VNIVERSITAT DE VALÈNCIA

Modelling of Neutron Star Magnetospheres

Petros Stefanou

Supervisors:
José Antonio Pons Botella
Pablo Cerdá-Durán

Doctoral Thesis

Programa de Doctorat en Física - 3126
Departament d'Astronomia i Astrofísica
Universitat de València
February 2024

Vniversitat de València

Modelización de Magnetosferas de Estrellas de Neutrones

Petros Stefanou

Directores:
José Antonio Pons Botella
Pablo Cerdá-Durán

Tesis Doctoral

Programa de Doctorat en Física - 3126
Departament d'Astronomia i Astrofísica
Universitat de València
Febrero 2024

Índice

Acknowledgements

First and foremost, I would like to thank my supervisor, Jose. For the all the guidance and help that he provided, from the most insignificant bureaucratic detail to the biggest and deepest questions about physics. For being caring and eager to assist me in all my problems. For being harsh when he needed to. For having answers to all my questions and for teaching me how to think as a physicist. For being a supervisor that I was glad to have. I would also like to thank my co-supervisor, Pablo. For always being available and helpful with all my problems and issues. For assisting me during my trips to Valencia and for the great collaboration that we had. For teaching me how to be a better coder and programmer and for his profound insight in physics. Both of them are amazing scientists, teachers and human beings. I am glad for the opportunity to collaborate with them and they have my eternal gratitude.

I would also like my fellow members in the Relativistic Astrophysics group in Alicante. Professors Juan Antonio and Pedro for the interesting discussions in the journal clubs and the group meetings. It was a pleasure and a lesson seeing the way they understand and practice physics. My friend Pantelis for his presence, his support and his help. Jorge Urban for the collaboration and the common journey through the challenges of AI. Antonio, Pelayo, Alberto and Jorge with whom I shared my office, for all the interesting discussions the common doubts, the shared frustrations and the great time we had.

I would like to thank all the administrative staff of UA and UV and especially Manel for taking care of the numerous issues that occurred. Mine was really a peculiar case and their help is appreciated.

Lastly, I would like to thank my girlfriend, Aldana, for being the treasure that she is. My brother, Alexandros, for making me a better person. My parents, Anna and Stefanos, for letting me follow my own path and supporting me unconditionally.

Introducción

Este es un resumen extendido de los resultados presentados en los trabajos Stefanou, Pons y Cerdá-Durán, 2023 (cit. en adelante como Paper I), Urbán et al., 2023 (cit. en adelante como Paper II) y Stefanou, Urbán y Pons, 2023 (cit. en adelante como Paper III). El tema central en los tres casos es el estudio de configuraciones libres de fuerzas (FF, de *force-free* en inglés) del campo magnético en las magnetosferas de estrellas de neutrones. Paper I trata sobre soluciones numéricas de magnetosferas con helicidad no nula (*twisted magnetospheres*) en 3D, particularizadas al caso de los magnetars, y utilizando un esquema numérico de diferencias finitas que implementa el método Grad-Rubin. Paper II explora la aplicabilidad de los nuevos métodos numéricos basados en redes neuronales llamados PINNs (de *Physics Informed Neural Networks*) en magnetosferas de magnetars para el caso axisimétrico. Paper III extiende el uso de las PINNs al problema más complejo de las magnetosferas de púlsares axisimétricos, donde los efectos de rotación, que se desprecian para los magnetars, son importantes. En este resumen, brindamos una breve introducción a las estrellas de neutrones y sus magnetosferas, explicamos la metodología aplicada en cada problema, presentamos los resultados más importantes obtenidos y concluimos con algunas observaciones finales.

Estrellas de Neutrones

Las estrellas de neutrones son los cadáveres estelares, el remanente de una estrella masiva después de su muerte violenta y explosiva, conocida como supernova. Son extremas en todos los aspectos de sus propiedades físicas. Todas las disciplinas de la física pueden encontrar un problema interesante para estudiar relacionado con las estrellas de neutrones, desde la astrofísica hasta la física nuclear y desde el electromagnetismo hasta la relatividad general. Esto las convierte en laboratorios celestiales ideales que nos permiten investigar la naturaleza en condiciones inalcanzables en la Tierra.

Poco después del descubrimiento del neutrón por Chadwick, 1932, Baade y Zwicky, 1934 propusieron que durante las explosiones de supernova, se podrían crear objetos pequeños pero extremadamente densos en el centro de la estrella en explosión. Sugirieron que la enorme presión alcanzada en el centro de la explosión sería suficiente para permitir una desintegración beta inversa, durante la cual los electrones y protones se combinan para producir neutrones y neutrinos. Los neutrinos podrían salir de la estrella llevándose una cantidad sustancial de energía, dejando atrás un objeto muy denso compuesto principalmente de neutrones. Llamaron a estos objetos "estrellas de neutrones".

La primera estrella de neutrones fue descubierta en 1967 por Jocelyn Bell (Hewish et al., 1968). Mientras trabajaba en un proyecto para detectar señales provenientes de cuásares, notó un pulso muy rápido e increíblemente constante proveniente de cierta dirección en el cielo. Después de un examen exhaustivo, todas las fuentes conocidas de emisión en radio, tanto galácticas, extragalácticas como construidas por el hombre, fueron eliminadas como posible origen de la señal. Bell había descubierto una nueva fuente de radio celestial. Poco después de la publicación del documento de descubrimiento, Pacini, 1967 y Gold, 1968, en trabajos independientes, identificaron a estos objetos como estrellas de neutrones magnetizadas en rotación, confirmando la predicción de Baade y Zwicky, 1934.

Desde entonces, se han detectado muchas más estrellas de neutrones, exhibiendo una fenomenología observacional muy rica. Su radiación cubre todas las bandas del espectro electromagnético, desde radio hasta rayos gamma, y presenta una gran variabilidad en términos de su dependencia temporal. Puede ser persistente, en forma de radiación de cuerpo negro permanente, periódica, en forma de pulsos estables, o transitoria, en forma de eventos explosivos impredecibles.

Las estrellas de neutrones son también los objetos visibles más compactos en el universo, solo superados por los agujeros negros. Su masa alcanza valores de entre 1,1 y 2 masas solares, concentradas en una región de unos 15 km de radio, tan pequeña como una ciudad media. Esto se traduce en densidades que comienzan desde $\rho \sim 10^{10} \mathrm{g}/\mathrm{cm}^3$ en la corteza hasta $\rho \sim 10^{15} \mathrm{g}/\mathrm{cm}^3$ en el núcleo, donde la densidad supera incluso la del núcleo atómico. Además, giran extremadamente rápido, con períodos de rotación en el rango $P \sim 10^{-3} - 10^{1}$ s que aumentan a un ritmo muy estable debido a la pérdida de momento angular por el torque ejercido por la magnetosfera en rotación. Por último, y más importante para las aplicaciones astrofísicas, presentan los campos magnéticos más fuertes encontrados en el universo, con valores estimados en el rango $B \sim 10^{11} - 10^{15}$ G en su superficie, un billón o más de veces más alto que el campo magnético de la Tierra.

A diferencia de las estrellas normales, las estrellas de neutrones no producen energía a través de procesos termo-nucleares en su núcleo. El remanente energéticos que heredaron de la explosión de supernova se almacenan en forma de energía térmica, rotacional y/o magnética. Este reservorio de energía se irradia lentamente, haciendo que la estrella de neutrones disminuya su velocidad de rotación y eventualmente se apague. Una excepción a esta regla general son las estrellas de neutrones que pertenecen a sistemas binarios, donde la energía y el momento angular puede transferirse hacia la estrella a través del acrecimiento de materia de su compañera.

Se pueden distinguir diferentes clases de estrellas de neutrones basadas en su fenomenología observacional, sus propiedades físicas y la forma de energía que alimenta su radiación. La categorización más importante se realiza con la ayuda del diagrama $P - \dot{P}$, un diagrama cuyos ejes son el período de la estrella de neutrones y su primera derivada temporal (ver figura 1). Esta tesis se centrará en dos de estas clases: magnetars (Paper I; Paper II) y púlsares (Paper III).

Los magnetars, como su nombre indica, son las estrellas de neutrones más magnetizadas, con campos superficiales del orden de $B \sim 10^{14} - 10^{15}$ G. Debido a sus enormes campos magnéticos, la disminución de la velocidad de rotación debido al frenado magnético es inmensa y pierden la mayor parte de su energía rotacional a

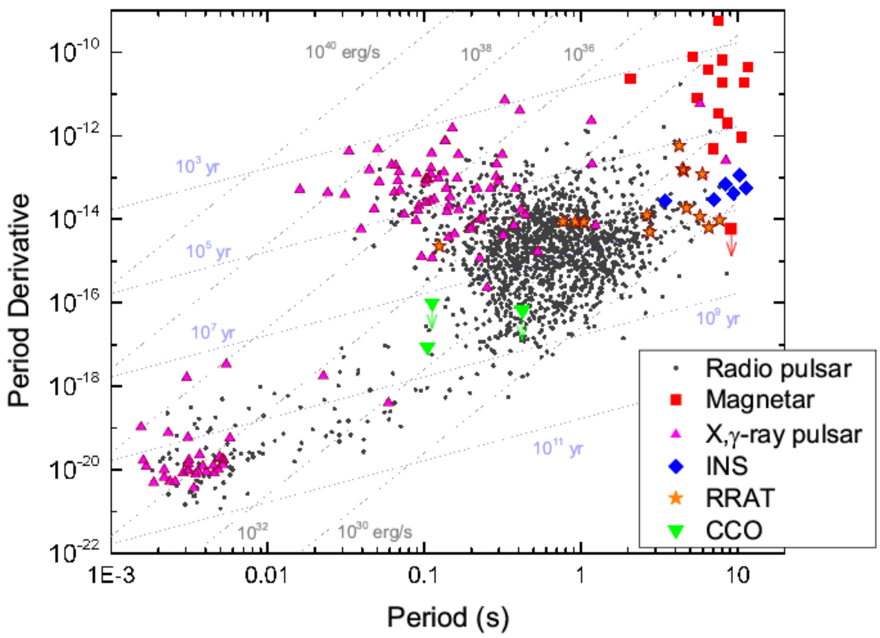


Figura 1: El diagrama $P - \dot{P}$ de la población de las estrellas de neutrones (Harding, 2013).

una edad relativamente joven. Esto los convierte en los rotadores más lentos entre todas las estrellas de neutrones, con períodos en el rango $P \sim 1 - 10$ s. Su fuente de energía es la disipación de sus enormes campos (Thompson y Duncan, 1995; Thompson y Duncan, 1996). Principalmente se los observa en rayos X a través de eventos transitorios de diferentes escalas de energía y tiempo (ráfagas, estallidos y fulguraciones gigantes, ver Mereghetti, Pons y Melatos, 2015). Se cree que esta actividad está asociada a un campo magnético retorcido, que sufre repentinamente una reconfiguración a gran escala liberando la energía almacenada (Beloborodov y Thompson, 2007).

Los púlsares son la clase más poblada de estrellas de neutrones. Tienen campos magnéticos superficiales típicamente del orden de $B \sim 10^{12}$ G y períodos en el rango $P \sim 10^{-3} - 10^{-1}$ s, lo que significa que la rotación juega un papel crucial. El motor que alimenta su actividad es la energía rotacional almacenada. Su característica más distintiva es la radiación pulsada que se puede observar en todas las bandas del espectro electromagnético. Es causada por la combinación de una emisión altamente anisotrópica en forma de haces estrechos y su rápida rotación.

Magnetosfera

Aunque el campo magnético a gran escala de una estrella de neutrones se considera predominantemente dipolar, todas las estrellas de neutrones están rodeadas de partículas cargadas que se mueven a lo largo del campo magnético, y estas corrientes implican desviaciones de un dipolo clásico. A la región, que gira con la estrella casi como un cuerpo rígido con velocidad angular Ω, se la llama magnetosfera. La corrotación deja de ser posible a una distancia $R_{\rm LC} = {}^{c}/_{\Omega}$, conocida como el radio del cilindro-luz, donde la velocidad tangencial de las partículas es igual a la velocidad de la luz. Más allá de esta distancia, las líneas del campo magnético deben abrirse. La magnetosfera es responsable de toda la radiación producida por las estrellas de neutrones, excepto la radiación térmica desde la superficie. La energía almacenada por la estrella se disipa, se convierte en energía cinética de las partículas y luego se irradia a través de varios mecanismos de emisión. La dinámica de la magnetosfera está dominada por el campo electromagnético. Todas las demás fuerzas, es decir, la gravedad, la presión del plasma y la inercia de las partículas, se consideran insignificantes en comparación con la fuerza de Lorentz, lo que resulta en una dinámica gobernada por la siguiente ecuación, conocida como la ecuación *force-free* (FF):

$$\boldsymbol{\nabla} \times \left(\boldsymbol{B} - \beta^2 \boldsymbol{B}_p\right) = \alpha \boldsymbol{B}, \tag{1}$$

donde $\beta = {}^{v}/_{c}$ es la velocidad de corrotación en unidades de la velocidad de la luz y $\boldsymbol{B}_p = \boldsymbol{B} - \boldsymbol{B}_\phi$. El parámetro α es una función escalar dada por

$$\alpha = \frac{4\pi}{c} \left(\boldsymbol{J} - \rho_e \boldsymbol{v}\right) \cdot \frac{\boldsymbol{B}}{B^2}, \tag{2}$$

que representa la proporción entre la componente de la corriente alineada con el campo, en el marco de referencia comóvil y la intensidad del campo magnético local (ver Philippov y Kramer, 2022). Se puede demostrar, utilizando la ecuación (1) y la condición de divergencia cero del campo magnético, que α debe obedecer la restricción

$$\boldsymbol{B} \cdot \nabla \alpha = 0, \tag{3}$$

que establece que α es constante a lo largo de cualquier línea de campo magnético particular.

Bajo el supuesto de axisimetría, la ecuación (1) puede escribirse en términos de dos funciones de flujo escalares $\mathcal{P}$ y $\mathcal{T}$ de la siguiente forma

$$\left(1 - \beta^2\right) \triangle_{GS} \mathcal{P} - 2\beta^2 \left(\frac{1}{r}\frac{\partial \mathcal{P}}{\partial r} + \frac{\cot\theta}{r^2}\frac{\partial P}{\partial \theta}\right) + T\frac{d\mathcal{T}}{d\mathcal{P}} = 0, \tag{4}$$

donde $\triangle_{GS}$ es el operador Grad-Shafranov, definido como

$$\triangle_{GS} \equiv \frac{\partial^2}{\partial r^2} + \frac{\sin\theta}{r^2}\frac{\partial}{\partial \theta}\left(\frac{1}{\sin\theta}\frac{\partial}{\partial \theta}\right). \tag{5}$$

Las funciones $\mathcal{P}$ y $\mathcal{T}$ deben cumplir con la restricción adicional

$$\nabla \mathcal{P} \times \nabla \mathcal{T} = 0, \tag{6}$$

lo que significa que una es función de la otra, es decir, $\mathcal{T} = \mathcal{T}(\mathcal{P})$.

Para los magnetars, β es muy pequeño debido a la rotación relativamente lenta, lo que significa que el cilindro de luz se encuentra mucho más allá de la región de interés. Por lo tanto, la ecuación (1) se puede simplificar como

$$\boldsymbol{\nabla} \times \boldsymbol{B} = \alpha \boldsymbol{B}, \tag{7}$$

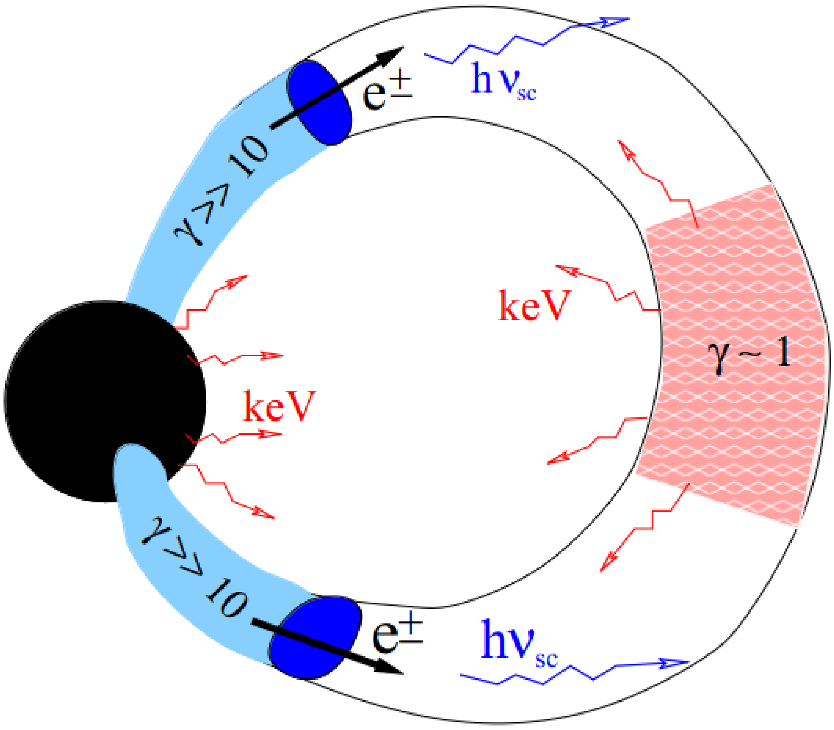


Figura 2: Representación esquemática de una magnetosfera de magnetar (Beloborodov, 2012).

lo que implica que la corriente siempre es paralela al campo magnético. Resolvemos esta forma de la ecuación en el Paper I y su versión axisimétrica (ecuación (4) con $\beta = 0$) en Paper II.

El campo magnético se desvía de la forma dipolar usual en una región relativamente cercana a la estrella debido a corrientes que se introducen en la magnetosfera desde el interior de la estrella. Allí, la evolución magneto-térmica en la corteza de la estrella desplaza gradualmente las huellas de las líneas del campo magnético. Las líneas se retuercen y se desarrolla un componente de campo magnético azimutal. La energía y la helicidad se transportan desde el interior hacia la magnetosfera. En algún punto crítico, tiene lugar una reconfiguración a gran escala del campo magnético, liberando la energía almacenada en forma de radiación de alta energía. Una representación esquemática de una magnetosfera de magnetar se muestra en la figura 2.

Por otro lado, en los púlsares la rotación es importante y la ecuación (1) debe considerarse completa. En Paper III, aunque nos limitaremos al caso axisimétrico (4). La magnetosfera se divide en dos regiones: una con líneas de campo que no cruzan el cilindro de luz y regresan a la estrella y otra con líneas de campo que sí cruzan el cilindro de luz y se abren hacia el infinito. La primera región puede considerarse libre de corriente con un campo magnético puramente poloidal, mientras que la segunda desarrolla un componente azimutal debido a la curvatura de las líneas de campo abiertas. Una característica muy importante de la magnetosfera del púlsar es la lámina de corriente que se desarrolla a lo largo de la última línea de campo cerrada (llamada separatrix) y a lo largo del ecuador magnético debido a la discontinuidad en el campo magnético. Una representación esquemática de una magnetosfera de un

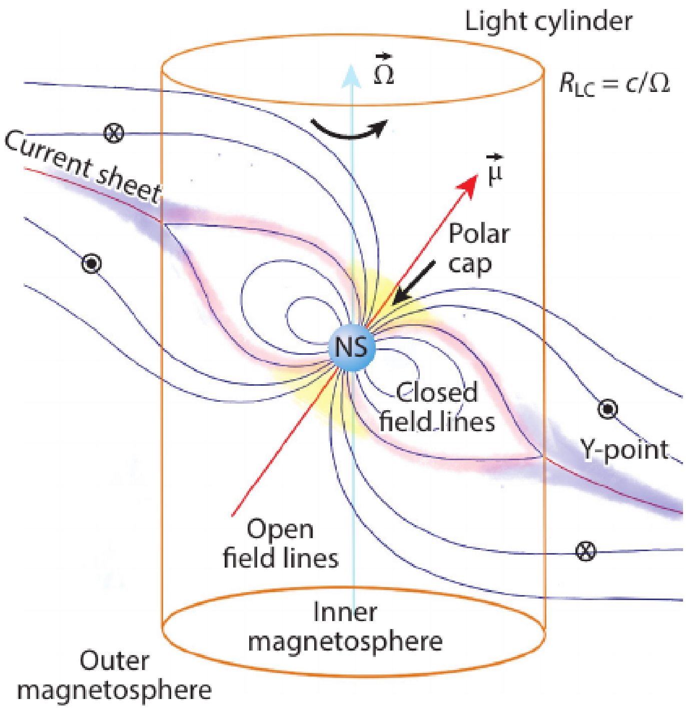


Figura 3: Representación esquemática de una magnetosfera de púlsar (Philippov y Kramer, 2022).

púlsar se muestra en la figura 3.

Metodología

El método Grad-Rubin

En Paper I, utilizamos un código numérico de diferencias finitas que implementa el método iterativo de Grad-Rubin (Grad y Rubin, 1958) para encontrar soluciones FF. La idea principal es descomponer el sistema de ecuaciones (3) y (7) en una parte hiperbólica y una parte elíptica y resolverlas mediante una iteración de punto fijo. La parte hiperbólica determina la función α resolviendo

$$\boldsymbol{B}^{(n)} \cdot \nabla \alpha^{(n+1)} = 0. \tag{8}$$

Aquí, el índice n denota el número de iteración. Para poder resolver el problema, necesitamos iniciar la iteración proporcionando una aproximación inicial para el campo magnético $\boldsymbol{B}^{(0)}$ y el valor superficial de α

$$\alpha^{(n+1)}\Big|_{S^+} = \alpha_S, \tag{9}$$

El problema está bien planteado solo si proporcionamos α en la parte de la superficie donde el campo magnético tiene cierta polaridad (por ejemplo, positiva), que es lo que denota el subíndice S^+. Una vez que α se determina en toda la magnetosfera, podemos resolver el campo magnético a través de la ecuación

$$\nabla \times \boldsymbol{B}^{(n+1)} = \alpha^{(n+1)} \boldsymbol{B}^{(n)}, \tag{10}$$

con un término fuente fijo $\alpha^{(n+1)}\boldsymbol{B}^{(n)}$ y sujeto a las condiciones de contorno

$$B_r(\theta,\varphi)\Big|_S = b_S(\theta,\varphi), \tag{11}$$

$$\boldsymbol{B}\Big|_\infty = 0. \tag{12}$$

Este procedimiento se repite iterativamente, actualizando α y $\boldsymbol{B}$ sucesivamente hasta que se satisfaga cierto criterio de convergencia. La parte hiperbólica se resuelve utilizando el método de las características (trazado de líneas de campo) mientras que la parte elíptica se resuelve utilizando un método de relajación (Jacobi) (Adsuara et al., 2016).

Physics-Informed Neural Networks (PINNs)

En Paper II y Paper III, empleamos PINNs (Lagaris, Likas y Fotiadis, 1997; Raissi, Perdikaris y Karniadakis, 2019), un enfoque muy prometedor basado en aprendizaje automático para resolver ecuaciones diferenciales en derivadas parciales (EDP) que ha surgido en los últimos años. En nuestro caso, lo aplicamos para encontrar soluciones de magnetosferas axisimétricas tamto de magnetars como de púlsares.

Las redes neuronales son aproximadores universales de funciones. Aceptan ciertas entradas, realizan operaciones simples pero no lineales en ellas que dependen de un conjunto de parámetros ajustables Θ y devuelven algunas salidas. El proceso de ajustar los parámetros Θ se llama entrenamiento de la red neuronal y se lleva a cabo con la ayuda de algoritmos de optimización avanzados. En el caso de PINNs, queremos que la red neuronal aproxime una función u que describe el estado de un sistema físico. Las entradas de la red son las coordenadas de puntos en el dominio de interés y, posiblemente, algunos parámetros físicos que determinan el problema. Durante el entrenamiento, los parámetros Θ se ajustan de manera que la solución aproximada minimice el residuo de la EDP que gobierna el sistema físico.

Un problema general de valores de frontera, sujeto a condiciones de contorno dadas, puede escribirse en términos de un operador diferencial no lineal $\boldsymbol{\mathcal{L}}$, como

$$\boldsymbol{\mathcal{L}}u(\boldsymbol{x};\Theta) - G(\boldsymbol{x}, u(\boldsymbol{x};\Theta)) = 0, \tag{13}$$

$$u|_{\partial\mathcal{D}} = f_b(\boldsymbol{x})|_{\boldsymbol{x}\in\partial\mathcal{D}}, \tag{14}$$

donde G es el término fuente, u es la solución del problema y $\boldsymbol{x}$ es un vector de coordenadas en algún dominio $\mathcal{D}$. Si u no es exacta, la ecuación anterior tendrá un residuo. Para cuantificar este residuo, definimos la llamada *función de pérdida* $\mathcal{J}$, utilizando el error cuadrático medio de la EDP evaluado en un conjunto grande de puntos distintos

$$\mathcal{J} = \frac{1}{N}\sum_{i=1}^{N} ||\boldsymbol{\mathcal{L}}u(\boldsymbol{x}_i;\Theta) - G(\boldsymbol{x}_i, u(\boldsymbol{x}_i;\Theta))||^2. \tag{15}$$

Las condiciones de contorno se tienen en cuenta definiendo la solución aproximada u como

$$u(\boldsymbol{x};\Theta) = f_b(\boldsymbol{x}) + h_b(\boldsymbol{x})\mathcal{N}(\boldsymbol{x};\Theta), \tag{16}$$

donde f_b es una función que satisface las condiciones de contorno, h_b es una función que se anula en la frontera y $\mathcal{N}$ es la salida de la red neuronal. El algoritmo de optimización intenta encontrar el conjunto de parámetros Θ que minimiza $\mathcal{J}$. Después del entrenamiento, $u(\boldsymbol{x};\Theta)$ sirve como un modelo sustituto de la verdadera solución de la EDP para cualquier punto en el dominio, incluso si no se utilizó durante el proceso de optimización.

Resultados

En cada uno de los tres artículos que forman este compendio, proporcionamos una serie de modelos distintos para ilustrar el efecto que tienen varios parámetros físicos en la magnetosfera. A continuación, presentamos los resultados más importantes.

Paper I

El problema de las magnetosferas de magnetar se ha estudiado extensamente en varios trabajos en la literatura para el caso axisimétrico (por ejemplo, Akgün et al., 2017 y referencias en él), que han establecido sus principales características y propiedades. Sin embargo, los resultados teóricos sobre la evolución del campo magnético interno de los magnetars (Lander y Gourgouliatos, 2019), así como las observaciones (Younes et al., 2022), sugieren que se forman bucles magnéticos en regiones localizadas, mientras que el resto de la magnetosfera permanece libre de corriente. Esta topología no puede ser capturada por simulaciones 2D axisimétricas, ya que están globalmente retorcidas. Presentamos, por primera vez, simulaciones 3D no axisimétricas de magnetosferas de magnetars estacionarias con bucles magnéticos locales. La corriente que fluye a lo largo de estos bucles atraviesa la superficie de la estrella de neutrones y la calienta. Por lo tanto, las áreas donde se encuentran las huellas de las líneas torcidas tendrán una temperatura mayor, en comparación con el resto de la superficie. Estas áreas se llaman, en consecuencia, áreas calientes.

Para modelar una magnetosfera con áreas calientes en la superficie, resolvemos la ecuación (7) utilizando el método Grad-Rubin. La solución estará determinada por los valores superficiales del campo magnético y la función FF α. Suponemos la siguiente expresión parametrizada para el valor superficial de α.

$$\alpha_S(\theta,\phi) = \alpha_0 e^{\frac{-(\theta-\theta_1)^2-(\phi-\phi_1)^2}{2\sigma^2}}. \tag{17}$$

Esto corresponde a un área con un perfil gaussiano, centrado en el punto (θ_1,ϕ_1) donde σ controla la anchura de la gaussiana. α_0 determina el valor máximo de α (en el centro de la gaussiana). Se asume que el campo magnético superficial es el de un dipolo. Las líneas de campo cuyo origen se encuentran dentro del área donde α no es cero desarrollarán un componente magnético azimutal y se torcerán, mientras que las líneas con origen fuera de dicha zona permanecerán sin corriente. Las líneas torcidas deben cerrarse nuevamente en la superficie, pero sus punto de retorno a la superficie deben estar desplazados (con respecto a su posición dipolar original), formando un segundo punto caliente de forma peculiar. Esto explica por qué la condición de contorno para α se impone solo en un extremo de las líneas, dejando que el otro se determine junto con la solución. Es imposible saber a priori cómo debería ser la forma del segundo punto caliente para que exista una solución consistente.

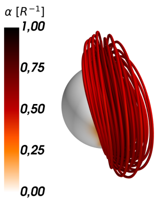
La Figura 4 muestra una serie de modelos para diferentes valores de los parámetros $\alpha_0, \theta_1, \sigma$. El color denota el valor de α, que es constante a lo largo de cada línea de campo en particular, pero difiere entre líneas. La Figura 5 muestra la dependencia de estos parámetros del exceso de energía E_e (diferencia relativa de la energía de un modelo con respecto a la energía de un dipolo puro), helicidad H y máxima torsión $T_{\alpha_{max}}$ (definimos la torsión de una línea de campo como $T_\alpha = \alpha L$, donde L es la longitud de la línea). La Figura 6 muestra la temperatura efectiva de la superficie de la estrella para un modelo particular.

Estos resultados nos llevan directamente a tres conclusiones importantes: 1) Las magnetosferas se desvían más de la configuración dipolar a medida que α_0 y σ aumentan y a medida que θ disminuye. Esto significa que la magnetosfera está más retorcida si el punto caliente está atravesado por corrientes más intensas, está más cerca del polo y cubre un área más grande. 2) La energía, la helicidad y la máxima torsión aumentan rápidamente a medida que la magnetosfera se vuelve más torcida. 3) La temperatura de los puntos calientes y la forma peculiar de "coma" del segundo punto caliente se vuelven más intensas con α mayores.

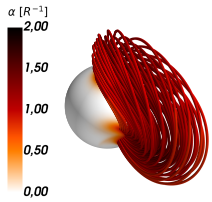
Aumentar los parámetros más allá de los valores de la tercera columna en la Figura 4, hace que el resolvedor de Grad-Rubin falle ya que no puede converger a una solución. Este efecto se ha estudiado antes en magnetosferas axisimétricas 2D (Akgün et al., 2017; Mahlmann et al., 2019) y es un indicador de que es probable que las soluciones con magnetosferas altamente retorcidas sean inestables y se espera que produzcan reconexiones que liberen la energía almacenada y restauren la magnetosfera a un estado con menos tensión. Nuestros cálculos sobre la cantidad de energía disponible en el campo magnético y la temperatura y geometría de la superficie de los puntos calientes están bastante cerca de los valores observados.

Paper II

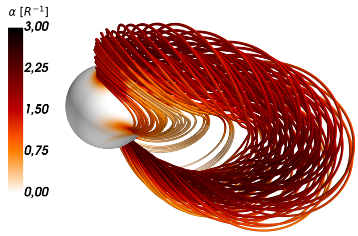
La modelización de las magnetosferas de magnetars a menudo se basa en expresiones parametrizadas como (17) para tener en cuenta la corriente que se introduce debido a procesos internos en la corteza. Un enfoque más realista sería resolver la evolución magneto-térmica en la corteza acoplada a una magnetosfera FF. Esto nos permitiría utilizar una distribución de corriente físicamente motivada en la superficie como condición de frontera para la magnetosfera. Además, nos permitiría utilizar



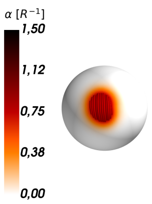
(a) $\alpha_0 = 1{,}0$



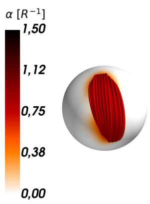
(b) $\alpha_0 = 2{,}0$



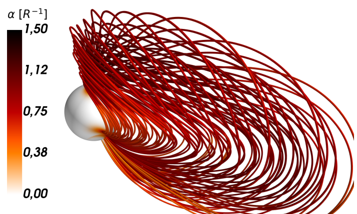
(c) $\alpha_0 = 3{,}0$



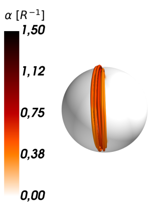
(d) $\theta_1 = 85^\circ$



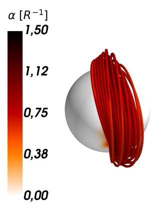
(e) $\theta_1 = 60^\circ$



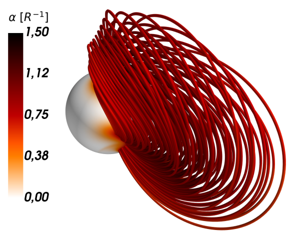
(f) $\theta_1 = 35^\circ$



(g) $\sigma = 0{,}05$



(h) $\sigma = 0{,}15$



(i) $\sigma = 0{,}25$

Figura 4: Muestras de magnetosferas con diferentes valores de α_0, θ_1 y σ. En cada fila, solo un parámetro varía mientras los demás permanecen fijos. En cada columna, de izquierda a derecha, la magnetosfera está más alejada de la solución potencial libre de corriente. El campo magnético está más retorcido para corrientes fuertes (α_0 grande), separación grande de puntos calientes (θ_1 pequeño) y área grande de punto caliente (alto σ). Para mayor claridad, solo se muestran las líneas con $\alpha > 0{,}5\alpha_0$. (Paper I).

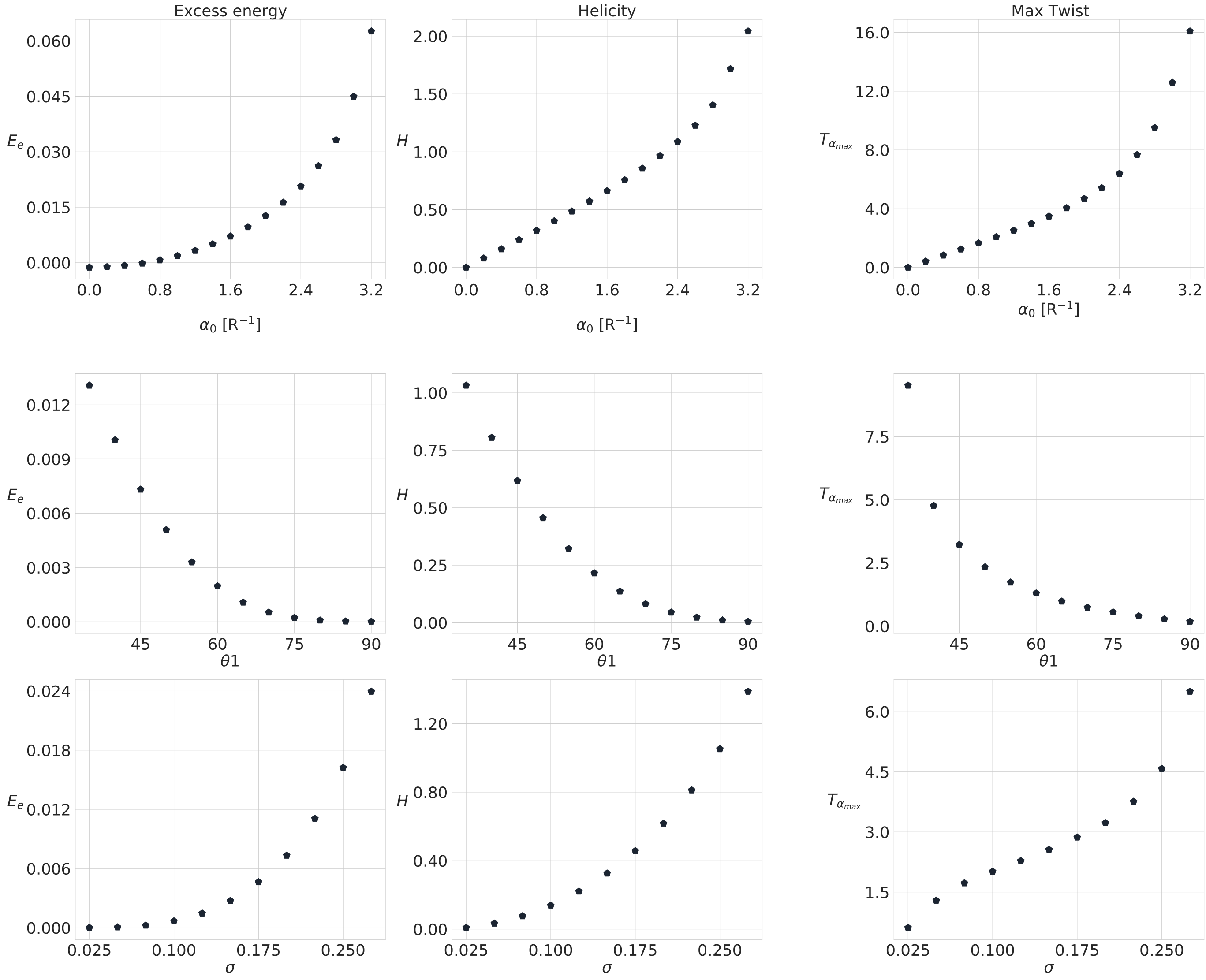


Figura 5: Dependencia de la energía en exceso normalizada E_e, la helicidad H y la máxima torsión $T_{\alpha_{max}}$ en los parámetros α_0 (primera línea), θ_1 (segunda línea) y σ (tercera línea). (Paper I).

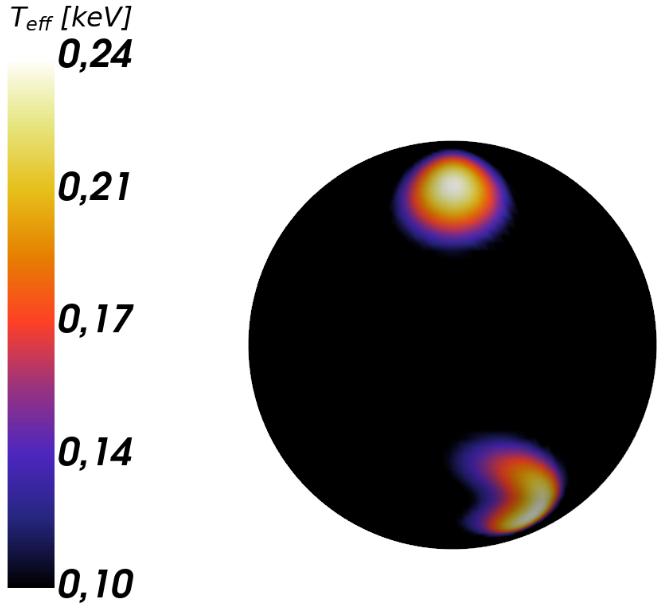


Figura 6: Temperatura efectiva de los puntos calientes en la superficie de la estrella. La barra de color muestra los valores de T_{eff} en unidades de keV. El valor más bajo de la barra de color corresponde a la temperatura superficial típica de una estrella de neutrones en ausencia de corrientes de paso. (Paper I).

condiciones de frontera de campo magnético FF para la evolución magneto-térmica dentro de la corteza, un caso escasamente explorado en la literatura (lo habitual es usar soluciones potenciales, sin corrientes). Lamentablemente, realizar simulaciones conjuntas de este tipo es prácticamente inviable debido a las propiedades físicas muy diferentes de la corteza y la magnetosfera (densidad de plasma, escalas de tiempo, etc.). Por otro lado, resolver la configuración magnetosférica en cada paso de la evolución magneto-térmica es extremadamente lento.

En un intento por superar estos problemas, exploramos la aplicabilidad de un resolvedor PINN para magnetosferas libres de fuerzas. Nos enfocamos en el caso axisimétrico que es más simple para ilustrar mejor la efectividad de este método antes de aplicarlo al problema 3D general. Cada modelo magnetosférico está determinado por su condición de frontera superficial y su término fuente. La característica innovadora clave que introducimos es que entrenamos la PINN para condiciones de frontera y términos fuente generales, proporcionando esta información como entrada adicional a la red en lugar de codificarla directamente en la función de pérdida. Las PINN tienen la ventaja de ser extremadamente rápidas en la evaluación de nuevas soluciones, porque la mayor parte del costo computacional se gasta en el proceso de entrenamiento. Una vez que la PINN está entrenada, cualquier solución puede ser evaluada en muy poco tiempo proporcionando un conjunto de coordenadas, las condiciones de frontera y el término fuente. Por lo tanto, el costo adicional de calcular la configuración magnetosférica en cada paso de tiempo durante la evolución interna es pequeño.

Es conveniente usar el sistema de coordenadas esféricas compactificadas, definido como

$$q = \frac{R}{r} \tag{18}$$

$$\mu = \cos\theta. \tag{19}$$

Este sistema de coordenadas ofrece algunas ventajas, la más importante de las cuales es que las condiciones de frontera en el infinito radial (representadas por un solo punto, $q = 0$) son mucho más fáciles de imponer. El operador diferencial $\mathcal{L}$ en (13) para este caso es el operador Grad-Shafranov dado en (5). El término fuente G está dado por $G = -T\frac{dT}{dP}$. Estamos entrenando la red para encontrar soluciones $\mathcal{P}(q, \mu; \Theta)$ de la ecuación diferencial parcial

$$q^2 \frac{\partial}{\partial q}\left(q^2 \frac{\partial \mathcal{P}}{\partial q}\right) + \left(1 - \mu^2\right) q^2 \frac{\partial^2 \mathcal{P}}{\partial \mu^2} \tag{20}$$

Las condiciones de frontera se imponen a través de las funciones f_b y h_b de la siguiente manera:

$$f_b(q, \mu) = q^n \left(1 - \mu^2\right) \sum_{l=1}^{l_{\max}} \frac{b_l}{l} P_l'(\mu) \tag{21}$$

$$h_b(q, \mu) = q(1 - q)(1 - \mu^2), \tag{22}$$

donde P_l son los polinomios de Legendre y la prima denota la derivada con respecto a μ. Se puede verificar fácilmente que, cuando $q = 0$ (infinito) o $\mu = \pm 1$ (eje), entonces $\mathcal{P}$ es cero como debería ser. Por otro lado, cuando $q = 1$ (superficie), $\mathcal{P}$ se da como combinación lineal de los polinomios de Legendre hasta el orden l_{max}, que es una elección racional para los campos magnéticos de estrellas de neutrones. Para el término fuente G, tenemos que especificar la forma funcional de $\mathcal{T}(\mathcal{P})$. Hemos elegido una forma cuadrática que depende de dos coeficientes s_1 y s_2 de manera que

$$\mathcal{T} = s_1 \mathcal{P} + s_2 \mathcal{P}^2. \tag{23}$$

La PINN se entrena para condiciones de frontera y términos fuente arbitrarios. Para ello, incluimos los coeficientes $\{b_l\}$, que determinan las condiciones de frontera, y s_1, s_2 que determinan el término fuente, junto con las coordenadas q, μ como entradas. Entrenamos para un gran número de valores aleatorios de estas entradas, cambiando periódicamente el conjunto de entrenamiento. Realizamos un estudio básico de hiperparámetros para decidir la configuración óptima para la PINN. Concluimos que una arquitectura totalmente conectada con 4 capas de 80 neuronas es suficiente para tener en cuenta la complejidad y la alta dimensionalidad de nuestro problema. Para evaluar la precisión de la PINN, primero resolvemos el caso libre de corriente $G = 0$, que es un campo potencial con solución analítica conocida. Las comparaciones muestran que la solución de la PINN tiene un error relativo global del orden de $\sim 10^{-4}$ respecto a la solución analítica, lo que se considera suficientemente preciso para los propósitos de este trabajo. La figura 7 muestra un modelo particular libre de corriente adquirido una vez que se ha entrenado la PINN.

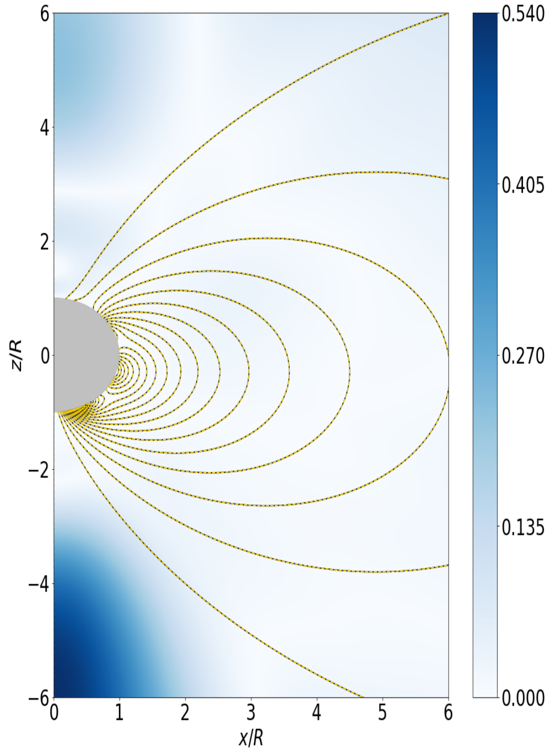


Figura 7: Mapa de colores del error relativo (en porcentaje) entre $\mathcal{P}$ y $\mathcal{P}_{\rm ex}$. Las líneas punteadas amarillas corresponden a los contornos de $\mathcal{P}$ mientras que las líneas sólidas negras corresponden a $\mathcal{P}_{\rm ex}$. Los coeficientes multipolares en f_b para este ejemplo particular son $b_1 = 1$, $b_{l\geq 2} = (-1)^{l+1}\, 0{,}6$ (Paper II).

Nuestro objetivo es acoplar modelos magnetosféricos FF con la evolución interna. Para ello, procedemos a entrenar la PINN para $G \neq 0$. Luego, usamos la red entrenada para evaluar rápidamente modelos magnetosféricos.

En cada paso de tiempo, usamos los valores del campo magnético radial en la corteza para calcular los coeficientes multipolares b_l y los valores de las funciones $\mathcal{P}$ y $\mathcal{T}$ para calcular los coeficientes s_1, Ss_2 ajustando la función (23). Luego, alimentamos estos valores a la PINN para evaluar un modelo magnetosférico correspondiente. Usamos los valores del campo magnético libre de fuerzas, según lo calculado por la PINN en la zona por encima de la superficie para imponer condiciones de frontera para el próximo paso de la evolución magneto-térmica. La evaluación de un modelo magnetosférico desde una red entrenada es muy rápida y el costo computacional adicional es pequeño. Este acoplamiento permite que las corrientes atraviesen la corteza y fluyan desde el interior hacia la magnetosfera, afectando la evolución interna. La figura 8 muestra la diferencia entre imponer condiciones de frontera libres de corriente (vacío) y FF.

Paper III

Las magnetosferas de los púlsares presentan complicaciones adicionales debido a las discontinuidades que aparecen en forma de láminas de corriente. Contopoulos, Kazanas y Fendt, 1999 lograron superar estas dificultades y resolver numéricamente el problema (4) axisimétrico por primera vez aplicando correcciones iterativas a la soluciones de ambos lados del cilindro de luz hasta que coincidieran. Desde enton-

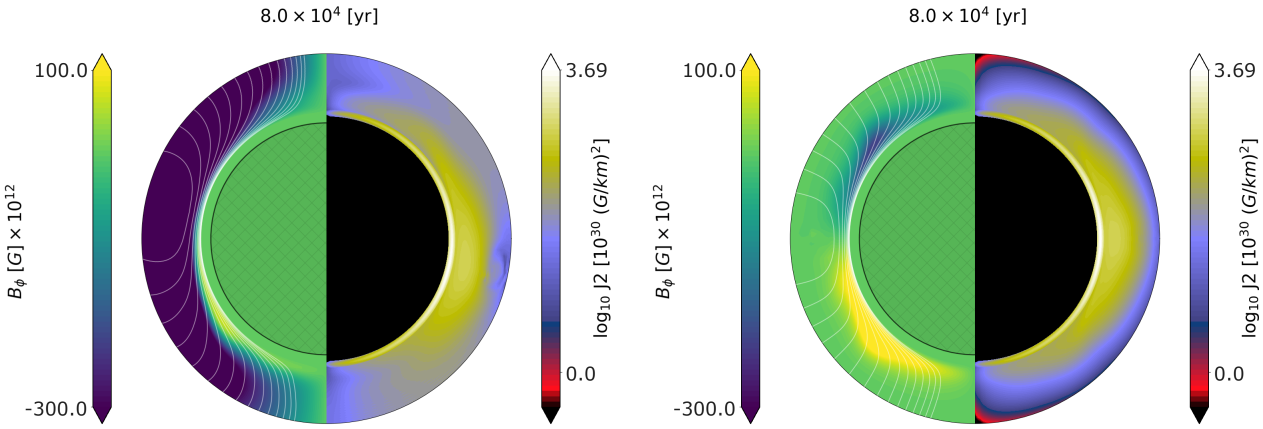


Figura 8: Una instantánea de la evolución del campo magnético y la corriente eléctrica a 80 kyr. En el hemisferio izquierdo, mostramos la proyección meridional de las líneas de campo magnético (líneas blancas) y el campo toroidal (colores). En el hemisferio derecho, mostramos el cuadrado del módulo de la corriente eléctrica, es decir, $|J|^2$ (nota la escala log). La corteza se ha ampliado por un factor de 8 con fines de visualización. Panel izquierdo: condiciones de contorno libres de fuerzas. Panel derecho: condiciones de contorno al vacío (Paper II.

ces, muchos autores han verificado y ampliado sus hallazgos. Reproducimos estos resultados, confirmamos la aptitud de las redes neuronales para manejar problemas que contienen discontinuidades y destacamos nuevas posibilidades usando nuestro resolvedor PINN. Este es un paso esencial para extender la metodología PINN al caso tridimensional, donde el problema estacionario (1) sigue sin resolverse.

Seguimos una metodología similar a la de Paper II. La principal diferencia en el problema de los púlsares es que las corrientes se introducen en la magnetosfera debido a la apertura de líneas de campo que cruzan el cilindro-luz y no a través de la superficie estelar. Esto significa que la función toroidal $\mathcal{T}(\mathcal{P})$ no puede suministrarse como una función parametrizada, sino que debe encontrarse de manera autoconsistente como parte de la solución para garantizar un cruce suave de las líneas de campo a través del cilindro-luz. Por lo tanto, el PINN debe tener dos salidas, $\mathcal{T}$ y $\mathcal{P}$. Hacemos esto separando nuestra red en dos subredes, una que toma como entrada las coordenadas q, μ y devuelve la función $\mathcal{P}$ y otra que toma como entrada la función $\mathcal{P}$ y devuelve la función $\mathcal{T}$.

Exploramos cuatro familias de soluciones: (i) la solución clásica, que tiene una lámina de corriente ecuatorial y una lámina de corriente en la separatrix que se intersectan en el cilindro-luz (en el llamado punto Y), (ii) soluciones similares a la solución clásica, pero con el punto Y ubicado en un punto arbitrario entre la superficie y el cilindro-luz, (iii) soluciones sin lámina de corriente ecuatorial que tienen una inversión suave del campo magnético poloidal de los dos hemisferios y (iv) soluciones donde el campo magnético no es dipolar. La flexibilidad y generalidad de los PINN nos brinda la oportunidad de tratar las láminas de corriente ecuatorial y de la separatrix de diferentes maneras, dependiendo del caso en cuestión.

La figura 9 muestra la forma funcional de $\mathcal{T}(\mathcal{P})$ adquirida para varias de estas modelos.

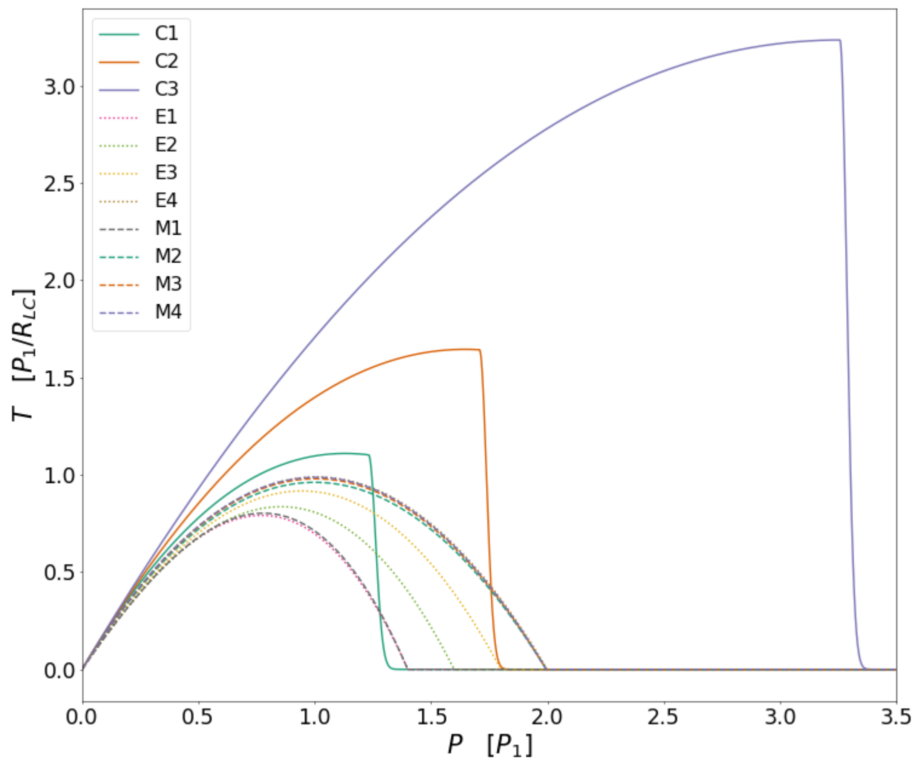


Figura 9: Las funciones $T(P)$ para los modelos estudiados en Paper III.

Conclusiones

En los tres documentos que conforman este compendio, estudiamos en detalle modelos de magnetosfera de estrellas de neutrones en la aproximación *force-free*.

En Paper I, presentamos, por primera vez, modelos magnetosféricos tridimensionales de magnetars con bucles magnéticos retorcidos localizados y conectados a puntos calientes en la superficie. Este tipo de configuraciones están fuertemente respaldada por datos observacionales y nuestro estudio contribuye a dar una explicación más cuantitativa del mecanismo que origina la actividad explosiva de los magnetars. Nuestros resultados proporcionan información importante sobre la naturaleza y propiedades geométricas de los puntos calientes y su conexión con la configuración magnetosférica general. Este es un paso importante hacia una modelización general y autoconsistente de la actividad de los magnetars.

Además, implementamos por primera vez una PINN para obtener una solución numérica. Se muestra que este método moderno es robusto y suficientemente preciso, pero también lo suficientemente flexible como para abordar muchos de los asuntos abiertos en problemas relacionados.

En Paper II presentamos una mejora en la implementación de PINNs que se entrena para condiciones de contorno y términos fuente arbitrarios. Esta idea novedosa nos permite acoplar la evolución magnetotérmica interior con una magnetosfera libre de fuerza exterior, permitiendo que las corrientes fluyan libremente a través de la corteza. Mostramos que esto puede afectar significativamente a la evolución interior, permitiendo nuevas y no exploradas familias de soluciones. Además, tiene la ventaja formal de que logramos imponer condiciones de contorno magnetosféricas de manera autoconsistente y no a través de parámetros definidos por el usuario.

En el Paper III hemos extendido el uso de las PINNs a las magnetosferas de púlsares incluyendo el efecto de la rotación. Mostramos que nuestro método es capaz de tratar patologías como la lámina de corriente y, además, proporcionamos una serie de familias de soluciones que capturan eficazmente todos los modelos encontrados en la literatura.

Nuestra comprensión es que las PINNs, a pesar de su corta edad como resolvedores de EDP, pueden ser una herramienta poderosa en muchas aplicaciones donde los métodos clásicos tienen limitaciones o fallan directamente. Tales casos incluyen el acoplamiento de diferentes dominios donde las simulaciones conjuntas son prohibitivas, o problemas con discontinuidades. Esperamos que a medida que el campo madure aún más, su eficiencia y precisión mejoren. Específicamente, su propiedad de escalar bien para dimensiones más altas los convierte en la opción más competitiva para problemas a gran escala.

Todo el trabajo presentado en esta tesis fue diseñado con el objetivo de allanar el camino para una modelización completa y autoconsistente en 3D de las magnetosferas de estrellas de neutrones. Creemos que este objetivo se logró tanto en términos de modelización física como de implementación numérica. Los proyectos en curso ya están mejorando las propiedades de convergencia y precisión de las PINNs y extendiendo la implementación al caso tridimensional completo para la magnetosfera del púlsar. Este es un problema aún abierto y sin solución. Los extensiones futuras de este proyecto se centrarán en estudiar más a fondo las propiedades de los modelos magnetosféricos en 3D y su acoplamiento con la evolución interior.

Referencias

Adsuara, J. E. et al. (sep. de 2016). "Scheduled Relaxation Jacobi method: Improvements and applications". En: *Journal of Computational Physics* 321, págs. 369-413. DOI: 10.1016/j.jcp.2016.05.053. arXiv: 1511.04292 [math.NA].

Akgün, T. et al. (oct. de 2017). "The force-free twisted magnetosphere of a neutron star – II. Degeneracies of the Grad–Shafranov equation". En: *Monthly Notices of the Royal Astronomical Society* 474.1, 625–635. ISSN: 1365-2966. DOI: 10.1093/mnras/stx2814. URL: http://dx.doi.org/10.1093/mnras/stx2814.

Baade, W. y F. Zwicky (mayo de 1934). "Cosmic Rays from Super-novae". En: *Proceedings of the National Academy of Science* 20, págs. 259-263. DOI: 10.1073/pnas.20.5.259.

Beloborodov, Andrei M. (2012). "ON THE MECHANISM OF HARD X-RAY EMISSION FROM MAGNETARS". En: *The Astrophysical Journal* 762.1, pág. 13. DOI: 10.1088/0004-637X/762/1/13. URL: https://dx.doi.org/10.1088/0004-637X/762/1/13.

Beloborodov, Andrei M. y Christopher Thompson (mar. de 2007). "Corona of Magnetars". En: *The Astrophysical Journal* 657.2, págs. 967-993. DOI: 10.1086/508917. URL: https://doi.org/10.1086%2F508917.

Chadwick, J. (feb. de 1932). "Possible Existence of a Neutron". En: *Nature* 129, pág. 312. DOI: 10.1038/129312a0.

Contopoulos, Ioannis, Demosthenes Kazanas y Christian Fendt (ene. de 1999). "The Axisymmetric Pulsar Magnetosphere". En: 511.1, págs. 351-358. DOI: 10.1086/306652. arXiv: astro-ph/9903049 [astro-ph].

Gold, T. (mayo de 1968). "Rotating Neutron Stars as the Origin of the Pulsating Radio Sources". En: *Nature* 218, págs. 731-732. DOI: 10.1038/218731a0.

Grad, H. y H. Rubin (1958). *Hydromagnetic equilibria and force-free fields*. Inf. téc. INIS-XU–021. United Nations (UN), págs. 190-197. URL: http://inis.iaea.org/search/search.aspx?orig_q=RN:39082408.

Harding, Alice K. (dic. de 2013). "The neutron star zoo". En: *Frontiers of Physics* 8, págs. 679-692. DOI: 10.1007/s11467-013-0285-0. arXiv: 1302.0869 [astro-ph.HE].

Hewish, A. et al. (feb. de 1968). "Observation of a Rapidly Pulsating Radio Source". En: *Nature* 217, págs. 709-713. DOI: 10.1038/217709a0.

Lagaris, I. E., A. Likas y D. I. Fotiadis (mayo de 1997). “Artificial Neural Networks for Solving Ordinary and Partial Differential Equations”. En: *arXiv e-prints*, physics/9705023, physics/9705023. arXiv: physics/9705023 [physics.comp-ph].

Lander, S K y K N Gourgouliatos (abr. de 2019). “Magnetic-field evolution in a plastically failing neutron-star crust”. En: *Monthly Notices of the Royal Astronomical Society* 486.3, págs. 4130-4143. ISSN: 0035-8711. DOI: 10.1093/mnras/stz1042. eprint: https://academic.oup.com/mnras/article-pdf/486/3/4130/28559240/stz1042.pdf. URL: https://doi.org/10.1093/mnras/stz1042.

Mahlmann, J F et al. (oct. de 2019). “Instability of twisted magnetar magnetospheres”. En: *Monthly Notices of the Royal Astronomical Society* 490.4, 4858–4876. ISSN: 1365-2966. DOI: 10.1093/mnras/stz2729. URL: http://dx.doi.org/10.1093/mnras/stz2729.

Mereghetti, Sandro, José A. Pons y Andrew Melatos (oct. de 2015). “Magnetars: Properties, Origin and Evolution”. En: *Space Sci. Rev.* 191.1-4, págs. 315-338. DOI: 10.1007/s11214-015-0146-y. arXiv: 1503.06313 [astro-ph.HE].

Pacini, F. (nov. de 1967). “Energy Emission from a Neutron Star”. En: *Nature* 216, págs. 567-568. DOI: 10.1038/216567a0.

Philippov, A. y M. Kramer (ago. de 2022). “Pulsar Magnetospheres and Their Radiation”. En: *ARA&A* 60, págs. 495-558. DOI: 10.1146/annurev-astro-052920-112338.

Raissi, M., P. Perdikaris y G. E. Karniadakis (feb. de 2019). “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations”. En: *Journal of Computational Physics* 378, págs. 686-707. DOI: 10.1016/j.jcp.2018.10.045.

Stefanou, Petros, Jose A. Pons y Pablo Cerdá-Durán (feb. de 2023). “Modelling 3D force-free neutron star magnetospheres”. En: *MNRAS* 518.4, págs. 6390-6400. DOI: 10.1093/mnras/stac3570. arXiv: 2211.08957 [astro-ph.HE].

Stefanou, Petros, Jorge F. Urbán y José A. Pons (nov. de 2023). “Solving the pulsar equation using physics-informed neural networks”. En: *MNRAS* 526.1, págs. 1504-1511. DOI: 10.1093/mnras/stad2840. arXiv: 2309.06410 [astro-ph.HE].

Thompson, Christopher y Robert C. Duncan (jul. de 1995). “The soft gamma repeaters as very strongly magnetized neutron stars - I. Radiative mechanism for outbursts”. En: *Monthly Notices of the Royal Astronomical Society* 275.2, págs. 255-300. ISSN: 0035-8711. DOI: 10.1093/mnras/275.2.255. eprint: https://academic.oup.com/mnras/article-pdf/275/2/255/18539652/mnras275-0255.pdf. URL: https://doi.org/10.1093/mnras/275.2.255.

— (dic. de 1996). “The Soft Gamma Repeaters as Very Strongly Magnetized Neutron Stars. II. Quiescent Neutrino, X-Ray, and Alfven Wave Emission”. En: *The Astrophysical Journal* 473.1, págs. 322-342. DOI: 10.1086/178147. URL: https://doi.org/10.1086/178147.

Urbán, Jorge F. et al. (sep. de 2023). “Modelling force-free neutron star magnetospheres using physics-informed neural networks”. En: *MNRAS* 524.1, págs. 32-42. DOI: 10.1093/mnras/stad1810. arXiv: 2303.11968 [astro-ph.HE].

Younes, George et al. (ene. de 2022). “Pulse Peak Migration during the Outburst Decay of the Magnetar SGR 1830-0645: Crustal Motion and Magnetospheric Untwisting”. En: *ApJ* 924.2, L27, pág. L27. DOI: 10.3847/2041-8213/ac4700. arXiv: 2201.05517 [astro-ph.HE].

Appendices

A1. Publicaciones

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **518,** 6390–6400 (2023)
Advance Access publication 2022 December 7

https://doi.org/10.1093/mnras/stac3570

Modelling 3D force-free neutron star magnetospheres

Petros Stefanou,[1★] Jose A. Pons[2] and Pablo Cerdá-Durán[1]

[1]Departament d'Astronomia i Astrofèsica, Universitat de Valéncia, Dr. Moliner 50, E-46100 Burjassot, Valéncia, Spain
[2]Departament de Fèsica Aplicada, Universitat d'Alacant, Ap. Correus 99, E-03080 Alacant, Spain

Accepted 2022 November 30. Received 2022 November 29; in original form 2022 November 7

ABSTRACT

Magnetars exhibit a variety of transient high-energy phenomena in the form of bursts, outbursts, and giant flares. It is a common belief that these events originate in the sudden release of magnetic energy due to the rearrangement of a twisted magnetic field. We present global models of a 3D force-free (FF) non-linear twisted magnetar magnetosphere. We solve the FF equations following the Grad–Rubin approach in a compactified spherical domain. Appropriate boundary conditions are imposed at the surface of the star for the current distribution and the magnetic field. Our implementation is tested by reproducing various known analytical as well as axisymmetric numerical results. We then proceed to study general 3D models with non-axisymmetric current distributions, such as fields with localized twists that resemble hotspots at the surface of the star, and we examine characteristic quantities such as energy, helicity, and twist. Finally, we discuss implications on the available energy budget, the surface temperature, and the diffusion time-scale, which can be associated with observations.

Key words: magnetic fields – stars: magnetars – stars: neutron.

1 INTRODUCTION

Magnetars are a special class of isolated neutron stars (NSs) characterized by their strong magnetic field, typically of the order of 10^{14} G or more. They are relatively young and slowly rotating with periods of the order of 1–10 s. This combination of characteristics suggests a very rapid spin-down of the star due to intense magnetic braking, placing these objects in the upper right corner of the $P\dot{P}$ diagram. Contrary to other classes of NSs, where rotation or accretion provides the necessary energy reservoir, the powering source of magnetar emission is the dissipation of their enormous magnetic field (Thompson & Duncan 1995, 1996). Magnetars are observed mainly in the X-ray and soft γ-ray bands and show a plethora of transient phenomena, highly variable and with unpredictable occurrence (Mereghetti, Pons & Melatos 2015; Kaspi & Beloborodov 2017). This transient activity is believed to be linked to the presence of a twisted, current-filled magnetosphere supported by the long-term evolution of the interior magnetic field (Beloborodov & Thompson 2007). In particular, the evolution of the magnetic field in the NS crust gradually displaces the footprints of the magnetic field lines, feeding the magnetosphere with energy and introducing currents in the twisted region. At some critical point, a large-scale reconfiguration of the magnetic field takes place, releasing the stored energy in the form of high-energy radiation. Whether the triggering mechanism of these events is originated in the star's crust, caused by a mechanical failure due to the build-up of magnetic stresses (Perna & Pons 2011), or is magnetospherically driven, where the large magnetospheric twist becomes unsustainable and reconnection takes place (Lyutikov 2003), is still a topic of investigation.

Nevertheless, an important task in revealing the physics of the magnetar bursting activity is the study of equilibrium solutions of the twisted magnetospheric field for a given field and current distribution at the surface of the star. Various works in the literature have attacked this problem in axial symmetry, by solving a Grad–Shafranov equation. Glampedakis, Lander & Andersson (2014) provided stationary axisymmetric solutions extending from the interior of the star to the magnetosphere. A similar effort was done by Pili, Bucciantini & Del Zanna (2015), but in a GRMHD framework. Kojima (2017) studied magnetospheric solutions considering effects of general relativity. Akgün et al. (2016, 2017) provided a detailed parametric study, varying the extent of the current-filled region, the non-linearity of the model, and the strength of the toroidal field. They showed that the magnetosphere could support twists up to ∼1.5 rad and could reach total energies up to ∼ 25 per cent more than the vacuum solution. A second interesting finding was that the Grad–Shafranov equation is degenerate and solutions with different total energies could exist for the same model parameters. By performing dynamical simulations on these models, Mahlmann et al. (2019) confirmed that only the lowest energy branch of these solutions is stable. Similar instabilities have been observed in a dynamically twisted magnetospheres for values of ∼1 rad (Parfrey, Beloborodov & Hui 2012, 2013; Carrasco et al. 2019).

In all these works, the general consensus is that beyond a maximum twist, no magnetospheric solution is achievable. This is a strong sign of the existence of some critical point where solutions may become unstable, leading to a sudden, large-scale reorganization of the magnetic field.

Some observational features suggest that the topology of the twisted magnetosphere is highly unlikely to be axisymmetric (Tiengo et al. 2013) and rather resembles the sun's coronal configuration, with localized, closed magnetic loops (Younes et al. 2022). Theoretical results on the evolution of the internal magnetic field also lean

★ E-mail: petrouil1@hotmail.com, petros.stefanou@uv.es

© 2022 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/518/4/6390/6881733 by Universidad de Alicante user on 26 February 2024

towards the same conclusion. For example, Lander & Gourgouliatos (2019) argue that the magneto-plastic evolution in the NS crust could produce magnetospheric twist that are local and non-axisymmetric.

The evolution of the magnetic field inside the star plays a key role in the eventual magnetospheric configuration. The deviation of the exterior magnetic field from the standard, current-free, dipole picture depends on the current distribution and the toroidal field at the surface, which, in turn, are a result of the magneto-thermal evolution of the magnetic field in the crust, under the effect of Hall drift and Ohmic dissipation (Viganò et al. 2013). Additionally, allowing currents to flow through the crust (i.e using a non-vacuum solution as a surface boundary condition for the magneto-thermal evolution) can significantly affect the internal magnetic field. Sadly, the vastly different physical conditions in the two regions, mainly the plasma density and the Alfvén crossing time, make it difficult to perform joint studies of the evolution of the magnetic field both inside and outside the star. As a possible solution, Akgün et al. (2017, 2018) performed 2D simulations of the magneto-thermal evolution of the NS crust coupled to a force-free (FF) magnetospheric solution of the Grad–Shafranov equation. In this work, we aim at giving a step forward in our understanding of magnetar magnetospheres by presenting the first 3D models in this context.

The following sections are organized as follows. In Section 2, we state the problem and we give an overview of the mathematical equations that describe it. In Section 3, we present the numerical method that we employ and go over the details of the numerical set-up. We perform tests for the reliability of our numerical implementation in Section 4. Sections 5 and 6 are dedicated to our results and an analysis of their most interesting features. We conclude with some final remarks and future suggestions in Section 7.

2 BASIC EQUATIONS

2.1 FF field

The magnetospheres of isolated NSs are generally considered to be dominated by the magnetic field. The physical conditions in the magnetosphere, i.e. low mass of the charge carriers, low density of the magnetospheric plasma, and non-relativistic co-rotational speed, allow us to neglect gravity, thermal pressure, and particle inertia as they are several orders of magnitude smaller than the **electromagnetic** forces. In addition, in the particular case of magnetars, due to their slow rotation, the light cylinder is expected to lie far away from the star's surface and the rotationally induced electric fields are not expected to play any significant role in the dynamics of the magnetosphere. With these assumptions, the equation that describes the force balance in the magnetosphere is reduced to simply demanding the magnetic force to be equal to zero

$$\boldsymbol{J} \times \boldsymbol{B} = 0, \tag{1}$$

where

$$\boldsymbol{J} = \frac{c}{4\pi} \nabla \times \boldsymbol{B} \tag{2}$$

is the electric current. In addition, the magnetic field needs to be divergence-free, as described by the solenoidal condition

$$\nabla \cdot \boldsymbol{B} = 0. \tag{3}$$

Equation (1) merely implies that the current is always parallel to the magnetic field

$$\nabla \times \boldsymbol{B} = \alpha(\boldsymbol{r}) \boldsymbol{B}, \tag{4}$$

where α is a parameter, usually called the FF parameter or the current parameter, associated with the strength of the twist (or equivalently, the toroidal component of the magnetic field in axisymmetry) in the magnetosphere. Equation (4) is the so-called FF equation that describes the magnetic field of a twisted magnetosphere in the FF regime. It is straightforward to show, from equations (3) and (4), that α has to obey the constraint

$$\boldsymbol{B} \cdot \nabla \alpha = 0, \tag{5}$$

which states that α is constant along any particular magnetic field line. In other words, the field lines lie on isocontours of the parameter α. We note that equation (5) introduces an implicit non-linearity to the problem through the dependence of α on the magnetic field. In general, three distinct cases of equation (4) can be studied:

(i) $\alpha = 0$: current-free field ($\boldsymbol{J} = 0$).
The equation reduces to $\nabla \times \boldsymbol{B} = 0$, which describes a potential field consisting of a linear combination of multipoles.

(ii) $\alpha =$ const. : linear FF field.
The equation reduces to the Helmholtz equation $\nabla^2 \boldsymbol{B} + a^2 \boldsymbol{B} = 0$, which accepts solutions in terms of the spherical Bessel functions.

(iii) $\alpha = \alpha(\boldsymbol{r})$: non-linear FF field.
The equation is non-linear. One has to use numerical methods in order to determine the solution.

The first two cases, while useful to give insight, do not correspond to magnetospheres that can appear in nature. Case (i) describes a current-free, vacuum magnetosphere which is not consistent with the well-established idea of the existence of a twist and cannot explain the magnetospheric emission observed from magnetars. Case (ii) results in solutions that do not decay with distance but are oscillatory at infinity and therefore conservation of energy cannot be enforced.

Case (iii) is the only one that is physically relevant and can describe a real magnetar magnetosphere. Solving this non-linear equation, especially in a 3D, non-axisymmetric set-up, is a challenging task. We describe the numerical method with which we attack this problem in Section 3.

2.2 Compactification

The absence of elements such as the light cylinder, an accretion disc, etc. highly favours the use of spherical coordinates for the description of the problem. However, since the physical solutions need to decay sufficiently fast at large distances, one has to place the outer boundary far away to make sure that the fields satisfy the required asymptotic solution. This usually comes at the cost of lowering the resolution, since achieving both is computationally expensive. In addition, special attention must be paid to the implementation of boundary conditions to avoid contamination of the solution from inaccuracies in the boundaries. These problems can severely hinder any attempt to achieve accurate solutions, especially in 3D, where the computational cost is an important factor.

An effective strategy is to compactify the radial coordinate by mapping the entire physical domain, from the stellar surface R up to infinity, to a sphere of radius R, with the origin as its centre. This is achieved by the following transformation:

$$r \in [R, \infty] \longrightarrow q \in [0, R], \tag{6}$$

which can be realized by the, rather simple, transformation formula

$$q = \frac{R^2}{r}, \tag{7}$$

while the angular coordinates θ and φ remain unchanged.

Downloaded from https://academic.oup.com/mnras/article/518/4/6390/6881733 by Universidad de Alicante user on 26 February 2024

One can easily check that in both systems the surface is given by the same value, $r = R \Leftrightarrow q = R$ and that infinity in the normal radial coordinate is represented by just one point, the origin, in the compact radial coordinate, $r = \infty \Leftrightarrow q = 0$.

By introducing this coordinate system and writing the equations in terms of q instead of r, we alleviate the difficulties of imposing boundary conditions at large distances. The price to pay is small as only minor changes are introduced in the equations (see the next section). With this approach, all the quantities that need to be specified at infinity and are expected to decay with radius can be simply put to zero at $q = 0$.

An additional advantage of this coordinate transformation is that quantities that are inversely proportional to some power of the radial coordinate $\propto r^{-p}$ now become polynomials in terms of the coordinate q. This makes finite differences approximations of their derivatives significantly more accurate.

Hereafter, we use compactified spherical coordinates as our coordinate system unless stated otherwise. In order to maintain a clear and simple notation, we assume that distances are measured in terms of the stellar surface so that $R = 1$. Therefore, $q \in [0, 1]$ and the surface of the star corresponds to $q = 1$.

3 METHOD

There exists a number of different numerical methods in the literature that can be used for solving equations similar to equation (4), which are mainly encountered in the area of solar physics for the reconstruction of coronal magnetic fields (see Wiegelmann & Sakurai 2021, for a detailed review on this topic). For our purposes, a suitable method is the so-called *Grad–Rubin* method.

3.1 General idea

The Grad–Rubin method was originally proposed by Grad & Rubin (1958) in the context of laboratory plasma physics but implemented for the first time by Sakurai (1981) and later by Amari, Boulmezaoud & Mikic (1999) to determine local magnetic field configurations in the solar corona.

The approach that we follow in this work is largely based on the one described in Amari et al. (1999, 2006, 2013) for the reconstruction of magnetic fields in the solar corona. In this approach, the magnetic field $\boldsymbol{B}$ is described in terms of a vector potential $\boldsymbol{A}$, given by

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}. \tag{8}$$

We choose a convenient gauge condition corresponding to the Coulomb gauge:

$$\nabla \cdot \boldsymbol{A} = 0, \; (q \leq 1) \tag{9}$$

$$\nabla_{ang} \cdot \boldsymbol{A}_{ang}\Big|_S = 0, \; (\text{surface}, \;\; q = 1), \tag{10}$$

where the subscript S denotes the surface of the NS at $q = 1$ and

$$\nabla_{ang} \cdot \boldsymbol{A}_{ang} = \frac{q}{\sin\theta}\frac{\partial}{\partial\theta}(A_\theta \sin\theta) + \frac{q}{\sin\theta}\frac{\partial A_\phi}{\partial\theta} \tag{11}$$

is the part of the divergence that involves only the angular derivatives.

Boundary conditions need to be provided at the surface of the NS. These are: (a) $b_S(\theta, \phi)$, the radial component of the magnetic field, and (b) $\alpha_S(\theta, \phi)$, the value of α at the surface. Condition (b), in particular, determines the region where currents flow, and must be implemented in a way that ensures that the two footprints of any given magnetic field line have the same value of α.

As argued in Bineau (1972), the existence of a solution is guaranteed for such boundary conditions if, in addition, the values of α are sufficiently small and the non-linearity is relatively weak.

In the numerical procedure devised by Grad and Rubin, the system of equations (4) and (5) is solved by means of a fixed-point iteration, in which one solves alternatively equation (5) for α with $\boldsymbol{B}$ fixed (hyperbolic part) and equation (4) for $\boldsymbol{B}$ with α fixed (elliptic part), until convergence to a global solution is achieved. We described next the numerical details of both parts separately.

As discussed in Section 2.2, our implementation uses compactified spherical coordinates (q, θ, φ). The range of our domain is $q \in [0, 1], \theta \in [0, \pi], \varphi \in [0, 2\pi]$, where we discretize all our equations in a grid $\{q_i, \theta_j, \phi_k\}$, with $i = 1,..., m, j = 1,..., n$, and $k = 1,..., o$. The surface and infinity are the only real boundaries on our domain, where physical boundary conditions need to be imposed. The boundaries on θ and φ are just internal boundaries, where care needs to be taken to ensure that the solution crosses them smoothly and there are no singularities that may cause divergence of the solution at the axis.

3.2 Hyperbolic part

For a given magnetic field $\boldsymbol{B}$, equation (5) implies that, whatever the solution of the FF equation, it has to fulfil that α is constant along magnetic field lines. In particular, at the two footprints of each magnetic field line the value of α has to be identical. This sets a constraint on the possible values of α at the surface (the boundary condition α_0). The problem is that this constraint cannot be set a priori, because it needs the knowledge of the magnetic field line structure, which is part of the solution that we want to obtain. Setting an arbitrary function α_0 could very easily lead to boundary conditions with no FF magnetospheric solution possible.

For this reason, α_S must be provided only at a region of a given magnetic polarity, i.e. either the region with $B_q > 0$ (positive polarity) or the region with $B_q < 0$ (negative polarity). To simplify the arguments, we always assume the former, since the other case is analogous.

In this case, an initial value problem for α can be written in the form

$$\boldsymbol{B}^{(n)} \cdot \nabla\alpha^{(n+1)} = 0, \; (\text{exterior}, \;\; q < 1) \tag{12}$$

$$\alpha^{(n+1)}\Big|_{S^+} = \alpha_S, \; (\text{surface}, \;\; q = 1), \tag{13}$$

where the exponent n denotes the fixed-point iteration number and the subscript S^+ denotes the part of the surface where $B_q > 0$.

For our purposes, the most suitable method for solving equation (12) is the *Characteristics* or *Field Line Tracing* method, because the characteristic curves of α are none other than the magnetic field lines.

Every grid point in the domain is threaded by a field line. If the only source of currents is the NS surface, then any field line can be traced back to the surface of the star where the values of α are known. Therefore, it is possible to build a numerical method that traces magnetic field lines from any grid point to the surface of the star and assigns the value of α there to the grid point. However, this procedure would be computationally expensive in general. Instead, we have devised a numerical method that is essentially equivalent but significantly more efficient computationally.

The magnetic field lines can be computed by integrating the equation

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\lambda} = \frac{\boldsymbol{B}}{||\boldsymbol{B}||}, \tag{14}$$

Downloaded from https://academic.oup.com/mnras/article/518/4/6390/6881733 by Universidad de Alicante user on 26 February 2024

where λ is a parametric variable corresponding to the arc length along the magnetic field line. Numerically, it is convenient to express and solve this equation in Cartesian coordinates to avoid the coordinate singularities at the axis associated with the compactified coordinates chosen in this work. Details on the transformations from spherical to Cartesian coordinates and vice versa are given in Appendix B.

The first part in the procedure is to compute the value of α at all grid points at the surface at q_1, a sphere right above the surface of the star. For this purpose, we integrate the magnetic field lines starting at every grid point with coordinates (q_1, θ_j, ϕ_k) in the direction opposite to $\boldsymbol{B}$. This integration is continued all along the magnetosphere until the surface is reached. Because we are following the field lines backwards, the landing point (footprint) necessarily has $B_q > 0$ and hence is part of S_+, where the values of α are known. In general, the footpoint will not coincide precisely with a grid point, but is possible to compute it as an interpolation of α_0 at the surface. As a result, we get the value of α at all points in q_1.

The second part of the procedure builds the solution in the rest of the magnetosphere ($i > 1$) computing it layer by layer in the radial direction. To compute α in a radial layer at q_{i+1}, from the values of α at q_i, we integrate a magnetic field line starting at each point $(q_{i+1}, \theta_j, \phi_k)$ in the direction towards the surface at q_i. This integration is very fast in general, because the distance between both surfaces is small. The landing point at q_i will not coincide in general with a grid point, but the value of α can be interpolated within the surface, and assigned to corresponding point at q_{i+1}. The procedure is continued until the solution at all grid points is known.

All the integrations (equation 14) are performed using the Adams predictor-corrector methods. All the necessary interpolations are preformed using B-splines of order $k = 2$.

3.3 Elliptic part

Once $\alpha^{(n+1)}$ is known everywhere, we proceed to compute the value of $\boldsymbol{B}^{(n+1)}$ or equivalently $\boldsymbol{A}^{(n+1)}$. The magnetic field can be calculated everywhere in the domain by solving the elliptic boundary value problem

$$\nabla \times \boldsymbol{B}^{(n+1)} = \alpha^{(n+1)} \boldsymbol{B}^{(n)} \tag{15}$$

$$B_q(\theta, \varphi)\Big|_S = b_S(\theta, \varphi), \tag{16}$$

$$\boldsymbol{B}\Big|_\infty = 0. \tag{17}$$

with a fixed source term $\alpha^{(n+1)}\boldsymbol{B}^{(n)}$.

We need to express this problem in terms of the vector potential in the selected gauge. Using equations (8) and (9) and some vector calculus identities, we can restate equation (15) as an equation for the vector potential

$$\nabla^2 \boldsymbol{A}^{(n+1)} = -\alpha^{(n+1)} \boldsymbol{B}^{(n)}. \tag{18}$$

Besides, by introducing a poloidal scalar field Φ, which, at the surface, satisfies the condition

$$\nabla^2_{ang}\Phi\Big|_S = -b_0(\theta, \varphi), \tag{19}$$

we can calculate the angular components of the vector potential at the surface as follows:

$$A_\theta\Big|_S = \tfrac{1}{\sin\theta}\tfrac{\partial\Phi}{\partial\phi} \tag{20}$$

$$A_\phi\Big|_S = -\tfrac{\partial\Phi}{\partial\theta}. \tag{21}$$

More details about these derivations can be found in Appendix A (see, in particular, equations A3 and A6). The radial component of the vector potential, from equations (9) and (10), needs to satisfy the Robin-type boundary condition

$$\left(2qA_q - q^2\frac{\partial A_q}{\partial q}\right)\Big|_S = 0. \tag{22}$$

At infinity, all components need to smoothly decay, so the boundary condition (17) is simply replaced by

$$\boldsymbol{A}\Big|_\infty = 0. \tag{23}$$

Equation (18) is a vector *Poisson* equation with a linear source term, the known current from the previous iteration step. The three components of (18) form a system of equations which is coupled, because of terms introduced by the action of the *Laplacian* on a vector in curvilinear coordinates. This is an inconvenience, but our problem was non-linear to begin with and we already had to deal with an iteration procedure for its solution. We can simply pass any additional, coupled terms to the right-hand side and solve a scalar *Poisson* equation for each component with a generalized source term. Expanding equation (18), we get a system of three equations:

$$\begin{aligned}
&\nabla^2 A_q^{(n+1)} - 2q^2 A_q^{(n+1)} \\
&= -\alpha^{(n+1)} B_q^{(n)} + 2q^2 \cot\theta\, A_\theta^{(n)} + 2q^2 \frac{\partial A_\theta^{(n)}}{\partial\theta} \\
&+ \frac{2q^2}{\sin\theta}\frac{\partial A_\varphi^{(n)}}{\partial\varphi},
\end{aligned} \tag{24}$$

$$\begin{aligned}
&\nabla^2 A_\theta^{(n+1)} - q^2 \frac{A_\theta^{(n+1)}}{\sin^2\theta} \\
&= -\alpha^{(n+1)} B_\theta^{(n)} - 2q^2 \frac{\partial A_q^{(n)}}{\partial\theta} + 2q^2 \frac{\cos\theta}{\sin^2\theta}\frac{\partial A_\varphi^{(n)}}{\partial\varphi},
\end{aligned} \tag{25}$$

$$\begin{aligned}
&\nabla^2 A_\varphi^{(n+1)} - q^2 \frac{A_\varphi^{(n+1)}}{\sin^2\theta} \\
&= -\alpha^{(n+1)} B_\varphi^{(n)} - \frac{2q^2}{\sin\theta}\frac{\partial A_q^{(n)}}{\partial\varphi} - 2q^2 \frac{\cos\theta}{\sin^2\theta}\frac{\partial A_\theta^{(n)}}{\partial\varphi}.
\end{aligned} \tag{26}$$

Here, the first term (the current) in the right-hand side of each equation is the actual physical source term. All the other terms in the right-hand side of the equations are by-products of the vector *Laplacian*. The second term in the left-hand side is also a by-product of the vector *Laplacian*, but we chose to keep it there for numerical implementation purposes.

All elliptic operators are solved using a Scheduled Relaxation Jacobi method (Adsuara et al. 2016). This iterative method has the simplicity of the Jacobi method but with convergence properties similar to the optimal Successive Over-Relaxation (SOR) method. It also has the advantage with respect to the SOR method of not having to optimize any overrelaxation parameter.

3.4 Divergence cleaning

To make sure that the gauge chosen in equation (9) is maintained through the iteration process and is not affected by numerical contamination, we introduce a divergence cleaning step after the calculation of the vector potential $\boldsymbol{A}$.

If, due to approximation and round-off errors, the divergence of $\boldsymbol{A}$ is different from zero, we redefine $\boldsymbol{A}$ by introducing a scalar field ξ as follows:

$$\boldsymbol{A}' = \boldsymbol{A} - \nabla\xi, \tag{27}$$

Downloaded from https://academic.oup.com/mnras/article/518/4/6390/6881733 by Universidad de Alicante user on 26 February 2024

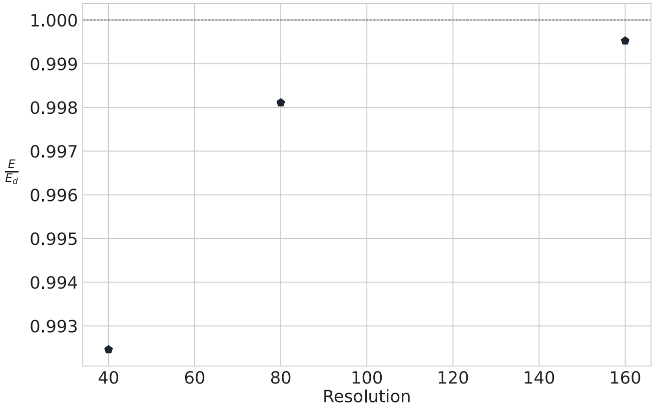


Figure 1. Normalized energy of the tilted dipole solution as a function of the resolution. The dotted line denotes the analytical solution. The labels of the x-axis show only the radial resolution for clarity.

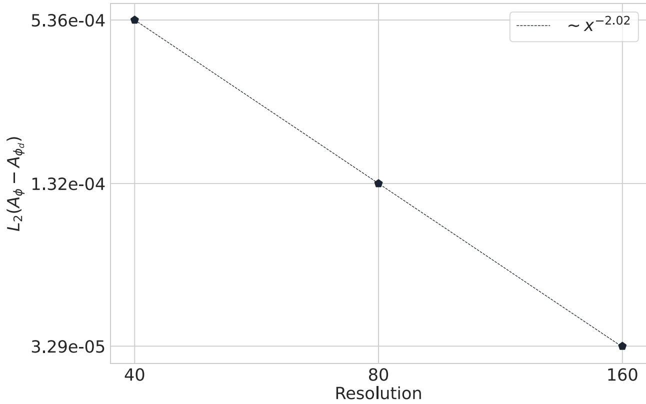


Figure 2. Convergence of the L_2 norm of the difference between numerical and analytical solution with resolution. Only the A_ϕ component is plotted. Both axes are in logarithmic scale. The dotted line corresponds to a power-law fit with index $p = 2.02$, which indicates that the convergence is of second order, as expected by our discretization.

and we demand that the divergence of the new vector potential $\boldsymbol{A}'$ satisfies the gauge $\nabla \cdot \boldsymbol{A}' = 0$. This is only true if

$$\nabla^2 \xi = \nabla \cdot \boldsymbol{A}, \tag{28}$$

that is, ξ is the solution of a Poisson equation with a source term equal to the old, non-zero divergence of the vector potential.

With this process, any numerically induced deviation from the selected gauge is compensated by subtracting the parts that contribute to it from the vector potential.

4 TESTS

Before proceeding to study astrophysically relevant models, we put our code to the test against known solutions. The following test cases will demonstrate that the code is well-behaved and reliable in capturing all the features of an FF non-axisymmetric magnetosphere in compactified spherical coordinates.

4.1 Test 1: tilted dipole

The first test aims at providing a benchmark for the code when solving for three-dimensional magnetic fields. In particular, a dipole magnetic field with its axis tilted by an angle β around the x-axis. No currents are introduced, so $\alpha = 0$ everywhere for this test. We focus on validating other aspects of the code, such as the implementation of the compactified coordinate system, the behaviour of the code close to pathological regions (axis, infinity), and the order of convergence. In addition, we calculate the energy and the helicity of the magnetic field to get an estimate of the numerical dissipation introduced.

Tilting a vector field $\boldsymbol{B}46010(\boldsymbol{r})$ is a two-step procedure. First, a rotation of the coordinates (q, θ, ϕ) by an angle $-\beta$ is performed, followed by a rotation of the components of the vector (v_q, v_θ, v_ϕ) by an angle β. The detailed process is described in Appendix B.

We calculate three different solutions for the tilted dipole, with different resolutions: $40 \times 20 \times 40$, $80 \times 20 \times 80$, and $160 \times 80 \times 160$. Fig. 1 shows the energy stored in the solution as a function of the resolution. It is calculated numerically by the integral

$$E = \int_{\mathcal{V}} \frac{B^2}{8\pi} \mathrm{d}V \tag{29}$$

over the entire volume of our domain and normalized to the analytical value for a purely dipolar solution

$$E_d = \frac{B_0^2 R^3}{3}, \tag{30}$$

where B_0 denotes the surface magnetic field at the equator and R is the radius of the star. The energy approaches nicely the expected value with increasing resolution. For the case with the highest resolution considered here, the difference between numerical and analytical value is less than 0.1 per cent. We define the helicity of the magnetic field as

$$H = \int \boldsymbol{A} \cdot \boldsymbol{B} \, \mathrm{d}V. \tag{31}$$

Helicity is of the order $\sim 10^{-15}$ for all resolutions and therefore very close to the theoretical value of $H_d = 0$. Fig. 2 shows the L_2 norm of the difference between the numerical and analytical solution. We only plot the A_ϕ component of the vector potential, since the other two components are zero and the comparison is not really informative. The convergence of the L_2 norm is of second order, as expected.

All the above metrics indicate that our numerical code is trustworthy and consistent when dealing with current-free solutions. In the next sections, we examine cases with $\alpha \neq 0$.

4.2 Test 2: axisymmetric FF field

Twisted magnetospheres in axial symmetry have been studied in detail in a number of works in the past. This test aims at reproducing some of the well-established results obtained in Akgün et al. (2017), using the Grad–Rubin method.

The two numerical schemes treat the FF problem quite distinctly. In Akgün et al. (2017), the variable of the problem is a poloidal function P and the current is introduced through a toroidal function $T(P)$, which is determined by three parameters P_c, σ, and s. These parameters control the extent of the current-filled region, the non-linearity of the model, and the strength of the current. Under the assumption of axisymmetry and the decomposition of the magnetic field into the scalars P and $T(P)$, equation (4) is reduced to the Grad–Shafranov equation for P. The vector potential $\boldsymbol{A}$ and the magnetic field $\boldsymbol{B}$ are then reconstructed from P and $T(P)$. In the Grad–Rubin method, equation (4) is solved as it is, and the variables of the problem are the vector potential $\boldsymbol{A}$, as well as the current parameter

Downloaded from https://academic.oup.com/mnras/article/518/4/6390/6881733 by Universidad de Alicante user on 26 February 2024

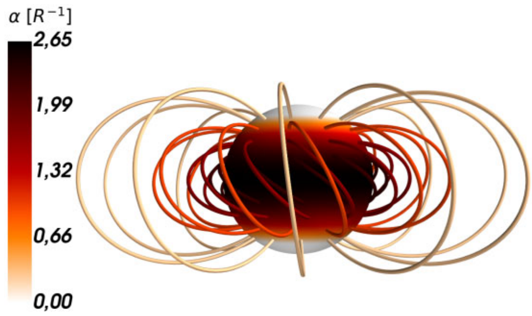


Figure 3. A reproduced example by the Grad–Rubin method, of an axisymmetric FF solution, akin to those studied in Akgün et al. (2017). Model parameters are: $P_c = 0.3405074$, $\sigma = 2$, and $s = 2$, and the solution shown corresponds to the lower energy branch.

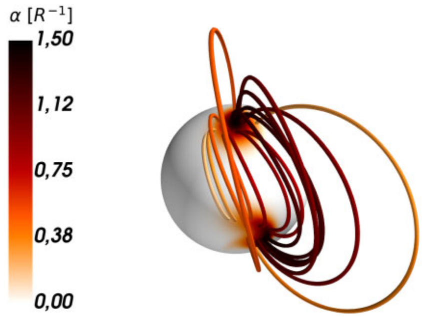


Figure 4. A selection of twisted field lines in an FF magnetosphere with two current-inducing hotspots. The colour code denotes the value of α, which is constant along any particular field line. Model parameters are: $\alpha_0 = 1.5, \theta_1 = 45^\circ$, $\phi_1 = 180^\circ$, and $\sigma = 0.2$.

α. In the nomenclature of Akgün et al. (2017), α is the same as $T'(P)$, where the prime denotes the derivative with respect to P (see also equation 8 in Akgün et al. 2016) and P is related to the azimuthial component of the vector potential through the relation $P = \frac{\sin\theta}{q} A_\phi$. It is straightforward to reproduce any of the models studied in Akgün et al. (2017) using our own implementation, by imposing at the surface the current distribution $\alpha = T'(P; P_c, \sigma, s)$ that corresponds to the parameters of that model. We remind that α has units of inverse length (we choose to express it in units of the star radius $[R^{-1}]$). For the purposes of this test, we choose the solution obtained for $P_c = 0.3405074$, $\sigma = 2$, and $s = 2$ and belongs to the lower energy branch. We will refer to this solution as the FF2D solution from now on.

As the two solutions have been calculated in different geometries and different grids, an interpolation of one of them in the grid of the other is needed in order to properly compare them. The FF2D solution has been calculated in a much finer, 2D 400×100 grid, while the resolution of our grid is at most $160 \times 80 \times 160$ for this test. We choose therefore to interpolate the FF2D solution and evaluate it in our own, coarser, grid. The 3D magnetic field configuration of the solution (as obtained by our code) is shown in Fig. 3.

Comparing the Grad–Rubin solution with the one obtained in Akgün et al. (2017), we see a converging pattern similar to the one shown in Figs 1 and 2. For the highest resolution used, the difference in energy is less than 1 per cent and in helicity less than 10 per cent. The L_2 norm of the difference of the A_ϕ component between the two solutions converges with a power-law index of $p = 1.23$. All these metrics indicate a less effective convergence than in the previous test. This is understandable and can be attributed to two factors. First, the full Grad–Rubin method with currents (hyperbolic and elliptic part) has a complicated set-up that involves iterations, integrations, interpolations, and various other steps that can affect the convergence order. Second, in this test, we are comparing two numerical solutions produced by different numerical schemes, in different grids and an interpolation is performed between them.

4.3 Test 3: tilted axisymmetric FF field

In order to test our code in a configuration as general as possible, we make one more step by tilting the FF2D solution of Section 4.2. This permits us to create a fully 3D, current-filled magnetospheric configuration, with twisted lines crossing the coordinate axes. Although this is technically still an axisymmetric solution around some axis, it is not axisymmetric in our coordinate system and therefore can serve as a valid, general 3D test with currents. Energy and helicity are compared to the values of un-tilted FF2D solution, as they should not be affected by rotation. Again, we get a difference of less than 1 per cent for the energy and less than 10 per cent for the helicity. Due to technical reasons, it is easier to compare the L_2 norm of the difference in B_ϕ component for this case, instead of A_ϕ. The convergence follows a power law of index $p = 1.21$, similar to the one for the un-tilted case.

5 RESULTS

Observations of magnetars suggest that the radial current distribution over the NS's surface is localized into specific regions, the *hotspots*, while the rest of the magnetosphere remains current-free. Current-filled lines emerge/close inside the hotspots, forming twisted magnetic loops. The resulting global magnetospheric configuration depends on the size, location, and shape of the spots, or the intensity of the current parameter α, among other variables. In this section, we systematically explore the dependence of magnetospheric models on these parameters.

5.1 Reference model

We begin by discussing a characteristic example of the family of models considered in this work: a magnetic dipole, locally twisted due to the presence of two current inducing hotspots. Only field lines whose footprints lie inside the spots are twisted, while the rest of the magnetosphere is barely affected by the presence of the localized current loop and remains current-free and purely poloidal. We impose a radial component of the magnetic field at the surface of the form

$$b_S(\theta, \phi) = 2\cos\theta, \tag{32}$$

corresponding to a dipole. The boundary conditions are completed by giving values of α in the region where $B_q > 0$, by the relation

$$\alpha_{S^+}(\theta, \phi) = \alpha_0 \mathrm{e}^{\frac{-(\theta-\theta_1)^2-(\phi-\phi_1)^2}{2\sigma^2}}. \tag{33}$$

This corresponds to a Gaussian profile, centred at the point (θ_1, ϕ_1) where σ controls the width of the Gaussian. α_0 determines the maximum of the parameter α (at the centre of the Gaussian).

In Fig. 4, we show a selection of current-filled, twisted lines connecting the two hotspots. The first immediately noticeable feature is the asymmetric 'comma-like' shape of the hotspot residing in the Southern hemisphere. We remind here that, to acquire a consistent solution, the surface values of α are given only for one polarity of the magnetic field. Thus, only one of the spots is fixed a priori

Downloaded from https://academic.oup.com/mnras/article/518/4/6390/6881733 by Universidad de Alicante user on 26 February 2024

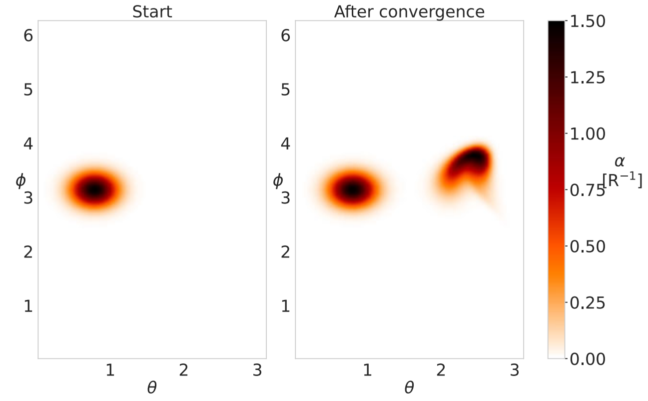


Figure 5. Surface values of α plotted in the $\theta - \phi$ plane. The left figure shows the prescribed values, only in the Northern hemisphere where $B_q > 0$. The right figure shows the values after the solution has converged, corresponding to the real boundary. Model parameters are the same as in Fig. 4.

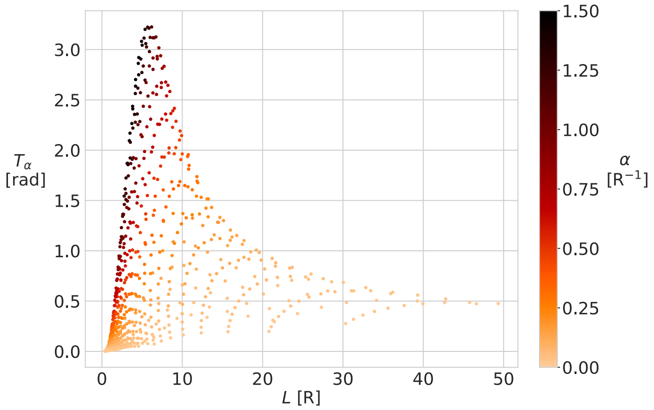


Figure 6. Twist of the magnetic field lines as a function of the total length. Each point corresponds to one field line of the reference model discussed in Section 5.1. The colour denotes the value of α labelling the line. Twist is measured in radians, using the definition (36). Lines with high α tend to have larger twists.

while the exact position and shape of the other is consistently determined. Fig. 5 compares the user-provided boundary condition for the first iteration (left) and the consistent boundary condition obtained after the solution has converged (right). We have tested that using the comma-like spot as boundary condition we recover the Gaussian-like profile on the other side, showing that the numerical solution is consistent and this is not a numerical artefact. Note that, interestingly, hotspots with a similar coma-like shape have been reported in the context of millisecond pulsars (Kalapotharakos et al. 2021), albeit created by a different mechanism (open field lines that cross the light cylinder, instead of closed twisted field lines).

5.2 Twist

The twist of a magnetic field line is defined mathematically as the integral of the torsion along the parametric curve $\boldsymbol{r}(\lambda)$ corresponding to the line:

$$T_w = \int_C \mathcal{T} \mathrm{d}\lambda, \tag{34}$$

where λ is the length along the curve and the torsion $\mathcal{T}$ is given by the following formula, which uses the Frenet–Serret unit vectors: the tangent $\hat{\boldsymbol{\tau}}$, the normal $\hat{\boldsymbol{n}}$, and the binormal $\hat{\boldsymbol{b}} = \hat{\boldsymbol{\tau}} \times \hat{\boldsymbol{n}}$,

$$\mathcal{T} = \frac{\mathrm{d}\hat{\boldsymbol{n}}}{\mathrm{d}\lambda} \cdot \hat{\boldsymbol{b}}. \tag{35}$$

Calculating the Frenet–Serret unit vectors is straightforward in the case of magnetic field lines, since $\hat{\boldsymbol{\tau}} = \frac{\boldsymbol{B}}{B}$ and $\hat{\boldsymbol{n}} = \frac{\mathrm{d}\hat{\boldsymbol{\tau}}/\mathrm{d}\lambda}{||\mathrm{d}\hat{\boldsymbol{\tau}}/\mathrm{d}\lambda||}$. The problem with the above definition is that the consecutive differentiations introduce numerical errors and noise which can be fairly significant. A simpler alternative definition (Berger & Prior 2006) is directly related to the parameter α

$$T_\alpha = \frac{1}{2} \int_C \alpha \mathrm{d}\lambda \tag{36}$$

and, since α is constant along a field line, the twist is simply given by

$$T_\alpha = \alpha L. \tag{37}$$

We refer the interested reader to appendix C in Liu et al. (2016) for an excellent and detailed discussion of the differences between the two definitions.

Fig. 6 shows the distribution of twists of a sample of field lines with footprints on the hotspot as calculated as a function of the total length L of the line in units of the stellar radius R (equation 36). The colour scale indicates the value of α that labels the line. Lines with footprints close to the centre of the spot are shown with a deep red colour and have higher values of the twist (forming the peak in the plot). A faded orange colour corresponds to lines emerging from the periphery of the hotspot, where α is relatively low and the line is only slightly twisted, but these are generally longer lines (forming the tail of the plot). The very long and twisted lines should be unstable beyond certain limit, as illustrated by the figure, where the absence of points occupying the upper right corner reflects the fact that a line cannot have simultaneously large values of α and L.

5.3 Dependence on α_0

The parameter α_0 controls the intensity of the current. As its value increases, the twist of the lines threaded by currents becomes stronger. In the first row of Fig. 7 we show samples of solutions with different values of α_0. In order to quantify the effect of α_0 on the magnetosphere, we examine the dependence of various important physical quantities on it.

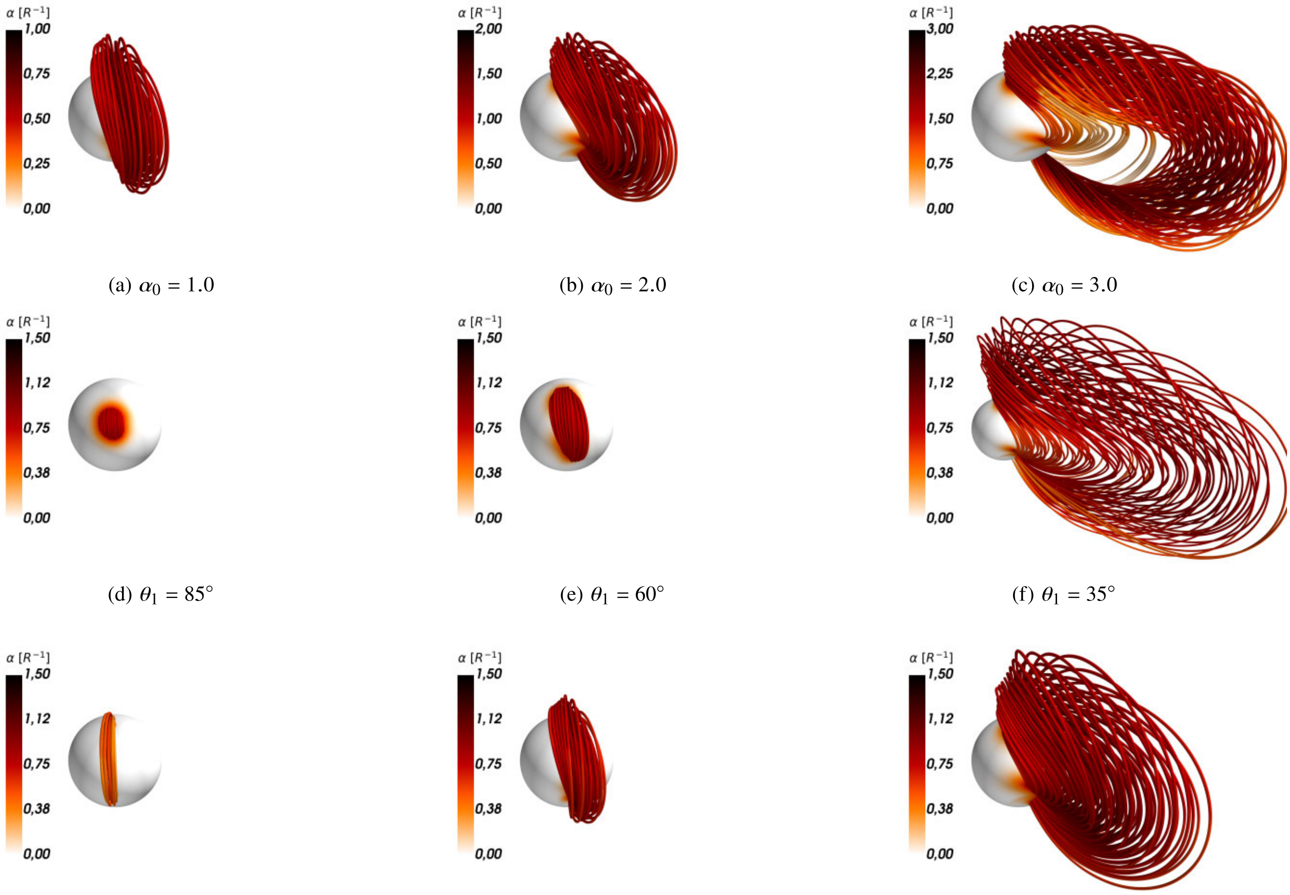
We define the normalized excess energy as

$$E_\mathrm{e} = \frac{E - E_\mathrm{d}}{E_\mathrm{d}}, \tag{38}$$

where E_d is given by equation (30). Considering that the dipole is the lowest energy configuration, this quantity gives an estimate of the available energy budget if the magnetosphere is eventually completely untwisted and relaxed.

The first row of Fig. 8 shows the dependence of the normalized excess energy E_e, the helicity H, and the maximum twist T_{α_max} (as calculated by equation 36) on α_0. As expected, all quantities increase as we get further and further away from the potential solution ($\alpha_0 = 0$). Beyond a value of $\alpha_0 \sim 3$, the numerical code is unable to converge to a solution. The non-existence of a solution can be a flag for the triggering of instabilities in the magnetosphere, as it becomes unable to support such intense currents and strong twists. Looking at Fig. 7(c), one can clearly see that magnetospheres with such high

Downloaded from https://academic.oup.com/mnras/article/518/4/6390/6881733 by Universidad de Alicante user on 26 February 2024



(a) $\alpha_0 = 1.0$ (b) $\alpha_0 = 2.0$ (c) $\alpha_0 = 3.0$

(d) $\theta_1 = 85^\circ$ (e) $\theta_1 = 60^\circ$ (f) $\theta_1 = 35^\circ$

(g) $\sigma = 0.05$ (h) $\sigma = 0.15$ (i) $\sigma = 0.25$

Figure 7. Samples of magnetospheres with different values of α_0, θ_1, and σ. In each row, only one parameter varies, while the others remain fixed to the reference model values. In each column, from left to right, the magnetosphere is farther away from the potential current-free solution. The magnetic field is more twisted for strong currents (high α_0), large hotspot separation (small θ_1), and large hotspot area (high σ). For clarity, only lines with $\alpha > 0.5\alpha_0$ are shown.

values of α threading sufficiently long lines are highly distorted and may not be stable.

5.4 Dependence on θ_1

The parameter θ_1 determines the position of the hotspot. A small value means that the spot is located close to the pole. If the initial, untwisted magnetic field is a dipole, it essentially controls the distance between hotspots and the total volume of the coronal flux tube (the farther apart the footprints, the longer the flux tube length and the larger the volume). Note that it is meaningless to consider values of $\theta_1 > 90^\circ$ for our particular untwisted magnetic field. Samples of solutions with different values of θ_1 are shown in the second row of Fig. 7. The second row in Fig. 8 shows the dependence of E_e, H, and $T_{\alpha_{max}}$ on θ_1. All three quantities increase as the hotspot's location approaches the magnetic pole. This is the expected behaviour, because the size of the current-filled region becomes larger. Changing the value of θ_1 seems to have a lesser impact on energy and helicity than increasing the intensity of the current α_0. However, it has a significant impact on the maximum twist. When current-filled lines, even with small values of α approach the pole, their twist is high due to their length and is difficult to acquire solutions. This can be explained by the fact that our assumptions for the FF regime demand that the region containing currents cannot extend up to large distances.

5.5 Dependence on σ

The parameter σ determines the area of the hotspot. We only consider cases with common σ for the θ and ϕ directions, which means that the hotspot is symmetric. A bigger value of σ means that more field lines are threaded with current. This affects not only the total volume of the current-filled region but also its extent, since a big σ means that lines close to the poles can have current. In the third row of Fig. 7 we show samples of solutions with different values of σ, while in the third row of Fig. 8 we show how energy, helicity, and maximum twist alter with increasing the size of the hotspot. As expected, models with large values of σ can store more energy, since the total volume of the twisted region increases.

6 DISCUSSION

In the models presented in the previous section, the excess energy stored in the magnetosphere can reach values up to $E_e \sim 6$ per cent of the energy of the potential solution. This is the maximum available energy that a twisted magnetosphere can release in a catastrophic

Downloaded from https://academic.oup.com/mnras/article/518/4/6390/6881733 by Universidad de Alicante user on 26 February 2024

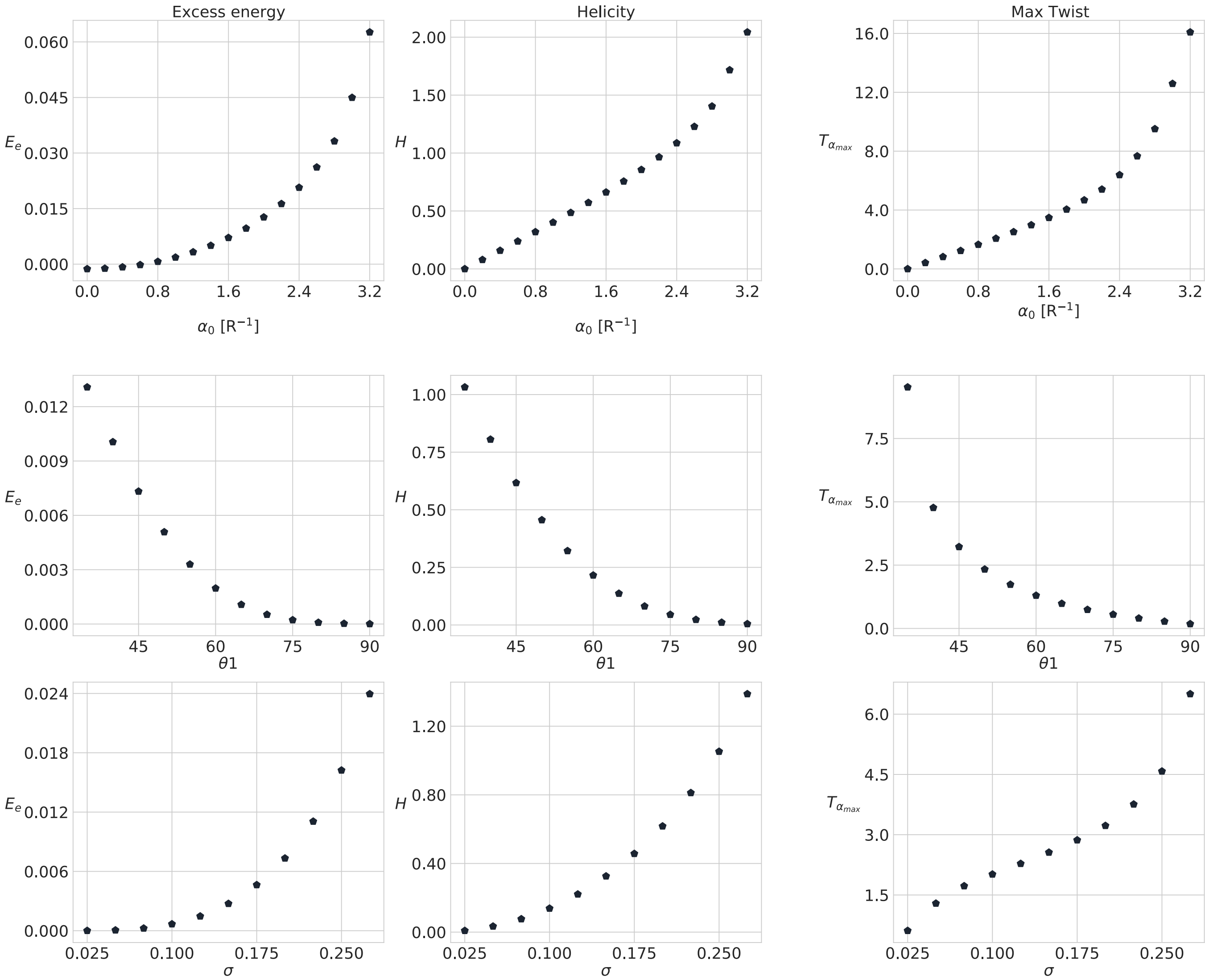


Figure 8. Dependence of the normalized excess energy E_e, the helicity H, and the maximum twist $T_{\alpha_{max}}$ on the parameters α_0 (first line), θ_1 (second line), and σ (third line).

event. The magnetic energy of a purely dipolar field is

$$E_d = 3.3 \times 10^{45} \text{ erg } B_{14}^2 \left(\frac{R}{10 \text{ km}}\right)^3, \quad (39)$$

where B_{14} is the field in units of 10^{14} G. If one considers typical values for a magnetar, i.e. a star of radius 10 km and a magnetic field of the order of 10^{13}–10^{15} G, E_e lies in the range of 2×10^{42} erg $< E_e < 2 \times 10^{46}$ erg, well in line with the typical energetics of magnetar flares and outbursts, but perhaps falls a little short in comparison to the highly energetic giant flares.

Note that an exhaustive analysis of the entire parameter space is outside the scope of this work. It is possible that a certain combination of α_0, θ_1, and σ, or even a different current profile would produce higher values of E_e. In particular, if we associate giant flares to very large-scale global events, a configuration with smaller θ_1 and large α would be indeed consistent with the most extreme cases.

However, it is highly unlikely that the results could differ drastically from what has been presented here, given the fact that finding more extreme solutions becomes increasingly more difficult. So we do not expect that hypothetical hypergiant or supergiant flares be possible. Observed giant flares seem to be already at the edge of the possibilities consistent with these family of models.

Concerning the twist, it tends to reach values significantly higher than the maximum values reported for axisymmetric magnetospheres (about 1.5 rad). This can be partly attributed to the fact that, in 3D, lines can freely twist and bend and even change direction without the restrictions imposed by axisymmetry.

An additional question relevant to the interpretation of magnetar observations is how these models differ from pure current-free magnetospheres in terms of the emission spectrum. A first relevant issue is that currents flowing through the star surface at the footprints of the twisted magnetic lines can heat the surface (hence the name hotspots).

One can estimate the temperature of the hotspot following the arguments in Akgün et al. (2018). Assuming that the current is predominantly dissipated in a thin layer below the star surface (the region with highest resistivity) and the energy released is radiated away as black body radiation from the surface of the star, the effective temperature can be calculated from (see equation 11 from Akgün

Downloaded from https://academic.oup.com/mnras/article/518/4/6390/6881733 by Universidad de Alicante user on 26 February 2024

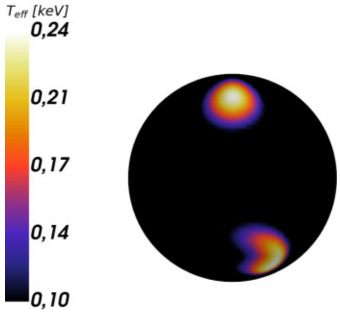


Figure 9. Effective temperature of the hotspots at the surface of the star as calculated by equation (40) for the reference model of Section 5.1. The colorbar shows the values of $T_{\rm eff}$ in keV units. The lowest value of the colorbar corresponds to the typical surface temperature of an NS in quiescence.

et al. 2018)

$$T_{\rm eff} \simeq 0.2\ {\rm keV} \left(\frac{\Delta r}{1\ {\rm m}}\right)^{\frac{1}{4}} \left(\frac{\eta}{10^4\ {\rm cm^2\ s^{-1}}}\right)^{\frac{1}{4}} B_{14}^{\frac{1}{2}} (\alpha R)^{\frac{1}{2}} \quad (40)$$

Considering Δr and η are of order 1 in the units given in the above expression, a magnetic field of $B_{14} \approx 1 - 10$ G and values of αR in the range 1–3, the effective temperature of the hotspot lies in the range of 0.1 keV $< T_{\rm eff} <$ 0.6 keV. Although this is a rough estimate, these values are on par with observations (Coti Zelati et al. 2017; Younes et al. 2022). Furthermore, our Gaussian profiles of α at the hotspot are consistent with the double BB models and the umbral/penumbral shape for the pulsed X-ray emission reported in Younes et al. (2022), with the hotter component coming from the centre of the Gaussian (say within 1σ) and the colder one from the periphery, where α and hence $T_{\rm eff}$ have decayed. Fig. 9 shows a surface map of $T_{\rm eff}$ that corresponds to the reference model of Section 5.1, in real units. Currents in the magnetosphere are not permanent, but rather gradually dissipate. The diffusion time-scale can be approximated by

$$\tau \simeq 3\ {\rm yr} (\alpha R)^{-2} \left(\frac{\eta}{10^4\ {\rm cm^2\ s^{-1}}}\right)^{-1}. \quad (41)$$

That means that stronger twists ($\alpha R > 1$) tend to live for a few months up to several years, while weaker twists last longer. This is in line with expectations from outburst observations.

Ideally, a joint simulation of the interior and the exterior of the star, where the magneto-thermal evolution of the field in the crust is coupled with an FF magnetosphere, would give us the exact location and shape of the hotspots, as well as a consistent way of estimating its temperature. Such extension is out of our present scope and reserved for future work.

7 CONCLUSION

In this work, we have presented a family of 3D FF twisted models for magnetar magnetospheres. We have considered a dipolar magnetic field and a localized current with a Gaussian profile at the surface of star. Our goal is to emulate the behaviour of the magnetosphere in the presence of current-inducing hotspots. We have studied the effect of the parameters controlling the strength and geometry of the current-filled region on physical quantities such as the energy, helicity, and twist of the magnetic field, the effective surface temperature of the hotspot, and the diffusion time-scale. Our results show good agreement with the observations and the expected properties of real magnetars. The next step is to consider a coupling of the magnetospheric solution with the internal magneto-thermal evolution so that the hotspot properties (values of α and geometry) could be determined self-consistently.

ACKNOWLEDGEMENTS

We acknowledge support from the Generalitat Valenciana (PROMETEO/2019/071) and the AEI grants PGC2018-095984-B-I00, PID2021-125485NB-C21, and PID2021-127495NB-I00. JAP acknowledges the Alexander von Humboldt Stiftung for a Humboldt Research Award. PCD acknowledges support from the *Ramon y Cajal* programme of the AEI (RYC-2015-19074).

DATA AVAILABILITY

All data produced in this work will be shared on reasonable request to the corresponding author.

REFERENCES

Adsuara J. E., Cordero-Carrión I., Cerdá-Durán P., Aloy M. A., 2016, J. Comput. Phys., 321, 369
Akgün T., Miralles J. A., Pons J. A., Cerdá-Durán P., 2016, MNRAS, 462, 1894
Akgün T., Cerdá-Durán P., Miralles J. A., Pons J. A., 2017, MNRAS, 474, 625
Akgün T., Cerdá-Durán P., Miralles J. A., Pons J. A., 2017, MNRAS, 472, 3914
Akgün T., Cerdá-Durán P., Miralles J. A., Pons J. A., 2018, MNRAS, 481, 5331
Amari T., Boulmezaoud T. Z., Mikic Z., 1999, A&A, 350, 1051
Amari T., Boulmezaoud T. Z., Aly J. J., 2006, A&A, 446, 691
Amari T., Aly J. J., Canou A., Mikic Z., 2013, A&A, 553, A43
Beloborodov A. M., Thompson C., 2007, ApJ, 657, 967
Berger M. A., Prior C., 2006, J. Phys. A: Math. Gen., 39, 8321
Bineau M., 1972, Comm. Pure Appl. Math., 25, 77
Carrasco F., Viganò D., Palenzuela C., Pons J. A., 2019, MNRAS, 484, L124
Coti Zelati F. et al., 2017, MNRAS, 471, 1819
Geppert U., Wiebicke H. J., 1991, A&AS, 87, 217
Glampedakis K., Lander S. K., Andersson N., 2014, MNRAS, 437, 2
Grad H., Rubin H., 1958, Technical report, Hydromagnetic equilibria and force-free fields. United Nations (UN), http://inis.iaea.org/search/search.aspx?orig_q = RN:39082408
Kalapotharakos C., Wadiasingh Z., Harding A. K., Kazanas D., 2021, ApJ, 907, 63
Kaspi V. M., Beloborodov A. M., 2017, ARA&A, 55, 261
Kojima Y., 2017, MNRAS, 468, 2011
Lander S. K., Gourgouliatos K. N., 2019, MNRAS, 486, 4130
Liu R. et al., 2016, ApJ, 818, 148
Lyutikov M., 2003, MNRAS, 346, 540
Mahlmann J. F., Akgün T., Pons J. A., Aloy M. A., Cerdá-Durán P., 2019, MNRAS, 490, 4858
Mereghetti S., Pons J. A., Melatos A., 2015, Space Sci. Rev., 191, 315
Parfrey K., Beloborodov A. M., Hui L., 2012, ApJ, 754, L12
Parfrey K., Beloborodov A. M., Hui L., 2013, ApJ, 774, 92
Perna R., Pons J. A., 2011, ApJ, 727, L51
Pili A. G., Bucciantini N., Del Zanna L., 2015, MNRAS, 447, 2821
Sakurai T., 1981, Sol. Phys., 69, 343
Thompson C., Duncan R. C., 1995, MNRAS, 275, 255
Thompson C., Duncan R. C., 1996, ApJ, 473, 322
Tiengo A. et al., 2013, Nature, 500, 312
Viganò D., Rea N., Pons J. A., Perna R., Aguilera D. N., Miralles J. A., 2013, MNRAS, 434, 123
Wiegelmann T., Sakurai T., 2021, Living Rev. Sol. Phys., 18, 1
Younes G. et al., 2022, ApJ, 924, L27

Downloaded from https://academic.oup.com/mnras/article/518/4/6390/6881733 by Universidad de Alicante user on 26 February 2024

APPENDIX A: POLOIDAL–TOROIDAL DECOMPOSITION

A commonly used practice to deal with the magnetic field in curvilinear coordinates is to decompose it in a poloidal and a toroidal component and express these components in terms of two scalar functions. Through the literature, various approaches can be found, regarding what the scalar functions represent, as well as notation for the quantities introduced. In this section, we go through the poloidal–toroidal decomposition equations in our notation (see Geppert & Wiebicke 1991 for a more detailed treatment of the subject).

The principle behind this process is to introduce two scalar functions $\Phi(\boldsymbol{r})$, $\Psi(\boldsymbol{r})$ and split the magnetic field to a purely poloidal and a purely toroidal part as follows:

$$\boldsymbol{B}_p = \nabla \times (-\boldsymbol{r} \times \nabla\Phi), \tag{A1}$$

$$\boldsymbol{B}_t = -\boldsymbol{r} \times \nabla\Psi, \tag{A2}$$

where $\boldsymbol{r} = r\hat{\boldsymbol{r}}$ is the position vector and we are working in spherical coordinates (r, θ, φ). From the definition of B_p one can readily deduce that the toroidal component of the vector potential is given by

$$\boldsymbol{A}_t = -\boldsymbol{r} \times \nabla\Phi. \tag{A3}$$

Any solenoidal field can be unambiguously represented from equations (A1) and (A2). Expanding equation (A1) with the help of vector calculus identities gives us

$$\boldsymbol{B}_p = -\boldsymbol{r}\nabla^2\Phi + \nabla\frac{\partial}{\partial r}(r\Phi). \tag{A4}$$

Toroidal fields are purely angular; they do not have a radial component. All the information about the radial magnetic field is incorporated in B_p. If we isolate only the radial component of equation (A4) we end up with an equation that completely describes the radial magnetic field

$$B_r = -r\nabla^2\Phi + r\frac{\partial^2}{\partial r^2}\Phi. \tag{A5}$$

One can see that the second term in the right-hand side in the above equation cancels with the term involving the radial derivative of the Laplacian. We are then left with

$$B_r = -r\nabla^2_{\rm ang}\Phi = -\nabla^2_{\rm ang}(r\Phi), \tag{A6}$$

where

$$\nabla^2_{ang} = \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2} \tag{A7}$$

is the angular part of the Laplacian.

APPENDIX B: ROTATING A VECTOR FIELD IN SPHERICAL COORDINATES

In this appendix, we explain the steps taken to perform a rotation of a vector field in spherical coordinates.

While a spherical geometry is ideal for simulating NS magnetospheres, spherical coordinates can often be cumbersome because of the non-constant nature of the unit vectors and the inherent divergence at the origin and at the axis. Tasks such as rotating a vector field are much easier to deal with in Cartesian coordinates.

The steps that need to be followed to rotate a vector field $\boldsymbol{B}$ by an angle β around the x-axis are as follows:

(i) Transform to Cartesian coordinates;
(ii) Rotate by an angle $-\beta$;
(iii) Transform back to spherical coordinates;
(iv) Express the vector field in the rotated coordinates;
(v) Transform the vector components to Cartesian;
(vi) Rotate by an angle β;
(vii) Transform the vector components back to spherical.

The spherical to Cartesian transformation is given by the relations

$$x = \tfrac{1}{q}\sin\theta\cos\phi \tag{B1}$$

$$y = \tfrac{1}{q}\sin\theta\sin\phi \tag{B2}$$

$$z = \tfrac{1}{q}\cos\theta. \tag{B3}$$

Rotation by an angle $-\beta$ around the x-axis is straightforward

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & \sin\beta \\ 0 & -\sin\beta & \cos\beta \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{B4}$$

Transformation back to spherical is given by the relations

$$q' = \left(x'^2 + y'^2 + z'^2\right)^{-\frac{1}{2}} \tag{B5}$$

$$\theta' = \arccos(z'q') \tag{B6}$$

$$\phi' = \arctan 2\,(y', x'), \tag{B7}$$

where arctan2 is the 2-argument arctangent that ensures that the correct and unambiguous value of ϕ is returned. By definition, arctan2 returns values in the interval $(-\pi, \pi]$. In order to map ϕ' to the desired interval $[0, 2\pi)$, we add 2π to all the negative values. Besides, one can easily see that $q' = q$, which is expected since the radius does not change with the rotation.

To rotate any scalar field, it is enough to express it in these new, rotated coordinates. However, for vector fields a rotation of the components is also needed. First, the vector field $\boldsymbol{B}$ is expressed in the rotated coordinates such that $B'_q = B_q(q', \theta', \phi')$, $B'_\theta = B_\theta(q', \theta', \phi')$, $B'_\phi = B_\phi(q', \theta', \phi')$. Then, we map the vector field to the Cartesian unit vectors

$$\begin{pmatrix} B'_x \\ B'_y \\ B'_z \end{pmatrix} = \begin{pmatrix} \sin\theta'\cos\phi' & \cos\theta'\cos\phi' & -\sin\phi' \\ \sin\theta'\sin\phi' & \cos\theta'\sin\phi' & \cos\phi' \\ \cos\theta' & -\sin\theta' & 0 \end{pmatrix}\begin{pmatrix} B'_q \\ B'_\theta \\ B'_\phi \end{pmatrix}, \tag{B8}$$

and perform a rotation of angle β

$$\begin{pmatrix} \tilde{B}_x \\ \tilde{B}_y \\ \tilde{B}_z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & \sin\beta \\ 0 & -\sin\beta & \cos\beta \end{pmatrix}\begin{pmatrix} B'_x \\ B'_y \\ B'_z \end{pmatrix}. \tag{B9}$$

The last step is to transform the rotated components that have been calculated in the rotated coordinate system back to spherical

$$\begin{pmatrix} \tilde{B}_q \\ \tilde{B}_\theta \\ \tilde{B}_\phi \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \end{pmatrix}\begin{pmatrix} \tilde{B}_x \\ \tilde{B}_y \\ \tilde{B}_z \end{pmatrix}. \tag{B10}$$

This paper has been typeset from a TEX/LATEX file prepared by the author.

Downloaded from https://academic.oup.com/mnras/article/518/4/6390/6881733 by Universidad de Alicante user on 26 February 2024

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY
MNRAS **524,** 32–42 (2023)
Advance Access publication 2023 June 16
https://doi.org/10.1093/mnras/stad1810

Modelling force-free neutron star magnetospheres using physics-informed neural networks

Jorge F. Urbán,[1] Petros Stefanou,[1,2★] Clara Dehman[3,4] and José A. Pons[1]

[1]*Departament de Física Aplicada, Universitat d'Alacant, Ap. Correus 99, E-03080 Alacant, Spain*
[2]*Departament d'Astronomia i Astrofísica, Universitat de Valéncia, Dr Moliner 50, E-46100, Burjassot, Valéncia, Spain*
[3]*Institute of Space Sciences (ICE-CSIC), Campus UAB, Carrer de Can Magrans s/n, E-08193, Barcelona, Spain*
[4]*Institut d'Estudis Espacials de Catalunya (IEEC), Carrer Gran Capitá 2–4, E-08034 Barcelona, Spain*

Accepted 2023 June 12. Received 2023 May 31; in original form 2023 March 18

ABSTRACT

Using physics-informed neural networks (PINNs) to solve a specific boundary value problem is becoming more popular as an alternative to traditional methods. However, depending on the specific problem, they could be computationally expensive and potentially less accurate. The functionality of PINNs for real-world physical problems can significantly improve if they become more flexible and adaptable. To address this, our work explores the idea of training a PINN for general boundary conditions and source terms expressed through a limited number of coefficients, introduced as additional inputs in the network. Although this process increases the dimensionality and is computationally costly, using the trained network to evaluate new general solutions is much faster. Our results indicate that PINN solutions are relatively accurate, reliable, and well behaved. We applied this idea to the astrophysical scenario of the magnetic field evolution in the interior of a neutron star connected to a force-free magnetosphere. Solving this problem through a global simulation in the entire domain is expensive due to the elliptic solver's needs for the exterior solution. The computational cost with a PINN was more than an order of magnitude lower than the similar case solved with a finite difference scheme, arguably at the cost of accuracy. These results pave the way for the future extension to three-dimensional of this (or a similar) problem, where generalized boundary conditions are very costly to implement.

Key words: magnetic fields – methods: numerical – stars: magnetars – stars: neutron – neural networks – physics-informed neural networks.

1 INTRODUCTION

Deep learning (DL) is a subset of techniques comprehended in machine learning that is fundamentally based on multilayered neural networks (NNs). In recent years, DL has been widely used to perform a large variety of tasks. Examples include (among many others) computer vision (to perform image classification; Traore, Kamsu-Foguem & Tangara 2018), face recognition (Lawrence et al. 1997) or medical diagnosis (Kugunavar & Prabhakar 2021), speech recognition (Chan et al. 2015), and robotics to emulate human-like walking and running or mobile navigation in pedestrian environments (Hayat & Mall 2013).

Physics-informed neural network (PINN; Raissi, Perdikaris & Karniadakis 2019) is a DL approach used to numerically approximate the solution of non-linear partial differential equations (PDEs). The original idea was born more than 20 yr ago (Lagaris, Likas & Fotiadis 1997), but the lack of the necessary computational resources made it complicated to put it into practice. In recent years, we account with graph-based automatic differentiation, as well as different frameworks that support computations in CPUs and GPUs, such as Tensorflow or Pytorch, and a dramatic increase in computational power. These factors, combined with a blooming interest in machine-learning applications in science, have given birth to this promising new field. PINNs have been used, among many other applications, in fluid dynamics (Cai et al. 2021), nuclear reactor dynamics (Schiassi et al. 2022), radiative transfer (Mishra & Molinaro 2021; Chen et al. 2022), and black-hole spectroscopy (Luna et al. 2022). PINNs incorporate the underlying physical laws that govern a system (the PDEs) and then optimize the NN so that the residual of the PDE is minimal. Unlike more traditional DL approaches in other fields, PINNs do not require large amounts of data – or any data at all – for the training of the NN.

Compared to classical finite-differences/finite-elements methods, PINNs still fall short in terms of efficiency and precision. However, they present some advantages as flexible, multipurpose PDE solvers. For example, the PINN formulation allows us to solve problems in arbitrary, unstructured meshes without using high-resolution, memory-consuming grids. In addition, once a PINN is trained for a general problem, the calculation of a new solution is swift and consists only of the few operations needed during the forward pass through the network. This is a potential advantage in speed compared to classical methods.

In this work, we assess the applicability of a PINN solver in elliptic problems. In particular, we focus on the problem of modelling force-free (FF) magnetospheres of neutron stars (NS) in the non-rotating, axisymmetric limit. There is a wealth of work in the literature formulating this problem in terms of the GS equation (Glampedakis,

★ E-mail: petros.stefanou@uv.es

© 2023 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/524/1/32/7199782 by Universidad de Alicante user on 26 February 2024

Lander & Andersson 2014; Pili, Bucciantini & Del Zanna 2015; Akgün et al. 2016; Akgün et al. 2017; Akgün et al. 2017; Kojima 2017; Akgün et al. 2018), which gives us the opportunity to make detailed comparisons and draw robust conclusions on the performance and generalizability of the PINN solver. In particular, we have compared our results with our previous work (Akgün et al. 2016), where a finite-difference scheme based on an iterative method solving a block-tridiagonal system was used.

The paper is organized as follows: in Section 2, we give a brief mathematical overview of the physics of NS magnetospheres. In Section 3, we describe in detail how the PINN solver is built. We present the solutions acquired by the PINN solver with error estimates in Section 4. In Section 5, we demonstrate PINN's capabilities through an astrophysical application. Section 6 is dedicated to discuss our main results and how to improve and generalize our approach to face more difficult problems in the near future.

2 MODELLING AXISYMMETRIC FF MAGNETOSPHERES

In the FF regime, and considering that the contribution of gravity, inertia, plasma pressure, and rotation in the dynamics of an NS magnetosphere is negligible compared to the magnetic force, the force-balance equation reduces to the simple form

$$(\nabla \times \boldsymbol{B}) \times \boldsymbol{B} = 0, \tag{1}$$

where $\boldsymbol{B}$ denotes the magnetic field. This regime is better suited for magnetar magnetospheres (Thompson & Duncan 1995, 1996), because magnetars are slow rotators and the absence of any rotationally induced electric fields is a very good approximation.

In axisymmetry, we can express the magnetic field in terms of a poloidal and a toroidal stream function $\mathcal{P}$ and $\mathcal{T}$. We will follow the notation and formalism as in Akgün et al. (2016). We refer the interested reader to that work for a more detailed description. In spherical coordinates (r, θ, ϕ) and in terms of the stream functions, the magnetic field reads

$$\boldsymbol{B} = \nabla\mathcal{P} \times \nabla\phi + \mathcal{T}\nabla\phi, \tag{2}$$

where $\nabla\phi = \frac{\boldsymbol{e}_\phi}{r\sin\theta}$ with $\boldsymbol{e}_\phi$ being the azimuthal unit vector. Substituting equations (2) into (1), the ϕ-component of the equation gives

$$\nabla\mathcal{P} \times \nabla\mathcal{T} = 0, \tag{3}$$

which simply states that $\mathcal{T} = \mathcal{T}(\mathcal{P})$ must be a function of $\mathcal{P}$ (or vice versa). The remaining components give us the so-called GS equation

$$\triangle_{\rm GS}\mathcal{P} + G(\mathcal{P}) = 0. \tag{4}$$

Here, $G(\mathcal{P}) = \mathcal{T}(\mathcal{P})\frac{d\mathcal{T}}{d\mathcal{P}}$ is the source term accounting for the presence of currents in the magnetosphere and $\triangle_{\rm GS}$ is the GS operator

$$\triangle_{\rm GS} \equiv r^2\sin^2\theta\,\nabla\cdot\left(\frac{1}{r^2\sin^2\theta}\nabla\right). \tag{5}$$

For convenience, we will use compactified spherical coordinates (see Stefanou, Pons & Cerdá-Durán 2023) (q, μ, ϕ), where $q = \frac{1}{r}$ and $\mu = \cos\theta$ instead of the usual (r, θ, ϕ). In this set of coordinates, the GS operator reads

$$\triangle_{\rm GS} \equiv q^2\partial_q\left(q^2\partial_q\right) + \left(1-\mu^2\right)q^2\partial_{\mu\mu}. \tag{6}$$

To solve equation (4), we must also provide BCs and the functional form of the source term, that is, $\mathcal{T}(\mathcal{P})$. The particular functional form is arbitrary and different choices are possible. In Akgün et al. (2016), they used

$$\mathcal{T}(\mathcal{P}) = s\,(\mathcal{P}-\mathcal{P}_c)^\sigma\,\Theta\,(\mathcal{P}-\mathcal{P}_c), \tag{7}$$

where Θ is the Heaviside function, and s, $\mathcal{P}_c$, and σ are parameters that control the relative strength of the toroidal and poloidal components, the region where the toroidal field is non-zero and the non-linearity of the model. However, this has the limitation that it assumes $\mathcal{P} > 0$. A possible generalization overcoming this constraint is

$$\mathcal{T}(\mathcal{P}) = s\,(|\mathcal{P}|-\mathcal{P}_c)^\sigma\,\Theta\,(|\mathcal{P}|-\mathcal{P}_c), \tag{8}$$

which allows for currents even for negative values of $\mathcal{P}$. We have explored different options but, for simplicity and the purpose of this paper, we will use a quadratic function for the astrophysical application in Section 5.2, defined as follows:

$$\mathcal{T}(\mathcal{P}) = s_1\mathcal{P} + s_2\mathcal{P}^2. \tag{9}$$

We impose BCs for $\mathcal{P}$ at the surface of the star ($q = 1$), at radial infinity ($q = 0$) and at the axis ($\mu = \pm 1$). Regularity and symmetry of the problem lead to

$$\mathcal{P}(q, \mu = \pm 1) = \mathcal{P}(q = 0, \mu) = 0.$$

In particular, one advantage of compactifying the radial coordinate (going from r to q) is to make it easier to impose BCs at radial infinity: rather than imposing a specific decay rate at large r, we can impose Dirichlet BCs at just one point ($q = 0$). This reduces unwanted numerical noise from the external boundary.

At the surface, we must provide the function $\mathcal{P}(\mu)$. Our implementation of BCs in an NN must be as general as possible but keep the number of parameters reasonably low. A reasonable and practical choice is to use some decomposition of the arbitrary function in terms of orthonormal polynomials. Considering the symmetry of our problem and that we are working with functions describing magnetic fields, the natural choice is to express $\mathcal{P}(q = 1)$ in terms of coefficients of a Legendre polynomial expansion. We use the following decomposition:

$$\mathcal{P}(q = 1, \mu) = \left(1-\mu^2\right)\sum_{l=1}^{l_{\max}}\frac{b_l}{l}P_l'(\mu), \tag{10}$$

where l is the order of the multipole ($l = 1$ corresponds to a dipole), P_l are the Legendre polynomials (not to be confused with $\mathcal{P}$, the poloidal flux function), and the prime denotes differentiation with respect to μ. Thus, the BC at the surface is completely determined by prescribing the b_l coefficients[1].

3 METHODOLOGY

3.1 Neural networks

NNs are universal approximators of mathematical functions (Hornik, Stinchcombe & White 1989). They are the result of compositions of simple but non-linear transformations at different layers. The way that these layers are interconnected indicates the NN *architecture*. Each layer contains a number of *neurons* that transform the inputs received from the previous layer and then pass the result to the next layer. This transformation is done in two basic steps. A linear combination of the inputs received and an evaluation through a non-linear *activation function*. Mathematically, this process can be expressed for each layer as follows:

$$\boldsymbol{a}^j = g(\mathbf{W}^j\boldsymbol{a}^{j-1} + \boldsymbol{b}^j), \tag{11}$$

[1]The $1/l$ normalization factor and the coefficients b_l in the expansion have been chosen to match the b_l coefficients used in Dehman et al. (2023).

Downloaded from https://academic.oup.com/mnras/article/524/1/32/7199782 by Universidad de Alicante user on 26 February 2024

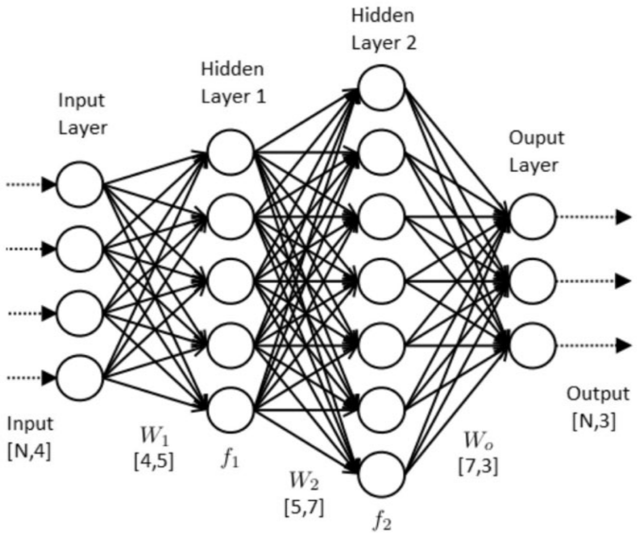


Figure 1. A schematic representation of a deep fully connected NN.

where $\boldsymbol{a}^j$ is a vector containing the values of the neurons in layer j, $\mathbf{W}^j$ and $\boldsymbol{b}^j$ denote the *weight* matrix and the *bias* vector of the layer, and g is the non-linear activation function. The most commonly used type of architecture is the fully connected neural network (FCNN). Each neuron is connected to all the neurons of the previous layer and connects to all the neurons of the next layer. The first layer of an FCNN is called the *input layer*, where user-provided information (the *inputs*) enter the network. The last layer is called the *output layer*. This is the final result (the *output*) of the network. Fig. 1 shows a schematic representation of an FCNN. The set Ω of the weights and biases of all the layers constitute the trainable parameters of an NN. They can be adjusted so that the relation between the input and the output approximates a desired function u. The process of adjusting the trainable parameters is called *training* and involves the following steps:

(i) The network is initialized with a random set of weights and biases.

(ii) A number of inputs $\boldsymbol{x}$ is fed to the network. The network propagates the inputs through its layers, transforms them according to equation (11) and produces some outputs $\tilde{u}$ (the *prediction* of the network), which depend on the inputs and on the trainable parameters, so that $\tilde{u} = \tilde{u}(\boldsymbol{x}; \Omega)$. This is called a *forward pass*.

(iii) The difference between prediction and desired function is calculated. This is done through a *loss function*, which is an average global measure of the deviation of $\tilde{u}(\boldsymbol{x}; \Omega)$ from $u(\boldsymbol{x})$.

(iv) The value of the loss function depends on all the trainable parameters. Therefore, one can adjust these parameters so that the loss is minimized and $\tilde{u}(\boldsymbol{x}; \Omega)$ approximates $u(\boldsymbol{x})$. This is done by applying small corrections to each weight and bias, dictated by their contribution to the value of the loss. This process is called *backpropagation*.

(v) The backpropagation is performed using *gradient descent* based techniques, i.e. by calculating the gradients of the loss function with respect to the trainable parameters and then updating the trainable parameters by applying corrections to them that are proportional to those gradients.

(vi) All these gradient calculations would be very costly if they were to be calculated numerically. However, in modern machine-learning frameworks, all operations during a forward pass are recorded in *graphs*. This allows gradients to be calculated very efficiently and to machine accuracy by inverting the recorded operations, a technique called *Automatic Differentiation* (AD) (see Baydin et al. (2017) for a detailed description of AD in machine learning).

(vii) These steps are repeated iteratively in order to minimize the value of the loss. When convergence is achieved, the network's prediction should be a good approximation to the desired function.

The above procedure can be implemented using Tensorflow (Abadi et al. 2016), a freely available software framework for machine learning. It provides us with the tools and resources needed to develop and implement various machine-learning algorithms and models. Particularly, AD is implemented by the application programming interface (API) *GradientTape*,[2] which automatically records all the operations made during the forward pass and generates the corresponding graph used for calculating the gradients. On the other hand, different gradient descent optimization algorithms are implemented in the API Keras.[3] In this work, we used the ADAM algorithm (Kingma & Ba 2014) for all the calculations.

A slightly different version of equation (11), where any hidden layer is only interconnected with its closest neighbours, are the so-called *residual blocks*. A residual block is formed if, before the evaluation of the activation function at a certain layer j, we add the output of the K_{th} previous layer (a *skip connection*). Mathematically, the output $\boldsymbol{a}^j$ of a residual block corresponding to the layer j could be written as follows

$$\boldsymbol{a}^j = g(\mathbf{W}^j \boldsymbol{a}^{j-1} + \boldsymbol{b}^j + \boldsymbol{a}^{j-K}). \tag{12}$$

Stacking many of these blocks together forms a *Residual Neural Network* (ResNet), introduced first by He et al. (2015). ResNets were introduced in DL to address the problem in which, by increasing the depth of the NN, the gradients used to update the trainable parameters tend to diminish during backpropagation, making it difficult for the network to learn effectively (see He et al. (2016) for a detailed explanation). As ResNets are not explored in the literature of PINNs and are straightforward to implement by extending an FCNN, we have additionally considered this architecture.

3.2 Physics-informed neural networks

The standard way of training an NN accounts with data that consist of values of the true solution in a given discrete set of points in the input domain. However, this approach demands a large number of training examples to build a reliable relation between the inputs and the outputs. In particular, for astrophysical systems, this method would rely on data obtained through a large set of observations. The key novelty in PINNs is the incorporation of information about the physical laws into the training process. This is accomplished by minimizing the residual of the PDEs that govern the system, instead of the difference between prediction and real data/exact solution. Using data in the loss function is optional and sometimes may facilitate the optimization, but in practise is not necessary.

Consider that the function u that we aim to obtain with the PINN satisfies the following boundary value problem

$$\mathcal{L}u(\boldsymbol{x}) - G(\boldsymbol{x}, u(\boldsymbol{x})) = 0, \tag{13}$$

$$u|_{\partial\mathcal{D}} = f_b(\boldsymbol{x})|_{\boldsymbol{x}\in\partial\mathcal{D}}, \tag{14}$$

[2] https://www.tensorflow.org/api_docs/python/tf/GradientTape
[3] https://keras.io/api/

Downloaded from https://academic.oup.com/mnras/article/524/1/32/7199782 by Universidad de Alicante user on 26 February 2024

where $\mathcal{L}$ is a general non-linear differential operator, G is a source term, and $\boldsymbol{x}$ is a vector of coordinates in some domain $\mathcal{D}$. The PDE is subject to some BCs f_b at the boundary $\partial\mathcal{D}$ of the domain. If $\tilde{u}$ is an approximation to an exact solution given by a PINN, then equation (13) will have a residual, that is, the right-hand side will not be exactly zero. The smaller this residual is, the closer $\tilde{u}$ be to an exact solution. Thus, the loss function is precisely a suitable norm of this residual, defined as

$$\mathcal{J} = ||\mathcal{L}\tilde{u}(\boldsymbol{x};\Omega) - G(\boldsymbol{x}, \tilde{u}(\boldsymbol{x};\Omega))||, \tag{15}$$

where the derivatives of the function $\tilde{u}$ with respect to the coordinates $\boldsymbol{x}$ are also computed with AD (with the Tensorflow *GradientTape* API).

The function $\tilde{u}$ should also satisfy the BCs (14). There are different approaches to implement them. The most commonly used is to add a term in the loss function consisting of a norm of the residual of equation (14), so that the residuals of both equations (13) and (14) are minimized simultaneously, as follows:

$$\mathcal{J} = c_1||\mathcal{L}\tilde{u}(\boldsymbol{x};\Omega) - G(\boldsymbol{x}, \tilde{u}(\boldsymbol{x};\Omega))|| + c_2||\tilde{u}(\boldsymbol{x};\Omega) - f_b(\boldsymbol{x})||_{\boldsymbol{x}\in\partial\mathcal{D}}, \tag{16}$$

where c_1 and c_2 are coefficients that control the relative weight of the two terms. Then, the solution of the boundary value problem is directly the output $\mathcal{N}$ of the PINN

$$\tilde{u}(\boldsymbol{x};\Omega) = \mathcal{N}(\boldsymbol{x};\Omega). \tag{17}$$

We believe that this is not the optimal way to impose BCs. The individual terms in equation (16) can differ significantly, which means that minimizing $\mathcal{J}$ does not guarantee that both the PDE and the BCs are satisfied with the same accuracy. In order to surpass this problem, the coefficients c_1, c_2 can be adjusted so that the relative contribution of the two terms is of the same order. This can be done, for example, by arbitrarily choosing these coefficients and adjusting them through a trial-and-error process or by calculating the *neural tangent kernel* of the network (Wang, Yu & Perdikaris 2022). We find that, overall, the need for fine tuning additional hyperparameters in this method is an inconvenience.

Instead, we opt for an approach inspired by Lagaris et al. (1997) (see also a similar one, based on distance functions in Sukumar & Srivastava 2022). In that approach, the loss function is given only by equation (15), but the approximate solution of the boundary value problem is written as

$$\tilde{u}(\boldsymbol{x};\Omega) = f_b(\boldsymbol{x}) + h_b(\boldsymbol{x})\mathcal{N}(\boldsymbol{x};\Omega), \tag{18}$$

where f_b is a smooth and (at least) twice differentiable function that satisfies the BCs (see equation 14), h_b is a smooth and twice differentiable function that defines the boundary ($h_b = 0$ at $\partial\mathcal{D}$ and $h_b \neq 0$ in $\mathcal{D}$), and $\mathcal{N}$ is the output of the network. This redefinition ensures that $\tilde{u}$ matches exactly the BCs at $\partial\mathcal{D}$ regardless of the value of $\mathcal{N}$. In the rest of the domain, $\tilde{u}$ should satisfy the PDE. The PINN takes care of this by adapting $\mathcal{N}$ during training. We stress again that in this approach $\mathcal{N}$ by itself does not satisfy the PDE (in contrast to equation 17), but the combination of f_b, h_b, and $\mathcal{N}$ does. The choice of f_b and h_b is not unique: variations of this functions can lead to slight changes in the convergence rate or the final value of the loss function (see the examples in Section 4.2.1 for a more quantitative description).

In this work, we attempt to generalize the PINN approach to build a PDE solver valid for different and varied BCs (and possibly, source terms). We want our network to learn how to approximate any particular solution for a given operator $\mathcal{L}$. This means that the information about the BCs should be part of the *input* of the network (along with the coordinates) and not hardcoded in the loss function or in the parametrization (18). During training, the network needs to process a large number of points $\boldsymbol{x}$ and a large number of f_b functions so that it can generalize and provide solutions of (13) for any point in the domain and any BC. Of course, f_b could, in principle, be an arbitrary continuous function. For this reason, it should be intelligently encoded into the network's input to keep the number of parameters small and manageable.

In the following section, we present tests for our PINN solver for the GS equation as described in Section 2. We omit hereinafter to state explicitly the dependence of all the functions on the training variables Ω.

4 TEST AND MODELS

4.1 Current-free GS equation

As a first test, we consider the GS equation (equation 4) without current, that is $G(\mathcal{P}) = 0$. We set up the various elements of the PINN solver as follows. The function h_b describing the boundary is given by

$$h_b(q, \mu) = q(1-q)(1-\mu^2). \tag{19}$$

The function f_b that satisfies the BCs, where $h_b = 0$, is given by

$$f_b(q, \mu) = q^n\left(1-\mu^2\right)\sum_{l=1}^{l_{\max}}\frac{b_l}{l}P_l'(\mu). \tag{20}$$

Notice that equation (20) differs from equation (10) by a factor q^n with $n > 0$. We include this factor to enforce BCs both at the surface ($q = 1$) and infinity ($q = 0$). If $n = 1$, this parametrization is the same as that used in Lagaris et al. (1997) considering essential BCs in a rectangle. However, we prefer to leave n as a free parameter that is used to give more or less weight to the solution close to the star or away from the surface. We performed a detailed study of the influence of the hyperparameters of the model, including n, in the following section. We must remark that other parametrizations are possible and in principle can be tuned to improve the results of each specific problem.

The input layer of the NN consists of the coordinates of the point where the solution will be evaluated (q, μ) and the coefficients b_l determining the BC at $q = 1$. For the physical applications in this paper, we expect the dipole ($l = 1$) component of the magnetic field to be dominant. Therefore, we normalize all other multipole coefficients by dividing them by b_1 and we reduce the range of the possible values of $b_{l>1}$ between -1 and 1. With this choice, we can omit the b_1 coefficient because it is reabsorbed in the normalization factor of the magnetic field strength.

Hereafter, we limit ourselves to $l_{\max} = 7$, which suffices for our purposes and is a good compromise between generality of solutions and ease of training. Increasing the number of multipoles adds complexity and it would require larger networks to achieve the desired accuracy. Thus, one training point is defined as

$$(q, \mu, b_2, b_3, b_4, b_5, b_6, b_7),$$

In each forward pass providing $\mathcal{N}(q, \mu, \{b_l\})$, the solution of the differential equation at each training point i is given by

$$\mathcal{P}_i = \left(1-\mu^2\right)\left[\sum_{l=1}^{l_{\max}}\frac{b_l}{l}q^n P_l'(\mu) + q(1-q)\mathcal{N}(q, \mu, \{b_l\})\right]. \tag{21}$$

Downloaded from https://academic.oup.com/mnras/article/524/1/32/7199782 by Universidad de Alicante user on 26 February 2024

Notice that the output of the network $\mathcal{N}$ depends both on the coordinates and on the BCs. This is the crux of our approach, as we want the network to be able to generalize for any BC (expressed in terms of the b_l coefficients).

The weights and biases of the network are optimized by minimizing the loss function averaged over a large sample of training points $i_{\max}$

$$\mathcal{J} = \frac{1}{i_{\max}} \sum_{i=1}^{i_{\max}} [\triangle_{\rm GS} \mathcal{P}_i]^2. \tag{22}$$

We consider training sets of size $i_{\max} = 10^4$. At first sight, this number might look large. However, we must stress that it is required by the high dimensionality of our problem. Our parameter space consists of two coordinates plus six coefficients to describe the BC (fixing $b_1 = 1$). Covering this eight-dimensional parameter space with only three points in each dimension would need $3^8 = 6561$ points, which makes evident the crucial difference between training to solve a PDE with fixed BCs or training with arbitrary BCs (formally, an infinite number of additional parameters). The training set is changed periodically every few thousand epochs in order to feed the network with as many points as possible.

Fig. 2(a) shows the evolution of the loss function (22) with the number of training epochs. For this particular model, we have chosen an FCNN architecture with 4 hidden layers and 80 neurons at each layer with $n = 5$ in f_b (see the next section for details on the choice of these hyperparameters). The activation function g (see equation 11) chosen for all the hidden layers is the tanh function. The periodic spikes correspond to renewals of the training set. They can be understood as a measure of the ability of a model to generalize to new, unseen points. Fig. 2(b) shows an example of the final result once the PINN has been trained. The black solid lines show the *exact* analytical solution which is uniquely determined by the b_l coefficients

$$\mathcal{P}_{\rm ex}(q, \mu) = \left(1 - \mu^2\right) \sum_{l=1}^{l_{\max}} \frac{b_l}{l} q^l P_l'(\mu), \tag{23}$$

while the yellow dashed lines show the solution acquired by the PINN ($\mathcal{P}$). They are indistinguishable at the figure scale. The colourmap indicates the relative difference between $\mathcal{P}$ and $\mathcal{P}_{\rm ex}$, at most ~0.5 per cent for this particular model.

The magnetic field components can be computed from $\mathcal{P}$ via AD. From equation (2), we have

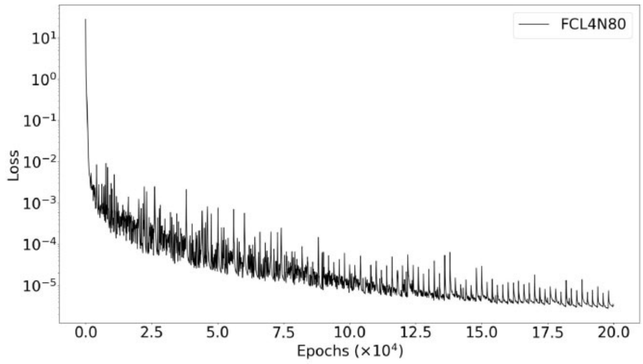
$$B_r = -q^2 \frac{\partial \mathcal{P}}{\partial \mu}, \tag{24}$$

$$B_\theta = \frac{q^3}{\sqrt{1-\mu^2}} \frac{\partial \mathcal{P}}{\partial q}. \tag{25}$$

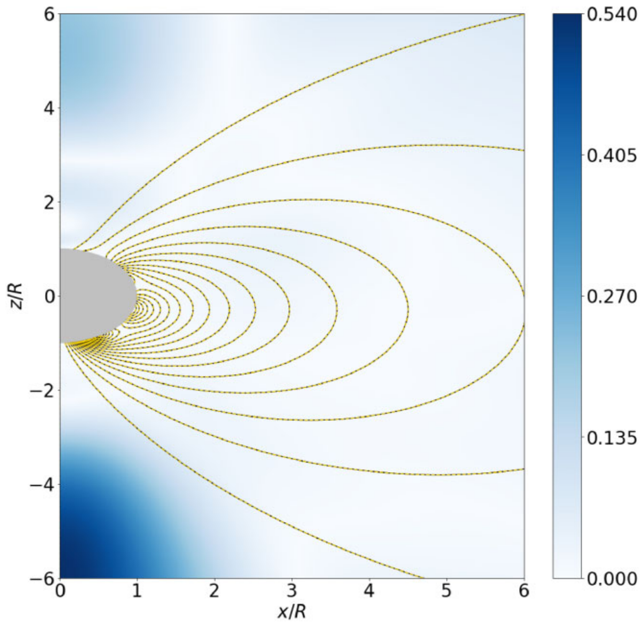
In finite difference schemes, one usually losses accuracy when taking numerical derivatives. To explore the performance of the PINN in this respect, we have computed different relative error norms of different orders (p) for $\mathcal{P}$, B_r, B_θ and for the magnetic field modulus $B = \sqrt{B_r^2 + B_\theta^2}$. Results are summarized in Table 1. These p-norms for a given order p are calculated for every variable as in Eivazi et al. (2022)

$$E_u = \frac{\|\tilde{u}(\boldsymbol{x}) - u(\boldsymbol{x})\|_p}{\|u(\boldsymbol{x})\|_p} \times 100, \tag{26}$$

Interestingly, the errors are of the same order of magnitude for the function $\mathcal{P}$ and its derivatives. This can be attributed to the fact that we train the PINN with a second-order PDE (the loss function involves second-order derivatives) and this includes additional information on the derivatives. Furthermore, using automatic differentiation is also an advantage over finite difference schemes, where accuracy depends on the resolution.



(a)



(b)

Figure 2. (a) Evolution of the loss function with the training epochs. The periodic spikes correspond to renewals of the training set. (b) Colourmap of the relative error (percentage) between $\mathcal{P}$ and $\mathcal{P}_{\rm ex}$. Yellow dashed lines correspond to contours of $\mathcal{P}$ while black solid lines correspond to $\mathcal{P}_{\rm ex}$. The multipole coefficients in f_b for this particular example are $b_1 = 1$, $b_{l \geq 2} = (-1)^{l+1} 0.6$.

Table 1. Relative error norms (percentage) between PINN and the exact solution. The numbers shown are averages of the respective norms of 100 new sets, consisting of 10^4 random points each. For each set, we compute the norms using equation (26).

	$E_{\mathcal{P}}$	E_{B_r}	E_{B_θ}	E_B
L_1 norm	0.017	0.017	0.025	0.015
L_2 norm	0.019	0.023	0.045	0.023

Downloaded from https://academic.oup.com/mnras/article/524/1/32/7199782 by Universidad de Alicante user on 26 February 2024

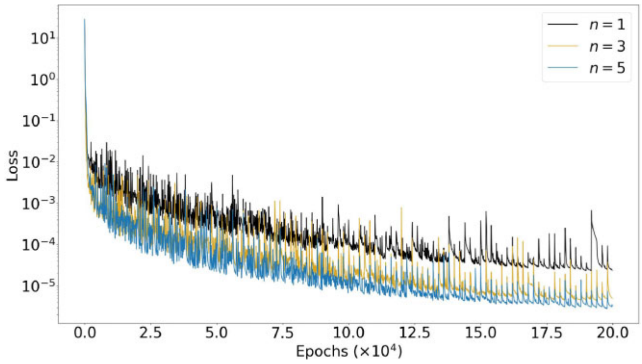


Figure 3. Evolution of the loss function with the training epochs for different values of the exponent n in f_b. The rest of the hyperparameters are as in Section 4.1.

Table 2. Relative error L_2 *norms* for different values of n. The results indicate a higher accuracy of the overall solution as n increases.

		$p=2$		
n	$E_{\mathcal{P}}$	E_{B_r}	E_{B_θ}	E_B
1	0.057	0.100	0.131	0.084
3	0.030	0.071	0.117	0.064
5	0.019	0.023	0.045	0.023

4.2 Influence of the PINN hyperparameters

We have performed a detailed study to measure the influence of various hyperparameters of our model. In particular, we have considered the following:

(i) Changes of the parametrization of the boundary (q^n power).
(ii) Number of neurons at each layer.
(iii) Number of hidden layers.
(iv) Resnet versus FC architectures.

We changed one hyperparameter at a time while keeping the rest fixed to the reference values of the previous section. The results of this study are presented separately in the following subsections.

4.2.1 Changes of the parametrization of the BCs

We begin by considering different values of the exponent in the q^n term in the boundary function (20). Fig. 3 shows the evolution of the loss with the training epochs for three different values of the exponent n, namely 1 (corresponding to the Lagaris parametrization), 3 and 5. As n increases, the impact of the surface BC becomes less important. In general, increasing n improves the convergence of the model and leads to more accurate solutions. Table 2 shows the relative error L_2 norms for the four quantities that we use to evaluate our results ($\mathcal{P}$, B_r, B_θ, B). All of them decrease with increasing n.

4.2.2 Number of neurons per hidden layer

Next, we explore the effect of the number of neurons per hidden layer N. Fig. 4 shows the evolution of the loss with the training epochs for $N = 20$, 40, 80. The number of neurons has a considerable impact on the convergence of each model. This is expected, because models with smaller N do not have enough free parameters to account for the complexity and variability of the solutions. Our results show that the loss reaches values that are smaller by a factor of 33 when doubling N from 20 to 40 and by a factor of 5 when doubling from 40 to 80. This is reflected, as well, in Table 3, where all quantities show a significant improvement in accuracy as N increases.

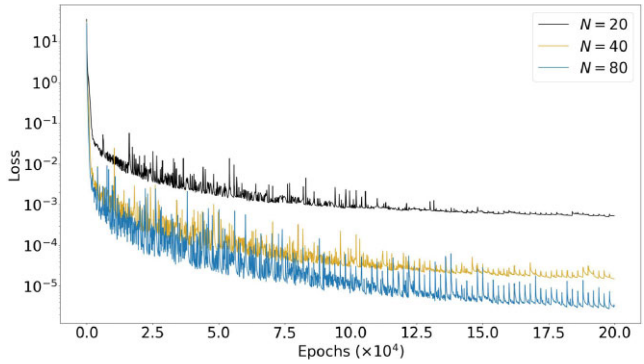


Figure 4. Evolution of the loss function with the training epochs for different values of the number of neurons per hidden layer N. The rest of the hyperparameters are as in Section 4.1.

Table 3. Relative error L_2 *norms* for different values of N. The results indicate a higher accuracy of the solution as N increases.

		$p=2$		
N	$E_{\mathcal{P}}$	E_{B_r}	E_{B_θ}	E_B
20	0.316	0.188	0.285	0.171
40	0.039	0.033	0.052	0.031
80	0.019	0.023	0.045	0.023

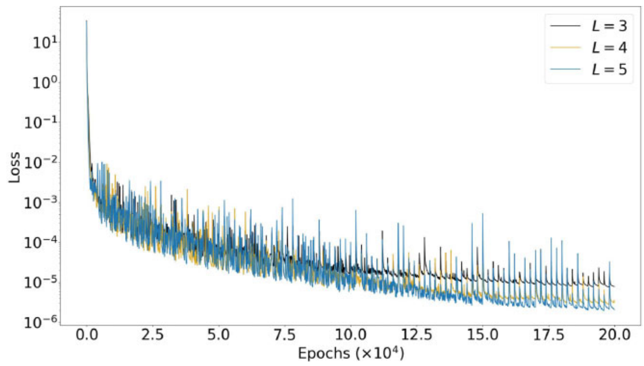


Figure 5. Evolution of the loss function with the training epochs for different values of the number of hidden layers L. The rest of the hyperparameters are as in Section 4.1.

4.2.3 Number of hidden layers

Following the same line of arguments, one could expect that increasing the number of hidden layers L would also lead to improved convergence and higher accuracy. However, our results show that adding more layers has a marginal impact on convergence and accuracy, or it can even lead to worse results for large networks. In other words, deeper networks are more prone to overfitting. Evidence of overfitting can be seen in Fig. 5 for $L = 5$. The spikes that correspond to renewals of the set of training points are much higher than expected, even at the later stages of training. This is, indeed, reflected in Table 4, where the model with $L = 5$ performs worse in terms of accuracy than the model with $L = 4$ because

Downloaded from https://academic.oup.com/mnras/article/524/1/32/7199782 by Universidad de Alicante user on 26 February 2024

Table 4. Relative error L_2 norms for different values of L. The results indicate that adding more layers does not improve the accuracy significantly and can lead to overfitting.

		$p = 2$		
L	$E_{\mathcal{P}}$	E_{B_r}	E_{B_θ}	E_B
3	0.030	0.029	0.050	0.028
4	0.019	0.023	0.045	0.023
5	0.014	0.025	0.064	0.027

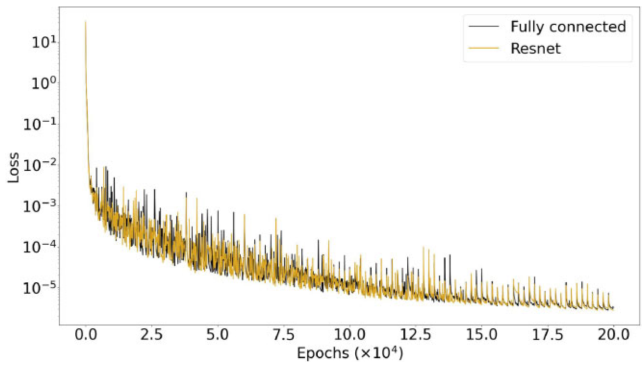


Figure 6. Evolution of the loss function with the training epochs for different NN architectures. The rest of the hyperparameters are as in Section 4.1.

Table 5. Relative error L_2 *norms* for different NN architectures. The results indicate that there is no appreciable difference for the two architectures considered.

		$p = 2$		
Model	$E_{\mathcal{P}}$	E_{B_r}	E_{B_θ}	E_B
FCL4N80	0.019	0.023	0.045	0.023
ResL4N80	0.018	0.019	0.034	0.018

it is overfitted to the training set and fails to generalize to unseen points. Considering, in addition, that training deeper networks is computationally expensive, we conclude that increasing too much the number of layers is not beneficial in terms of accuracy or convergence.

4.2.4 *Resnet versus fully connected*

Lastly, we consider two different types of NN architectures, FC architecture and ResNet architecture. Results are summarized in Fig. 6 and Table 5. No appreciable differences can be detected between the two models in convergence or overall accuracy.

4.3 FF GS equation

Once we have assessed the performance of our approach in the vacuum case by comparing our results with the analytical solutions, we now turn to the general case ($G(\mathcal{P}) \neq 0$).

The configuration of the PINN solver for this case is similar to the one described in the current-free case. f_b, h_b, n, $l_{\max}$, $i_{\max}$, N, L, architecture, number of epochs and optimizer are the same as in Section 4.1. There are two main differences: (a) the loss function and (b) the input. The loss function now includes the non-zero source term $G(\mathcal{P})$ which accounts for the presence of currents and is given

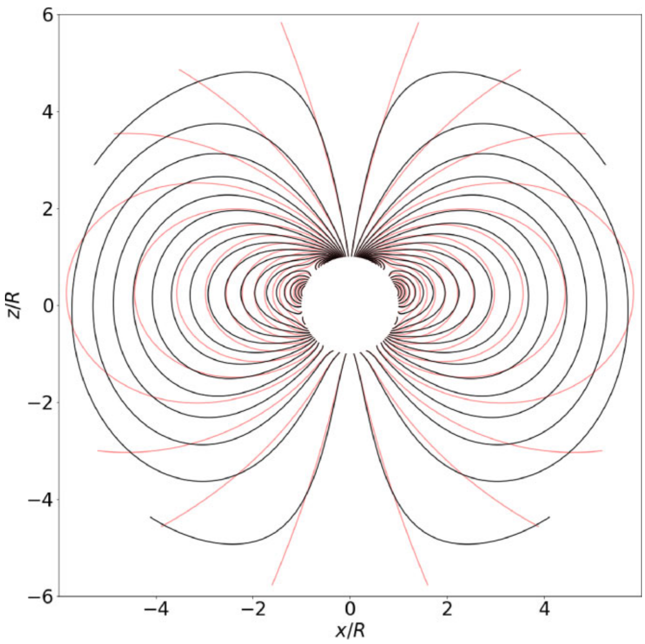


Figure 7. Field lines for the current-free (red) and FF (black) cases. The multipole coefficients at the surface are $b_{l>1} = 0.5$ for both. For the FF case, the coefficients in expression (9) for $\mathcal{T}(\mathcal{P})$ are $s_1 = 0.2$, $s = 0.4$.

by

$$\mathcal{J} = \frac{1}{i_{\max}} \sum_{i=1}^{i_{\max}} [\triangle_{\rm GS}\mathcal{P}_i - G(\mathcal{P}_i)]^2. \quad (27)$$

The input must include information about the functional form of $G(\mathcal{P})$ or equivalently $\mathcal{T}(\mathcal{P})$. In Section 2 we modelled $\mathcal{T}$ to be a quadratic function of $\mathcal{P}$ using two parameters, s_1 and s_2 (see equation (9)). Therefore, the input of the PINN must be extended to include these parameters. The PINN is trained to provide solutions for any value of s_1 and s_2 in the same sense that it is trained to provide solutions for any value of the multipole coefficients defining the BC. The input for the general FF case is

$(q, \mu, b_2, b_3, b_4, b_5, b_6, b_7, s_1, s_2)$.

Fig. 7 shows, for reference, a comparison between a current-free and an FF magnetic field, where we observe notorious difference in the structure of the field lines.

We note that there can be regions in the input parameter space where mathematical solutions do not exist (see Akgün et al. 2018; Mahlmann et al. 2019 for a detailed discussion). This is reflected in the evolution of the loss function, which fluctuates around values of order unity, meaning that there are no solutions that can minimize the residual of the PDE. Nevertheless, the PINN will return approximate solutions. It is up to the user to carefully evaluate the validity and accuracy of the results. A sufficiently low final value of the loss function is a first filter in this regard.

The lack of analytical solutions in the general case makes it difficult to estimate errors, which is a fundamental part of any scientific analysis. Despite the abundant literature on PINNs as PDE solvers in recent years, a systematic and consistent way for measuring errors and deciding on the quality of the provided approximate solutions is still lacking. Here, we adopt the following approach:

Downloaded from https://academic.oup.com/mnras/article/524/1/32/7199782 by Universidad de Alicante user on 26 February 2024

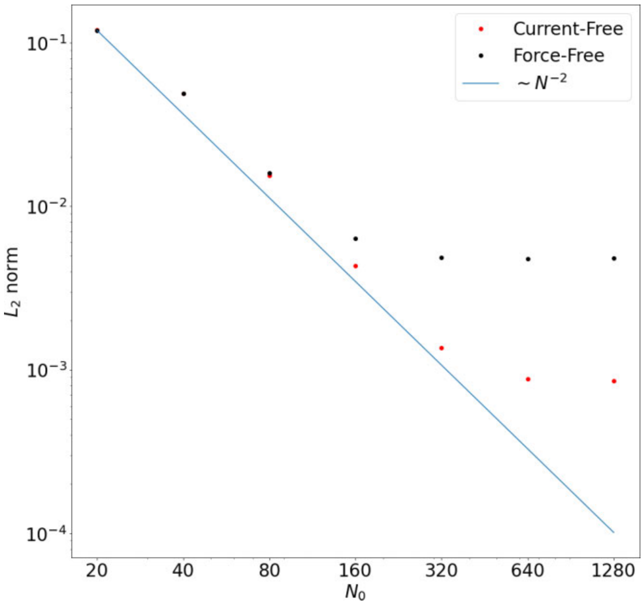


Figure 8. The L_2-norm of the discretized GS equation for the current-free (red) and FF (black) cases as a function of the resolution N_0.

(i) After our NN is trained, we create a regular grid (q_i, μ_j), where $i, j = 0,..., N_0$, and evaluate the PINN solution $\mathcal{P}$ at all the points with a forward pass.

(ii) Then, we discretize equation (4) using a second-order finite difference scheme, which gives us a different evaluation of the residual $\epsilon_{\rm FD}$ using the same function values evaluated in the previous step.

(iii) If $\mathcal{P}$ was the 'exact' solution of the PDE, the second-order residual $\epsilon_{\rm FD}$ would decrease with increasing resolution as $\sim N_0^{-2}$, where N_0 is the number of grid points (assuming both dimensions have the same resolution). In reality, $\mathcal{P}$ is only an approximate solution with an intrinsic error $\epsilon_{\rm NN}$ inherited from the quality and accuracy of the PINN. Therefore, $\epsilon_{\rm FD}$ will follow this power law only up to the point where the PINN approximation error $\epsilon_{\rm NN}$ starts to dominate the discretization error $\epsilon_{\rm FD}$.

Fig. 8 illustrates this behaviour. We plot the L_2-norm of the discretized GS equation for both the current-free $(G(\mathcal{P}) = 0)$ and FF cases $(G(\mathcal{P}) \neq 0)$ as a function of the number of grid points N_0. In both cases the L_2-norm drops as $\sim N_0^{-2}$ until it reaches a plateau which signalizes that $\epsilon_{\rm NN} > \epsilon_{\rm FD}$. We expect that, at worst, $\epsilon_{\rm NN}$ will be of the order of the square root of the loss function, because in equations (22) and (27) $\mathcal{J}$ is precisely L_2^2. In other words, when calculating $\epsilon_{\rm NN}$ for a particular example, with fixed multipole coefficients and source terms, we expect the error to be of the order $\sqrt{\mathcal{J}}$, within a factor of a few. We note that we obtain errors of the same order of magnitude for both cases, with a factor of $\sim$5 less for the vacuum. We expect a slightly higher error when we introduce the current term $G(\mathcal{P})$, because we increase the dimensionality of the problem, and we also introduce non-linear terms into the differential equation.

5 APPLICATION TO THE MAGNETOTHERMAL EVOLUTION OF NEUTRON STARS

Our astrophysical scenario of interest is the long-term evolution of magnetic fields in NSs. The evolution of the system is governed by two coupled equations: the heat diffusion equation and the induction equation (see the review by Pons & Viganò 2019 for more details). They must be complemented with a detailed specification of the local microphysics (neutrino emissivity, heat capacity, thermal and electrical conductivity) and the structure of the star, usually assumed as fixed throughout the NS's life. Our NS background model is a $1.4 M_\odot$ NS built with the Sly4[4] equation of state (Douchin & Haensel 2001). We use the two-dimensional (2D) magnetothermal code (latest version in Viganò et al. 2021) developed by our group suitably modified to implement the external BCs using the PINN (trained as described in the previous section) to assess its performance and potential. In particular, this implementation allows us to quickly switch and compare between vacuum BCs and the barely explored FF BCs. To our knowledge, only the work by Akgün et al. (2018) has presented results from simulations that included the effect of a magnetosphere threaded by currents. They had to implement a costly elliptic solver as a BC which slowed down the code considerably. The PINN implementation should, in principle, be much easier to change, efficient, and generalizable.

5.1 Current-free magnetospheric BCs

We begin by considering a crustal-confined magnetic field topology and vacuum BCs (no electrical currents circulating in the envelope and across the surface). We enforce BCs via multipole expansion of the radial magnetic field at the surface as described in Pons, Miralles & Geppert (2009) and Pons & Viganò (2019).

The coefficients of the multipole expansion can be computed from the radial component of the magnetic field at the surface of the star as follows:

$$b_l = \frac{2l+1}{2(l+1)} \int_0^\pi B_r(R,\theta) P_l(\cos\theta) \sin\theta \, d\theta. \tag{28}$$

During the evolution, we calculate at each time-step the b_l coefficients using equation (28). In the classical approach, we reconstruct the values of B_r and B_θ in the external ghost cells, explicitly

$$B_r = \sum_{l=1}^{l_{\max}} b_l \, (l+1) \, P_l \, (\cos\theta) \left(\frac{R}{r}\right)^{l+2}, \tag{29}$$

$$B_\theta = -\sin\theta \sum_{l=1}^{l_{\max}} b_l \, P_l' \, (\cos\theta) \left(\frac{R}{r}\right)^{l+2}, \tag{30}$$

where R is the radius of the NS. For conciseness, we refer to this procedure by the nomenclature OLD.

In the new PINNs approach, we use the b_l coefficients obtained from the Legendre decomposition as inputs to the PINN. The latter returns values of the poloidal flux function $\mathcal{P}$, or any required component of the magnetic field by taking derivatives (equations 24 and 25). Obviously, in this case (vacuum BCs), the PINN approach does not represent any advantage because we already know how to build the analytical solution. However, we want to ensure that the results of the simulations do not show any undesirable effects before moving on to a more complex case.

To compare the different employed techniques described above, we run axisymmetric crustal-confined magnetic field simulations using the 2D magnetothermal code (Viganò et al. 2021) with a grid of 99 angular points (from pole to pole) and 200 radial points. The initial field has a poloidal component of 10^{14} G (value at the pole and consists of a sum of a dipole $(b_1 = 1)$, a quadrupole $(b_2 = 0.6)$ and an octupole $(b_3 = 0.3)$. The initial toroidal quadrupolar component

[4]https://compose.obspm.fr/

Downloaded from https://academic.oup.com/mnras/article/524/1/32/7199782 by Universidad de Alicante user on 26 February 2024

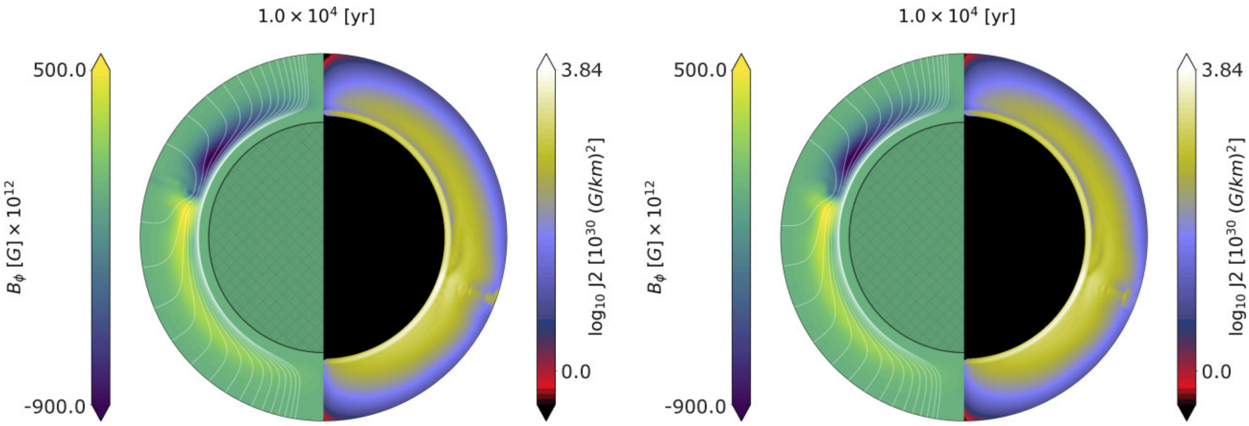


Figure 9. A snapshot of the magnetic field evolution and the electric current at 10 kyr, obtained using OLD (left-hand panel) and PINN (right-hand panel). In the left hemisphere, we show the meridional projection of the magnetic field lines (white lines) and the toroidal field (colours). In the right hemishere, we display the square of the modulus of the electric current, i.e. $|J|^2$ (note the log scale). The crust has been enlarged by a factor of 8 for visualization purposes.

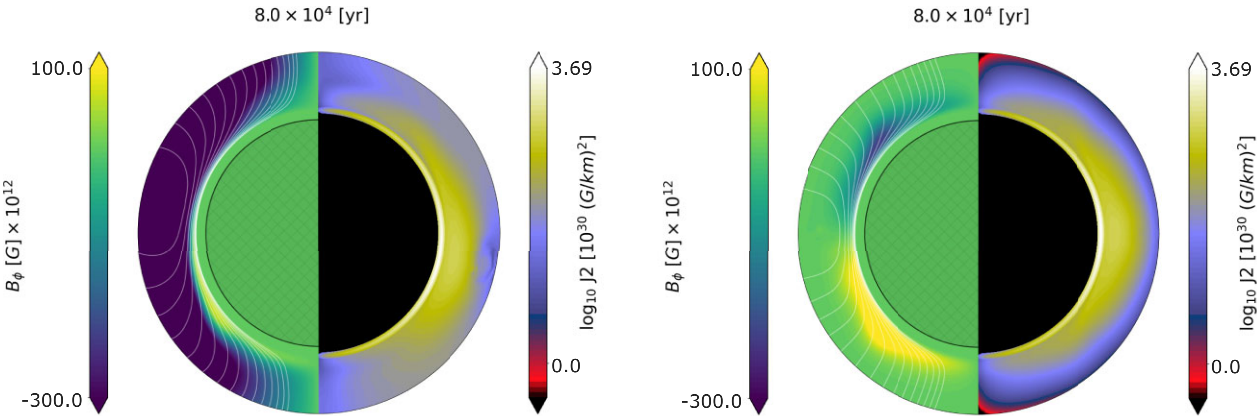


Figure 10. Same as Fig. 9. A snapshot of the magnetic field evolution and the electric current at 80 kyr. Left-hand panel: FF BCs. Right-hand panel: vacuum BCs.

has also a maximum initial value of 10^{14} G. The maximum number of multipoles is fixed to $l_{\max} = 7$ for PINN and to $l_{\max} = 50$ for OLD. The objective behind the use of different $l_{\max}$ is to assess the impact of truncating the multipole number when using the PINN.

The results of the comparison at $t = 10$ kyr are displayed in Fig. 9. On the left (right), we show the magnetic field profiles obtained with OLD (PINN) BCs. The overall evolution of the magnetic field and the electric current is very similar. Slight differences appear due to the multipolar truncation in the PINN case. We must note that, if the same maximum number of multipoles is set for both systems, we obtain almost identical results.

5.2 FF magnetospheric BCs

To couple the internal field evolution with an FF magnetosphere PINN solver, we must extend the vacuum case (Section 5.1) with additional steps. In our magnetothermal evolution code, we impose the external BCs by providing the values of the magnetic field components in two radial ghost cells for every angular cell. In the vacuum case, once the multipolar decomposition of the radial field over the surface is known, the solution in the ghost cells can be built analytically. However, in the general case, one must solve an elliptic equation in a different grid that must be extended far from the surface to properly capture to asymptotic behaviour at long distances. This process is very costly because it must be repeated tens of thousands of time-steps as the interior field evolves. In this situation, having trained the PINN, allows us to use it as a fast tool to provide required values of the solution in the ghost cells. We proceed as follows: First, at each evolution time-step, we must know the toroidal function $\mathcal{T}(\mathcal{P})$. For simplicity, in this application we use a quadratic function. At each time step we fit the values obtained from the internal evolution one cell below the surface. The fit provides the coefficients of the quadratic interpolation s_1 and s_2 defined in equation (9). Next, as described in Section 4.3, s_1 and s_2 are provided as additional input parameters to the forward pass. The PINN returns the poloidal flux function $\mathcal{P}$ and the components of the magnetic field needed at the ghost cells of the magnetothermal evolution code. With this information, the internal evolution can proceed to the next time-step.

We assume an initial FF magnetic field with a poloidal component of 3×10^{14} G at the polar surface and a maximum toroidal field of 3×10^{14} G. To understand the impact of the different BCs, we consider in one case FF BCs (the left-hand panel of Fig. 10) and in the other case vacuum BCs (the right-hand panel of Fig. 10). The results of the comparison are illustrated by two snapshots at $t = 80$ kyr of the evolution with the same initial model. We note that the initial FF magnetic field allows current sheets to thread the star's surface. A distinct magnetic field evolution is clearly observed if we apply one type of BCs or the other. The FF BCs (left) result in a stronger

Downloaded from https://academic.oup.com/mnras/article/524/1/32/7199782 by Universidad de Alicante user on 26 February 2024

toroidal dipole close to the surface and slightly displaced towards the north. The stronger toroidal component compresses the poloidal field lines closer to the poles. In contrast, for vacuum BCs, the poloidal field lines retain certain symmetry with respect to the equator, and the dominant toroidal component is now quadrupolar and concentrated at the crust/core interface, as shown in the right-hand panel of Fig. 10. The distribution of the electric current in the stellar crust is also different. Enforced by the vacuum BCs, current tends to vanish around the poles and close to the surface. This is similar to what was observed in Fig. 9 although the initial field topology is different. Instead, with FF BCs, currents near the surface are not forced to vanish. In the left-hand panel, the slightly more yellowish region in the Northern hemisphere and mid-latitudes indicates that significant current flows into the magnetosphere. This difference in current configurations would have important implications in the observed temperature distribution, as discussed in Akgün et al. (2018). We will address, in future works, a more detailed exploration of the effect of BCs since our purpose here is to illustrate with a few examples the potential of our approach.

6 CONCLUSIONS

Using PINNs to obtain a solution of a particular boundary value problem is, up to date, far more computationally expensive and arguably less accurate than using classical methods. The drawbacks are related to the training process, which involves the minimization of a high-dimensional loss function. Once a PINN is trained for a given boundary value problem, its utility is limited because it would be necessary to re-train to generate new solutions with different BCs.

The functionality of PINNs to real physical problems would become significantly better if their flexibility and adaptability can be increased. In this work, we explore this idea by training our PINN for general BCs and source terms, expressed through appropriate coefficients (a limited number of them) that enter as additional inputs in the network. Of course, this makes the training process computationally more expensive, but the evaluation of new generic solutions is very fast. In our study, the coverage of the parameter space is not exhaustive because we have used very limited computational resources (a personal computer), and our purpose is to show that this proof-of-concept works and can already be applied to some physical problems, even by non-experts in computer science. If necessary, it is straightforward to adapt our implementation for the specific needs of other applications (e.g. adding more multipoles in the boundary if we need to capture smaller scales, or extending the parametrization of the BCs). We are aware that it is possible to drastically improve the efficiency of our computations by employing GPU clusters or enriching our algorithms with advanced machine-learning techniques and that will be required, for example, in the extension to three-dimesion (3D) of this work.

We have also explored various configurations for our network through a basic hyperparameter space study. We conclude that the most impactful element is the number of neurons per layer. On the contrary, making the network deeper by adding more layers is not beneficial, beyond a reasonable minimum. Furthermore, the way that the solution is parametrized to always satisfy the BCs is important, indicating that the human orientation in some choices (as opposed to a zero-like approach) is still critical for physics problems. We also found that ResNet architectures do not offer any advantage for the kind of applications that we are dealing with. We emphasize that, from a theoretical point of view, for any continuous function there is always an NN that is able to approximate it to arbitrary precision (see the *Universal Approximation Theorem*; Cybenko 1989; Leshno et al. 1993; Pinkus 1999). However, the theorem does not specify the properties said network (e.g. its size, architecture, or the activation functions) nor if it is feasible to build and train it with current computational capabilities, but merely state that this NN exists. For this reason, in general, there is not guarantee that this function can be represented by another NN, and therefore during the training process the loss function will stagnate to a finite value. Moreover, the presence of non-linear activation functions in conjunction with the compositional structure of NNs renders the associated optimization problem non-convex. Therefore, the gradient descent based algorithm used during the training process may get stuck in a saddle point or in a local minimum, instead of moving towards the global one. All these issues affect unavoidably the convergence of the NNs. In classical numerical approaches, the dependence of the convergence on the implementation properties (order of the scheme, resolution, etc.) is known and well studied. On the other hand, in the DL field there is sparse mathematical evidence to instruct us how to build and train a network to achieve a more accurate result, although it is currently subject to active research (e.g. De Ryck, Lanthaler & Mishra 2021; De Ryck & Mishra 2022; De Ryck, Jagtap & Mishra 2022; Mishra & Molinaro 2022).

Nevertheless, we cannot ignore that NNs have proven to be very good function approximators in very distinct areas, even though perfect convergence is currently unlikely to be achieved for the practical reasons mentioned above. For this particular work, our results show that the PINN solutions are relatively accurate, reliable, and well behaved. For the current-free GS equation, where comparisons with an analytical solution can be made, we found relative differences of typically less than 1 per cent. For the force–free GS equation, we propose a method for estimating the error through the use of a finite difference discretization scheme. Our analysis shows that solutions are accurate up to a point that we associate to the intrinsic approximation error of the PINN. This approach is straightforward to implement and self-consistent and could be set as the standard procedure for estimating errors of PINN-based PDE solvers in general. Even if PINNs fall short in terms of precision compared to classical PDE solvers, they can still be used in conjunction with them in various cases. For example, many iterative solvers rely on good initial guesses to converge. A PINN can provide such an initial guess to be subsequently refined by a classical method.

The most interesting case where PINNs overcome the capabilities of classical methods are physical systems with two or more domains that are governed by vastly different physical conditions and time-scales. An example of such a case is the magnetothermal evolution in the interior of an NS that is connected to an FF magnetosphere. Solving this problem through a global simulation in the entire domain is very costly due to the needs of the elliptic solver for the exterior solution. On the contrary, PINNs provide a very effective way of imposing BCs at the interface of the two domains. Once a PINN is trained, accurate enough magnetospheric solutions in a few points (ghost cells of the interior evolution code) can be swiftly computed at each time-step without adding an excessive amount of computational cost. We have shown that for the well-tested case of vacuum BC, the PINN is accurate enough to satisfactorily reproduce the reference results obtained by the exact spectral decomposition approach. Nevertheless, it is about two times slower. As proof of concept, to demonstrate the potentiality of our approach, we also presented results for the less explored FF magnetospheric BCs. In this case, the computational cost was more than an order of magnitude smaller than the similar problem solved with classical methods. Indeed, we obtain solutions that allow currents that thread the NS's surface and flow into the magnetosphere, giving rise to a new family of internal field

Downloaded from https://academic.oup.com/mnras/article/524/1/32/7199782 by Universidad de Alicante user on 26 February 2024

configurations. We reserve a more detailed analysis of the physical properties of these solutions in 2D, and the more physically relevant extension to 3D, for future works.

ACKNOWLEDGEMENTS

We acknowledge the support through the grant PID2021-127495NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by the European Union, and the Astrophysics and High Energy Physics programme of the Generalitat Valenciana ASFAE/2022/026 funded by MCIN and the European Union NextGenerationEU (PRTR-C17.I1). JFU is supported by the predoctoral fellowship UAFPU21-103 funded by the University of Alicante. CD is supported by the ERC Consolidator Grant 'MAGNESIA' no. 817661 (PI: N. Rea) and has the partial support of NORDITA. This work has been carried out within the framework of the doctoral program in Physics of the Universitat Autònoma de Barcelona and it is partially supported by the programme Unidad de Excelencia María de Maeztu CEX2020-001058-M.

DATA AVAILABILITY

All data produced in this work will be shared on reasonable request to the corresponding author.

REFERENCES

Abadi M. et al., 2016 OSDI, USENIX Association, p. 265, https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi
Akgün T., Cerdá-Durán P., Miralles J. A., Pons J. A., 2017, MNRAS, 472, 3914
Akgün T., Miralles J. A., Pons J. A., Cerdá-Durán P., 2016, MNRAS, 462, 1894
Akgün T., Cerdá-Durán P., Miralles J. A., Pons J. A., 2017, MNRAS, 474, 625
Akgün T., Cerdá-Durán P., Miralles J. A., Pons J. A., 2018, MNRAS, 481, 5331
Baydin A. G., Pearlmutter B. A., Radul A. A., Siskind J. M., 2017, J. Mach. Learn. Res., 18, 5595
Cai S., Mao Z., Wang Z., Yin M., Karniadakis G. E., 2021, Acta Mechanica Sinica, 37, 1727
Chan W., Jaitly N., Le Q. V., Vinyals O., 2015, preprint (arXiv:1508.01211)
Chen X., Jeffery D. J., Zhong M., McClenny L., Braga-Neto U., Wang L., 2022, preprint (arXiv:2211.05219)
Cybenko G., 1989, Math. Control Signals Syst., 2, 303
De Ryck T., Mishra S., 2022, Adv. Comput. Math., 48, 79
De Ryck T., Lanthaler S., Mishra S., 2021, Neural Netw., 143, 732
De Ryck T., Jagtap A. D., Mishra S., 2022, preprint (arXiv:2203.09346)
Dehman C., Viganò D., Pons J. A., Rea N., 2023, MNRAS, 518, 1222
Douchin F., Haensel P., 2001, A&A, 380, 151
Eivazi H., Tahani M., Schlatter P., Vinuesa R., 2022, Phys. Fluids, 34, 075117
Glampedakis K., Lander S. K., Andersson N., 2014, MNRAS, 437, 2
Hayat S., Mall R. N., 2013, IJERT, 2
He K., Zhang X., Ren S., Sun J., 2015, preprint (arXiv:1512.03385)
He K., Zhang X., Ren S., Sun J., 2016, preprint (arXiv:1603.05027)
Hornik K., Stinchcombe M., White H., 1989, Neural Netw., 2, 359
Kingma D. P., Ba J., 2014, preprint (arXiv:1412.6980)
Kojima Y., 2017, MNRAS, 468, 2011
Kugunavar S., Prabhakar C. J., 2021, Visual Comput. Ind. Biomed. Art, 4,12
Lagaris I. E., Likas A., Fotiadis D. I., 1997, preprint (physics/9705023)
Lawrence S., Giles C., Tsoi A. C., Back A., 1997, IEEE Trans. Neural Netw., 8, 98
Leshno M., Lin V. Y., Pinkus A., Schocken S., 1993, Neural Netw., 6, 861
Luna R., Calderón Bustillo J., José Seoane Martínez J., Torres-Forné A., Font J. A., 2023, Phys. Rev. D, 107, 064025
Mahlmann J. F., Akgün T., Pons J. A., Aloy M. A., Cerdá-Durán P., 2019, MNRAS, 490, 4858
Mishra S., Molinaro R., 2021, J. Quant. Spectrosc. Radiative Trans., 270, 107705
Mishra S., Molinaro R., 2022, IMA J. Num. Anal., 43, 1
Pili A. G., Bucciantini N., Del Zanna L., 2015, MNRAS, 447, 2821
Pinkus A., 1999, Acta Numer., 8, 143
Pons J. A., Viganò D., 2019, Living Rev. Comput. Astrophys., 5, 3
Pons J. A., Miralles J. A., Geppert U., 2009, A&A, 496, 207
Raissi M., Perdikaris P., Karniadakis G. E., 2019, J. Comput. Phys., 378, 686
Schiassi E., De Florio M., Ganapol B. D., Picca P., Furfaro R., 2022, Ann. Nucl. Energy, 167, 108833
Stefanou P., Pons J. A., Cerdá-Durán P., 2023, MNRAS, 518, 6390
Sukumar N., Srivastava A., 2022, Comput. Methods in Appl. Mech. Eng., 389, 114333
Thompson C., Duncan R. C., 1995, MNRAS, 275, 255
Thompson C., Duncan R. C., 1996, ApJ, 473, 322
Traore B. B., Kamsu-Foguem B., Tangara F., 2018, Ecological Inform., 48, 257
Viganò D., Garcia-Garcia A., Pons J. A., Dehman C., Graber V., 2021, Comput. Phys. Commun., 265, 108001
Wang S., Yu X., Perdikaris P., 2022, J. Comput. Phys., 449, 110768

This paper has been typeset from a TEX/LATEX file prepared by the author.

© 2023 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/524/1/32/7199782 by Universidad de Alicante user on 26 February 2024

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **526,** 1504–1511 (2023)
Advance Access publication 2023 September 18

https://doi.org/10.1093/mnras/stad2840

Solving the pulsar equation using physics-informed neural networks

Petros Stefanou,[1,2]★ Jorge F. Urbán[1] and José A. Pons[1]

[1]*Departament de Física Aplicada, Universitat d'Alacant, Ap. Correus 99, E-03080 Alacant, Spain*
[2]*Departament d'Astronomia i Astrofísica, Universitat de València, Dr Moliner 50, E-46100 Burjassot, València, Spain*

Accepted 2023 September 12. Received 2023 September 11; in original form 2023 August 1

ABSTRACT

In this study, Physics-Informed Neural Networks (PINNs) are skilfully applied to explore a diverse range of pulsar magnetospheric models, specifically focusing on axisymmetric cases. The study successfully reproduced various axisymmetric models found in the literature, including those with non-dipolar configurations, while effectively characterizing current sheet features. Energy losses in all studied models were found to exhibit reasonable similarity, differing by no more than a factor of three from the classical dipole case. This research lays the groundwork for a reliable elliptic Partial Differential Equation solver tailored for astrophysical problems. Based on these findings, we foresee that the utilization of PINNs will become the most efficient approach in modelling three-dimensional magnetospheres. This methodology shows significant potential and facilitates an effortless generalization, contributing to the advancement of our understanding of pulsar magnetospheres.

Key words: magnetic fields – stars: neutron – pulsars.

1 INTRODUCTION

Physics-Informed Neural Networks (PINNs; Lagaris, Likas & Fotiadis 1997; Raissi, Perdikaris & Karniadakis 2019) is a relatively new but very promising family of Partial Differential Equation (PDE) solvers based on Machine Learning (ML) techniques. This method is suitable to obtain solutions of PDEs describing the physical laws of a given system by taking advantage of the very successful modern ML frameworks and incorporating physical knowledge. In recent years, PINN-based solvers have been used to solve problems in a great variety of fields: fluid dynamics (Cai et al. 2021), turbulence in supernovae (Karpov et al. 2022), radiative transfer (Korber et al. 2023), black hole spectroscopy (Luna et al. 2023), cosmology (Chantada et al. 2023), large scale structure of the Universe (Aragon-Calvo 2019), galaxy model fitting (Aragon-Calvo & Carvajal 2020), inverse problems (Pakravan et al. 2021), and many more.In our previous work (Urbán et al. 2023, Paper I hereafter), we presented a PINN solver for the Grad–Shafranov equation, which describes the magnetosphere of a slowly rotating neutron star endowed with a strong magnetic field (a magnetar) in the axisymmetric case. In that paper, we demonstrated the ability of the network to be trained for various boundary conditions and source terms simultaneously. In this work, our purpose is to extend our PINN approach to the more general – and more challenging – case of rapidly rotating neutron stars (pulsars). The rotating case presents new interesting challenges related to the presence of current sheets in the magnetosphere. Our implementation is able to deal with these pathological regions sufficiently well, demonstrating its potential for problems where classical methods struggle.

Pulsar magnetospheres have been studied extensively in the last 25 yr with various approaches, each with its advantages and limitations. Contopoulos, Kazanas & Fendt (1999) were the first to solve the axisymmetric, time-independent problem, a result that was later confirmed and improved by Gruzinov (2005) and Timokhin (2006). Spitkovsky (2006) used a full magnetohydrodynamic time-dependent code and was able to acquire solutions for aligned and oblique rotators. Solutions for arbitrary inclination were also obtained by Pétri (2012) who used a time-dependent pseudo-spectral code. More recent approaches involve large-scale particle-in-cell simulations that include the influence of accelerated particles (Cerutti et al. 2015; Philippov & Spitkovsky 2018). All of these, and many related, works have improved our understanding of the pulsar magnetosphere. However, some important questions remain still without an answer.

The structure of the paper is the following. After a brief summary of the relevant equations and boundary conditions in Section 2, we will describe the PINN method in Section 3, with emphasis in the new relevant details with respect to the magnetar problem (Paper I). In Section 4 we present our results, showing that we are able to reproduce the well-established axisymmetric results encountered in the literature, but also indicating that new, possibly unexplored solutions can be encountered. We summarize our most relevant conclusions and discuss possible improvements and future extensions in Section 5.

2 PULSAR MAGNETOSPHERES

Our aim is to extend our previous results from Paper I to find numerical solutions of the pulsar equation, that can be written as follows:

$$\nabla \times \left(\boldsymbol{B} - \beta^2 \boldsymbol{B}_p\right) = \alpha \boldsymbol{B}, \tag{1}$$

where $\beta = v/c = |\Omega \times \boldsymbol{r}|/c = \varpi/R_{\rm LC}$ is the co-rotational speed in units of the speed of light, with $R_{\rm LC} = c/\Omega$ being the light-cylinder (LC) radius and $\boldsymbol{B_p} = \boldsymbol{B} - \boldsymbol{B}_\phi$ is the magnetic field perpendicular

★ E-mail: petros.stefanou@uv.es

© 2023 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/526/1/1504/7276630 by Universidad de Alicante user on 26 February 2024

to the direction of rotation. Here, α is a scalar function given by

$$\alpha = \frac{4\pi}{c}\left(\boldsymbol{J} - \rho_e \boldsymbol{v}\right) \cdot \frac{\boldsymbol{B}}{B^2}, \tag{2}$$

representing the ratio between the field-aligned component of the current in the corotating frame and the local magnetic field strength. We refer to the comprehensive and thorough recent review by Philippov & Kramer (2022; and references therein) for a complete historical overview of magnetospheric physics and the mathematical derivation of the equations.

As in Paper I, we focus in the axisymmetric case and, for convenience, we use compactified spherical coordinates (q, μ, ϕ) where $q = 1/r$, $\mu = \cos\theta$. In these coordinates, any axisymmetric magnetic field can be written in terms of a poloidal and a toroidal scalar stream function P and T as

$$\boldsymbol{B} = \frac{q}{\sqrt{1-\mu^2}}\left(\nabla P \times \hat{\phi} + T\hat{\phi}\right). \tag{3}$$

Here, P is related to the magnetic flux and is constant along magnetic field lines. Plugging this expression into equation (1) and taking the toroidal component, we get

$$\nabla P \times \nabla T = 0, \tag{4}$$

which means that T is only a function of P ($T = T(P)$) and, therefore, is also constant along magnetic field lines. The poloidal component of equation (1) gives the well-known pulsar equation (Michel 1973; Scharlemann & Wagoner 1973), which in our coordinates reads

$$\left(1-\beta^2\right)\Delta_{\rm GS} P + 2\beta^2 q^2 \left(q\partial_q P + \mu\partial_\mu P\right) + G(P) = 0, \tag{5}$$

where $G(P) = TT'$ and $\Delta_{\rm GS}$ is the Grad–Shafranov operator, given by

$$\Delta_{\rm GS} \equiv q^2\partial_q\left(q^2\partial_q\right) + \left(1-\mu^2\right)q^2\partial_{\mu\mu}. \tag{6}$$

Notice that in the limiting case where $\beta = 0$, we ignore the rotationally induced electric field and equation (5) is reduced to the Grad–Shafranov equation. Hereafter, we will use the following shorthand notation for the extended operator

$$\Delta_{{\rm GS}\beta} \equiv \left(1-\beta^2\right)\Delta_{\rm GS} + 2\beta^2 q^2\left(q\partial_q + \mu\partial_\mu\right). \tag{7}$$

A crucial difference in the modelling of magnetar and pulsar magnetospheres is the source of poloidal currents TT'. In magnetars, currents are assumed to be injected into the magnetosphere due to the magnetic field evolution in the neutron star's (NS) crust, which twists the magnetosphere (Akgün et al. 2018). Therefore, the source term in equation (5) is given either as a user-parametrized model (as for example in Akgün et al. 2016 for the 2D problem or in Stefanou, Pons & Cerdá-Durán 2023 for the 3D problem) or in a self-consistent manner by coupling the interior evolution with the magnetosphere. We adopted the latter approach in Paper I (see the astrophysical application in section 5 of that paper), where a series of magnetospheric steady-state solutions were obtained for each time-step of the internal magnetothermal evolution.

In the pulsar magnetosphere, however, the source of current is the LC. Lines that cross the LC have to open up and bend, giving rise to azimuthial fields and currents. We refer to the region with open field lines as the open region. On the other hand, lines that do not cross the LC, but turn back to the surface, co-rotate rigidly with the the star. We refer to this region as the closed region. The source term TT' must be determined self-consistently to ensure smooth crossing of the lines along the LC. In particular, at the LC where $\beta = 1$, equation (5) takes the simple form

$$-q^2\left(q\partial_q P + \mu\partial_\mu P\right) = 2B_z = TT', \tag{8}$$

which places a constraint on the possible functions $T(P)$. This makes equation (5) complicated to solve, as two functions have to be determined simultaneously. In order to do so, additional physical boundary conditions have to be imposed (see next subsection).

Through this work, we measure distances in units of the radius of the star R and magnetic fields in units of the surface magnetic field strength at the equator B_0 (note that the surface magnetic field strength at the poles is $2B_0$). In these units, the magnetic flux is measured in units of the total poloidal flux $P_0 = B_0R^2$. It is convenient to additionally define the total magnetic flux carried by field lines that cross the LC in a non-rotating dipole $P_1 = P_0(R/R_{\rm LC})$, which will be useful in what follows. Finally, the toroidal stream function is measured in units of B_0R.

2.1 Boundary conditions

We assume that the magnetic field at the surface of the star $P(q = 1)$ is known. For example, in the case of a dipole

$$P(q=1, \mu) = B_0(1-\mu^2) \tag{9}$$

but we will also explore other options. The last closed field line, called separatrix, will be labelled by $P = P_c$. It marks the border between open and closed (current-free) regions. The value P_c corresponds to the total magnetic flux that crosses the LC. The point where this line meets the equator is the called Y-point and it should lie at a radius r_c, somewhere between the surface and the LC (see e.g. Timokhin 2006 for a detailed discussion). Far away from the surface, field lines should become radial, resembling a split monopole configuration (Michel 1973). Inside the closed region, the magnetosphere is current-free and purely poloidal. No toroidal fields are developed, so that

$$T(P > P_c) = 0 \tag{10}$$

by definition.

In the classical model, a current sheet should develop along the separatrix, supported by the discontinuity of the toroidal magnetic field between closed ($B_\phi = 0$) and open ($B_\phi \neq 0$) regions. Another current sheet should exist along the equator and beyond the Y-point, supported by the opposed directions of the magnetic field lines between the two hemispheres. The current sheet forms the return current that accounts for the current supported by out-flowing particles along the open field lines and closes the current circuit.

All this motivated the seminal works to impose equatorial symmetry by defining a numerical domain only on one hemisphere and imposing another boundary condition $P = P_c$ at the equator. To reproduce these models, we will impose that along the equator and beyond the Y-point

$$P(q < q_c, \mu = 0) = P_c. \tag{11}$$

Here, $q_c = 1/r_c$ is the corresponding location of the Y-point in compactified coordinates. However, later works have relaxed this assumption (Contopoulos, Kalapotharakos & Kazanas 2014), proposing a more general solution without a separatrix and with a transition region close to the equator. We will explore both cases in our results section.

2.2 Total energy and radiated power

The total energy of the electromagnetic field in the magnetosphere is given by

$$\mathcal{E} = \frac{1}{8\pi}\int (B^2 + E^2)\mathrm{d}V. \tag{12}$$

Downloaded from https://academic.oup.com/mnras/article/526/1/1504/7276630 by Universidad de Alicante user on 26 February 2024

It is convenient to describe the electromagnetic energy content of a given model in terms of the excess energy of a particular magnetospheric solution with respect to the non-rotating dipole,

$$\Delta\mathcal{E} = \frac{\mathcal{E} - \mathcal{E}_{\rm d}}{\mathcal{E}_{\rm d}}, \tag{13}$$

where

$$\mathcal{E}_{\rm d} = \frac{1}{3} B_0^2 R^3 \tag{14}$$

is the total magnetic energy of a non-rotating dipole.

The total power radiated away by a rotating magnetosphere can be calculated by integrating the Poynting flux over a sphere far away from the star

$$\dot{\mathcal{E}} = \frac{c}{4\pi} \int (\boldsymbol{E} \times \boldsymbol{B}) \cdot \hat{\boldsymbol{r}} {\rm d}\omega = -\frac{c}{R_{\rm LC}} \int_0^{P_{\rm c}} T {\rm d}P, \tag{15}$$

where ω is the solid angle. Again, it is convenient to measure the relative difference of the radiated power with respect to the classical order-of-magnitude estimate

$$\dot{\mathcal{E}}_{\rm d} = \frac{B_0^2 R^6}{R_{\rm LC}^4} c, \tag{16}$$

and hereafter we will express $\dot{\mathcal{E}}$ in units of $\dot{\mathcal{E}}_{\rm d}$.[1]

3 NETWORK STRUCTURE AND TRAINING ALGORITHM

In Paper I, we used a PINN to calculate approximate solutions of the axisymmetric Grad–Shafranov equation. We refer the reader to that paper for a more detailed description of the generic implementation. In this section, we will briefly summarize the main points and highlight the novelties introduced to adapt our solver to the pulsar case. The principal changes introduced aim at enforcing the physical constraint [boundary conditions, or the $T(P)$ requirement] by construction instead of leaving the job to minimize additional terms in the loss function, which usually does not reach the required accuracy.

We consider solutions in points $(q, \mu) \in \mathcal{D}$ in a two-dimensional domain $\mathcal{D}$. We denote by $\partial\mathcal{D}$ the boundary of this domain. To account for the equatorial constraint in equation (11), we will consider the equatorial line beyond $r_{\rm c}$ as part of $\partial\mathcal{D}$. In order to ensure that P depends on the coordinates while T is solely a function of P, we design a network structure consisting of two sub-networks that are trained simultaneously. The output of each sub-network depends only on its corresponding input.

The first sub-network takes the coordinates (q, μ) as input and returns as output a function that we denote by N_P. Then, P is calculated using

$$P(q, \mu) = f_{\rm b}(q, \mu) + h_{\rm b}(q, \mu) N_{\rm P}(q, \mu; \Theta). \tag{17}$$

where $f_{\rm b}$ can be any function in $\mathcal{D}$ that satisfies the corresponding boundary conditions at $\partial\mathcal{D}$. $h_{\rm b}$ is an arbitrary function representing some measure of the distance to the boundary, that must vanish at $\partial\mathcal{D}$. Both user-supplied functions $f_{\rm b}$ and $h_{\rm b}$ depend only on the coordinates and are unaffected by the PINN. There is some freedom to decide their specific form, as long as they retain certain properties (e.g. they are sufficiently smooth) and they have the desired behaviour at the boundary. The only part that is adapted during training is $N_{\rm P}$ through its dependence on the trainable parameters Θ. With this *parametrization* (or *hard enforcement*) approach, we ensure that boundary conditions are exactly satisfied by construction. This differs from the other usual approach, consisting of adding more terms related to the boundary conditions in the loss function. We will specify the particular form of the functions $f_{\rm b}(q, \mu)$ and $h_{\rm b}(q, \mu)$ in the next section when we discuss different cases.

Next, the second sub-network takes P as input and returns $N_{\rm T}$ as output. This automatically enforces that $T = T(P)$, required by equation (4). The new network output $N_{\rm T}(P; \Theta)$ is used to construct the function $T(P)$ as follows:

$$T(P) = g(P) N_{\rm T}(P; \Theta), \tag{18}$$

where g is another user-supplied function to include other physical restrictions. For example, we can use a step function to treat a possible discontinuity of T at the border between open and closed regions (equation 10). Another advantage of this approach is that T' is calculated from T via automatic differentiation in the second sub-network.

A subtle point in the pulsar problem is the determination of the critical value $P_{\rm c}$ that separates the open and closed regions. We have explored two different approaches: (a) $P_{\rm c}$ is self-determined by the network and (b) $P_{\rm c}$ is fixed to some value. In the first case, $P_{\rm c}$ is considered an extra output of the first sub-network, closely connected to P, resembling the approach taken by classical solvers for this problem. In contrast, the second method involves fixing $P_{\rm c}$ as a hyperparameter to a pre-determined value, leaving the network to discover a solution that aligns with this fixed value. This introduces the necessity to loosen some of the other imposed constraints, like the position of the Y-point, and encourages the exploration of solution properties from a fresh standpoint.

During the network training process, both P and T [and possibly $P_{\rm c}$, if we follow the case (a) above] are obtained simultaneously by minimizing the following loss function

$$\mathcal{L} = \omega_1 \mathcal{L}_{\rm PDE} + \omega_2 \sigma(P_{\rm c}), \tag{19}$$

where

$$\mathcal{L}_{\rm PDE} = \frac{1}{N} \sum_{(q,\mu)\in\mathcal{D}} \left[\Delta_{{\rm GS}\beta} P(q, \mu) + G(P(q, \mu))\right]^2. \tag{20}$$

Here, N represents the size of the training set, and ω_1 and ω_2 are adjustable parameters. In approach (b), $\omega_2 = 0$, whereas in approach (a), ω_2 is adjustable. Additionally, σ denotes the standard deviation of the $P_{\rm c}$ values for the points in the training set. $P_{\rm c}$ is a global magnetospheric value that should not depend on the coordinates. However, during training, there is no guarantee that this value will be the same for all the points that are considered. To ensure that the values of $P_{\rm c}$ at arbitrary points are as close to each other as possible, we minimize the standard deviation $(\sigma(P_{\rm c}))$ of the values of $P_{\rm c}$ at all the points of the training set by including an additional term in the loss function (19). This inclusion guarantees that the network's output for $P_{\rm c}$ remains constant and independent of the coordinates (q, μ). A schematic representation of our network's structure can be found in Fig. 1.

We consider a fully connected architecture for both sub-networks consisting of four hidden layers for the first one and two hidden layers for the second one. All layers have 40 neurons. Our training set consists of $N = 5000$ random points $(q, \mu) \in \mathcal{D}$, which are periodically changed every 500 epochs in order to feed the network with as many distinct points as possible. The total number of epochs

[1]We should stress that the (unfortunately) often used $\sin^2\theta$ dependence with the inclination angle in $\dot{\mathcal{E}}_{\rm d}$ is unphysical. It misleads to conclude that an aligned rotator does not emit, but it actually does radiate in a similar amount (but a factor of about 2) than the orthogonal rotator.

Downloaded from https://academic.oup.com/mnras/article/526/1/1504/7276630 by Universidad de Alicante user on 26 February 2024

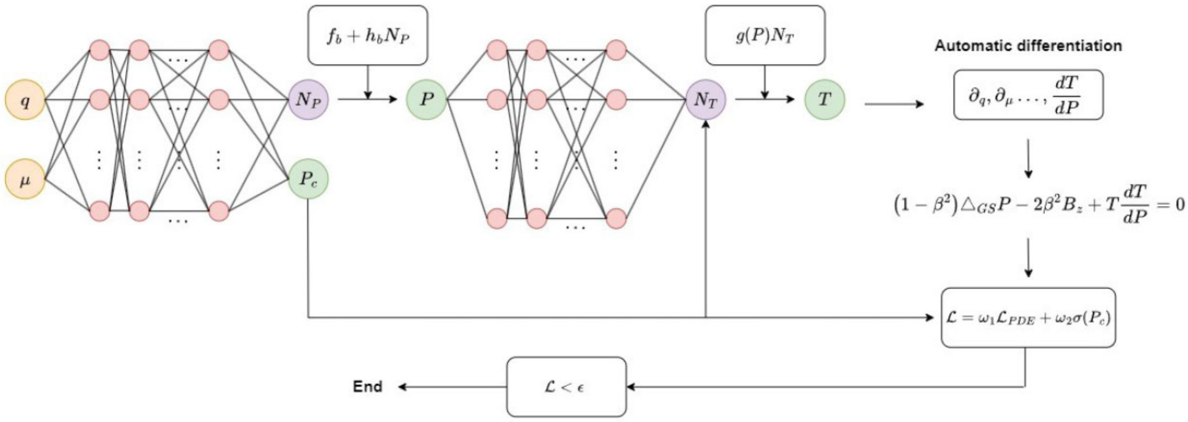


Figure 1. A sketch of the network structure. Two sub-networks are employed to ensure that $P = P(q, \mu)$ and $T = T(P)$.

is 35 000. This number may seem big at first sight, but is necessary because of the amount of points considered. In order to minimize the loss function, we use the ADAM optimisation algorithm (Kingma & Ba 2014) with an exponential learning rate decay. We have also considered in this work the idea of introducing trainable activation functions, suggested in PINNs originally by Jagtap, Kawaguchi & Karniadakis 2020. For each hidden layer k, the linear transformation performed at that layer is multiplied by a trainable parameter c_k. We found that this practice can accelerate convergence but a more rigorous study is out of the scope of this paper.

4 MAGNETOSPHERIC MODELS

In this section, we present the results obtained for various cases and under different physical assumptions. Our analysis encompasses the successful reproduction of all distinct axisymmetric models found in literature, which include:

(i) The classical solutions (Contopoulos et al. 1999; Gruzinov 2005),
(ii) The family of solutions with varying locations of the Y-point (Timokhin 2006),
(iii) The improved solution where the separatrix current sheet is smoothed out (Contopoulos et al. 2014),
(iv) The non-dipolar solutions (Gralla, Lupsasca & Philippov 2016).

Each of these models presents unique challenges and characteristics, and we will delve into the outcomes achieved for each one. In particular, the overall magnetospheric configuration, the poloidal flux at the separatrix P_c, the functions $T(P)$ and $G(P)$, and the energy losses $\dot{\mathcal{E}}$ agree with the previous works.

4.1 Hard enforcement of boundary conditions

As discussed earlier, we construct P according to equation (17) to fulfill the boundary conditions. For the cases (i) and (ii) we have employed an f_b with the following form:

$$f_b(q,\mu) = (1-\mu^2)\left[P_c + q(1-P_c)\frac{\mathrm{ReLU}^3\left(1-\frac{q_c}{q}\right)}{(1-q_c)^3}\right], \tag{21}$$

where ReLU (x) is the *Rectified Linear Unit* function, which returns zero if $x < 0$ and x if $x > 0$[2], $q_c = 1/r_c$ is the position of the Y-point and n is a free positive parameter.

Furthermore, we have chosen the following h_b function:

$$h_b(q,\mu) = (1-\mu^2)(1-q)\sqrt{\mathrm{ReLU}^3\left(1-\frac{q_c}{q}\right) + \mu^2}, \tag{22}$$

where the reader can check that h_b is zero at the boundary. The terms with the ReLU functions enforce that $P = P_c$ at the equator when $q < q_c$, and the third power ensures that P and its first and second derivatives are all continuous.

The parametrization for $T(P)$ requires the additional function g. We use a gaussian[3] transition beginning at $P = P_c$:

$$g(P) = \begin{cases} 1 & P \le P_c \\ e^{-\frac{(P-P_c)^2}{2(\delta P)^2}} & P > P_c \end{cases}, \tag{23}$$

where δP is a small number that controls the width of the current sheet, where the transition of T from a finite value to zero takes place. Note that g and its first derivative are both continuous at the separatrix, so T and T' are well defined.

For the cases (iii) and (iv), the condition $P = P_c$ at the equator is lifted and consequently, the parametrization functions must be modified. Our choice for these cases is:

$$f_b(q,\mu) = f_1(q)\sum_{l=1}^{l_{\max}} \tfrac{b_l}{l} P_l'(\mu) + f_2(q)(1-|\mu|) \tag{24}$$

$$h_b(q,\mu) = q(1-q)(1-\mu^2) \tag{25}$$

$$g(P) = -P\left(\mathrm{ReLU}\left(1-\tfrac{|P|}{P_c}\right)\right). \tag{26}$$

Adjustable weighting functions, denoted as f_1 and f_2, are utilized to control the relative importance of the two terms. To be specific, f_1 should be significant near the surface but diminish at considerable distances from the star, while f_2 should vanish close to the surface but be dominant at $q \ll 1$. At intermediate distances, approximately around the LC, both f_1 and f_2 should be much smaller than h_b to ensure that the neural network contribution in equation (17) dominates.

In order to allow for more versatile configurations beyond the standard dipole representation, the surface boundary condition is

[2] Alternative definitions of ReLU(x) are $x\mathcal{H}(x)$ ($\mathcal{H}$ being the Heaviside step function) and max$(0, x)$.
[3] Any other function that asymptotically approaches the Heaviside step function could be used instead.

Downloaded from https://academic.oup.com/mnras/article/526/1/1504/7276630 by Universidad de Alicante user on 26 February 2024

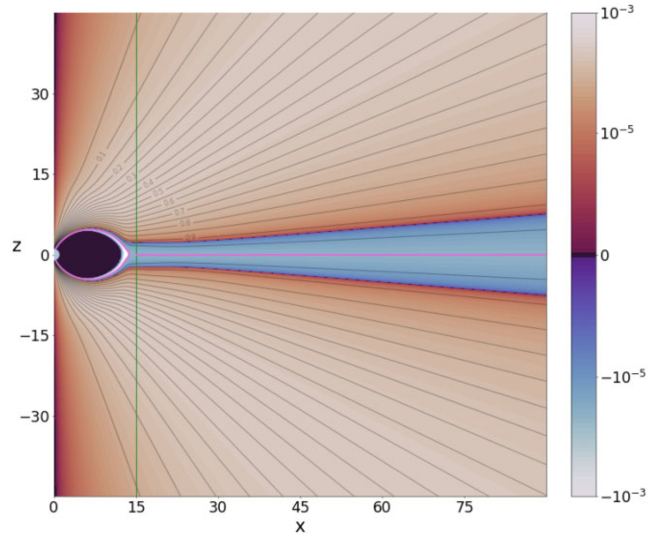


Figure 2. The classical axisymmetric pulsar magnetosphere. Fade black lines: magnetic field lines as contours of P. Thick pink line: separatrix and equatorial current sheet, where $P = P_c$. Vertical green line: LC. Colourmap: the source current $G(P)$. Labels on the field lines are in units of P_c. Colourbar is in symmetrical logarithmic scale. z and x axes are in units of the stellar radius R. The bulk of the return current flows along the current sheet, with a small percentage flowing along open field lines.

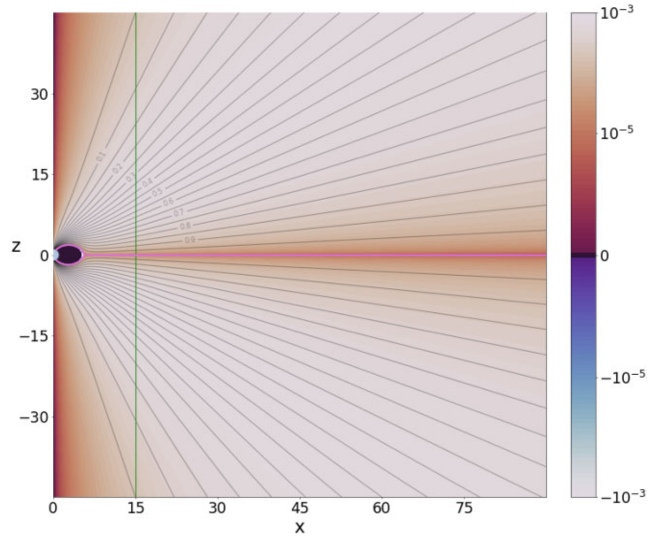


Figure 3. Same as Fig. 2 but with the Y-point positioned at $r_c = 0.4R_{LC}$. The region where the return current flows through open field lines in negligible.

Table 1. Model parameters for the different magnetospheric solutions.

Id.	b_l	r_c/R_{LC}	δP	P_c/P_1	P_∞/P_1	$\Delta\mathcal{E}$	$\dot{\mathcal{E}}$
C1	1,0,0	0.992	0.002	1.23	–	0.048	0.96
C2	1,0,0	0.7	0.002	1.71	–	0.089	1.91
C3	1,0,0	0.4	0.002	3.25	–	0.319	7.08
E1	1,0,0	–	–	1.40	0.77	0.024	0.39–0.73
E2	1,0,0	–	–	1.60	0.86	0.027	0.47–0.89
E3	1,0,0	–	–	1.80	0.95	0.031	0.57–1.10
E4	1,0,0	–	–	2.00	1.01	0.035	0.66–1.32
M1	1,−2,2	–	–	1.40	0.78	0.002[a]	1.94–2.58
M2	1,−1,1	–	–	2.00	1.00	0.009[a]	0.64–1.28
M3	1,−2,2	–	–	2.00	1.01	0.003[a]	1.93–2.57
M4	1,−3,3	–	–	2.00	1.01	0.002[a]	1.94–2.58

[a]*Note.* for the cases with multipolar content, excess energy has been calculated w.r.t a non-rotating magnetic field with the same multipolar content $\mathcal{E}_{\rm mul} = \frac{B_0^2 R^3}{2}\sum_l b_l^2 \frac{l+1}{2l+1}$.

presented as a linear combination of magnetic multipoles. To impose a current-free region, the function g is employed, and while it does not necessarily indicate a current sheet, it is chosen to be at least quadratic. This ensures that T' and consequently G experience at least one change of sign, allowing the current circuit to close.

4.2 Classical solutions

The solutions obtained in Contopoulos et al. (1999); Gruzinov (2005) will be referred to as the *classical solutions*. Fig. 2 illustrates contour lines of the poloidal flux function P, with the colourmap representing the source current $G = TT'$. All the well-known characteristics of the pulsar magnetosphere are observed in these solutions: The open field lines extend beyond the LC and stretch towards infinity, eventually adopting a split monopole configuration at considerable distances, and a current sheet forms at the equator due to the magnetic field's reversal between the North and South hemispheres.

In the region where field lines have $P > P_c$, both the toroidal magnetic field and poloidal current are zero, except for the narrow transition zone $[P_c, P_c + \delta P]$ located just inside the separatrix. Beyond the separatrix, the blue region illustrates the small portion of the return current that flows back to the star along open field lines. In contrast, the nearly white region just inside the separatrix represents the substantial portion of the return current that flows along the current sheet. We summarize in Table 1 the model parameters of different solutions. The first line corresponds to the classical model depicted in Fig. 2.

To showcase the accuracy of our findings, we present in Fig. 4a how well our solution aligns with the pulsar equation (5). The colour map illustrates the absolute error of the pulsar equation for our model, revealing remarkably low values within the bulk of the domain ($\lesssim 10^{-5}$). As expected, the error is slightly larger in proximity to the separatrix, a region of discontinuity. None the less, this discrepancy does not impede our solver from effectively approximating the solution throughout the rest of the domain, nor does it significantly affect the solution far from these regions.

In general, we anticipate the maximum error to be of the order of $\sim\sqrt{\mathcal{L}}$ since it corresponds to the error of the PDE for a considerable set of random points. Indeed, as depicted in Fig. 4b, at the conclusion of the training process, the loss reaches values around $\sim 10^{-7}$, confirming this assumption. The prominent spikes observed in the figure correspond to the periodic changes of the training set of points. However, it is noticeable that as the training epochs progress, the spikes diminish, indicating that the network has learned to generalize to new points without compromising accuracy. The small fluctuations between the spikes, should be interpreted as the variance in the approximation to the solution. As the network adapts its parameters to acquire the solution, it cannot simultaneously reconcile all the training points. Consequently, it fluctuates around a mean instead of finding a single minimum value.

With the same PINN, it is straightforward to produce solutions with different positions of the Y-point, as in Timokhin (2006) simply by varying the parameter q_c in equation (21). Fig. 3 shows an example of such a solution, where the Y-point lies at a distance $r_c = 0.4R_{LC}$. In this case, the totality of the return current flows through the current

Downloaded from https://academic.oup.com/mnras/article/526/1/1504/7276630 by Universidad de Alicante user on 26 February 2024

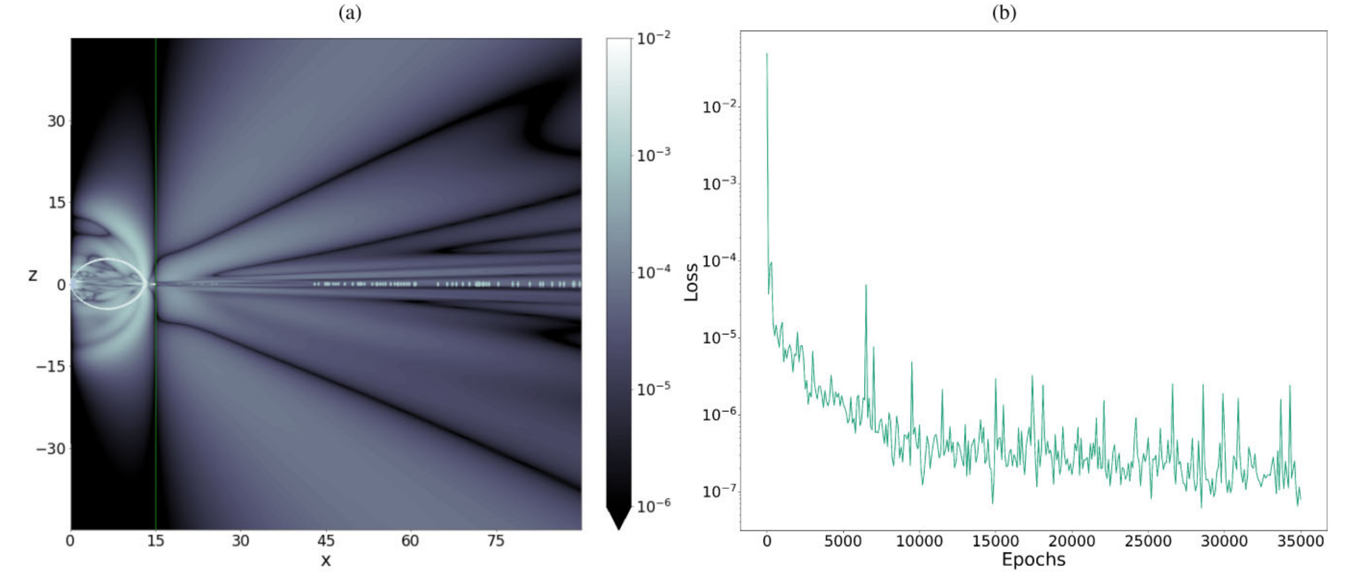


Figure 4. (a) Colourmap of the residuals of the pulsar equation. Colourbar is in log scale. (b) Evolution of the loss function with the training epochs. Only the value of the loss every 100 epochs is plotted for clarity. Big spikes correspond to changes in the training set. Small fluctuations can be interpreted as the variance of the PDE residual.

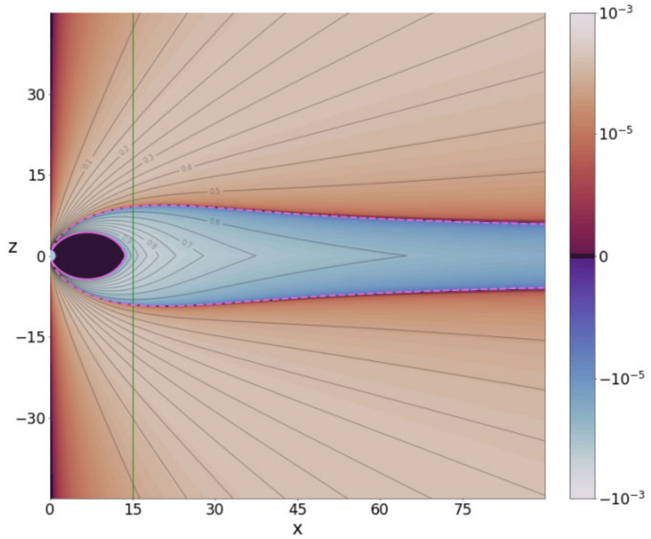


Figure 5. Same as Fig. 2 but for the case of non-fixed equatorial boundary condition. Field lines that cross the LC can close through the current sheet. The solid pink line indicates the line $P = P_c = 1.4P_1$, whereas the dashed pink line corresponds to $P = P_\infty = 0.77P_1$.

sheet. Interestingly, we observe that the luminosity for $r_c = 0.4R_{LC}$ is an order of magnitude larger than the classical solution (see Table 1).

As discussed in Timokhin (2006), both the luminosity and the total energy stored in the magnetosphere increase with decreasing r_c, so the magnetosphere will generally try to achieve the configuration with the minimum energy, that is, with r_c as close as possible to the LC. Therefore, although the configurations with a small r_c are mathematically sound and very interesting from the astrophysical point of view (much larger luminosity), they are probably short-lived and less frequent in nature than the standard configuration.

4.3 Non-fixed boundary condition at the equator

In most previous studies that used classical methods to tackle this problem, it was a common practice to solve it in just one hemisphere and setting boundary conditions at the equator (equation 11), where a kind of 'jump' is expected to occur. This approach is reasonable, but it limits the range of magnetospheric solutions that can be obtained. The solutions have equatorial symmetry (in addition to axisymmetry) and end up having a specific configuration: a dipole magnetic field at the surface and a Y-point configuration where the equatorial and separatrix current sheets meet.

However, by utilizing PINNs and taking advantage of their local and flexible nature for imposing constraints and boundaries, we can create solutions with fewer (only physically relevant, rather than mathematical) requirements. In this section, we introduce solutions where boundary conditions are applied only at the surface and at infinity, without restricting the equator as part of the solution domain.

At infinity, we simply demand that the solution approaches a split monopole configuration (last term in equation 24) with a specific value (denoted by $P = P_\infty$) at the equator:

$$P(\mu, q = 0) = P_\infty(1 - |\mu|).$$

Since the equator is free from any boundary conditions, some other constraint must be imposed to distinguish between a possibly infinite class of different solutions. We decide to set the value of P_c beforehand. This value separates the regions with and without electric currents. This approach is just as valid and perhaps more versatile than other constraints, like pinning down the position of the Y-point. As for P_∞, we do not fix its value; instead we let the network figure it out. The value of P_∞ separates the regions with currents of different sign.

In Fig. 5, we present a typical solution for these boundary conditions. This magnetospheric configuration bears resemblance to the one obtained in Contopoulos et al. (2014), where a substantial number of field lines that cross the LC close inside the equatorial current sheet. However, we have arrived to this result by following a different prescription. They enforced specific boundary conditions at the equator to ensure that the perpendicular component of the

Downloaded from https://academic.oup.com/mnras/article/526/1/1504/7276630 by Universidad de Alicante user on 26 February 2024

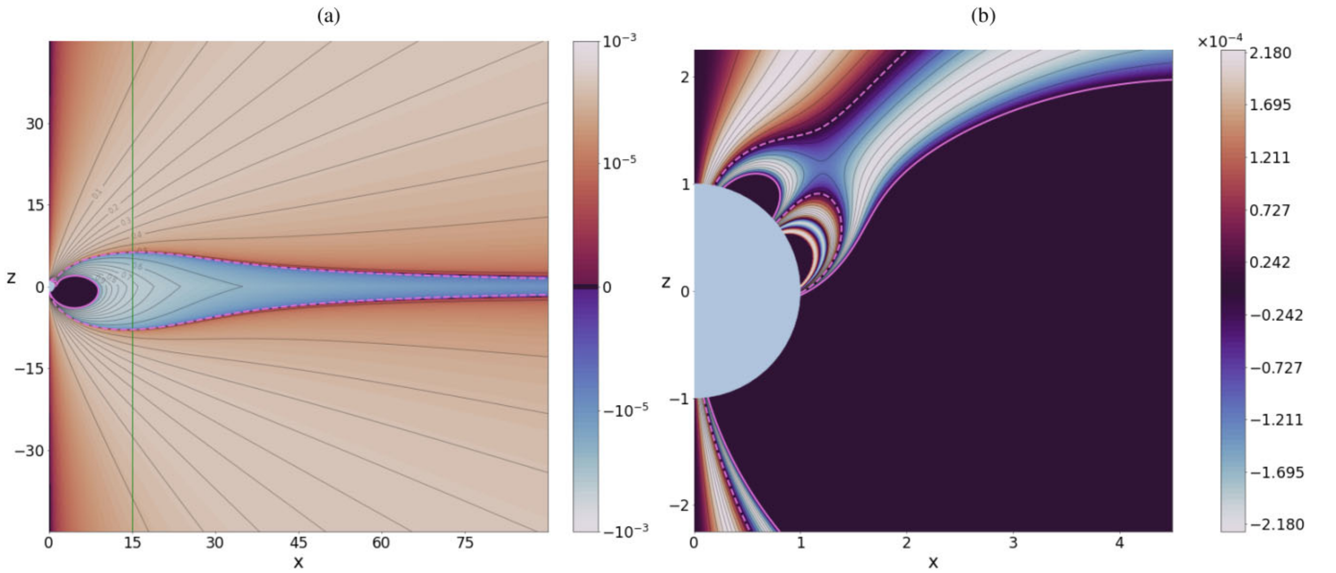


Figure 6. (a) Same as Fig. 5 but for a non-dipolar surface magnetic field. (b) Close up to highlight the multipolar content of the magnetic field close to the star. Notice that (a) is in logarithmic scale, while (b) is in linear scale.

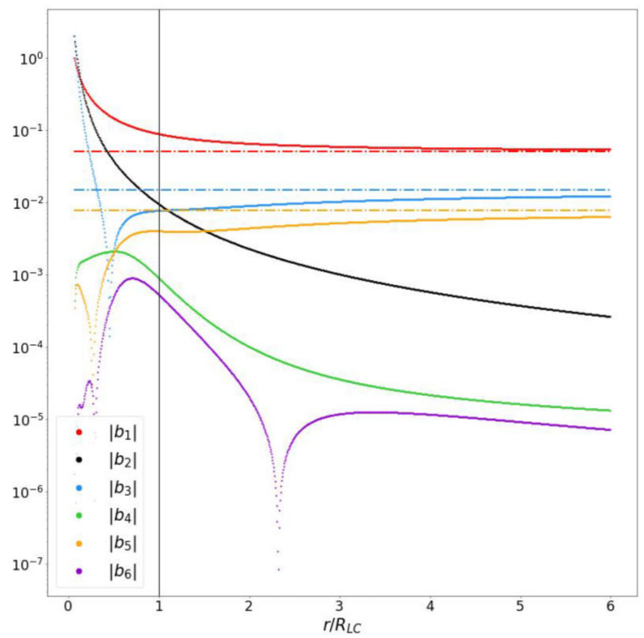


Figure 7. Coefficients b_l of the P polynomial expansion as a function of the radial distance r for the non-dipolar case. Dashed dotted lines correspond to the even coefficients of the split monopole expansion, being zero the odd ones.

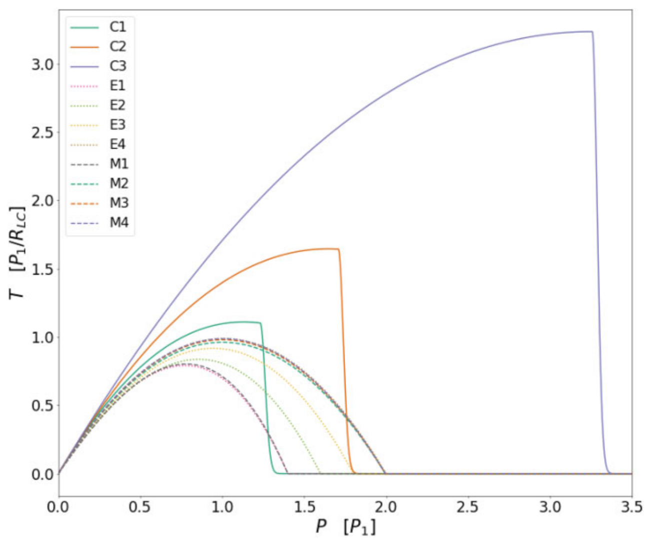


Figure 8. The functions $T(P)$ for the models of Table 1.

Lorentz force applied to the equatorial current sheet becomes zero. On the other hand, our approach involved imposing that the solution becomes a split monopole with a certain magnetic flux at infinity while leaving the equator unrestricted.

An interesting generalization of this set of solutions involves applying non-dipolar surface boundary conditions. Instead of sticking to the basic dipole case, the surface magnetic field can be a combination of various magnetic multipoles (Gralla et al. 2016). This option is not feasible in the classical solution because having contributions of even multipoles breaks the equatorial symmetry. Nevertheless, our solver can handle this situation without any issues, because it is working throughout the entire domain. As we move to significant distances from the star, we anticipate the multipole solution to gradually converge towards the classical dipole solution.

Fig. 6 shows an example of such a case. In particular, we consider a combination of a dipole, a quadrupole, and an octupole, with the corresponding coefficients in equation (24) being $b_1 = 1$, $b_2 = -2$ and $b_3 = 2$. At first glance, when observing the magnetosphere on a larger scale of tens of LC, it appears quite similar to the previous one. However, upon closer inspection of the zoomed-in right panel, the complex and intricate structure in the innermost region of the magnetosphere becomes evident.

To better understand the structure of this last model, in Fig. 7 we show the absolute value of the first six multipole weights computed over spheres at different radii. We see how the even $l = 2$ multipole quickly decreases with distance. Other even multipoles grow in the inner region, but they also become vanishingly small at a distance of a few LC. On the contrary, the odd multipoles approach the asymptotic

Downloaded from https://academic.oup.com/mnras/article/526/1/1504/7276630 by Universidad de Alicante user on 26 February 2024

values corresponding to the split monopole at radial infinity ($P \propto 1 - |\mu|$), shown with dashed lines in the figure.

5 DISCUSSION AND FINAL REMARKS

Table 1 summarizes the various models studied in this work. We use the letter C to denote solutions acquired following the classical approach, the letter E for solutions without an equatorial constraint, and the letter M for solutions with multipolar content. For the first three models, P_c is determined by the PINN, while for the rest of them is fixed to a pre-determined value. P_∞, when present, is always determined by the PINN.

The excess energy for all the models is of the order of a few per cent. Models with no equatorial restriction tend to have lower energies, indicating that they are the preferred configuration in nature. With the exception of the model C3, the luminosity for all models does not vary from the classical dipole case by more than a factor of three.

In Fig. 8, we present the function $T(P)$ for all the models of Table 1. The curves corresponding to models with fixed boundary conditions at the equator exhibit a steep transition to zero, which is modelled by the current sheet (equation 23) and begins at $P = P_c$. In contrast, the models with no restrictions at the equator smoothly reach the value $P = P_c$ through a continuous transition from positive (orange area in Figs 5, 6) to negative current (blue area in Figs 5, 6). Importantly, these models do not develop a current sheet to close the current circuit.

The area under each of the curves represents the luminosity (see equation 15), which explains why the model with a lower r_c exhibits a significantly higher energy loss rate, as it has more open lines carrying Poynting flux to infinity. Additionally, the presence or absence of equatorial restrictions leads to distinct behaviours in the models. In the models without equatorial restrictions, not all the Poynting flux carried by lines crossing the LC escapes to infinity. Instead, up to 50 per cent of the total pulsar spin-down energy flux (the area between P_∞ and P_c) remains confined in the equatorial current sheet. This trapped energy could potentially be dissipated in particle acceleration and high-energy electromagnetic radiation within a few times the LC radius, as was first pointed out in Contopoulos et al. (2014). However, the exact fraction of power that can be locally dissipated and reabsorbed, as opposed to the fraction that is genuinely lost and contributes to the spin-down, remains unclear. In the last column of Table 1, we have included the expected range of values for such models.

In this study, extending the results of our previous paper for slowly rotating magnetar magnetospheres, our success lies in the proficient application of PINNs to obtain numerical results for a diverse range of pulsar magnetospheric models, meticulously exploring various cases under different physical assumptions. Our extensive analysis encompasses the accurate reproduction of several distinct axisymmetric models found in the scientific literature. The decision to focus solely on axisymmetry was a deliberate one, aiming at affirming the remarkable ability of PINNs to effectively capture and characterize the intricate peculiarities exhibited by pulsar magnetospheres, including the features of the current sheet. Throughout our investigation, we purposefully accounted for a rich diversity of models, incorporating non-dipolar configurations.

This work serves as a stepping stone towards the development of a robust and trustworthy general elliptic PDE solver, specifically tailored to address the challenging complexities of this and related astrophysical problems. Looking ahead, our research can be naturally extended to the three-dimensional magnetospheric case, a promising prospect that holds the potential for reaching a deeper understanding of the underlying physics governing pulsar magnetospheres.

ACKNOWLEDGEMENTS

We acknowledge the support through the grant PID2021-127495NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by the European Union, the Astrophysics and High Energy Physics programme of the Generalitat Valenciana ASFAE/2022/026 funded by MCIN and the European Union NextGenerationEU (PRTR-C17.I1) and the Prometeo excellence programme grant CIPROM/2022/13. JFU is supported by the predoctoral fellowship UAFPU21-103 funded by the University of Alicante.

DATA AVAILABILITY

All data produced in this work will be shared on reasonable request to the corresponding author.

REFERENCES

Akgün T., Miralles J. A., Pons J. A., Cerdá-Durán P., 2016, MNRAS, 462, 1894
Akgün T., Cerdá-Durán P., Miralles J. A., Pons J. A., 2018, MNRAS, 481, 5331
Aragon-Calvo M. A., 2019, MNRAS, 484, 5771
Aragon-Calvo M. A., Carvajal J. C., 2020, MNRAS, 498, 3713
Cai S., Mao Z., Wang Z., Yin M., Karniadakis G. E., 2021, Acta Mech. Sin., 37, 1727
Cerutti B., Philippov A., Parfrey K., Spitkovsky A., 2015, MNRAS, 448, 606
Chantada A. T., Landau S. J., Protopapas P., Scóccola C. G., Garraffo C., 2023, Phys. Rev. D, 107, 063523
Contopoulos I., Kazanas D., Fendt C., 1999, ApJ, 511, 351
Contopoulos I., Kalapotharakos C., Kazanas D., 2014, ApJ, 781, 46
Gralla S. E., Lupsasca A., Philippov A., 2016, ApJ, 833, 258
Gruzinov A., 2005, Phys. Rev. Lett., 94, 021101
Jagtap A. D., Kawaguchi K., Karniadakis G. E., 2020, J. Comput. Phys., 404, 109136
Karpov P. I., Huang C., Sitdikov I., Fryer C. L., Woosley S., Pilania G., 2022, ApJ, 940, 26
Kingma D. P., Ba J., 2014, preprint(arXiv:1412.6980)
Korber D., Bianco M., Tolley E., Kneib J.-P., 2023, MNRAS, 521, 902
Lagaris I. E., Likas A., Fotiadis D. I., 1998, IEEE Transactions on Neural Networks, 9, 987
Luna R., Calderón Bustillo J., Seoane Martínez J. J., Torres-Forné A., Font J. A., 2023, Phys. Rev. D, 107, 064025
Michel F. C., 1973, ApJ, 180, L133
Pakravan S., A. Mistani P., Aragon-Calvo M. A., Gibou F., 2021, J. Comput. Phys., 440, 110414
Pétri J., 2012, MNRAS, 424, 605
Philippov A., Kramer M., 2022, ARA&A, 60, 495
Philippov A. A., Spitkovsky A., 2018, ApJ, 855, 94
Raissi M., Perdikaris P., Karniadakis G. E., 2019, J. Comput. Phys., 378, 686
Scharlemann E. T., Wagoner R. V., 1973, ApJ, 182, 951
Spitkovsky A., 2006, ApJ, 648, L51
Stefanou P., Pons J. A., Cerdá-Durán P., 2023, MNRAS, 518, 6390
Timokhin A. N., 2006, MNRAS, 368, 1055
Urbán J. F., Stefanou P., Dehman C., Pons J. A., 2023, MNRAS, 524, 32 (Paper I)

This paper has been typeset from a TEX/LATEX file prepared by the author.

© 2023 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/526/1/1504/7276630 by Universidad de Alicante user on 26 February 2024