo, E. (2013). Theta-frequency resonance at the cerebellum input stage improves spike timing on the millisecond time-scale. *Frontiers in Neural Circuits*, 7, 64.

Hanuschkin, A., Kunkel, S., Helias, M., Morrison, A., Diesmann, M. (2010). A general and efficient method for incorporating precise spike times in globally time-driven simulations. *Frontiers in Neuroinformatics*, 4, 113.

Holland, J.H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, USA.

Jörntell, H., Ekerot, C.F. (2006). Properties of somatosensory synaptic integration in cerebellar granule cells in vivo. *Journal of Neuroscience*, 26(45), 11786–11797.

Kim, J., Yoo, S. (2019). Software review: DEAP (Distributed Evolutionary Algorithm in Python) library. *Genetic Programming and Evolvable Machines*, 20(1), 139–142.

Kruskal, W.H., Wallis, W.A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), 583–621.

Lange, W. (1975). Cell number and cell density in the cerebellar cortex of man and some other mammals. *Cell and Tissue Research*, 157(1), 115–124.

Lindfield, G., Penny, J. (2017). *Introduction to Nature-Inspired Optimization*. Academic Press, USA.

Marić, M., Stanimirović, Z., Djenić, A., Stanojević, P. (2014). Memetic algorithm for solving the multilevel uncapacitated facility location problem. *Informatica*, 25(3), 439–466.

Marín, M., Sáez-Lara, M.J., Ros, E., Garrido, J.A. (2020). Optimization of efficient neuron models with realistic firing dynamics. The case of the cerebellar granule cell. *Frontiers in Cellular Neuroscience*, 14, 161.

Masoli, S., Rizza, M.F., Sgritta, M., Van Geit, W., Schürmann, F., D'Angelo, E. (2017). Single neuron optimization as a basis for accurate biophysical modeling: the case of cerebellar granule cells. *Frontiers in Cellular Neuroscience*, 11, 71.

Mathworks (2021). Kruskal-Wallis test. https://www.mathworks.com/help/stats/kruskalwallis.html. [Last access: March, 2021].

Moscato, P. (1989). *On evolution, search, optimization, genetic algorithms and martial arts: towards memetic algorithms*. Technical report, Caltech concurrent computation program, C3P Report.

Nair, M., Subramanyan, K., Nair, B., Diwakar, S. (2014). Parameter optimization and nonlinear fitting for computational models in neuroscience on GPGPUs. In: *2014 International Conference on High Performance Computing and Applications (ICHPCA)*. IEEE, pp. 1–5.

Naud, R., Marcille, N., Clopath, C., Gerstner, W. (2008). Firing patterns in the adaptive exponential integrate-and-fire model. *Biological Cybernetics*, 99(4–5), 335.

Plesser, H.E., Diesmann, M., Gewaltig, M.O., Morrison, A. (2015). *NEST: The Neural Simulation Tool*. Springer, New York, USA, pp. 1849–1852. 978-1-4614-6675-8. https://doi.org/10.1007/978-1-4614-6675-8_258.

Price, K., Storn, R.M., Lampinen, J.A. (2006). *Differential Evolution: A Practical Approach to Global Optimization*. Springer Science & Business Media, Germany.

Puertas-Martín, S., Redondo, J.L., Pérez-Sánchez, H., Ortigosa, P.M. (2020). Optimizing electrostatic similarity for virtual screening: a new methodology. *Informatica*, 31(4), 821–839.

Rao, R.V. (2016). Applications of TLBO algorithm and its modifications to different engineering and science disciplines. In: *Teaching Learning Based Optimization Algorithm*. Springer, Switzerland, pp. 223–267.

Rao, R.V., Savsani, V.J., Vakharia, D.P. (2012). Teaching–learning-based optimization: an optimization method for continuous non-linear large scale problems. *Information Sciences*, 183(1), 1–15.

Redondo, J.L., Arrondo, A.G., Fernández, J., García, I., Ortigosa, P.M. (2013). A two-level evolutionary algorithm for solving the facility location and design (1| 1)-centroid problem on the plane with variable demand. *Journal of Global Optimization*, 56(3), 983–1005.

Salhi, S. (2017). *Heuristic Search: The Emerging Science of Problem Solving*. Springer, Switzerland.

Schmahmann, J.D. (2019). The cerebellum and cognition. *Neuroscience Letters*, 688, 62–75.

Shopova, E.G., Vaklieva-Bancheva, N.G. (2006). BASIC—A genetic algorithm for engineering problems solution. *Computers & Chemical Engineering*, 30(8), 1293–1309.

Solis, F.J., Wets, R.J.B. (1981). Minimization by random search techniques. *Mathematics of Operations Research*, 6(1), 19–30.

Storn, R., Price, K. (1997). Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4), 341–359.

Venkadesh, S., Komendantov, A.O., Listopad, S., Scott, E.O., De Jong, K., Krichmar, J.L., Ascoli, G.A. (2018). Evolving simple models of diverse intrinsic dynamics in hippocampal neuron types. *Frontiers in Neuroinformatics*, 12, 8.

Waghmare, G. (2013). Comments on "A note on Teaching–Learning-Based Optimization algorithm". *Information Sciences*, 229, 159–169.

Williams, R.W., Herrup, K. (1988). The control of neuron number. *Annual Review of Neuroscience*, 11(1), 423–453.

N.C. Cruz is a researcher at the Supercomputing – Algorithms (SAL) Research Group at the University of Almería, Spain. He obtained his PhD from the University of Almería. His publications can be found on https://publons.com/researcher/1487279/nc-cruz/. His research interests include high-performance computing, global optimization and applications. Personal web page: http://hpca.ual.es/~ncalvo/.

M. Marín is a predoctoral student at the department of Biochemistry and Molecular Biology at the University of Granada. She is currently pursuing her PhD within the Applied Computational Neuroscience group at the Research Centre for Information and Communications Technologies. She is a young researcher participating in the international project Human Brain Project (HBP). Some of her publications are available at https://publons.com/researcher/3213167/milagros-marin/. Her interdisciplinary research interests are located between Health and Biochemistry (Cerebellum, Molecular Biology and Biomedicine) and Information and Communication Technologies (Computational Neuroscience and Bioinformatics). Personal web page: http://acn.ugr.es/people/mmarin/.

J.L. Redondo is an assistant professor at the Informatics Department at the University of Almería, Spain. She obtained her PhD from the University of Almería. Her publications can be found on https://www.scopus.com/authid/detail.uri?authorId=35206862500. Her research interests include high performance-computing, global optimization and applications. Personal web page: https://sites.google.com/ual.es/jlredondo.

E. M. Ortigosa is an assistant professor at the Computer Architecture and Technology Department at the University of Granada, Spain. She received her PhD degree in computer engineering from the University of Málaga, Spain. She has participated in the creation of the spin-off company Seven Solutions, S.L. It is an EBT (Technology-Based Company) that has received numerous awards. Her research interests include computational neuroscience and efficient network simulation methods, bioinformatics, and hardware implementation of digital circuits for real time processing in embedded systems. Personal web page: https://atc.ugr.es/informacion/directorio-personal/eva-martinez-ortigosa.

P.M. Ortigosa is a full professor at the Informatics Department at the University of Almería, Spain. She obtained her PhD from the University of Málaga, Spain. Her publications can be found on https://www.scopus.com/authid/detail.uri?authorId=6602759441. Her research interests include high-performance computing, global optimization and applications. Personal web page: https://sites.google.com/ual.es/pmortigosa.

4. On the use of a multimodal optimizer for fitting neuron models. Application to the cerebellar granule cell

Authors	**M. Marín***, N. C. Cruz, E. M. Ortigosa, M. J. Sáez-Lara, J. A Garrido & R. R. Carrillo		
Journal	*Frontiers in Neuroinformatics*		
Year	2021		
Volume	15		
Pages	17		
DOI	10.3389/fninf.2021.663797		
Impact Factor (JCR 2019)	2.649		
Categories	Mathematical & Computational Biology:	Ranking: 13/59	**(Q1)**
	Neurosciences:	Ranking: 164/272	(Q3)

frontiers in Neuroinformatics

ORIGINAL RESEARCH
published: 03 June 2021
doi: 10.3389/fninf.2021.663797

On the Use of a Multimodal Optimizer for Fitting Neuron Models. Application to the Cerebellar Granule Cell

Milagros Marín[1], Nicolás C. Cruz[2], Eva M. Ortigosa[3], María J. Sáez-Lara[1], Jesús A. Garrido[3] and Richard R. Carrillo[3]*

[1]Department of Biochemistry and Molecular Biology I, University of Granada, Granada, Spain, [2]Department of Informatics, University of Almería, ceiA3, Almería, Spain, [3]Department of Computer Architecture and Technology—CITIC, University of Granada, Granada, Spain

OPEN ACCESS

Edited by:
Eilif Benjamin Muller,
University of Montreal, Canada

Reviewed by:
Sergio Mauro Solinas,
ETH Zurich, Switzerland
Werner Van Geit,
École Polytechnique Fédérale de Lausanne, Switzerland
Salvador Dura-Bernal,
SUNY Downstate Medical Center, United States

***Correspondence:**
Milagros Marín
mmarin@ugr.es

Received: 03 February 2021
Accepted: 13 April 2021
Published: 03 June 2021

Citation:
Marín M, Cruz NC, Ortigosa EM, Sáez-Lara MJ, Garrido JA and Carrillo RR (2021) On the Use of a Multimodal Optimizer for Fitting Neuron Models. Application to the Cerebellar Granule Cell. Front. Neuroinform. 15:663797. doi: 10.3389/fninf.2021.663797

This article extends a recent methodological workflow for creating realistic and computationally efficient neuron models whilst capturing essential aspects of single-neuron dynamics. We overcome the intrinsic limitations of the extant optimization methods by proposing an alternative optimization component based on multimodal algorithms. This approach can natively explore a diverse population of neuron model configurations. In contrast to methods that focus on a single global optimum, the multimodal method allows directly obtaining a set of promising solutions for a single but complex multi-feature objective function. The final sparse population of candidate solutions has to be analyzed and evaluated according to the biological plausibility and their objective to the target features by the expert. In order to illustrate the value of this approach, we base our proposal on the optimization of cerebellar granule cell (GrC) models that replicate the essential properties of the biological cell. Our results show the emerging variability of plausible sets of values that this type of neuron can adopt underlying complex spiking characteristics. Also, the set of selected cerebellar GrC models captured spiking dynamics closer to the reference model than the single model obtained with off-the-shelf parameter optimization algorithms used in our previous article. The method hereby proposed represents a valuable strategy for adjusting a varied population of realistic and simplified neuron models. It can be applied to other kinds of neuron models and biological contexts.

Keywords: granule cell, cerebellum, neuron model, optimization, adaptive exponential integrate-and-fire, multimodal evolutionary algorithm

Abbreviations: AdEx, Adaptive exponential integrate-and-fire; EA, Evolutionary algorithm; GLIF, Generalized leaky integrate-and-fire; GA, Genetic algorithm; GoC, Golgi cell; GrC, Granule cell; HH, Hodgkin-Huxley; PSO, Particle swarm optimization; I-F, Intensity-frequency.

INTRODUCTION

Large-scale neural network simulations composed of thousands or millions of neurons are useful for better understanding brain information processing primitives. Simplified single-neuron models of low computational cost and based on a few parameters have been proposed to reproduce neuronal firing patterns to encode and decode the information contained in electrophysiological recordings (Izhikevich, 2004; Shan et al., 2017; Marín et al., 2020). These models are required to meet efficiency and biological realism for hypothesizing the functional impact of relevant neuron properties within large-scale simulations. Since most simplified models (e.g., the integrate-and-fire neuron model family) contain abstract parameters that prevent direct adjustment with the biological counterpart biophysical features, optimization algorithms represent an attractive approach for precisely setting the parameters in this kind of neuron models (Druckmann et al., 2007; Jolivet et al., 2008; Friedrich et al., 2014; Pozzorini et al., 2015). However, accurately fitting their parameters to reproduce biological data can be considered a challenging optimization problem that still remains partially unsolved.

The cerebellum is a major center of the nervous system involved in fine motor control, somatosensory processing, and many other non-motor tasks (Schmahmann, 2019). One of the cerebellar neuron types, the granule cells (GrCs) are the most abundant neurons in the human brain (Lange, 1975; Williams and Herrup, 1988). The GrCs are thought to regulate the information transmission through the main afferent system of the cerebellum (Jörntell and Ekerot, 2006). Experimental recordings have characterized two of their main firing features, such as regular repetitive firing and latency to the first spike under injected step currents (D'Angelo et al., 1995, 1998, 2001; Masoli et al., 2017). Also, previous findings suggest intrinsic spiking resonance (as enhanced bursting activity during low-frequency sinusoidal current injections) preferentially in the theta-frequency band (around 5–12 Hz *in vitro* recordings of cerebellar GrCs in rodents; D'Angelo et al., 2001). This complex behavior has been proposed to strengthen the transmission of information in the cerebellar input layer (D'Angelo et al., 2001, 2009; Gandolfi et al., 2013). The definition of cerebellar GrC models that replicate these complex patterns represents an initial step towards understanding the functional role of resonance in information processing and the involvement of the GrCs in the synchronization and learning of the cerebellum.

The relevance of heterogeneity in the population of neurons of the same type with variances in their properties has been highlighted in computational experimentation (Lengler et al., 2013; Migliore et al., 2018). However, the benefits of high variance in terms of biodiversity of neurons in the signal processing of the brain remain largely unexplored. The variances in the neuron properties were demonstrated to enhance the speed, responsiveness and robustness of the spiking neuron networks. Thus, the intrinsic variability of neurons in the brain is proposed to crucially change the network dynamics and could have a role in information processing. The generation of heterogeneous populations of spiking neurons whose properties are closely matched with biological data is of utmost necessity as a first-step in the demonstration of this novel assumption.

As the complexity of neuron models and the available computational power have increased, the use of different optimization algorithms for tuning this kind of simple models has also grown (Van Geit et al., 2008). Consequently, there have been used optimizers for tuning the parameters of computationally efficient neuron models and reproducing certain biological behaviors in previous works. Some authors opt for algorithms with a solid mathematical component, such as the Sequential Quadratic Programming (SQP) method used to tune the modified generalized leaky integrate-and-fire (E-GLIF) model of a cerebellar Golgi cell (GoC; Geminiani et al., 2018). Other examples are the Downhill simplex method and L-BFGS-B, which are included in the open-source optimization framework "Optimizer" (Friedrich et al., 2014). However, the use of optimizers relying on randomness and nature-inspired principles with generic and minimal mathematical components (Lindfield and Penny, 2017) is also very popular among authors (Van Geit et al., 2008). For instance, the referred "Optimizer" framework offers Evolutionary algorithms (EAs) and Simulated Annealing too (Friedrich et al., 2014). The "BluePyOpt' framework also relies on multi-objective EAs such as Non-dominated Sorting Genetic Algorithm-II (NSGA-II), Multi-Objective Covariance Matrix Adaptation Evolution Strategy (MO-CMA-ES), and IBEA (Van Geit et al., 2016). Similarly, Nair et al. (2015) fits the AdEx model of a cerebellar GrC using Particle Swarm Optimization (PSO), and Masoli et al. (2017) opts for the IBEA Genetic Algorithm (GA) to tune the detailed Hodgkin-Huxley (HH) model of a cerebellar GrC with the maximum ionic conductances.

Our previous work (Marín et al., 2020) proposed a tuning procedure based on traditional GAs (EAs based on basic genetic operators, such as crossover and mutation) for creating an adaptive exponential integrate-and-fire (AdEx) model of the cerebellar GrC. We proposed a complex objective function defined by the inherent properties mentioned above and measured as the accumulated distance between the *in vitro* recordings and the simulated responses of the neuron model that is being tuned. Finally, we selected and proposed a GrC model (a specific set of parameters of an AdEx generic neuron model) as the result of the process.

According to the previous literature review and independently of their class, the most used optimization strategies are either multi-objective (which by definition return a set of candidate solutions considering several criteria concurrently) or single-objective yet aimed at converging to a single optimal solution. However, the application of multimodal optimizers (Sareni and Krähenbühl, 1998; Jelasity et al., 2001; Shir et al., 2010) does not seem to be popular even though it has been found that it is possible to find real neurons and neuron models with very similar behavior but different parameters (Van Geit et al., 2008). Multi-Objective algorithms require working in parallel with different objective functions. They can find large sets (Pareto fronts) of equally valuable configurations at the expense of higher conceptual complexity than single-objective methods. On the

contrary, standard single-objective methods work in a simpler background and aim at converging to a single solution, but this behavior can be problematic since evaluations rely on models and some of them might not be as valid as estimated. In this context, the use of a multimodal optimizer arises as a mid-term solution between the algorithms designed for converging to a single solution and those considering different objective functions and returning a set of equally valuable options for an expert to decide. More precisely, a multimodal algorithm will focus on a single objective function, but it will also identify different equivalent solutions that should be ultimately filtered by an expert on the neuron model. Moreover, taking into account potential problems such as noise in the experimental data, low model accuracy, and degenerate cases of the selected objective function, experts might prefer some promising solutions over those theoretically better (strictly in terms of the referred objective function).

The present workflow aims to overcome the intrinsic limitations of the current optimization approaches by identifying a sparse population of different optimal solutions in the search space for a single objective function integrating various features. In this regard, this article extends the methodology presented in Marín et al. (2020) proposing an alternative optimization component based on a multimodal EA for building realistic and computationally efficient neuron models. We base our proposal on the same complex parametric optimization problem as in Marín et al. (2020): optimizing cerebellar GrC models that replicate the essential firing properties of the biological cell, which are essentially the decrease of latency to the first spike and spike frequency increase when the injected step-current intensity is raised [intensity-frequency (I-F) curves], and spiking resonance at the theta-frequency band during sinusoidal current injections. The final population of candidate solutions has to be analyzed and evaluated according to the biological plausibility and their objective to the target features. In order to illustrate the value of this approach, we explore the resulting diversity of the population of cerebellar GrC models and their functional spiking dynamics. Our results show the variability of plausible sets of values that this type of neuron can adopt underlying these complex characteristics.

The rest of the article is structured as follows: "Methodological Workflow" section describes the methodological workflow proposed in this article. "Materials and Methods" section explains the neuron model whose parameters must be tuned, the corresponding optimization problem, and the multimodal optimizer. "Results" section presents the results achieved and the spiking dynamics simulated by the selected neuron configurations. Finally, "Discussion" section contains the conclusions and states some possible future work lines.

METHODOLOGICAL WORKFLOW

In this section, we present the structure of the proposed optimization workflow. **Figure 1** briefly depicts the workflow chart of the methodology. The course of action runs as follows:

Data Preparation

Firstly, the user has to select the particular firing properties of interest of the cell under study extracted from *in vitro* or *in vivo* electrophysiological recordings under specific stimulation protocols. These features (also named objectives) and their protocols define the target function that will drive the selection during the optimization procedure.

Optimization Algorithm

The objective or fitness function, the optimization algorithm parameters and the neuron model parameters (all described in the sections below) correspond to the initial set-up of the optimization architecture. The multimodal algorithm performs an exploration of the parameter space using a simplified (point-neuron) model template. The neuron models with specific sets of parameters which are obtained during the search (also named candidate solutions) are evaluated according to the objective function. Those model configurations with the lowest total score and different enough from the rest are selected in each iteration and passed to the next optimization iterations during the optimization process. The output of this stage is a sparse population of candidate solutions that correspond to different sets of parameters that stand out in their zone of the search space.

Selection and Interpretation of the Candidate Solutions

Once the algorithm has selected different promising model configurations, the user will validate the most suitable neurons among them. This selection runs in accordance with those neuron models which show biological plausibility of their parameters and reproduce with high realism the firing behavior of the real neuron.

MATERIALS AND METHODS

In order to demonstrate the potential of the optimization workflow, we applied this methodology to the case of the cerebellar granule cells (GrCs) as a proof-of-concept. This approach allows us to validate the sparse population of candidate solutions obtained and according to the features defined in the existing literature. This section starts by describing the computational neuron model, defining the optimization problem to solve and the experimental pieces of single-cell recordings. It ends with the technical details for simulation reproducibility and data analysis.

Neuron Model Structure

Since GrCs have a compact electrotonic structure (Silver et al., 1992; D'Angelo et al., 1995; Delvendahl et al., 2015), single-compartment modeling is appropriate. One of the widely used computationally-efficient neuron models is the adaptive exponential integrate-and-fire (AdEx) model (Brette and Gerstner, 2005), but other types of point-neuron models could be considered in this first stage of the workflow (**Figure 1**). The AdEx model is capable of reproducing a diversity of neuronal dynamics customizing a few parameters (Naud et al., 2008). Its realism and great computational efficiency have been supported

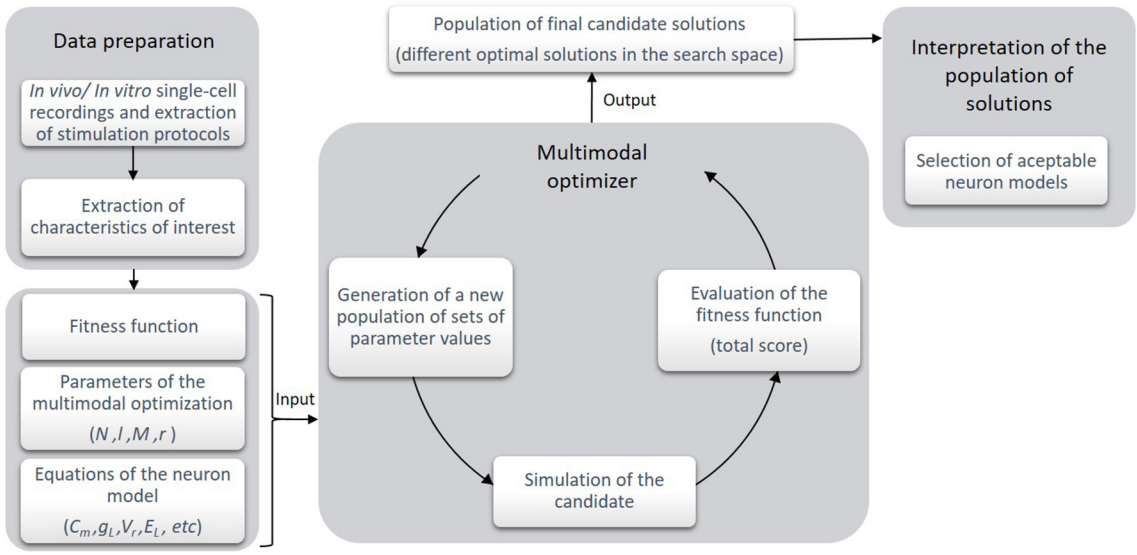


FIGURE 1 | Stages of the optimization framework. At the first stage, the experimental recordings of the biological cells are obtained and the characteristics of interest are selected and calculated. Then, according to these data, a fitness (or objective) function is built and the parameters of the algorithm and the neuron model are selected. At the second stage, the execution of the multimodal optimizer takes place. The optimization process consists in generating a number of candidates (i.e., sets of parameter values) which are simulated and evaluated according to the objective function. Finally, as the output of this stage, a population of different candidate solutions is returned in a single execution of the algorithm. These selected candidates are illustrative of some local optima in the search space of the objective function. At this point, the expert interprets if the candidate solutions are as numerically valid as their scores estimate. Quantitatively and qualitatively, the best-ranked candidate solutions are selected as the neuron models proposed.

by several comparisons with detailed models and experimental recordings (Brette and Gerstner, 2005; Jolivet et al., 2008; Naud et al., 2008; Nair et al., 2015; Marín et al., 2020). Accurately fitting the model with respect to experimental measurements is not straightforward. The adaptation state variable of the AdEx model allows good fitness with different firing modes (e.g., regular discharge, bursting, delayed spiking, etc.) depending on specific parameters values (Jolivet et al., 2008; Naud et al., 2008). However, its nonlinearity makes the optimization of its parameters, challenging.

The AdEx model consists of only two coupled differential equations and a reset condition.

$$C_m \frac{dV}{dt} = -g_L (V - E_L) + g_L \Delta_T \, exp\left(\frac{V - V_T}{\Delta_T}\right) + I(t) - w \quad (1)$$

$$\tau_w \frac{dw}{dt} = a (V - E_L) - w \quad (2)$$

$$\text{if } V > V_{peak} \text{ then } V \leftarrow V_r \text{ and } w \leftarrow w + b \quad (3)$$

Equation (1) describes the evolution of the first state variable, namely membrane potential (V), during current injection ($I(t)$). Equation (2) describes the evolution of the second state variable, namely adaptation current (w). When the current $I(t)$ drives V beyond the threshold potential (V_T), then the exponential term of the slope factor (Δ_T) in equation (1) dominates the action potential until V reaches the reset threshold potential (V_{peak}). Then, the reset condition (3) determines that V is instantaneously set to V_r and w is increased a fixed amount named b. Both equations (1) and (2) contain 10 free parameters that can be optimized in order to minimize an arbitrary objective function (namely the difference of the obtained neural model behavior with respect to a "desired behavior" for instance, reproducing firing characteristics of cell recordings). These parameters are: the total leak conductance (g_L), the leak reversal potential (E_L) and the membrane capacitance (C_m) in equation (1) that model the passive membrane mechanisms; the parameters Δ_T and V_T in the exponential term of equation (1) model the spike generation and shape; the time constant parameter (τ_w), the subthreshold adaptation (a) and the spike-triggered adaptation parameter (b) define the evolution of the state variable w in equation (2); the parameters V_{peak} and V_r that drive the reset condition as mentioned above. These parameters have been set within fixed ranges to constrain the exploring process (**Table 1**). The membrane potential was initially set to the same value as the leak reversal potential ($V_{init} = E_L$).

Model Context and Problem Definition

Selection of Features

In order to obtain a neuron model that replicates the behavior of the cerebellar GrCs we have selected some features which quantify some of the most characteristic firing properties of this neuron type: (1) mean frequency through repetitive firing discharge under direct current stimulation [equation (4)];

TABLE 1 | Model parameter ranges established for the search space of the optimization process.

Parameters	Min. value	Max. value	Parameters	Min. value	Max. value
C_m	0.1 pF	5.0 pF	V_T	−60 mV	−20 mV
Δ_T	1 mV	1,000 mV	a	−1 nS	1 nS
E_L	−80 mV	−40 mV	b	−1 pA	1 pA
V_r	−80 mV	−40 mV	g_L	0.001 nS	10.0 nS
V_{peak}	−20 mV	20 mV	τ_w	1 ms	1,000 ms

(2) latency to the first spike under direct current stimulation [equation (5)]; and (3) burst frequency in response to different sinusoidal current stimulation (stimulation with different oscillation frequencies) [equation (6)]. These features will be combined into a single objective or fitness function to be considered by the selected multimodal evolutionary optimizer.

Parameter Optimization

The parameter optimization is carried out minimizing the fitness function by weighting the difference of these quantified features with a reference taken from real electrophysiological recordings. Thus, the objective function is defined as the weighted sum of the scores of the specific features (*feature_score*) related to the spiking features, according to equations (4, 5 and 6). The definition of the objective function that contains all these features is extracted from Marín et al. (2020).

$$feature\_score_{Mean\ frequency} = \sum_{i=1}^{n} \left[abs(MF_{sim_i} - MF_{exp_i}) \cdot w_{Mean\,frequency} \right] \quad (4)$$

$$feature\_score_{First-spike\,latency} = \sum_{i=1}^{n} \left[abs\left(LAT_{sim_i} - LAT_{exp_i}\right) \cdot w_{First-spike\,latency} \right] \quad (5)$$

$$feature\_score_{Burst\,frequency} = \sum_{i=1}^{n} \left[abs(BF_{sim_j} - BF_{exp_j}) \cdot w_{Burst\ frequency} \cdot \left(std(BF_{sim_j}) + 1 \right) \right] \quad (6)$$

The feature score (score function of each feature or objective) is calculated as the absolute distance (*abs*) between the feature values extracted from the *in vitro* electrophysiological recordings (exp_i) and the feature values extracted from the simulated traces from the neuron model (sim_i). This is multiplied by the weight associated with each feature ($w_{Mean\ frequency}$, $w_{First-spike\ latency}$ and $w_{Burst\ frequency}$). The weights of the burst frequency ($w_{Burst\ frequency}$) and the mean frequency ($w_{Mean\ frequency}$) were set to 1 as they both were measured in hertz (Hz) and show values in comparable scales. The weight of the first-spike latency ($w_{First-spike\ latency}$) was weighted to 1,000 as it was measured in seconds (s). Hence, the algorithm equally weights 1 Hz-error at mean frequency feature, 1 ms-lag at first-spike latency and 1 Hz-error at burst frequency. However, the feature score can be modified (if enhancing the focus on a particular feature with respect to the others) or extended if some extra aspect is to be taken into consideration. For instance, a penalization was used in the definition of the burst frequency score ($feature_{score\ Burst\ frequency}$) to assure the stability of bursts [as proposed in Marín et al. (2020); Equation (6)].

Feature Measurement

The experimental recordings of the repetitive discharge are the mean frequency (defined as the number of spikes divided by the stimulation time) during 1-s length step-current injections of 10, 16 and 22 pA. The latency to the first spike is defined as the time the neuron takes to elicit its first spike upon current stimulation. Both features are extracted from *in vitro* patch-clamp recordings performed from acute cerebellar slices of a population of cerebellar GrCs (Masoli et al., 2017). The spiking resonance in the theta-frequency range is a complex behavior determined by the burst frequency [as the inverse of the average inter-spike interval (ISI) of the output neuron] during each stimulation cycle. Then, the average burst frequencies are measured throughout 10 consecutive cycles of sinusoidal stimulation. As it occurred in the *in vitro* recordings, we have set the burst frequency to zero when one or no spike per cycle has been obtained in the simulated neurons. The stimulation protocol consists of sinusoidal current injections with 6-pA and 8-pA amplitudes, sustained by a 12-pA offset during 22.5 s. They generate spike bursts in correspondence with the positive phase of the stimulus (sinusoidal phase of 270°). These features are extracted from *in vitro* patch-clamp recordings performed from acute cerebellar slices of a single cerebellar GrC (D'Angelo et al., 2001). It is worth mentioning that since the number of cells differ in both reference sources (the former from a population of cells and the latter from a single cell), and for the sake of equality, we selected the mean feature value of the mean frequency and the first-spike latency as a target type of neuron. On the other hand, the reference data points for the resonance frequency used in the fitness function are based on a single neuron measurement.

Optimization Method

As introduced, the problem stated in "Model Context and Problem Definition" section has been addressed with a multimodal optimizer, i.e., an optimization algorithm designed to concurrently obtain multiple different global and local solutions to a problem (Sareni and Krähenbühl, 1998; Shir et al., 2010). It is the Universal Evolutionary Global Optimizer (UEGO) proposed by Jelasity et al. (2001).

As can be deduced from its name, UEGO is an evolutionary optimization algorithm (Lindfield and Penny, 2017), so it works with a population of solutions and simulates their Darwinian evolution to progressively achieve better solutions. However, it belongs to the memetic category of EAs (Moscato, 1989; Molina et al., 2011). This kind of

method is characterized by promoting the autonomous behavior of candidate solutions as self-evolving agents in conjunction with the underlying evolutionary environment. Thus, in practical terms, a memetic method combines a generic evolutionary stage of global scope with a replaceable local search component. UEGO meets this requirement, which makes it highly adaptable to different optimization problems (Ortigosa et al., 2007; Redondo, 2009; Cruz et al., 2018).

The population of UEGO consists of different species, which is a fundamental concept for this method. Species are not plain candidate solutions as it occurs with standard GAs. Instead, every species combines a feasible and ranked candidate solution with an assigned radius around it in the search space. The radius is defined as a Euclidean distance to study the separation between different candidate solutions, i.e., to assess their similarity. Thus, a species defines a (hyper)sphere in the search space, and it is treated as an exploration window to center the independent local search component. **Figure 2A** depicts a sample species for a hypothetical optimization problem of two variables. As can be seen, the species represents both a candidate solution and a region in the search space on which the local search will focus. Since the referred example assumes two variables, the species can be easily visualized as circumferences. This is not the case for the problem at hand because the search space has 10 dimensions, i.e., the parameters to fit, and species will be hyperspheres in a 10-dimensional Euclidean space. Nonetheless, the underlying idea remains unaltered, and species will be generally depicted as circumferences for practical reasons. **Figure 2B** shows the structure of every species for the target problem and how UEGO perceives it.

As an algorithm, UEGO focuses on managing a population of different species, which defines its evolutionary part. It executes the steps shown in Algorithm 1. They are summarized next for the sake of self-completeness, but the interested reader is referred to Jelasity et al. (2001) and Ortigosa et al. (2001) for further details.

The algorithm takes the following parameters as input: (1) the maximum number of species to keep in the population (M); (2) the maximum number of evaluations of the objective function (N); (3) the minimum radius to keep between different species (r); and (4) the number of search levels or full cycles (l). After preliminary experimentation, these parameters have been set to $M = 100$, $N = 10{,}000{,}000$, $r = 0.7$, and $l = 50$ for the present study. The maximum number of species and function evaluations agree with the reference values proposed by Ortigosa et al. (2001), radii below 0.7 resulted in too many almost-equivalent solutions in this case, and the number of levels was progressively increased up

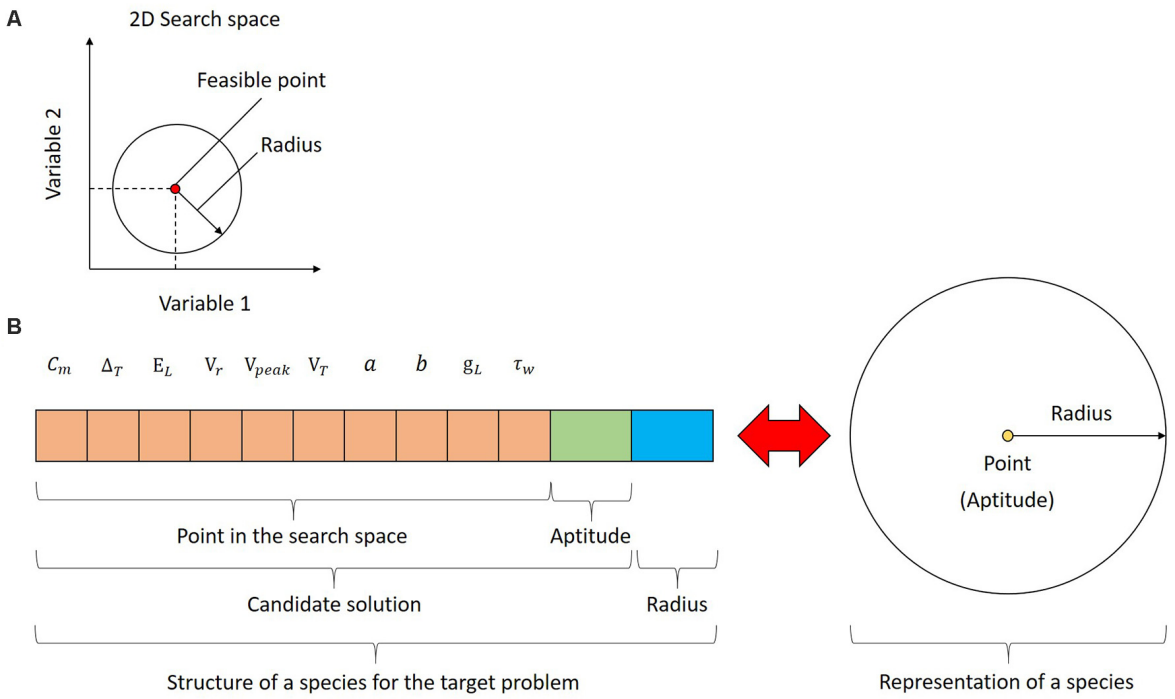


FIGURE 2 | Species in UEGO. **(A)** Species in the search space for a hypothetical bi-dimensional problem. The candidate solution is a feasible point, and it is linked to a particular radius around it to define a region to center the search. Both define a circumference in a two-dimensional Euclidean space, but the concept can be extended to any dimensions through hyperspheres. **(B)** Structure of a species for the problem at hand (on the left) and its conceptual meaning (on the right). As shown on the left, in practical terms, a species consists of three parts: (1) a feasible point in the search space with a component for each optimization variable (colored in orange); (2) the fitness of the referred point according to the objective function (colored in green); and (3) the radius linked to this species (colored in blue). UEGO will use the referred information to build a hypersphere from every species, and they will be ultimately treated as exploration windows in the search space. This aspect is shown on the right side of the figure using a plain circumference due to the impossibility of showing the corresponding hypersphere.

to ensure that the search lasts enough, and the best-performing solutions are competitive.

Having gathered the input, the first step of UEGO is the initialization of its population. For this purpose, it randomly selects a point in the search space, evaluates it, and assigns the first radius to it. By definition, the radius of the first species is equal to the diameter of the search space, which is computed as the Euclidian distance between the lower and upper bounds. Therefore, the region defined by the initial species covers the whole search space, and no solution will be unreachable. After that, the radii assigned to species at creation, and hence the region that they cover, will decrease. They do in geometrical progression with the number of levels until the last one, which is linked to the minimum radius specified by the user (see **Figure 3A**). This strategy of progressively reducing the mobility at search is known as cooling in the field of Optimization, and it is inspired by the process of annealing metal (Lindfield and Penny, 2017). It promotes exploration at the beginning to find the best zones of the search space, avoids premature stagnation, and forces convergence at the end.

The second step in Algorithm 1 shows the memetic nature of UEGO. Namely, it consists of launching the local search component. This stage, also seen in step 8, independently affects every species in the population, but at this point there only exists the initial one. As introduced, local search is treated as an isolated component, and the method selected is briefly described at the end of this section. It is only required to start at the center of the species and find a better point in its region after several movements. Theoretically, the local search algorithm is limited by the region of the species, i.e., the radius linked to its starting point. Thus, no single step made by the optimizer in a given species can be larger than the radius. However, every time that the local search component finds a better point in the region, it becomes the new center of that species, so they are countinously moving in the search space. **Figure 3B**

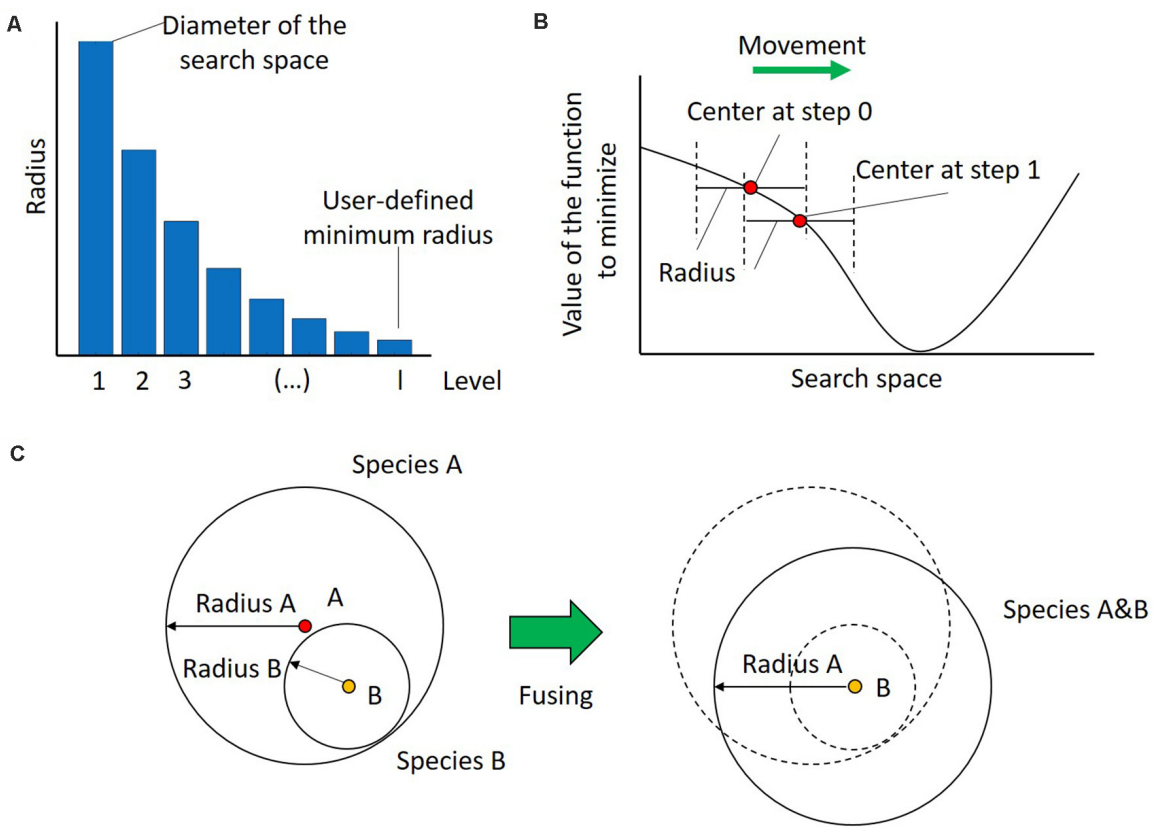


FIGURE 3 | Dynamics in UEGO. **(A)** Evolution of the radius length linked to the species created at every level or cycle. The radius of the first level, i.e., that assigned to the initial species, is equal to the diameter of the search space. Subsequent radii decrease in geometrical progression until the last one, which corresponds to the minimum radius defined by the user to consider different solutions. By proceeding this way, species of the first cycles have more mobility, which promotes exploration of the search space, and those of the last levels have more exploitation, which helps convergence. **(B)** Movement of a species in a hypothetical single-dimension search space after an iteration of the local search method. The whole species is moved after finding a better center. Notice that the new point falls within the original region defined by the species, but its new position also moves the search focus. **(C)** Fusing species. On the left two overlapping species, A and B, with radii Radius A and Radius B, respectively. On the right, the species A&B that results from their fusion. This new species keeps the largest radius of both, i.e., Radius A, which aims to keep the regions to explore as broad as possible to avoid premature convergence. Additionally, take into account that since the center of Species A&B is that of B, it can be concluded that Species B has a better, i.e., lower value of the cost function. Finally, notice how the new species has slightly moved from the original zones, so it will be possible to explore a new zone of the search space.

illustrates this concept. Aside from defining the initial point and the maximum step size for the local search method, UEGO also controls its computational budget, i.e., its number of function evaluations. The details about how UEGO distributes the total function evaluations allowed are out the scope of this work, but they are covered in depth in Jelasity et al. (2001) and Ortigosa et al. (2001). Regardless, the principle followed is to allow more function evaluations for later search levels, when it should be more interesting to explore the promising regions previously found. Also, notice that the local search algorithm should try to save function evaluations by early terminating when it finds itself unable to locate a better solution, so not all the allowed function evaluations might be consumed.

The two previous steps form the first level of search for UEGO. For this reason, the loop in Step 3 counts from 2. Notice that if the number of levels were set to 1 by the user, the algorithm would be mainly equivalent to launching the local search method from a random point. The result would hence be the single species after being locally optimized. Nonetheless, this situation is mostly theoretical. A standard configuration of UEGO is expected to execute several levels of search or cycles. Each of them consists of Steps 4–9.

Step 4 defines the computation of the radius that will be assigned to any new species created at the current level. As introduced, they decrease in geometrical progression. This step also involves determining the number of function evaluations that can be consumed for creating and locally optimizing species in Steps 5 and 8, respectively. The budget for creation is always three times the maximum number of species allowed, but that of local optimization increases with the number of levels as summarized above. See Jelasity et al. (2001) and Ortigosa et al. (2001) for further information.

Step 5 is where UEGO tries to increase its population. It first divides the creation budget among the existing species to calculate how many points will be allowed to evaluate. After that, within the region defined by every species, the algorithm randomly takes the permitted number of candidate solutions. Then, the points of every species are systematically paired with each other, and their middle points are evaluated. If the solution at any of the middle point is worse than that at its extremes, both members of the pair define new species. This is done under the assumption that they are on different sub-areas in the search space. The radius assigned to these new species will be the one that corresponds to the current level, which should be lower than any previous one. By proceeding this way, multiple new species will appear within the limits of every existing one and focusing on smaller regions to concentrate the search. Additionally, notice that UEGO will update the center of the initial species if any of the candidate points considered in their regions is a better solution, so they can move.

Step 6 scans the current population to check if the center of any pair of species is closer to each other than the radius of the current level. The goal is to avoid spending too much computing time separately exploring the same region. Species that overlap according to this criterion are fused into a single one. More specifically, the center of the resulting one is that which represents a better solution. The radius will be the largest one of them, which aims to keep the scope of search as broad as possible to avoid premature convergence. **Figure 3C** depicts this step assuming that species B is better than A, but it has a shorter radius. Notice that this definition ensures that there will always be a species whose radius covers the whole search space derived from the initial one, so it is always possible to reach any point in the search space.

Step 7 checks the length of the current population. If there are more species than allowed by the user through parameter M, those with the shortest radius are removed until the population size is in the valid range again. The removal criterion is aligned with the previous idea of maintaining species that allow escaping from low-performing local optima. The last two steps are both procedures already described. Namely, Step 8 will independently launch the local search component from every existing species, which will make them move around the search space. Step 9 rescans the population to identify and fuse any species that overlap after having been moved. The UEGO algorithm ends with step 11 by returning the surviving species. According to the process described, they are expected to be different promising solutions. The separation between them in the search space, i.e., degree of difference, will be the minimum radius defined by the user at least. Therefore, as intended, the users of this method and framework (**Figure 1**) have several options to study (final candidate selection) in case those solutions with the best numerical fitness do not appropriately meet the qualitative requirements that can be further analyzed in a subsequent stage.

Regarding the local search component previously referred to, the SASS or Solis and Wets' method (Solis and Wets, 1981; Molina et al., 2011) has been used. It is a stochastic hill-climber that starts at the center of the given species and randomly decides a direction to move. The amplitude of every jump cannot exceed the radius of the species, and it is scaled depending on the number of positive (improving) and negative (non-improving) movements. This optimizer has been selected because it does not require any specific properties of the objective function. Besides, it has already been successfully used within UEGO (Ortigosa et al., 2007; Redondo, 2009). The configuration of this method is the recommended one. Namely, movements are made by adding a normally-distributed random perturbation vector with a standard deviation between 1e-5 and 1, starting at the upper bound and ultimately rescaled by the radius of the species. The standard deviation is doubled after five consecutive successful movements or halved after three consecutive failed ones. Notice that the local search method will terminate after 32 consecutive failed or discarded movements, no matter the remaining budget.

Data Analysis

Multidimensional Scaling

To further illustrate the multimodal distribution of the different optimal solutions, we have applied the Classical Multidimensional Scaling (MDS) method (using scikit-learn Python library; Pedregosa et al., 2011). The distribution of the parameter values of the solutions through n dimensions (in our case, n = 10 parameters that define a neuron model, also named candidate solution) is denoted as landscape. Using MDS,

the differences among landscapes were visualized as distances in the bi-dimensional plane. The input vector of distance to the MDS is calculated as a simple Euclidean distance between landscapes (as in other analysis works, such as in Rongala et al., 2018). Given a distance dot matrix, this algorithm recovers a representation of D-dimensional coordinates of data (in our case, $D = 2$ dimensions). This method allows studying the different landscapes chosen during the algorithm execution and represents their values in a 10-dimensional space embedded in a *2D* plot.

Algorithm 1.- UEGO Algorithm

Input: M, N, r, l	//Max. species, max. evaluation, min. radius, levels
1.- Initialize_List_Of_Species	//Create the first species
2.- Optimize_Species	//Launch the local search on it
3.- **for** i = 2 to l **do**:	//Main loop (Steps 1 and 2 are define the first level)
4.- Compute_Level_Config	//Manage the use of function evaluations and radii
5.- Create_Species	//Create species in the zones of the exising ones
6.- Fuse_Species	//Avoid that species overlap each other at this level
7.- Shorten_Species_List	//Remove species if there are more than allowed
8.- Optimize_Species	//Launch the local search on every existing species
9.- Fuse_Species	//Avoid that species overlap each other at this level
10.- **end for**	
11.- Return_Surviving_Species	//The remaining species become the set of results

Technical Details for Reproducibility

Python (version 2.7.12) and MATLAB (version 2018b) implementations were used to launch the second stage of the workflow (the exploration processes of the multimodal algorithm). The proposed pipeline allows simulating the neuron models and calculating the features scores through the Python-NEST environment (Python Software Foundation Python 2.7.12, 2016; van Rossum, 1995) and NEST simulator 2.14.0 (Peyser et al., 2017) and evaluating and exploring different candidate solutions in optimization cycles through Python-MATLAB implementations. After considering 10 independent executions with different seeds, UEGO executes 50,000 function evaluations on average. The one selected for further analysis in "Results" section used 47,951, which approximately results in 32 h of run time in the execution platform. In the last stage of the workflow, the reproduction and validation of the resulting neuron models (candidate solutions) were analyzed using Python-NEST scripts. The Figures were generated using Matplotlib (version 2.2.5; Hunter, 2007; Caswell et al., 2020) and Numpy (version 1.16.6; Harris et al., 2020) libraries. The simulations were run with an Intel Core i7–4790 processor with 4 cores and 32 Gb of RAM. The source code and data are available in this public repository: https://github.com/MilagrosMarin/Multimodal-optimization-for-fitting-cerebellar-GrCs.

RESULTS

Analytical Results

The results achieved by the optimization stage consist of a population of up to 100 candidate solutions, which is a user-given parameter. More accurately, the multimodal optimizer tries to find different yet promising parameter sets in the search space, and it uses a minimum user-given separation radius for this purpose (see the "Materials and Methods" section). For this purpose, through the search, the optimizer manages a population of feasible configurations that are distributed over the search space, can produce new ones, move, and absorb others when they are considered to represent the same parameter set. Thus, the number of candidate solutions that ultimately survive as results for further consideration by an expert might vary. Since the optimizer is stochastic, its results might vary between different executions, so it has been launched 10 times as mentioned in "Technical Details for Reproducibility" section. **Supplementary Figure 1** shows that the results are similar between executions in terms of search space coverage and overall numerical quality. The one selected for further analysis in this section resulted in a final population of 25 different candidate solutions.

These candidate solutions are minimized to the target multi-feature objective function and their scores are represented in **Figure 4**. The candidate solutions show different combinations of feature adjustments in order to reach the minimal total score, which reveals a well-balanced definition of the multi-feature objective function. Not unexpectedly, the spiking resonance feature contributed the most to the score (pink bars in **Figure 4**; as this is the feature with the highest number of *in vitro* reference points). Although they were selected by their well-ranked solutions, some of them (from 18 to 25) have low-performing configurations for some of the considered properties (mostly for the feature of latency to the first spike, green bars in **Figure 4**).

The different solutions in the search space expected to be returned by the algorithm are visualized using the MDS algorithm, evidencing the multimodality of the search space (**Figure 5A**). The candidate solutions with the lowest scores (under 250 units, which are the solutions from 1 to 12 and in warmer reddish colors) correspond to different high-quality solutions. They are sparsely located along the 2-dimensional display (zoom of the most representative candidates, and colored based on the total score of each candidate solution, in **Figure 5A**). The parameter configurations that define each of these solutions explored a variety of values from within their boundaries (e.g., V_{reset}, E_L, τ_w), which means that the algorithm successfully went over the parameter space (i.e., landscapes; **Figure 5B**).

Reproduction of Spiking Dynamics

In the section above, the candidate solutions of most interest (from 1 to 12) showed quantitative fitting to the supra-threshold characteristics defined in the multi-feature objective function with minimized scores. In this section, the top-ranked solutions are qualitatively analyzed regarding their accuracy in capturing this intrinsic excitability of cerebellar GrCs, i.e., firing discharge with a mean frequency increased whereas latency to the first spike

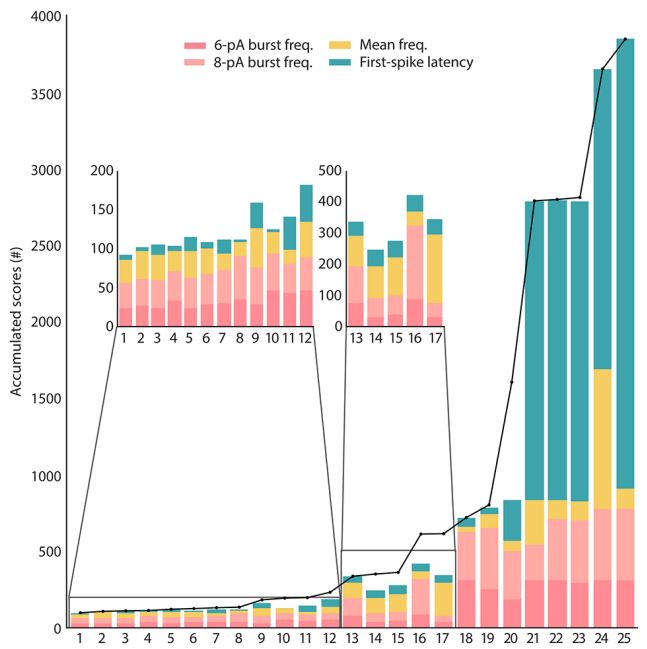


FIGURE 4 | Scores of the candidate solutions. The total scores obtained from minimizing the features of the candidates according to the objective function are represented in black dots joined together by a black line. These total scores are unwrapped by the scores of each single feature (feature scores) for every candidate solution and represented in stacked bars. The feature calculations are explained in "Model Context and Problem Definition" section. The feature of burst frequency (pink bar plots) has been divided based on the sinusoidal stimulation amplitude of 6 pA or 8 pA (the standard deviation has not been included in the stacked bar).

firing decreased under injected currents and spiking resonance in the theta range under sinusoidal currents.

Although the top-ranked solutions achieved lower score values, this fact might not imply reproducing the complex firing dynamics of the neuron. The case of the spiking resonance is an appropriate example of this possibility: although the experimental points are suitably adjusted to the graphic curve, the neuron responses are larger when the neuron behavior is extrapolated to higher sinusoidal frequencies. That is, as mentioned in "Feature Measurement" in "Materials and Methods" section, the burst frequencies generated with stimulation frequencies beyond 10 Hz fell to zero because the *in vitro* measurements contained either one or no spikes. Since there are no points of burst frequencies in higher frequencies in the biological measurements, we have avoided including them in the objective function.

The whole subset of interesting candidates (from 1 to 12) reproduced the complex spiking behaviors within the limits of our electrophysiological observations. The parameter values obtained for the set of solutions are contained in **Table 2**. Candidates 1–8 manage to reproduce all the three spiking features mentioned above (**Figure 6**). The spiking resonance curves (left plots in **Figure 6**) were successfully replicated (with preferred frequencies around 6–20 Hz) by the whole subset of candidates. The intensity-frequency (I-F) plots (middle plots in **Figure 6**) were almost linear between 0 and 100 Hz and the latencies to the first spike were also replicated (right plots in **Figure 6**) as in the biological piece of evidence (D'Angelo et al., 2001).

With respect to the qualitative adjustment of the best-ranked solutions, the top-four candidates reproduce all the three spiking behaviors. More specifically, candidate 2 reproduces the spiking behaviors according to experimental registers of real cells (**Figure 6A**). Other best-ranked candidates, such as 1, 3 and 4, also reproduce qualitatively these behaviors as the reference reports, but with resonance curves (around 5–20 Hz, as seen in left plots in **Figure 6A**) slightly shifted out of the concrete theta band of *in vitro* cerebellar GrCs (around 6–12 Hz in D'Angelo et al., 2001).

Regarding the quantitative comparison of these best-ranked solutions (the distance of the feature values from the experimental measurements defined in the integrated objective function), candidate 2 obtained the highest score for the concrete points of the mean frequency feature (yellow bar in the dashed box in **Figure 6A**), but the lowest score for the first-spike latency feature (green bar in the dashed box in **Figure 6A**). That is, candidates 1, 3 and 4 obtained lower scores for the mean frequency feature than candidate 2.

This fact together with the shifted resonant curves in higher preferred frequencies may indicate an incompatibility of both firing properties (i.e., the repetitive spike discharge and the spiking resonance), within the AdEx models, as we previously hypothesized in Marín et al. (2020). Thus, the GrC behavior complexity in reproducing these features, being beyond the capabilities of these AdEx models with a single parameter configuration (GrCs have different functioning modes).

Similarly, candidates 5, 6, 7 and 8 reproduce all the features but with the resonance curves in slightly higher preferred frequencies (around 6–15 Hz; left plot in **Figure 6B**). In addition, candidates 5, 6 and 7 showed larger initial latencies (around 100, 80 and 300 ms, respectively) than the candidates mentioned before (i.e., 1,2,3,4 and 8, with initial latencies around 50–60 ms) which are closer to the experimental recordings used as reference (right plots of **Figure 6B**). Example traces generated from two of the best candidate solutions are shown in **Figure 6C**. In particular, the left plots of **Figure 6C** show the generation of spike bursts clustered in triplets or longer bursts in time slots corresponding to the positive phase of the sinusoidal current, as described in the reference report (D'Angelo et al., 2001). It is worth mentioning that the ISI during the burst duration (i.e., the oscillatory burst frequency defined in the fitness function) is very similar to that from real traces. In addition, the repetitive spike discharge, experimentally evidenced in cerebellar GrCs (D'Angelo et al., 1995, 2001), generated by these candidates are shown in the right plots of **Figure 6C**.

DISCUSSION

The present study illustrates the application of a novel optimization framework to the case of the cerebellar GrCs: the automated identification of different and promising configurations of the neuron model parameters to reproduce

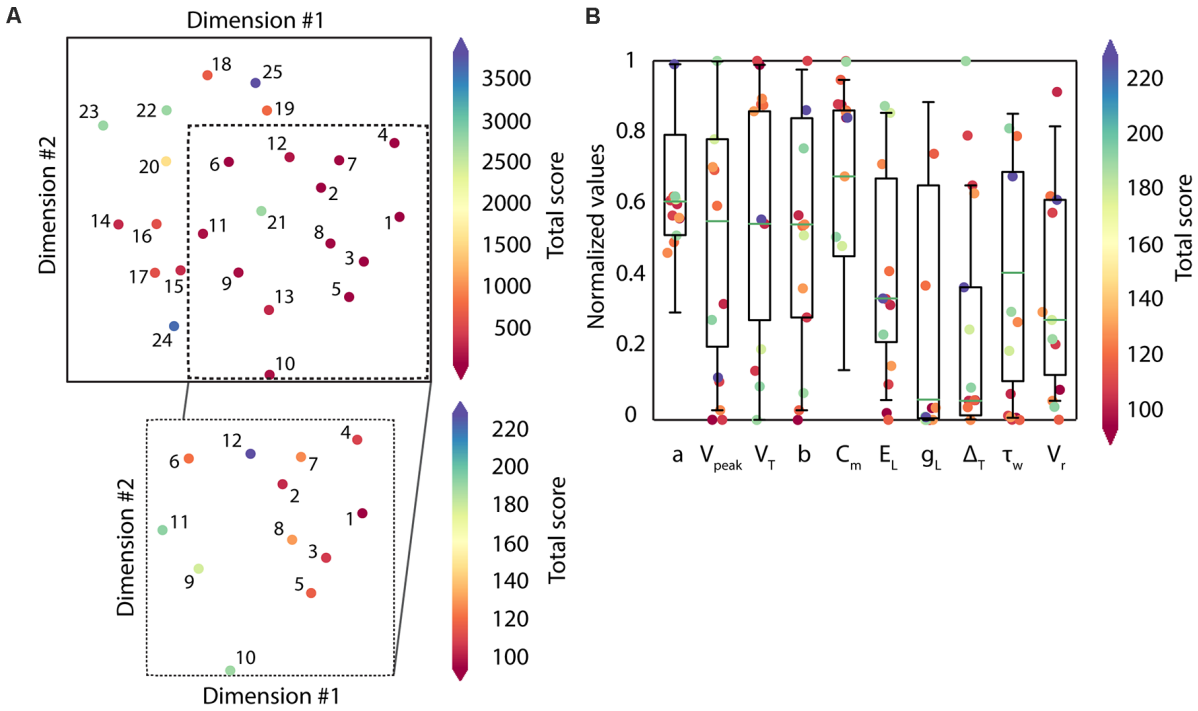


FIGURE 5 | Distribution of the candidate solutions and their respective parameter values. **(A)** Representation of the population of solutions (n = 25) using Classical Multidimensional Scaling (MDS). The colormap shows the total score of each candidate solution, red-colored being the solutions with the lowest scores (best solutions). In zoom, the most representative candidates (those with scores under 250 units; n = 12). **(B)** Parameter distribution of the candidate solutions. Each dot corresponds to the parameter value that defines every candidate solution with the lowest scores (those with scores under 250 units; n = 12). The colormap reflects the total score obtained by each parameter. The boxes correspond to the interquartile ranges (IQR) between the first (Q1) and the third (Q3) quartiles. The green line represents the median among the values obtained for each parameter. The whiskers correspond to the 5- and 95- percentile respectively.

TABLE 2 | Parameter configurations of the best-ranked candidate solutions and their total score.

Candidate solution	Parameter configuration										
	a (nS)	V_{peak} (mV)	V_T (mV)	b (pA)	C_m (pF)	E_L (mV)	g_L (nS)	Δ_T (mV)	τ_w (ms)	V_r (mV)	Total score
1	0.123	−19.981	−20.446	−0.999	4.226	−79.225	0.333	55.881	7.138	−76.638	93.992
2	0.202	−7.078	−38.149	0.140	4.400	−67.194	0.001	54.382	73.441	−43.458	102.906
3	0.223	−19.984	−54.527	−0.429	4.408	−66.527	0.003	653.468	1.700	−71.568	106.522
4	0.139	−15.731	−20.000	1.000	4.998	−76.037	0.001	792.296	12.717	−56.917	108.708
5	0.244	7.841	−24.830	−0.946	4.741	−79.985	7.413	55.820	1.039	−79.972	116.652
6	−0.010	3.859	−24.929	0.081	4.324	−63.421	3.742	36.480	790.811	−55.069	121.169
7	−0.069	−18.914	−24.208	0.089	4.303	−51.501	0.342	1.116	273.243	−77.872	126.610
8	0.126	8.171	−25.613	−0.267	3.423	−73.981	0.005	631.309	8.875	−67.973	130.303

complex spiking behavior through multimodal algorithms for an expert to decide. The solutions produced by the multimodal optimization process represent a valuable analysis tool that facilitates better understanding of how certain neural model properties are supported by a specific parameter configuration. Two challenges were addressed: (1) the optimization of efficient neuron models that allow the replication of complex dynamics such as the spiking resonance in the theta frequency band while maintaining other typical GrC dynamics such as the regular repetitive firing and the spike timing; and (2) the generation of a diverse population of neuron models with widely explored configurations in sparse local minima. These challenges were addressed by optimizing a multi-feature fitness function defined with the distinctive characteristics of cerebellar GrCs. In this case, the spiking resonance in the theta-frequency band of the GrCs is a complex behavior believed to improve the information processing in the cerebellum (D'Angelo et al., 2001, 2009; Gandolfi et al., 2013). The mean frequency of repetitive firing and the spike timing (latency to the first spike) are the main properties of the GrC used to measure their intrinsic excitability. Addressing this approach through the optimization workflow resulted in the full-fledged exploration of a population of efficient neuron models that sufficiently reproduce highly realistic dynamics. Finally, it is also important to take into consideration a validation of the neuronal firing dynamics in order to analyze in detail the behavior of the obtained neuron

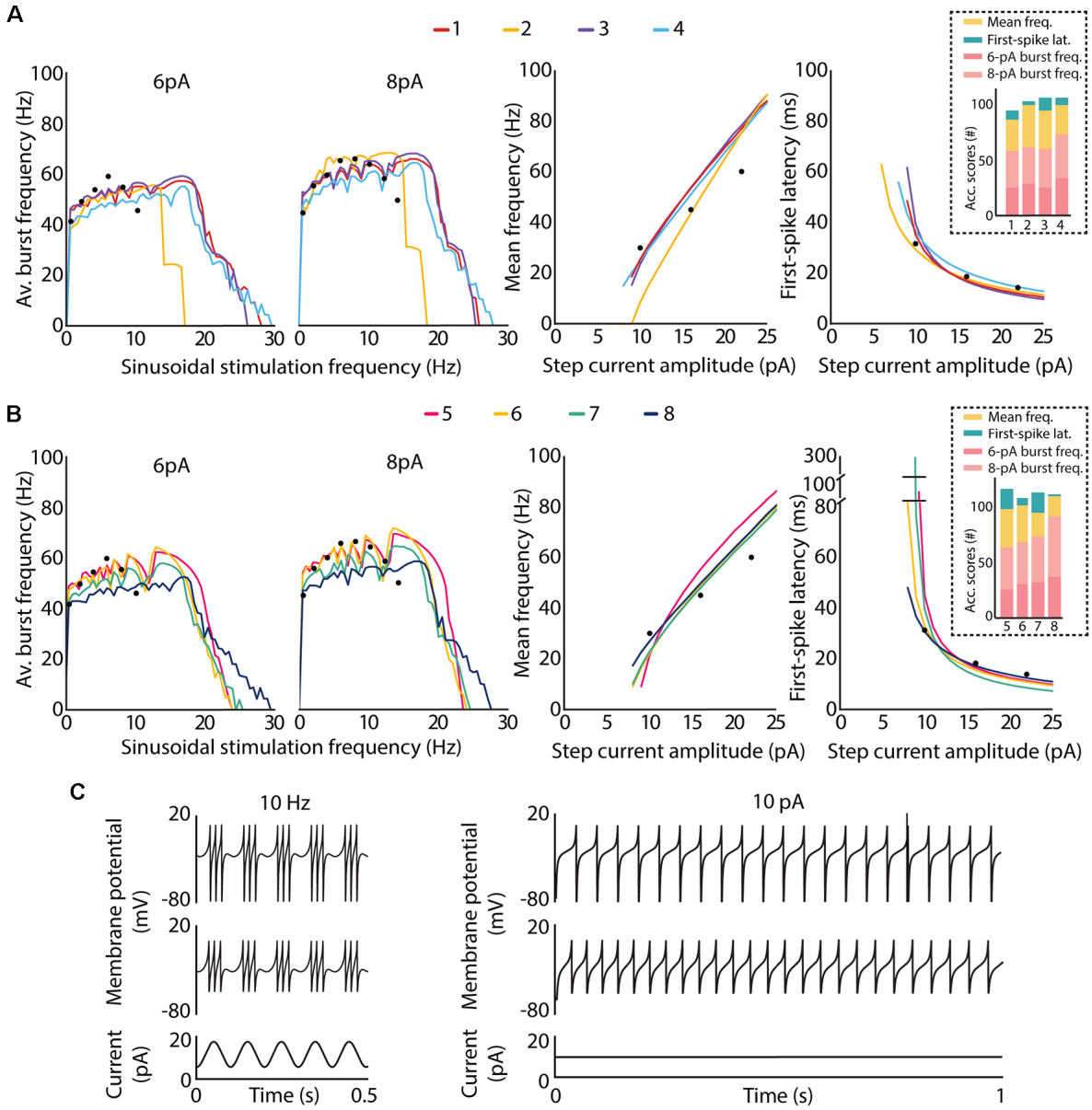


FIGURE 6 | Intrinsic excitability of the top-ranked neuron models of cerebellar GrCs. The optimal candidate solutions are qualitatively analyzed according to the features defined in the objective function, i.e., spiking resonance in the theta range under sinusoidal stimulation (left plots), mean frequency (middle plots) and latency to the first spike (right plots), the last two under direct current stimulation. Black dots represent the experimental data used as reference in the optimization process. **(A)** Spiking dynamics of the top-four candidate solutions of the final population. Their accumulated scores of each feature are represented as bars in the dashed box. **(B)** Spiking dynamics of the candidate solutions ranked 5–8 of the final population. Their accumulated scores for each feature are represented as bars in the dashed box. **(C)** Example traces from two of the best candidate solutions (top plots from candidate 5 and middle plots from candidate 6) show: on the left, spike bursts from both neurons under sinusoidal current injection of 10-Hz frequency, 6-pA amplitude and 12-pA offset, as it is represented in the bottom plot; on the right, repetitive spike discharge from both neurons during a step-current injection of 10 pA, which is represented in the bottom plot.

models and how the parameter diversity can be steered to adapt the model to specific purposes or studies. The selected neuron models are presented as efficient tools that can formulate biological network hypotheses and shed some light on future neuroscientific research.

To solve neuron model tuning problems, it is possible to opt for methods based on robust mathematical principles whenever the objective function has some properties, such as being expressed by a particular type of analytical formulation and being differentiable. For instance, the point-neuron model of a cerebellar GoC proposed by Geminiani et al. (2018) was modified in order to optimize part of its parameters using a SQP algorithm from spike voltage traces under input current steps. The SQP algorithm uses differential calculus in locating

the optimum points and allows to simultaneously minimize the objective function and the constraint function. An alternative to these methods are the EAs, such as GAs, and the PSO, which allow solving parameters tuning problems that classical methods might fail for multidimensional non-linear systems, such as the AdEx model. These algorithms provide high flexibility, universality (being able to be applied to different cases) and proved to be fast and efficient strategies to take into consideration for fitting neuron models (Cachón and Vázquez, 2015; Van Geit et al., 2016; Shan et al., 2017). This is the case of the optimization of an AdEx model of a cerebellar granule cell (GrC) and a Golgi cell (GoC) proposed in Nair et al. (2015). The fitness function measured the similarity between spike trains from spiking traces. However, the PSO algorithm was modified since all the solutions of the search did not result in a feasible solution due to the non-linear dynamics of the AdEx equations. In our previous study (Marín et al., 2020), we optimized an AdEx neuron model of a cerebellar GrC based on specific features (not whole traces) from *in vitro* recordings using "simple GA". In Marín et al. (2020), we proposed a single final candidate solution as the best approximation of the multi-feature fitness function of the cerebellar GrCs. However, in this work we take a step further in finding and fitting multiple neuron model configurations in a single run based on such a complex fitness function. This allows a detailed analysis of how neuron properties are supported by specific parameter configurations.

The objective of the present study is not to promote UEGO as the most effective algorithm in plain values of the objective function. Instead, the aim is to define an alternative framework that relies on this multimodal method for gathering and studying heterogeneous model configurations with independence of strictly being the best ranked. However, notice that UEGO can numerically compete with the results achieved by the GA used in the reference work (Marín et al., 2020). For the sake of completeness, the mean results of the GA proposed in Marín et al. (2020), which was the initial option for solving the problem at hand, have been compared to the mean results of the best-ranked solutions of the UEGO execution described in this article (**Table 3**). The referred GA took 30,000 function evaluations, but UEGO executes 50,000 on average with the configuration proposed, which is almost twice. For this reason, the number of cycles of the genetic method has been doubled to increase its exploration possibilities and take comparable computational effort. While the GA allows obtaining a unique best solution (low score), a multimodal algorithm such as UEGO allows generating multiple candidate solutions that reproduce reasonably well the neuron behaviors with wider parameter configurations (**Figure 7**).

Our aim is to provide a set of feasible, promising, and well-distributed solutions that result from numerical optimization for an expert to select the most appropriate one. The multimodal algorithms allow adjusting the exploration and consequent extraction of more than one candidate solution through the parameter landscapes (understood as the "space" of possible parameter values that a solution can take after the optimization process). An advantage of using a multimodal optimization method is that it results in a population of candidate solutions which is diverse in terms of parameters values as they best fit the target features at different areas in the parameter space. This allows using the candidate solutions as the substrate for a detailed parameter analysis with respect to the neuron model desired properties. In addition, the algorithm can adjust the parameter exploration according to different working ranges (wide exploration within the parameter boundaries). This population of final candidate solutions characterizes the behavior of the neuronal dynamics across the parameter space, i.e., how neuronal dynamics change as the parameters are modified (they can complement or conflict with each other towards optimizing a multi-feature fitness function, finding trade-offs among parameters in a solution). The expert is able to match the best parameter set ups towards optimizing one specific feature or another, or rather select a parameter set up to fit all the different target features at the same time (avoiding one feature to dominate against other ones within the combined cost function). The exploration and extraction of a diverse population of solutions facilitate the analysis process of how specific parameters ranges help to adjust particular features (the algorithm might perform an unbalanced adjustment of features, focusing more on some of them and distances to others). The multimodal optimization, in spite of being a more specific and robust engine (and so, more laborious), implies a suitable alternative for detailed exploration and analysis of the neuronal dynamics against the single candidate solution obtained by other simpler algorithms which might lose biological information necessary for the subsequent study.

TABLE 3 | Comparative table of best solutions from UEGO and regular GA.

Method	a (nS)	V_{peak} (mV)	V_T (mV)	b (pA)
UEGO	0.12 ± 0.10	−3.06 ± 14.25	−34.48 ± 14.67	−0.22 ± 0.59
GA	0.34 ± 0.09	−1.69 ± 11.24	−28.96 ± 7.23	0.48 ± 0.20
	C_m (pF)	E_L (mV)	g_L (nS)	Δ_T (mV)
UEGO	4.23 ± 0.72	−64.87 ± 12.61	1.18 ± 2.35	353.29 ± 357.56
GA	3.49 ± 0.45	−48.59 ± 4.95	0.63 ± 0.48	15.53 ± 7.33
	τ_w (ms)	V_r (mV)	Total score	
UEGO	166.41 ± 236.07	−66.93 ± 11.04	127.43 ± 30.13	
GA	346.13 ± 177.24	−69.26 ± 7.06	107.61 ± 5.04	

The table shows the mean values and standard deviations of each neuron parameter from the best candidate solutions of UEGO (candidate solutions from 1 to 10) and each of the single best solutions from 10 independent executions of regular GA (with different seeds). Note that the dispersion of the parameters is considerably wider in UEGO.

Future Implications

The novel workflow presented here constitutes a flexible and versatile tool that can be generally applied to this level of complexity with other commonly used point-neuron models, such as the GLIF or integrate-and-fire neuron models, and with other types of spiking dynamics, as long as the electrophysiological data is available. The multimodal optimization algorithm is only led by the value of the objective function, but this approach does not determine the goodness of the solution (the minimized score), although it is capable of exploring a biodiverse population of solutions according

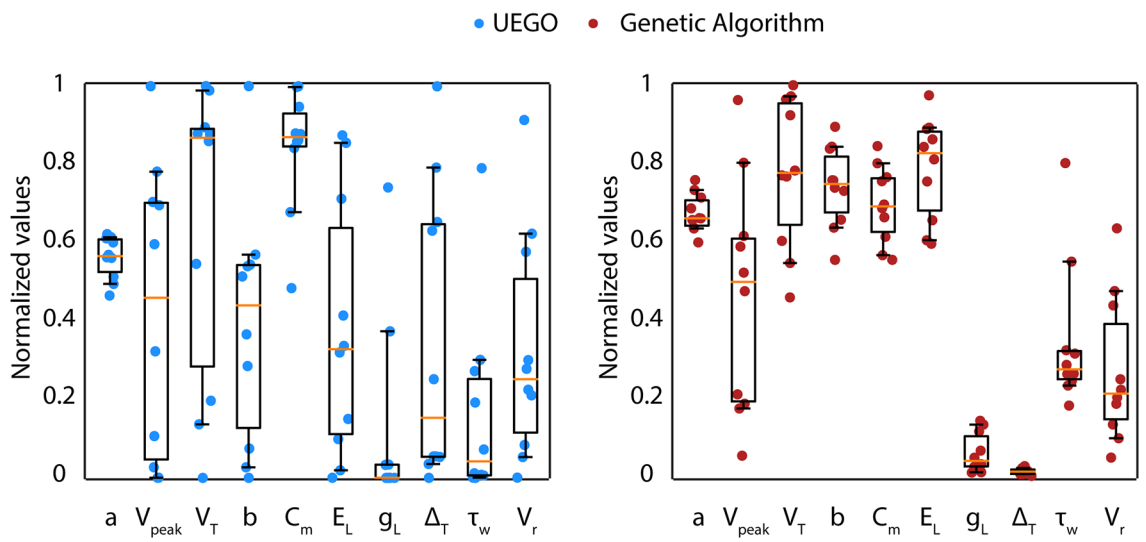


FIGURE 7 | Parameter distribution of candidate solutions from UEGO and GA. Comparison between the parameter values obtained from the best-ranked candidate solutions (from 1 to 10) in the execution of UEGO (left—blue circles) and from the best single solution of 10 independent executions of the GA (right—red circles). Each dot corresponds to the parameter value that defines every candidate solution. The boxes correspond to the IQR between the first (Q1) and the third (Q3) quartiles. The orange line represents the median among the values obtained for each parameter. The whiskers correspond to the 5- and 95- percentile respectively.

to pre-optimized solutions with interrelated parameters. The pre-optimization allows filtering the solutions according to numerically promising configurations. This facilitates the analysis of the parameter space in relation to the desired neuron properties. The post-optimization is based on the decision of the user. This proposal is an automatization of the population diversity of plausible neuron models for complex spiking behaviors.

In our results we already have seen certain biodiversity in the parameter configurations of the final population that can lead to a specific behavior shown by biological cells. If the heterogeneity of GrCs is a real fact in the biology of granule cells, then it could be also reflected in the variability of neuronal dynamics of the neuron models that reproduce the same target behavior (Lengler et al., 2013; Migliore et al., 2018). Regarding the biological data used as reference, in this article we generate a heterogeneous population of neurons based on different parameter configurations of the AdEx model and mimicking the neuronal behavior extracted from biological data. In future work, it would be of outstanding interest to optimize from a population of real cerebellar neurons that show variations in the target behaviors so that diversity can be explicitly captured. The construction of a multi-objective fitness function, compounded by several error functions that all have to be optimized simultaneously, could be a future extension of the presented workflow in order to analyze the Pareto front of all the possible parameter configurations. This would allow exploring the direct relationships among parameters and single features.

Concluding Remarks

In this article, we present a novel and robust optimization framework integrating a multimodal algorithm that co-optimizes the spiking resonance in the theta-frequency band, the repetitive spiking discharge and the latency to the first spike in efficient models of cerebellar GrCs. The validity of the framework is confirmed by analyzing the electrophysiological predictions of the biological characteristics. The proposed methodology will be reflected as ease-of-use through the following workflow, even though a multimodal algorithm usually requires high knowledge of the field and it is difficult to use for non-expert users. The UEGO algorithm exhibits its strength in adapting to the complex data structure associated with the neuron dynamics. The optimization workflow helps to easily generate a population of functional neuron models. In addition, employing a multimodal algorithm plays a key role in the proposed workflow to help the exploration of different local minima. The outcomes of the optimization study show promising results that successfully establish the solution repository considering multiple features in the function. Such results are verified by presenting the spiking resonance, repetitive firing and timing curves and the dominated solutions. According to the analytical results, the candidate solutions exhibit a consonant relationship between the features, meaning that the algorithm does not need to make a decision to balance the trade-off benefits (equilibrated distributions). The efficient models and features obtained in this work are mainly to demonstrate the feasibility of the proposed optimization workflow. It can be easily modified by other types of point-neuron models (such as GLIF) or other neuron characteristics in future work. The application to the case of cerebellar GrCs implies taking a step further towards advanced exploration of candidate solutions. It facilitates the evaluation of models based on different neuronal parameters which represent various internal neuronal mechanisms to achieve the target spiking behaviors defined in a complex fitness function.

DATA AVAILABILITY STATEMENT

The methodology and datasets presented in this study can be found in an online repository. The repository can be found at: https://github.com/MilagrosMarin/Multimodal-optimization-for-fitting-cerebellar-GrCs.

AUTHOR CONTRIBUTIONS

MM and NC: study design. MM, MS-L, and EO: literature and database search. NC and EO: methodology. MM, JG, MS-L, and RC: analysis and interpretation of results. MM, NC, and JG: writing of the article. All the results included in this article are part of MM's PhD thesis. All authors contributed to the article and approved the submitted version.

FUNDING

This article integrates work from authors from different research groups and has been funded by the EU Grant HBP (H2020 SGA3. 945539), the Spanish Ministry of Economy and Competitiveness (RTI2018-095993-B-I00), the national grant INTSENSO (MICINN-FEDER-PID2019-109991GB-I00), the regional grants of Junta de Andalucía (CEREBIO: JA FEDER P18-FR-2378, P18-RT-1193, and A-TIC-276-UGR18), and the University of Almería (UAL18-TIC-A020-B).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fninf.2021.663797/full#supplementary-material.

REFERENCES

Brette, R., and Gerstner, W. (2005). Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J. Neurophysiol.* 94, 3637–3642. doi: 10.1152/jn.00686.2005

Cachón, A., and Vázquez, R. A. (2015). Tuning the parameters of an integrate and fire neuron *via* a genetic algorithm for solving pattern recognition problems. *Neurocomputing* 148, 187–197. doi: 10.1016/j.neucom.2012.11.059

Caswell, T. A., Droettboom, M., Lee, A., Hunter, J., Firing, E., Stansby, D., et al. (2020). *matplotlib/matplotlib v2.2.5.* Available online at: https://zenodo.org/record/1202077#.YHqgMmc71v0.

Cruz, N. C., Redondo, J. L., Álvarez, J. D., Berenguel, M., and Ortigosa, P. M. (2018). Optimizing the heliostat field layout by applying stochastic population-based algorithms. *Informatica* 29, 21–39. doi: 10.15388/informatica.2018.156

D'Angelo, E., De Filippi, G., Rossi, P., and Taglietti, V. (1995). Synaptic excitation of individual rat cerebellar granule cells in situ: evidence for the role of NMDA receptors. *J. Physiol.* 484, 397–413. doi: 10.1113/jphysiol.1995.sp020673

D'Angelo, E., De Filippi, G., Rossi, P., and Taglietti, V. (1998). Ionic mechanism of electroresponsiveness in cerebellar granule cells implicates the action of a persistent sodium current. *J. Neurophysiol.* 80, 493–503. doi: 10.1152/jn.1998.80.2.493

D'Angelo, E., Koekkoek, S. K. E., Lombardo, P., Solinas, S., Ros, E., Garrido, J. A., et al. (2009). Timing in the cerebellum: oscillations and resonance in the granular layer. *Neuroscience* 162, 805–815. doi: 10.1016/j.neuroscience.2009.01.048

D'Angelo, E., Nieus, T., Maffei, A., Armano, S., Rossi, P., Taglietti, V., et al. (2001). Theta-frequency bursting and resonance in cerebellar granule cells: experimental evidence and modeling of a slow K^+-dependent mechanism. *J. Neurosci.* 21, 759–770. doi: 10.1523/JNEUROSCI.21-03-00759.2001

Delvendahl, I., Straub, I., and Hallermann, S. (2015). Dendritic patch-clamp recordings from cerebellar granule cells demonstrate electrotonic compactness. *Front. Cell. Neurosci.* 9:93. doi: 10.3389/fncel.2015.00093

Druckmann, S., Banitt, Y., Gidon, A. A., Schürmann, F., Markram, H., and Segev, I. (2007). A novel multiple objective optimization framework for constraining conductance-based neuron models by experimental data. *Front. Neurosci.* 1, 7–18. doi: 10.3389/neuro.01.1.1.001.2007

Friedrich, P., Vella, M., Gulyás, A. I., Freund, T. F., and Káli, S. (2014). A flexible, interactive software tool for fitting the parameters of neuronal models. *Front. Neuroinform.* 8:63. doi: 10.3389/fninf.2014.00063

Gandolfi, D., Lombardo, P., Mapelli, J., Solinas, S., and D'Angelo, E. (2013). Theta-frequency resonance at the cerebellum input stage improves spike timing on the millisecond time-scale. *Front. Neural Circuits* 7:64. doi: 10.3389/fncir.2013.00064

Geminiani, A., Casellato, C., Locatelli, F., Prestori, F., Pedrocchi, A., and D'Angelo, E. (2018). Complex dynamics in simplified neuronal models: reproducing golgi cell electroresponsiveness. *Front. Neuroinform.* 12:88. doi: 10.3389/fninf.2018.00088

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., et al. (2020). Array programming with NumPy. *Nature* 585, 357–362. doi: 10.1038/s41586-020-2649-2

Hunter, J. D. (2007). Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* 9, 90–95. doi: 10.1109/mcse.2007.55

Izhikevich, E. M. (2004). Which model to use for cortical spiking neurons? *IEEE Trans. Neural Netw.* 15, 1063–1070. doi: 10.1109/TNN.2004.832719

Jelasity, M., Ortigosa, P. M., and García, I. (2001). UEGO, an abstract clustering technique for multimodal global optimization. *J. Heuristics* 7, 215–233. doi: 10.1023/A:1011367930251

Jolivet, R., Kobayashi, R., Rauch, A., Naud, R., Shinomoto, S., and Gerstner, W. (2008). A benchmark test for a quantitative assessment of simple neuron models. *J. Neurosci. Methods* 169, 417–424. doi: 10.1016/j.jneumeth.2007.11.006

Jörntell, H., and Ekerot, C.-F. (2006). Properties of somatosensory synaptic integration in cerebellar granule cells *in vivo*. *J. Neurosci.* 26, 11786–11797. doi: 10.1523/JNEUROSCI.2939-06.2006

Lange, W. (1975). Cell number and cell density in the cerebellar cortex of man and some other mammals. *Cell Tissue Res.* 157, 115–124. doi: 10.1007/BF00223234

Lengler, J., Jug, F., and Steger, A. (2013). Reliable neuronal systems: the importance of heterogeneity. *PLoS One* 8:e80694. doi: 10.1371/journal.pone.0080694

Lindfield, G., and Penny, J. (2017). *Introduction to Nature-Inspired Optimization.* 1st Edn. Academic Press. Available online at: https://www.elsevier.com/books/introduction-to-nature-inspired-optimization/lindfield/978-0–12-803636–5. Accessed January 27, 2021.

Marín, M., Sáez-Lara, M. J., Ros, E., and Garrido, J. A. (2020). Optimization of efficient neuron models with realistic firing dynamics. The case of the cerebellar granule cell. *Front. Cell. Neurosci.* 14:161. doi: 10.3389/fncel.2020.00161

Masoli, S., Rizza, M. F., Sgritta, M., Van Geit, W., Schürmann, F., and D'Angelo, E. (2017). Single neuron optimization as a basis for accurate biophysical modeling: the case of cerebellar granule cells. *Front. Cell. Neurosci.* 11:71. doi: 10.3389/fncel.2017.00071

Migliore, R., Lupascu, C. A., Bologna, L. L., Romani, A., Courcol, J.-D., Antonel, S., et al. (2018). The physiological variability of channel density in hippocampal CA1 pyramidal cells and interneurons explored using a unified data-driven modeling workflow. *PLoS Comput. Biol.* 14:e1006423. doi: 10.1371/journal.pcbi.1006423

Molina, D., Lozano, M., Sánchez, A. M., and Herrera, F. (2011). Memetic algorithms based on local search chains for large scale continuous optimisation problems: MA-SSW-Chains. *Soft Comput.* 15, 2201–2220. doi: 10.1007/s00500-010-0647-2

Moscato, P. (1989). *On Evolution, Search, Optimization, Genetic Algorithms and Martial Arts: Towards Memetic Algorithms.* Available online

at: https://citeseer.ist.psu.edu/moscato89evolution.html. Accessed January 27, 2021.

Nair, M., Subramanyan, K., Nair, B., and Diwakar, S. (2015). "Parameter optimization and nonlinear fitting for computational models in neuroscience on GPGPUs," in *Proceedings of the 2014 International Conference on High Performance Computing and Applications (ICHPCA)* (Bhubaneswar, India: IEEE), 22–24. doi: 10.1109/ICHPCA.2014.7045324

Naud, R., Marcille, N., Clopath, C., and Gerstner, W. (2008). Firing patterns in the adaptive exponential integrate-and-fire model. *Biol. Cybern.* 99, 335–347. doi: 10.1007/s00422-008-0264-7

Ortigosa, P. M., García, I., and Jelasity, M. (2001). Reliability and performance of UEGO, a clustering-based global optimizer. *Journal of Global Optimization* 19, 265–289. doi: 10.1023/A:1011224107143

Ortigosa, P. M., Redondo, J. L., García, I., and Fernández, J. J. (2007). A population global optimization algorithm to solve the image alignment problem in electron crystallography. *J. Glob. Optim.* 37, 527–539. doi: 10.1007/s10898-006-9060-x

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.

Peyser, A., Sinha, A., Vennemo, S. B., Ippen, T., Jordan, J., Graber, S., et al. (2017). *NEST 2.14.0.* Available online at: http://dx.doi.org/10.5281/zenodo.882971.

Pozzorini, C., Mensi, S., Hagens, O., Naud, R., Koch, C., and Gerstner, W. (2015). Automated high-throughput characterization of single neurons by means of simplified spiking models. *PLoS Comput. Biol.* 11:e1004275. doi: 10.1371/journal.pcbi.1004275

Python Software Foundation Python 2.7.12. (2016). *Python Lang. Ref. Version 2.7.* Available online at: http://www.python.org.

Redondo, J. L. (2009). *Solving Competitive Location Problems via Memetic Algorithms. High Performance Computing Approaches.* Thesis. Almería: Universidad Almería.

Rongala, U. B., Spanne, A., Mazzoni, A., Bengtsson, F., Oddo, C. M., and Jörntell, H. (2018). Intracellular dynamics in cuneate nucleus neurons support self-stabilizing learning of generalizable tactile representations. *Front. Cell. Neurosci.* 12:210. doi: 10.3389/fncel.2018.00210

Sareni, B., and Krähenbühl, L. (1998). Fitness sharing and niching methods revisited. *IEEE Trans. Evol. Comput.* 2, 97–106. doi: 10.1109/4235.735432

Schmahmann, J. D. (2019). The cerebellum and cognition. *Neurosci. Lett.* 688, 62–75. doi: 10.1016/j.neulet.2018.07.005

Shan, B., Wang, J., Zhang, L., Deng, B., and Wei, X. (2017). Fitting of adaptive neuron model to electrophysiological recordings using particle swarm optimization algorithm. *Int. J. Mod. Phys. B* 31, 1–15. doi: 10.1142/s0217979217500230

Shir, O. M., Emmerich, M., and Bäck, T. (2010). Adaptive niche radii and niche shapes approaches for niching with the CMA-ES. *Evol. Comput.* 18, 97–126. doi: 10.1162/evco.2010.18.1.18104

Silver, R. A., Traynelis, S. F., and Cull-Candy, S. G. (1992). Rapid-time-course miniature and evoked excitatory currents at cerebellar synapses in situ. *Nature* 355, 163–166. doi: 10.1038/355163a0

Solis, F. J., and Wets, R. J. B. (1981). Minimization by random search techniques. *Math. Operations Res.* 6, 19–30. doi: 10.1287/moor.6.1.19

Van Geit, W., De Schutter, E., and Achard, P. (2008). Automated neuron model optimization techniques: a review. *Biol. Cybern.* 99, 241–251. doi: 10.1007/s00422-008-0257-6

Van Geit, W., Gevaert, M., Chindemi, G., Rössert, C., Courcol, J.-D., Muller, E. B., et al. (2016). BluePyOpt: leveraging open source software and cloud infrastructure to optimise model parameters in neuroscience. *Front. Neuroinform.* 10:17. doi: 10.3389/fninf.2016.00017

van Rossum, G. (1995). *Python Tutorial, May 1995. CWI Rep. CS-R9526, 1–65.* Available online at: https://www.narcis.nl/publication/RecordID/oai%3Acwi.nl%3A5008.

Williams, R. W., and Herrup, K. (1988). The control of neuron number. *Annu. Rev. Neurosci.* 11, 423–453. doi: 10.1146/annurev.ne.11.030188.002231

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Chapter 4

General Discussion

This chapter shows a summary of the main contributions presented in this thesis and a proposal of future work.

4.1 Revisiting the thesis objectives

We specify the journal articles included in this thesis that have addressed the initially-planned specific objectives, as follows:

1. Deepening in the biochemical and physiological knowledge of the brain	•Marín et al. (2019) •Marín et al. (2020)
2. Development of a methodology for simulation and simplification of neuron models	•Marín et al. (2020) •Marín et al. (2021) •Cruz et al. (2021)
3. Development of computationally efficient neuron models that integrate nonlinear cellular dynamics	•Marín et al. (2020) •Marín et al. (2021) •Cruz et al. (2021)
4. Study of the resonance at the cellular level in cerebellar granule cells and evaluation of its computational capabilities in the brain	•Marín et al. (2020) •Marín et al. (2021) •Cruz et al. (2021)
5. Development of a model of cerebellar granule cell integrating nonlinear cellular dynamics (such as the spiking resonance)	•Marín et al. (2020) •Marín et al. (2021) •Cruz et al. (2021)
6. Evaluation of the capabilities and efficiency in the information transmission of the granular layer with nonlinear cellular models	•Future work

Due to time constraints, the specific aims regarding the evaluation of the learning capabilities in the information processing of nonlinear cellular properties are the next steps to address in our future research (as we detail in the next section of this chapter "**4.3 Future Work**").

It is also worthy to mention that this thesis included an additional experimentation regarding a preliminary exploration of the connection between nonlinear neuronal dynamics

and their potential clinical manifestations and genetic causes. Particularly, this thesis has aimed two complementary specific objectives:

7. Deepening in the knowledge of cerebellar diseases. Particularly, exploring the possible neurological diseases that could allow connecting the different levels of genetics, electrophysiology, neuron level and network level.	•Marín et al. (2019)
8. Development of a methodology for the identification of the most relevant genes and manifestations of a neurological complex disease and its potential pharmacological targets. This aim will allow extracting features reproducible at a computational level.	•Marín et al. (2019)

4.2 Main contributions

The preliminary experimentation of this thesis is headed for a better understanding of the etiopathogenesis and treatment of neurological diseases. Our first main contribution is the presentation of a semi-automatic integrative workflow that identify the most common polygenic influences and their associated manifestations (and still based on similar molecular mechanisms) throughout a set of diseases, contributing to the overall understanding of their pathomechanisms. This contribution allows shedding some light on the up-to-date knowledge of pathomechanisms of neurological disorders and related manifestations, and their diagnosis and potential therapies. Particularly, the set of complex disorders to analyze is the case of channelopathies.

Methodologically, the proposed workflow mines multi-dimensional databases and platforms based on systems biology approaches, as protein-protein interaction networks. This pipeline aims to be easy-to-use for a non-expert user in systems biology. It also allows extracting the most relevant genes of complex disease and interpreting its biological processes and possible comorbidities with other diseases.

This contribution establishes a new bridge-builder among fields that extract as productive results as other traditional search systems (such as exhaustive or systematic searches) and software tools for functional annotation. Moreover, the proposed workflow even

extracts information that *a priori* does not seem relevant when the starting point is a very large group of genes in disease.

The second main contribution that this thesis has achieved is directed towards the understanding of the brain operation with respect to the impact that potential intrinsic dynamics of neurons could take part on them. Thus, we contributed by developing a methodology for the simulation and simplification of mathematical neuron models. Specifically, the aim of the methodology is to create computationally efficient neuron models that reproduce non-linear cellular dynamics. That is, to find suitable sets of model parameters in order to capture specific firing dynamics under different stimulation protocols (including different spiking properties in the setup of the algorithms used). These models will facilitate further *in silico* simulations of large-scale microcircuits to better understand the computational role of the suprathreshold dynamics of the cell on a large scale.

In this regard, we have studied automatic parameter exploration strategies in order to develop simplified neuron models based on the AdEx generic model template. Considering both biological relevance and computational efficiency, the proposed strategies aim to capture essential aspects of single-neuron processing. We have also compared alternative and sophisticated optimization methods widely used in other fields such as Engineering.

The third main contribution in this thesis is focused on the application of these methodologies to the case of cerebellar granule cells (GrCs) in order to replicate the most essential properties of the biological cell. That is, the neuronal dynamics that are key for the frequency and timing of firing patterns in the neuronal code. This contribution involves studying different mathematical definitions of the neuron electroresponsiveness (named objective functions) experimentally evidenced in previous studies to conduct the optimization of this spiking activity.

Finally, we have contributed by studying different mathematical definitions of the neuron electroresponsiveness (in what is called objective functions) experimentally evidenced in previous studies to conduct the optimization of the spiking activity. More specifically, the part of the thesis has been directed towards the study of a potential cellular mechanism in the somatosensory information representations and processing at the cerebellar cortex, i.e., the intrinsic resonance in the theta-frequency band of the cerebellar GrCs. Thus, the last main contribution is the generation of different computational neuron models of cerebellar GrCs replicating this complex behavior. We have complemented this contribution by also presenting heterogeneous populations of GrC models that replicate the intrinsic variances in their properties.

4.3 Future work

As future work, it is possible to define new goals in the two main research lines covered.

Regarding the presentation of an integrative methodology to identify relevant genes in complex diseases, long-term applications are directed towards the exploration of the dysfunctional dynamics of neurons. Particularly, those cell dynamics resulted from the alteration of these relevant genes. The replication of the resulting behavior of these alterations would allow simulating models of disease at cellular and network levels.

With respect to the study of the functional implication of the theta-frequency-band resonance for information processing at the cerebellar cortex, the first step after this thesis is to address the evaluation of the computational capabilities of this complex behavior. For this aim, we will simulate a neuron network layer composed of several cerebellar GrCs using the neuron models proposed in this thesis. We expect to demonstrate that the intrinsic resonance in theta-band of GrCs regulates the information transfer by enhancing the signal transmission and learning capabilities in the granular layer, as hypothesized in previous works (D'Angelo et al., 2001, 2009; Solinas et al., 2007; Gandolfi et al., 2013; Garrido et al., 2016). Thus, the integration of realistic firing regimes in the neuron models presented (such as bursting and resonance) into complex learning tasks would improve the understanding of the mechanisms of learning.

4.4 Conclusions

1. We have shown the usefulness of a semi-automatic integrative workflow that successfully mines current available databases and platforms based on protein-protein interaction networks applied to channelopathies. This workflow has demonstrated to be able to produce as significant results as a non-automatic research system but in a more efficiently way, functioning as a bridge-builder among fields. It is also able to extract information that *a priori* might not seem relevant when the starting point is a very large group of genes potentially related to a disease.

2. We have presented an automatic optimization strategy based on traditional GAs for the development of computationally efficient neuron models that reproduce realistic firing properties under different stimulation protocols. We have proposed cerebellar GrC models that suitably predict the main suprathreshold dynamics, such as the spiking resonance at the theta-frequency band, the repetitive spiking discharge and the latency to the first spike. These models provide both efficiency and biological plausibility,

facilitating further studies about the computational role of these cells in the framework of large-scale neuron networks.

3. The previous optimization context has been further studied to find more effective methods. Four different meta-heuristics optimizers, some of them successfully used in diverse problems of other fields, have been compared to the GA originally proposed. According to the results, they outperform quantitatively and qualitatively the reference method. The resulting model gains in biological realism, featuring higher temporal accuracy of the first spike than that obtained in the previous article. This finding supports the relevance of using an effective and efficient optimization engine in the referred methodology.

4. We have also studied the existence of multiple sub-optimal solutions using a multimodal optimization algorithm that can keep track of the different regions. This work has been presented as a novel and robust optimization framework integrating a multimodal algorithm applied to the same challenge of optimizing efficient cerebellar GrC models. The proposed workflow is reflected as ease-of-use even though a multimodal algorithm usually requires high knowledge of the field and it could be difficult to use for non-expert users. This workflow have explored different local minima and easily generated a population of functional neuron models that achieve successful biological realism in the selected target features using different parameter configurations. Thus, multimodal optimization represents an efficient optimization strategy for the exploration of the parameter space of these kind of neuron model.

5. According to the analytical results, the resulting neuron models exhibit a significant relationship between the features, meaning that the algorithm achieves a good trade-off between the different target features. The efficient models and features obtained in this work demonstrate the viability of the proposed optimization workflow, which can be customized for other types of point-neuron models (such as GLIF) or other neuron features. It facilitates the evaluation of models based on different neuronal parameters that represent various internal neuronal mechanisms to achieve the target spiking behaviors defined in a complex fitness function.

4.5 Conclusiones

1. Hemos demostrado la utilidad de un flujo de trabajo integrador semi-automático que explota con éxito las bases de datos y plataformas disponibles en la actualidad basadas en redes de interacción proteína-proteína aplicadas a las canalopatías. Este flujo de trabajo ha demostrado ser capaz de producir resultados tan significativos como un

sistema de búsqueda no automático, pero de una manera más eficiente, funcionando como un constructor de puentes entre campos de investigación. También es capaz de extraer información que a priori podría no parecer relevante cuando el punto de partida es un grupo muy grande de genes potencialmente relacionados con una enfermedad.

2. Hemos presentado una estrategia de optimización automática basada en GAs (algoritmos genéticos) tradicionales para el desarrollo de modelos de neurona computacionalmente eficientes y que reproducen propiedades de disparo realistas bajo diferentes protocolos de estimulación. Hemos propuesto modelos de GrCs cerebelares que predicen adecuadamente las principales dinámicas supra-umbrales, como la resonancia de disparo en la banda de frecuencia theta, la descarga repetitiva de disparo y la latencia al primer disparo. Estos modelos proporcionan tanto eficiencia como plausibilidad biológica, facilitando nuevos estudios sobre el papel computacional de estas células en el marco de las redes neuronales a gran escala.

3. El contexto de optimización anterior se ha estudiado más a fondo para encontrar métodos más eficaces. Se han comparado cuatro optimizadores meta-heurísticos diferentes, algunos de ellos utilizados con éxito en diversos problemas de otros campos de investigación, con el algoritmo genético (GA) propuesto anteriormente. Según los resultados, superan cuantitativa y cualitativamente al método de referencia. El modelo resultante gana en realismo biológico, presentando una mayor precisión temporal al primer disparo que aquélla obtenida en el artículo anterior. Este hallazgo apoya la relevancia de utilizar un motor de optimización eficaz y eficiente en la metodología referida.

4. También hemos estudiado la existencia de múltiples soluciones sub-óptimas utilizando un algoritmo de optimización multimodal que puede rastrear las diferentes regiones. Este trabajo ha sido presentado como un marco de optimización novedoso y robusto que integra un algoritmo multimodal aplicado al mismo reto de optimizar modelos eficientes de GrCs cerebelares. El flujo de trabajo propuesto destaca por la facilidad de uso a pesar de que un algoritmo multimodal suele requerir un alto conocimiento de la materia y podría ser difícil de utilizar para usuarios no expertos. Este flujo de trabajo ha explorado diferentes mínimos locales y ha generado fácilmente una población de modelos de neurona funcionales que consiguen un logrado realismo biológico en las características objetivo seleccionadas utilizando diferentes configuraciones de parámetros. Así, la optimización multimodal representa una estrategia de optimización eficiente para la exploración del espacio de parámetros de este tipo de modelos neuronales.

5. En función de los resultados analíticos, los modelos de neurona resultantes muestran una relación significativa entre las características, lo que significa que el algoritmo logra un buen equilibrio entre las diferentes características objetivo. Los modelos de neurona eficientes y las características obtenidas en este trabajo demuestran la viabilidad del flujo de trabajo de optimización propuesto, que puede personalizarse para otro tipo de modelos de neurona punto (como el modelo GLIF) u otras características neuronales. Este flujo de trabajo facilita la evaluación de modelos basados en diferentes parámetros neuronales que representan varios mecanismos neuronales internos para lograr los comportamientos de disparo objetivo definidos en una función fitness compleja.

Bibliography

Albus, J. S. (1971). A theory of cerebellar function. *Math. Biosci.* 10, 25–61. doi:10.1016/0025-5564(71)90051-4.

Applied Computational Neuroscience Research Group Available at: http://acn.ugr.es/.

Armano, S., Rossi, P., Taglietti, V., and D'Angelo, E. (2000). Long-term potentiation of intrinsic excitability at the mossy fiber-granule cell synapse of rat cerebellum. *J. Neurosci.* 20, 5208–16. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10884304 [Accessed February 15, 2019].

Ashida, R., Cerminara, N. L., Brooks, J., and Apps, R. (2018). *Principles of organization of the human cerebellum: macro- and microanatomy*. 1st ed. Elsevier B.V. doi:10.1016/B978-0-444-63956-1.00003-5.

Bareš, M., Apps, R., Avanzino, L., Breska, A., D'Angelo, E., Filip, P., et al. (2019). Consensus paper: Decoding the Contributions of the Cerebellum as a Time Machine. From Neurons to Clinical Applications. *Cerebellum* 18, 266–286. doi:10.1007/s12311-018-0979-5.

Barranca, V. J., Johnson, D. C., Moyher, J. L., Sauppe, J. P., Shkarayev, M. S., Kovačič, G., et al. (2013). Dynamics of the exponential integrate-and-fire model with slow currents and adaptation. *J Comput Neurosci* 37, 161–180. doi:10.1007/s10827-013-0494-0.

Brette, R., and Gerstner, W. (2005). Adaptive Exponential Integrate-and-Fire Model as an Effective Description of Neuronal Activity. *J. Neurophysiol.* 94, 3637–3642. doi:10.1152/jn.00686.2005.

Buzsáki, G. (2006). *Rhythms of the brain*. New York, NY, US: Oxford University Press doi:10.1093/acprof:oso/9780195301069.001.0001.

Computational Models of Brain and Behavior (2017). John Wiley & Sons, Ltd doi:10.1002/9781119159193.

Courtemanche, R., Chabaud, P., and Lamarre, Y. (2009). Synchronization in primate cerebellar granule cell layer local field potentials: basic anisotropy and dynamic changes during active expectancy. *Front. Cell. Neurosci.* 3, 6. doi:10.3389/neuro.03.006.2009.

Coveney, P. V., Dougherty, E. R., and Highfield, R. R. (2016). Big data need big theory too. in *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 20160153. doi:10.1098/rsta.2016.0153.

Cruz, N. C., Marín, M., Redondo, J. L., Ortigosa, E. M., and Ortigosa, P. M. (2021). A Comparative Study of Stochastic Optimizers for Fitting Neuron Models. Application to the Cerebellar Granule Cell. *Informatica* 0, 1–22. doi:10.15388/21-infor450.

D'Angelo, E., Koekkoek, S. K. E., Lombardo, P., Solinas, S., Ros, E., Garrido, J. A., et al. (2009). Timing in the cerebellum: oscillations and resonance in the granular layer. *Neuroscience* 162, 805–815. doi:10.1016/j.neuroscience.2009.01.048.

D'Angelo, E., Mazzarello, P., Prestori, F., Mapelli, J., Solinas, S., Lombardo, P., et al. (2011). The cerebellar network: From structure to function and dynamics. *Brain Res. Rev.* 66, 5–15. doi:10.1016/j.brainresrev.2010.10.002.

D'Angelo, E., Nieus, T., Maffei, A., Armano, S., Rossi, P., Taglietti, V., et al. (2001). Theta-Frequency Bursting and Resonance in Cerebellar Granule Cells: Experimental Evidence and Modeling of a Slow K + -Dependent Mechanism. *J. Neurosci.* 21, 759–770. doi:10.1523/JNEUROSCI.21-03-00759.2001.

D'Angelo, E., Solinas, S., Garrido, J. A., Casellato, C., Pedrocchi, A., Mapelli, J., et al. (2013a). Realistic modeling of neurons and networks: towards brain simulation. *Funct. Neurol.* 28, 153. doi:10.11138/FNeur/2013.28.3.153.

D'Angelo, E., Solinas, S., Mapelli, J., Gandolfi, D., Mapelli, L., and Prestori, F. (2013b). The cerebellar Golgi cell and spatiotemporal organization of granular layer activity. *Front. Neural Circuits* 7, 93. doi:10.3389/fncir.2013.00093.

Das, A., and Narayanan, R. (2017). Theta-frequency selectivity in the somatic spike-triggered average of rat hippocampal pyramidal neurons is dependent on HCN channels. *J. Neurophysiol.* 118, 2251–2266. doi:10.1152/jn.00356.2017.

De Carlos, J. A., and Borrell, J. (2007). A historical reflection of the contributions of Cajal and Golgi to the foundations of neuroscience. *Brain Res. Rev.* 55, 8–16. doi:10.1016/j.brainresrev.2007.03.010.

Delvendahl, I., Straub, I., and Hallermann, S. (2015). Dendritic patch-clamp recordings from cerebellar granule cells demonstrate electrotonic compactness. *Front. Cell. Neurosci.* 9, 93. doi:10.3389/fncel.2015.00093.

Di Silvestre, D., Brambilla, F., Scardoni, G., Brunetti, P., Motta, S., Matteucci, M., et al. (2017). Proteomics-based network analysis characterizes biological processes and pathways activated by preconditioned mesenchymal stem cells in cardiac repair mechanisms. *Biochim. Biophys. Acta - Gen. Subj.* 1861, 1190–1199. doi:10.1016/j.bbagen.2017.02.006.

Diaz-Beltran, L., Cano, C., Wall, D., and Esteban, F. (2013). Systems Biology as a Comparative Approach to Understand Complex Gene Expression in Neurological Diseases. *Behav. Sci.*

(Basel). 3, 253–272. doi:10.3390/bs3020253.

Dugué, G. P., Brunel, N., Hakim, V., Schwartz, E., Chat, M., Lévesque, M., et al. (2009). Electrical Coupling Mediates Tunable Low-Frequency Oscillations and Resonance in the Cerebellar Golgi Cell Network. *Neuron* 61, 126–139. doi:10.1016/j.neuron.2008.11.028.

Eccles, J. C. (1973). *The understanding of the brain*. , ed. McGrow-Hill Available at: https://psycnet.apa.org/record/1973-28323-000 [Accessed May 7, 2021].

Eccles, J. C., Ito, M., and Szentágothai, J. (1967). *The Cerebellum as a Neuronal Machine*. Springer Berlin Heidelberg doi:10.1007/978-3-662-13147-3.

Ferreira González, I., Urrútia, G., and Alonso-Coello, P. (2011). Systematic Reviews and Meta-Analysis: Scientific Rationale and Interpretation. *Rev. Española Cardiol. (English Ed.* 64, 688–696. doi:10.1016/j.rec.2011.03.027.

Fox, D. M., Tseng, H. A., Smolinski, T. G., Rotstein, H. G., and Nadim, F. (2017). Mechanisms of generation of membrane potential resonance in a neuron with multiple resonant ionic currents. *PLoS Comput. Biol.* 13, 1–30. doi:10.1371/journal.pcbi.1005565.

Friedrich, P., Vella, M., Gulyás, A. I., Freund, T. F., and Káli, S. (2014). A flexible, interactive software tool for fitting the parameters of neuronal models. *Front. Neuroinform.* 8, 63. doi:10.3389/fninf.2014.00063.

Gandolfi, D., Lombardo, P., Mapelli, J., Solinas, S., and D'Angelo, E. (2013). Theta-Frequency Resonance at the Cerebellum Input Stage Improves Spike Timing on the Millisecond Time-Scale. *Front. Neural Circuits* 7, 1–16. doi:10.3389/fncir.2013.00064.

Garrido, J. A., Luque, N. R., D'Angelo, E., and Ros, E. (2013a). Distributed cerebellar plasticity implements adaptable gain control in a manipulation task: a closed-loop robotic simulation. *Front. Neural Circuits* 7, 159. doi:10.3389/fncir.2013.00159.

Garrido, J. A., Luque, N. R., Tolu, S., and D'Angelo, E. (2016). Oscillation-Driven Spike-Timing Dependent Plasticity Allows Multiple Overlapping Pattern Recognition in Inhibitory Interneuron Networks. *Int. J. Neural Syst.* 26, 1650020. doi:10.1142/S0129065716500209.

Garrido, J. A., Ros, E., and D'Angelo, E. (2013b). Spike Timing Regulation on the Millisecond Scale by Distributed Synaptic Plasticity at the Cerebellum Input Stage: A Simulation Study. *Front. Comput. Neurosci.* 7. doi:10.3389/fncom.2013.00064.

Geminiani, A., Casellato, C., Antonietti, A., D'Angelo, E., and Pedrocchi, A. (2018). A Multiple-Plasticity Spiking Neural Network Embedded in a Closed-Loop Control System to Model Cerebellar Pathologies. *Int. J. Neural Syst.* 28, 1750017. doi:10.1142/S0129065717500174.

Gewaltig, M. O., Morrison, A., and Plesser, H. E. (2012). NEST by example: An introduction to the neural simulation tool NEST. *Comput. Syst. Neurobiol.*, 533–558. doi:10.1007/978-94-007-3858-4_18.

Golgi, C. (1906). The Neuron Doctrine: theory and Facts. *Nobel Lect. Physiol. Med.*, 190–217.

Available at: https://www.nobelprize.org/prizes/medicine/1906/golgi/lecture/ [Accessed May 7, 2021].

Goñi, J., Esteban, F. J., de Mendizábal, N. V., Sepulcre, J., Ardanza-Trevijano, S., Agirrezabal, I., et al. (2008). A computational analysis of protein-protein interaction networks in neurodegenerative diseases. *BMC Syst. Biol.* 2, 52. doi:10.1186/1752-0509-2-52.

Guevara, E. E., Hopkins, W. D., Hof, P. R., Ely, J. J., Bradley, B. J., and Sherwood, C. C. (2021). Comparative analysis reveals distinctive epigenetic features of the human cerebellum. *PLOS Genet.* 17, e1009506. doi:10.1371/journal.pgen.1009506.

Hartmann, M. J., and Bower, J. M. (1998). Oscillatory Activity in the Cerebellar Hemispheres of Unrestrained Rats. *J. Neurophysiol.* 80, 1598–1604. doi:10.1152/jn.1998.80.3.1598.

Herculano-Houzel (2010). Coordinated scaling of cortical and cerebellar numbers of neurons. *Front. Neuroanat.* 4, 12. doi:10.3389/fnana.2010.00012.

Hines, M. L., and Carnevale, N. T. (1997). The NEURON Simulation Environment. *Neural Comput.* 9, 1179–1209. doi:10.1162/neco.1997.9.6.1179.

Human Brain Project (HBP) - Brain Simulation Available at: https://www.humanbrainproject.eu/en/brain-simulation/.

Hutcheon, B., and Yarom, Y. (2000). Resonance, oscillation and the intrinsic frequency preferences of neurons. *Trends Neurosci.* 23, 216–222. doi:10.1016/S0166-2236(00)01547-2.

Ito, M. (2014). "Cerebellar Microcircuitry☆," in *Reference Module in Biomedical Sciences* (Elsevier). doi:10.1016/b978-0-12-801238-3.04544-x.

Kim, J. B. (2014). Channelopathies. *Korean J Pediatr* 57, 1–18. doi:10.3345/kjp.2014.57.1.1.

Knupp, K., and Brooks-Kayal, A. R. (2017). "Channelopathies," in *Swaiman's Pediatric Neurology* (Elsevier), 405–411. doi:10.1016/B978-0-323-37101-8.00049-7.

Lengler, J., Jug, F., and Steger, A. (2013). Reliable Neuronal Systems: The Importance of Heterogeneity. *PLoS One* 8, 80694. doi:10.1371/journal.pone.0080694.

Maex, R., and De Schutter, E. (1998). Synchronization of Golgi and granule cell firing in a detailed network model of the cerebellar granule cell layer. *J. Neurophysiol.* 80, 2521–2537. doi:10.1152/jn.1998.80.5.2521.

Magistretti, J., Castelli, L., Forti, L., and D'Angelo, E. (2006). Kinetic and functional analysis of transient, persistent and resurgent sodium currents in rat cerebellar granule cells in situ: An electrophysiological and modelling study. *J. Physiol.* 573, 83–106. doi:10.1113/jphysiol.2006.106682.

Marín, M., Cruz, N. C., Ortigosa, E. M., Sáez-Lara, M. J., Garrido, J. A., and Carrillo, R. R. (2021). On the Use of a Multimodal Optimizer for Fitting Neuron Models. Application to the

Cerebellar Granule Cell. *Front. Neuroinform.* 15, 17. doi:10.3389/fninf.2021.663797.

Marín, M., Esteban, F. J., Ramírez-Rodrigo, H., Ros, E., and Sáez-Lara, M. J. (2019). An integrative methodology based on protein-protein interaction networks for identification and functional annotation of disease-relevant genes applied to channelopathies. *BMC Bioinformatics* 20, 565. doi:10.1186/s12859-019-3162-1.

Marín, M., Sáez-Lara, M. J., Ros, E., and Garrido, J. A. (2020). Optimization of Efficient Neuron Models With Realistic Firing Dynamics. The Case of the Cerebellar Granule Cell. *Front. Cell. Neurosci.* 14, 161. doi:10.3389/fncel.2020.00161.

Martin-Sanchez, F., and Verspoor, K. (2014). Big Data in Medicine Is Driving Big Changes. *Yearb. Med. Inform.* 23, 14–20. doi:10.15265/IY-2014-0020.

Masquelier, T., Hugues, E., Deco, G., and Thorpe, S. J. (2009). Oscillations, Phase-of-Firing Coding, and Spike Timing-Dependent Plasticity: An Efficient Learning Scheme. *J. Neurosci.* 29, 13484–13493. doi:10.1523/jneurosci.2207-09.2009.

Medini, C., Nair, B., D'Angelo, E., Naldi, G., and Diwakar, S. (2012). Modeling Spike-Train Processing in the Cerebellum Granular Layer and Changes in Plasticity Reveal Single Neuron Effects in Neural Ensembles. *Comput. Intell. Neurosci.* 2012, 1–17. doi:10.1155/2012/359529.

Meng, X., Li, J., Zhang, Q., Chen, F., Bian, C., Yao, X., et al. (2020). Multivariate genome wide association and network analysis of subcortical imaging phenotypes in Alzheimer's disease. *BMC Genomics* 21, 1–12. doi:10.1186/s12864-020-07282-7.

Migliore, R., Lupascu, C. A., Bologna, L. L., Romani, A., Courcol, J.-D., Antonel, S., et al. (2018). The physiological variability of channel density in hippocampal CA1 pyramidal cells and interneurons explored using a unified data-driven modeling workflow. *PLOS Comput. Biol.* 14, e1006423. doi:10.1371/journal.pcbi.1006423.

Mitoma, H., Manto, M., and Gandini, J. (2020). Recent advances in the treatment of cerebellar disorders. *Brain Sci.* 10. doi:10.3390/brainsci10010011.

Musgaard, M., Paramo, T., Domicevica, L., Andersen, O. J., and Biggin, P. C. (2018). Insights into channel dysfunction from modelling and molecular dynamics simulations. *Neuropharmacology* 132, 20–30. doi:10.1016/j.neuropharm.2017.06.030.

Naud, R., Marcille, N., Clopath, C., and Gerstner, W. (2008). Firing patterns in the adaptive exponential integrate-and-fire model. *Biol. Cybern.* 99, 335–347. doi:10.1007/s00422-008-0264-7.

Naveros, F., Garrido, J. A., Carrillo, R. R., Ros, E., and Luque, N. R. (2017). Event- and time-driven techniques using parallel CPU-GPU co-processing for spiking neural networks. *Front. Neuroinform.* 11, 7. doi:10.3389/fninf.2017.00007.

Naveros, F., Luque, N. R., Garrido, J. A., Carrillo, R. R., Anguita, M., and Ros, E. (2015). A Spiking Neural Simulator Integrating Event-Driven and Time-Driven Computation Schemes Using

Parallel CPU-GPU Co-Processing: A Case Study. *IEEE Trans. Neural Networks Learn. Syst.* 26, 1567–1574. doi:10.1109/TNNLS.2014.2345844.

Necchi, D., Lomoio, S., and Scherini, E. (2008). Axonal abnormalities in cerebellar Purkinje cells of the Ts65Dn mouse. *Brain Res.* 1238, 181–188. doi:10.1016/J.BRAINRES.2008.08.010.

Pappalardo, F., Rajput, A.-M., and Motta, S. (2016). Computational modeling of brain pathologies: the case of multiple sclerosis. *Brief. Bioinform.*, bbw123. doi:10.1093/bib/bbw123.

Peek, N., Holmes, J. H., and Sun, J. (2014). Technical Challenges for Big Data in Biomedicine and Health: Data Sources, Infrastructure, and Analytics. *Yearb. Med. Inform.* 23, 42–47. doi:10.15265/IY-2014-0018.

Pellerin, J.-P., and Lamarre, Y. (1997). Local Field Potential Oscillations in Primate Cerebellar Cortex During Voluntary Movement. *J. Neurophysiol.* 78, 3502–3507. doi:10.1152/jn.1997.78.6.3502.

Peyser, A., Sinha, A., Vennemo, S. B., Ippen, T., Jordan, J., Graber, S., et al. (2017). NEST 2.14.0. doi:10.5281/ZENODO.882971.

Plaza-Florido, A., Altmäe, S., Esteban, F. J., Cadenas-Sanchez, C., Aguilera, C. M., Einarsdottir, E., et al. (2020). Distinct whole-blood transcriptome profile of children with metabolic healthy overweight/obesity compared to metabolic unhealthy overweight/obesity. *Pediatr. Res.*, 1–8. doi:10.1038/s41390-020-01276-7.

Prieto, A., Prieto, B., Ortigosa, E. M., Ros, E., Pelayo, F., Ortega, J., et al. (2016). Neural networks: An overview of early research, current frameworks and new challenges. *Neurocomputing* 214, 242–268. doi:10.1016/j.neucom.2016.06.014.

Ramon Cajal, S. (1894). The Croonian lecture.—La fine structure des centres nerveux. *Proc. R. Soc. London* 55, 444–468. doi:10.1098/rspl.1894.0063.

Rancz, E. A., Ishikawa, T., Duguid, I., Chadderton, P., Mahon, S., and Häusser, M. (2007). High-fidelity transmission of sensory information by single cerebellar mossy fibre boutons. *Nature* 450, 1245–1248. doi:10.1038/nature05995.

Reeber, S. L., Otis, T. S., and Sillitoe, R. V. (2013). New roles for the cerebellum in health and disease. *Front. Syst. Neurosci.* 7. doi:10.3389/fnsys.2013.00083.

Roggeri, L., Rivieccio, B., Rossi, P., and D'Angelo, E. (2008). Tactile stimulation evokes long-term synaptic plasticity in the granular layer of cerebellum. *J. Neurosci.* 28, 6354–6359. doi:10.1523/JNEUROSCI.5709-07.2008.

Ros, E., Carrillo, R., Ortigosa, E. M., Barbour, B., and Agís, R. (2006). Event-driven simulation scheme for spiking neural networks using lookup tables to characterize neuronal dynamics. *Neural Comput.* 18, 2959–2993. doi:10.1162/neco.2006.18.12.2959.

Rotstein, H. G. (2017). Spiking resonances in models with the same slow resonant and fast

amplifying currents but different subthreshold dynamic properties. *J. Comput. Neurosci.* 43, 243–271. doi:10.1007/s10827-017-0661-9.

Sánchez-Valle, J., Tejero, H., Fernández, J. M., Juan, D., Urda-García, B., Capella-Gutiérrez, S., et al. (2020). Interpreting molecular similarity between patients as a determinant of disease comorbidity relationships. *Nat. Commun.* 11, 1–13. doi:10.1038/s41467-020-16540-x.

Sánchez-Valle, J., Tejero, H., Ibáñez, K., Portero, J. L., Krallinger, M., Al-Shahrour, F., et al. (2017). A molecular hypothesis to explain direct and inverse co-morbidities between Alzheimer's Disease, Glioblastoma and Lung cancer. *Sci. Rep.* 7, 4474. doi:10.1038/s41598-017-04400-6.

Sathyanesan, A., Zhou, J., Scafidi, J., Heck, D. H., Sillitoe, R. V., and Gallo, V. (2019). Emerging connections between cerebellar development, behaviour and complex brain disorders. *Nat. Rev. Neurosci.* 20, 298–313. doi:10.1038/s41583-019-0152-2.

Schmahmann, J. D. (2019). The cerebellum and cognition. *Neurosci. Lett.* 688, 62–75. doi:10.1016/j.neulet.2018.07.005.

Schorge, S. (2018). Channelopathies go above and beyond the channels. *Neuropharmacology* 132, 1–2. doi:10.1016/j.neuropharm.2018.02.011.

Solinas, S., Forti, L., Cesana, E., Mapelli, J., De Schutter, E., and D'Angelo, E. (2007). Fast-reset of pacemaking and theta-frequency resonance patterns in cerebellar golgi cells: simulations of their impact in vivo. *Front. Cell. Neurosci.* 1, 4. doi:10.3389/neuro.03.004.2007.

Solinas, S., Nieus, T., and D'Angelo, E. (2010). A realistic large-scale model of the cerebellum granular layer predicts circuit spatio-temporal filtering properties. *Front. Cell. Neurosci.* 4, 12. doi:10.3389/fncel.2010.00012.

Spillane, J., Kullmann, D. M., and Hanna, M. G. (2016). Genetic neurological channelopathies: molecular genetics and clinical phenotypes. *J. Neurol. Neurosurg. Psychiatry* 87, 37–48. doi:10.1136/jnnp-2015-311233.

Stimberg, M., Brette, R., and Goodman, D. F. M. (2019). Brian 2, an intuitive and efficient neural simulator. *Elife* 8. doi:10.7554/eLife.47314.

Stoilova-McPhie, S., Ali, S., and Laezza, F. (2013). Protein-Protein Interactions as New Targets for Ion Channel Drug Discovery. *Austin J. Pharmacol. Ther.* 1, 1–6.

Van Geit, W., Achard, P., and De Schutter, E. (2007). Neurofitter: a parameter tuning package for a wide range of electrophysiological neuron models. *Front. Neuroinform.* 1, 1. doi:10.3389/neuro.11.001.2007.

Van Geit, W., De Schutter, E., and Achard, P. (2008). Automated neuron model optimization techniques: a review. *Biol. Cybern.* 99, 241–251. doi:10.1007/s00422-008-0257-6.

Van Geit, W., Gevaert, M., Chindemi, G., Rössert, C., Courcol, J.-D., Muller, E. B., et al. (2016). BluePyOpt: Leveraging Open Source Software and Cloud Infrastructure to Optimise

Model Parameters in Neuroscience. *Front. Neuroinform.* 10, 17. doi:10.3389/fninf.2016.00017.

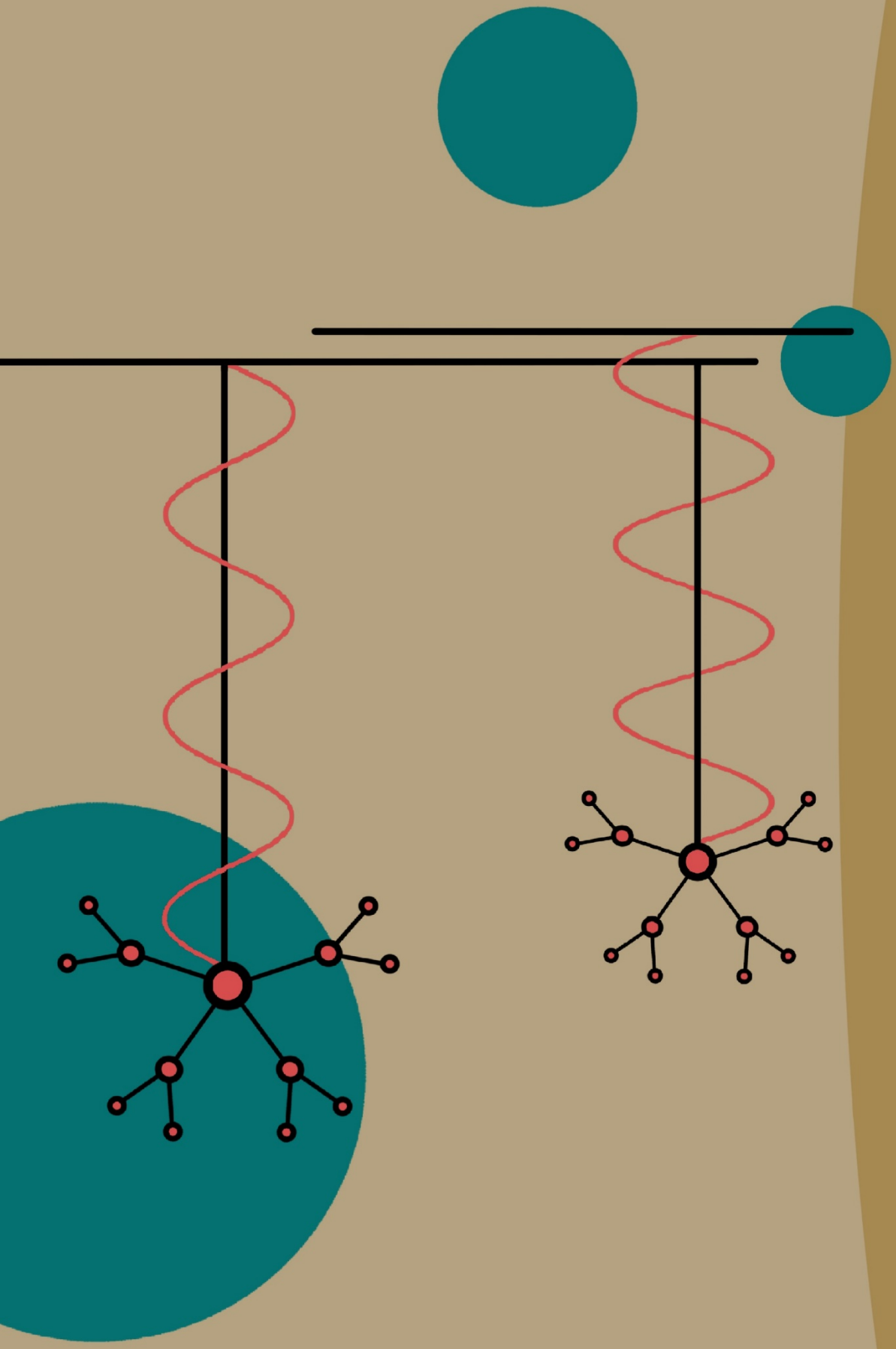
Venkadesh, S., Komendantov, A. O., Listopad, S., Scott, E. O., De Jong, K., Krichmar, J. L., et al. (2018). Evolving Simple Models of Diverse Intrinsic Dynamics in Hippocampal Neuron Types. *Front. Neuroinform.* 12, 8. doi:10.3389/fninf.2018.00008.

Wang, G., Jia, Y., Ye, Y., Kang, E., Chen, H., Wang, J., et al. (2021). Identification of key methylation differentially expressed genes in posterior fossa ependymoma based on epigenomic and transcriptome analysis. *J. Transl. Med.* 19, 1–14. doi:10.1186/s12967-021-02834-1.

Wang, H., Sun, M. J., Chen, H., Zhang, J., Zhang, L. Bin, Zhang, W. W., et al. (2019). Spontaneous recovery of conditioned eyeblink responses is associated with transiently decreased cerebellar theta activity in guinea pigs. *Behav. Brain Res.* 359, 457–466. doi:10.1016/j.bbr.2018.11.030.

Welsch, S., and Welsch, U. (2006). Sobotta-Atlas Histologie. *Zytologie, Histol. mikroskopische Anat.* 6, 676.

Yang, C., Li, C., Wang, Q., Chung, D., and Zhao, H. (2015). Implications of pleiotropy: challenges and opportunities for mining Big Data in biomedicine. *Front. Genet.* 6, 229. doi:10.3389/fgene.2015.00229.



UNIVERSIDAD DE GRANADA

DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR I

CITIC