

Diseño y optimización de un robot de combate

Metodología, resultados y límites de una optimización en simulación

COMBATBOTS

combatbots.ai

Resumen. Se presenta el diseño y la optimización del arma de un robot de combate de la clase 500 g mediante simulación de sólidos rígidos, y la metodología que hace que esa medida sea defendible. El instrumento es un banco de impacto contra una diana instrumentada de masa e inercia conocidas, con el arma accionada **por par** y no por velocidad, paso de integración fijado por estudio de convergencia y el ángulo de contacto **marginalizado** sobre 32 repeticiones. Sobre esa base se rediseña el tambor: el criterio que gobierna no es la masa, que no resulta ser la restricción activa, sino su distribución radial. La masa efectiva en la punta sube de 28,6 a 43,8 g y la energía entregada de $1,80 \pm 0,12$ a $3,53 \pm 0,33$ J, un 96 %, confirmado con 32 repeticiones tras descartar el sesgo de selección del cribado. El diseño resultante gana los diez combates de un torneo reglamentario. Se reportan además tres resultados metodológicos con consecuencias para cualquier optimización de este tipo: la dispersión inducida por el ángulo de contacto (factor 3,7) supera a la diferencia entre diseños; la métrica sustituta del banco no ordena igual que el resultado en arena; y cuatro de siete configuraciones de contacto declaran un ganador distinto. Ninguna cifra se ha validado contra ensayos físicos, y se detalla el protocolo que lo haría.

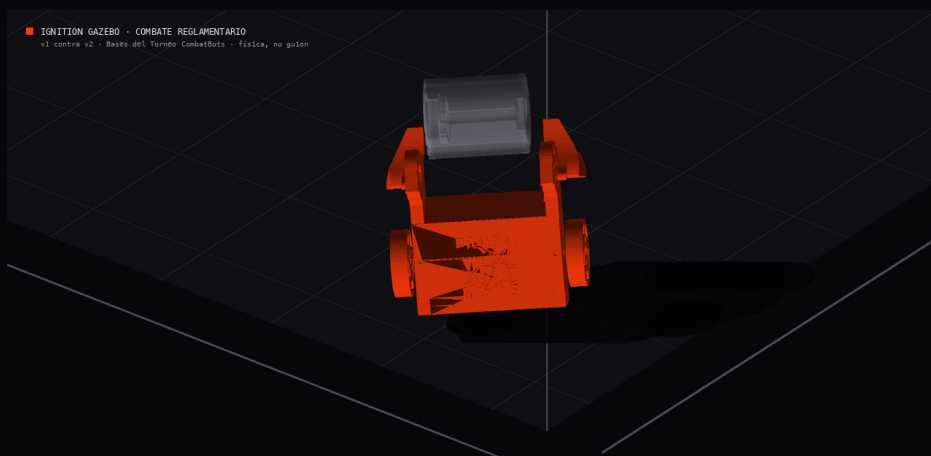
Palabras clave: robótica de combate, optimización basada en simulación, dinámica de contacto, diseño de rotores, validación de simuladores, brecha simulación-realidad

1 Introducción

El arma dominante en la robótica de combate ligera es el rotor de inercia: una masa que gira a gran velocidad y entrega en el impacto la energía que ha ido acumulando. Diseñarla bien es un problema de optimización con restricciones —masa total limitada por reglamento, tensión de aro limitada por el material, tiempo de arranque limitado por el motor— y con una función objetivo incómoda de medir, porque sólo se manifiesta en un choque de milisegundos.

Este artículo presenta el diseño y la optimización del arma de un robot de la clase 500 g enteramente en simulación, y la metodología que hace que el resultado sea defendible. El objetivo declarado es la **energía transferida a un blanco normalizado**, y el resultado es una mejora del 96 %: de $1,80 \pm 0,12$ J a $3,53 \pm 0,33$ J, confirmada con 32 repeticiones. El diseño resultante gana los diez combates de un torneo bajo reglamento [5].

La contribución tiene dos partes, y la segunda vale más que la primera. La primera es el diseño: **el criterio que gobierna no es la masa sino su distribución radial**,



Fotograma de un clip. Vídeo en

<https://combatbots.ai/research/banco-de-impacto/media/torneo-golpe.mp4>

Figura 1. El golpe mayor del torneo, a $8\times$ cámara lenta: el diseño optimizado descarga $55,2\text{ J}$ en una sola mordida, el 85 % de la energía almacenada en su rotor.

y se muestra que en esta clase el límite reglamentario de peso no es la restricción activa. La segunda es metodológica: qué hace falta para que una medida simulada de impacto signifique algo. Aquí se documentan cuatro requisitos —accionamiento por par, paso de integración convergido, marginalización del ángulo de contacto y calibración contra la pieza real— cada uno de los cuales, omitido, produce un artefacto **mayor que la mejora que se busca medir**.

De la validación del instrumento salen además tres resultados con consecuencias para cualquier optimización de este tipo, y se reportan como tales: la dispersión inducida por el ángulo de contacto supera a la diferencia entre diseños; la métrica sustituta del banco no ordena igual que el resultado en competición; y cuatro de siete configuraciones de contacto declaran un ganador distinto sobre los mismos modelos. Ninguna cifra se ha validado contra ensayos físicos; el protocolo que lo haría se detalla al final.

2 Trabajo relacionado

El diseño de robots de combate apenas existe como objeto de investigación. El rastreo bibliográfico devuelve trabajos docentes de grado, centrados en el diseño conceptual y la fabricación más que en la optimización [17], y un manual técnico de comunidad que sigue siendo la referencia del campo dos décadas después [23]. No hemos encontrado ningún trabajo previo que aplique optimización paramétrica basada en simulación de sólidos rígidos al arma de un robot de esta clase. Ése es el hueco que este artículo ocupa, y conviene decir que es un hueco pequeño: la escasez de literatura mide sobre todo la marginalidad académica del dominio.

Y conviene preguntarse por qué ese vacío existe, porque la comunidad técnica es grande, activa y produce conocimiento sofisticado. La explicación no parece ser la ética. La academia es en efecto sensible a la investigación con connotación armamentística —el boicot a una universidad por un centro de investigación en defensa, en 2018, lo demostró— pero aquella reacción se dirigía a sistemas **letales y autónomos**, y se levantó cuando la institución aclaró que no los desarrollaría. Un robot de arena, radiocontrolado y de entretenimiento, no cumple ninguno de esos criterios. El debate sobre doble uso en robótica abierta existe y es serio [27], pero no hemos encontrado documentación de que revistas o comités rechacen trabajos sobre combate deportivo por esa razón; afirmarlo sería conjetura.

Lo que sí está documentado es más prosaico y tiene que ver con **cómo nace una competición**. RoboCup se diseñó explícitamente como un “standard problem” de investigación [18], con banco de pruebas compartido y actas indexadas, y genera desde entonces producción académica medible [25]. El combate robótico nació como espectáculo. Esa diferencia de origen fija los incentivos de todos los participantes: no hay problema formulado como referencia común, no hay plataforma de publicación, y sus constructores no tienen carrera académica que alimentar. A eso se suma una barrera estructural de toda la robótica —la dificultad de reproducir experimentos ajenos, que la literatura de bancos de prueba documenta con crudeza [33]— agravada aquí porque cada robot es una pieza única.

El conocimiento, en consecuencia, no falta: **circula fuera del circuito revisado por pares**, en manuales de comunidad [23], wikis y foros. Es funcional para quien compite e invisible para quien busca en una base de datos académica. Este artículo intenta trasladar una parte de ese conocimiento al formato que permite discutirlo.

El análogo físico maduro es el volante de inercia. Un tambor y un rotor de almacenamiento comparten el problema de repartir masa contra un límite de tensión, y la literatura de almacenamiento cinético lo tiene formalizado en términos de energía específica y topología de rotor [21]. La analogía es útil pero **no es directa**, y la diferencia importa: un volante optimiza la energía *almacenada* por unidad de masa, lo que empuja hacia perfiles de tensión uniforme; un arma de combate optimiza la energía *transferida* en el impacto, que depende de la inercia vista desde el punto de contacto. Los dos objetivos no coinciden, y nuestros datos lo muestran (§5).

La métrica correcta viene de la mecánica del impacto. Cross derivó la masa efectiva en el punto de golpeo de un cuerpo extendido, $m_{ef} = M/(1 + r^2/k^2)$, para bates y raquetas [10], y en robótica de seguridad la masa efectiva junto con la velocidad son los dos parámetros que gobiernan la severidad de un impacto [16]. Aplicarla como función objetivo del diseño de un arma es, hasta donde sabemos, una traslación nueva.

Sobre el instrumento, el modelado de contacto rígido frente a blando y el papel del coeficiente de restitución están revisados desde hace tiempo [15], y la validación de simuladores contra impactos reales medidos es un problema activo: se han evaluado varios simuladores sobre impactos instrumentados [1] y se ha cuantificado un error medio del 3,1 % entre el salto de velocidad simulado y el medido [34]. Esa cifra es la vara con la que habrá que medir la validación física pendiente de este trabajo.

Y una delimitación necesaria. Existe una tradición larga de codiseño de morfología y control, desde las criaturas evolucionadas en simulación [30] y su fabricación automática [22] hasta el aprendizaje conjunto de diseño y política [28]. **Este trabajo no es codiseño:** el control se mantiene fijo —el mismo piloto autómatas para todos los diseños— y lo que se optimiza es sólo la morfología del arma. La comparación es pertinente, pero el alcance es más estrecho.

3 El banco de impacto

3.1 Qué mide, y contra qué

El banco [4] enfrenta un robot a una **diana**: un ortoedro de 454 g —la masa de un rival de la clase— de $100 \times 100 \times 60$ mm, siempre en el mismo sitio.

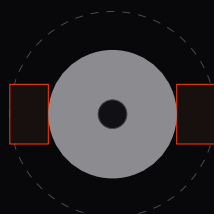
Que la diana sea un ladrillo y no un robot es la primera decisión, y es deliberada. Un rival de verdad tiene ruedas y geometría que engancha, así que dos armas distintas producirían impactos en sitios distintos y la comparación mediría **dos cosas a la vez**. Un ortoedro de masa e inercia conocidas hace que toda la diferencia medida salga del arma.

La secuencia la gobierna un plugin sobre tiempo de simulación: el arma coge revoluciones con el robot quieto, a los 0,60 s entra la tracción, hacia los 0,85 s se produce el impacto, y el registro termina a 1,60 s.

De todo lo que se registra, la métrica que decide es la energía que se lleva la diana. Y hay una segunda que no decide pero vigila: la que **pierde** el tambor, $\frac{1}{2}I(\omega_0^2 - \omega_1^2)$. Las dos tienen que moverse juntas, porque dependen de cosas distintas: la de la diana sale de cómo el solucionador reparte el impulso —la parte menos fiable de cualquier simulador de sólidos rígidos, con límites de predicción que no se arreglan calibrando mejor [13]— y la caída de revoluciones sale de la conservación del momento angular. Cuando divergen, lo que se está midiendo es el solucionador.

V1 línea de base: el kit tal cual está

SECCIÓN DEL TAMBOR



Ø35.74 · punta 28.58 · pared 13.9

Ø CUERPO	35.74 mm
Ø BARRIDO	57.16 mm
PARED	13.87 mm
DIENTES	2 · ancho 16.6 mm
MASA EFECTIVA	28.6 g

ALZADO A ESCALA · LA DIANA EN SU POSICIÓN DE ENSAYO



diana · 454 g

MASA TOTAL	426 g
VELOCIDAD DE PUNTA	40.9 m/s
ENERGÍA ALMACENADA	23.97 J

Figura 2. La línea de base. A la izquierda, la sección del tambor de v1: el barrido punteado llega a 28,6 mm, pero el cuerpo (el aro gris) es pequeño y la masa vive cerca del eje. A la derecha, el montaje del ensayo a escala, con la diana en su posición fija. Las cotas completas de todos los diseños están en la **tabla A.1**.

V2 tambor anular Ø54 vaciado, punta a 38 mm

SECCIÓN DEL TAMBOR



MASA TOTAL	426 g
VELOCIDAD DE PUNTA	54.4 m/s
ENERGÍA ALMACENADA	64.84 J

Figura 3. El mismo par de vistas para v2. El cuerpo del tambor pasa de Ø36 a Ø54 y la pared adelgaza de 13,9 a 6,2 mm: la misma masa, más lejos del eje. La masa efectiva en la punta sube de 28,6 a 43,8 g y la energía almacenada de 24,0 a 64,8 J.

3.2 El arma se mueve por par, no por velocidad

Ésta es la decisión del banco, y salió de un fallo.

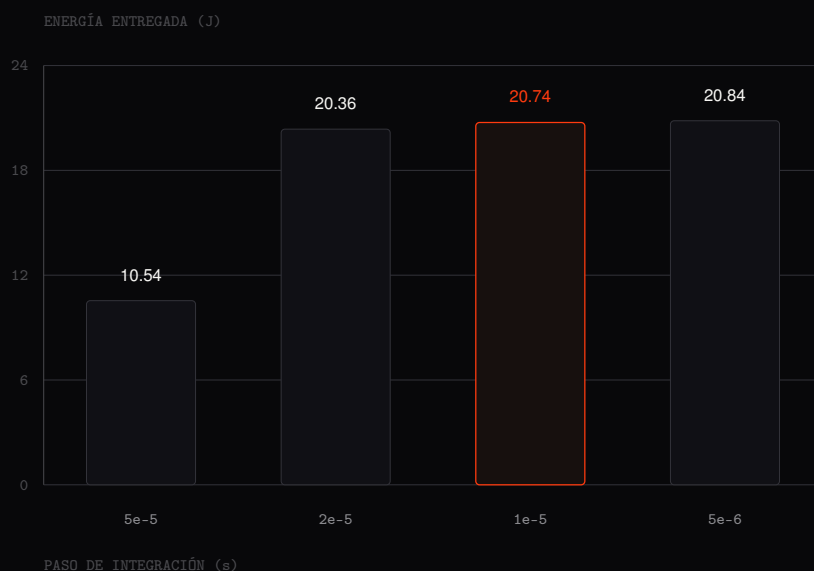
Los controladores que trae el simulador imponen la velocidad de la junta: el motor de física inyecta el par que haga falta, sin techo. En un tambor eso significa que **durante el impacto el arma no pierde ni una revolución y entrega energía ilimitada**. Un tambor con el doble de inercia mide igual que uno con la mitad, porque la energía no sale de la inercia sino del controlador.

El controlador por defecto convertía el arma en una fuente infinita de energía. Con él, medir la inercia del tambor era medir el simulador.

Por eso hay un plugin propio en C++ con modelo de motor de continua y límite de corriente. No es una mejora de realismo: sin él, el experimento no tiene sentido. Es un caso extremo de algo ya descrito en robótica evolutiva: los diseños que mejor puntúan en simulación tienden a explotar los fenómenos peor modelados [19].

3.3 El paso de integración sale de un estudio, no de un criterio

A 1.430 rad/s la punta del tambor va a 41 m/s. Con el paso habitual de 10^{-4} s avanza 4,1 mm por paso: el arma **atraviesa la diana sin detectar contacto**.



A $5 \cdot 10^{-5}$ la punta avanza 2,05 mm por paso y atraviesa la diana.

$1 \cdot 10^{-5}$ y $5 \cdot 10^{-6}$ coinciden en el 0,5 %: converge ahí.

Figura 4. Energía entregada frente al paso de integración. No es una curva sobre el arma: es una curva sobre el simulador.

Con el paso mal elegido, el arma parecería la mitad de buena de lo que es. Y el error no avisa: la simulación corre, escribe números y no se queja.

3.4 La fase del diente se marginaliza, no se elimina

El simulador es determinista: la misma orden da el mismo resultado al bit. Pero a 1.348 rad/s y con dos dientes pasa un diente cada 2,3 ms, y el robot avanza a un metro por segundo. Mover la diana 2,3 mm cambia con qué diente y con qué ángulo se produce el contacto.

Ésta es la figura que más dice del artículo entero.



Figura 5. Las 32 tiradas de cada finalista. La caja es la media con su error estándar; los extremos, el recorrido completo.

Sobre v1, con 32 tiradas (tabla A.2), la energía entregada va de 0,65 a 2,41 J: un factor 3,7 entre la peor y la mejor, siendo el mismo robot, el mismo golpe y el mismo simulador. En el ganador del bucle el recorrido es de 1,25 a 10,98 J, factor 8,8; en uno de los candidatos llega a 17,9.

Dicho de otra forma: **la variación dentro de un diseño es varias veces mayor que la diferencia entre diseños**. Cualquier comparación basada en una sola tirada es ruido con formato de dato.

La tentación es fijar la fase, y sería un error: en un combate nadie elige con qué diente pega. Así que no se elimina, se **marginaliza** — ocho posiciones en el bucle, treinta y dos en la confirmación, y se toma la media con su error estándar. Cuántas réplicas dedicar a cada candidato frente a cuántos candidatos explorar es un problema con teoría propia [3]; aquí se resolvió a ojo.

3.5 El modelo paramétrico se calibra contra la pieza real

El generador describe el tambor con sólidos elementales; la pieza real tiene nervios y vaciados. Si el modelo no reprodujera a v1, la comparación mediría la diferencia entre dos **modelos** y no entre dos **diseños**. El ajuste iguala a la vez la masa (86,31 g) y la inercia ($2,3211 \cdot 10^{-5}$ kg·m²) medidas sobre la malla real, con residuo del 0,00 % en ambas.

Y al calibrar salieron dos cosas sobre v1 que nadie había mirado. Primera: el “tambor Ø57” de v1 **no es un tambor Ø57**. Sólo el 1,4 % de sus vértices pasa de 25 mm de

radio y la mediana está en 14,5 mm. Segunda, y peor porque es un error nuestro: la simulación de v1 **sobrestimaba la inercia de su arma un 21 %**, porque el generador repartía la masa no descrita en CAD de forma homogénea y los 18 g del tornillo pasante van **en el eje**, donde no aportan inercia. De 104 g de tambor, el rival nota 28,6.

3.6 Veinte alteraciones, dos que funcionaron

Con el instrumento en pie, la primera búsqueda se hizo a mano: ocho rondas, cada una con una hipótesis explícita —dónde va la masa, cuánta, revoluciones, dientes, radio de impacto, altura del golpe— y de dos a cuatro candidatos por ronda. Veinte ensayos en total.



Cada barra es un candidato, contra el diseño vigente en su ronda.

En rojo, el único que ganó su ronda. Cinco de siete rondas no produjeron ninguna mejora.

Figura 6. Cada barra es un candidato frente al diseño vigente en su ronda. Las cotas completas de los veinte, con dispersión y caída de revoluciones, están en la **tabla A.3**.

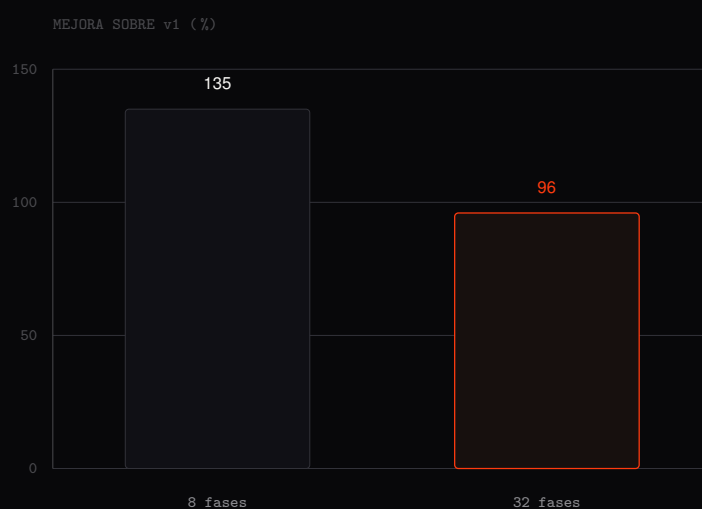
El resultado importa menos que su forma. **Cinco de las siete rondas no produjeron ninguna mejora**, y dos alteraciones que parecían obvias salieron claramente peor: subir a seis dientes costó un 55 % de energía, y ampliar el presupuesto de masa a 500 g con un tambor más largo, un 16 %. Casi todo lo que se prueba empeora, que es la forma normal de una búsqueda de diseño y no la que suelen contar los informes.

Dos rondas sí funcionaron, y entre las dos hacen el salto entero: **llevar el cuerpo del tambor a Ø54 (+107 %)** y **la punta a 38 mm (+13,5 %)**. La primera es la que confirma la tesis de la sección anterior: no se añadió masa, se movió.

Merece la pena mirar una fila que no ganó. En la ronda de revoluciones, subir a 3S y 11,4 V dio +17,8 % — pero con una dispersión del 182 %, la mayor de toda la tanda. Un candidato con esa varianza no es una mejora: es una tirada afortunada esperando a que la remidan.

3.7 La maldición del ganador

Ocho fases bastan para ordenar candidatos, pero no para publicar una cifra. Quedarse con el máximo de un conjunto ruidoso **lo sobreestima siempre**. Smith y Winkler lo llamaron *la maldición del optimizador*: al elegir por el valor estimado más alto, el valor real del elegido es sistemáticamente menor que su estimación, porque la selección premia el error favorable [31]. Es el mismo mecanismo que hace que ajustar hiperparámetros contra el conjunto de validación sobreestime lo ganado [2].



El mismo diseño, remedido. La cifra que se publica es la segunda.

Quedarse con el máximo de un conjunto ruidoso lo sobreestima siempre.

Figura 7. La mejora de v2 sobre v1, medida dos veces sobre el mismo diseño.

Cuarenta puntos porcentuales de la mejora inicial eran el método, no el arma.

4 Del kit al diseño nuevo

4.1 El espacio de diseño

El espacio de diseño tiene una decena de parámetros: radio del cuerpo del tambor, radio de la punta, anchura, número de dientes, espesor de pared, diámetro de rueda, vía, altura del arma y geometría del morro.

Un barrido de 33.544 diseños válidos dejó un hallazgo que contradice la intuición de la categoría: **la masa no es la restricción activa**. El frente de mejores diseños es el mismo con 426 g que con 500 g. Lo que limita no es el peso permitido sino el tiempo

que el arma tarda en coger revoluciones, $\tau = I \cdot R/k^2$. Los gramos que sobran van al chasis y a las ruedas.

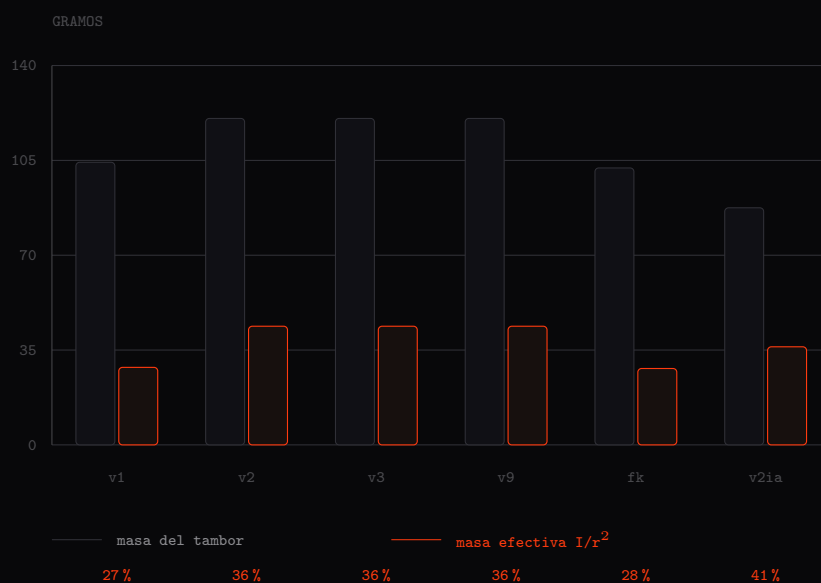


Todos a la misma escala. La masa efectiva no sigue al diámetro: sigue a cuánta masa hay lejos del eje.

Figura 8. Los tambores de todos los contendientes en sección y a la misma escala. De v1 a v2 el diámetro del cuerpo crece de 35,7 a 54 mm y la pared adelgaza de 13,9 a 6,2: la misma masa, más lejos del eje.

Y ahí está el criterio de diseño de un arma de inercia, que no es el peso sino **dónde está el peso**. La magnitud que decide es la masa efectiva en el punto de golpeo, I/r^2 : la inercia que el rival nota. No es una métrica inventada aquí — es la que Cross formaliza para bates y raquetas [10] y la que en robótica de seguridad gobierna, junto con la velocidad, la severidad de un impacto [16]. Lo que aquí se hace es usarla como **función objetivo de diseño** en lugar de como métrica de riesgo.

El contraejemplo lo pone el rival externo. ForceKnight tiene **el tambor de mayor diámetro de todos, Ø60**, y una de las masas efectivas más bajas, 28,2 g — porque su pared es de 26 mm y casi toda esa masa está pegada al eje. Un tambor grande y macizo rinde menos que uno menor y hueco: v2, con Ø54 y 6,2 mm de pared, entrega un 55 % más.



Debajo, qué fracción de la masa del tambor nota el rival.

De los 120 g de tambor de v2, el rival nota 44.

Figura 9. Masa del tambor frente a masa efectiva. Dos tambores de peso casi idéntico entregan masas efectivas muy distintas, y ninguno llega a la mitad.

4.2 El bucle automático

A partir de v2, la búsqueda se automatizó. Cada iteración propone cambios, criba con ocho fases, remide al elegido con 32, lo lleva a la arena contra el mejor anterior y contra el rival externo, y decide [7]. La parada exige **dos cosas a la vez**: tres fallos consecutivos y haber batido al rival externo. Corrió nueve iteraciones y paró por reloj sin cumplir ninguna.

Iteración	Hipótesis	Energía (J)	Δ padre	Resultado	Minutos
v3	ruedas 60 via 150	3.61	+2.2 %	mejora	67
v4	monodiente 38	4.18	+15.8 %	fallo	33
v5	ruedas 72 rampa 112	3.53	-2.2 %	no concluyente	56
v6	boca 88 ruedas 72 rampa 112	3.64	+0.7 %	fallo	38
v7	estabilidad 500 morro 97	4.19	+16.0 %	no concluyente	53
v8	tambor 100 pared 7.5	2.72	-24.6 %	fallo	30
v9	vientres panza 6 rampa 112	3.44	-4.7 %	no concluyente	59
v10	rampa 112 anillo fino 3.5	3.66	+1.3 %	fallo	49
v11	rampa 112 alcance 163	3.02	-16.4 %	fallo	33



Relleno: la ronda se dio por mejora. Vacío: fallo o no concluyente.

Figura 10. Las nueve iteraciones, con v1, v2 y el rival externo como referencias.

Mírese la fila de **v4**: la energía de banco más alta de todo el bucle, 4,18 J, un 15,8 % sobre su padre — y declarada **fallo**, porque en la arena no batió al campeón. El banco y el juez ya estaban discrepando en la segunda iteración, y nadie lo leyó como lo que era.

5 La arena

5.1 El torneo: v2 contra v1

Con v2 cerrado, se enfrentó a v1 bajo el reglamento: diez combates de hasta 120 s con parada por KO, la misma arena, el mismo piloto automático y la puntuación del §3 —agresión, control y daño, máximo 17 puntos por combate.

Combate	Duración	KO	Puntos v1	Puntos v2	Golpe mayor v1	Golpe mayor v2
c00	12.24 s	v1	6	11	19.7 J	52.1 J
c01	16.95 s	v1	6	11	18.7 J	48.5 J
c02	23.95 s	v1	7	10	19.9 J	45.4 J
c03	18.01 s	v1	7	10	35.6 J	54.4 J
c04	19.43 s	v1	7	10	20.7 J	54.1 J
c05	18.28 s	v1	7	10	19.0 J	45.5 J
c06	29.32 s	v1	7	10	20.8 J	50.4 J
c07	21.61 s	v1	7	10	20.9 J	48.8 J
c08	16.62 s	v1	6	11	21.2 J	51.8 J
c09	19.73 s	v1	7	10	20.9 J	55.2 J
Total		10–0	67	103		

v2 gana los diez. Ninguno se decide por inmovilidad: **los diez KO son por destrucción acumulada**. El combate más corto dura 12,24 s y el más largo 29,32 s.

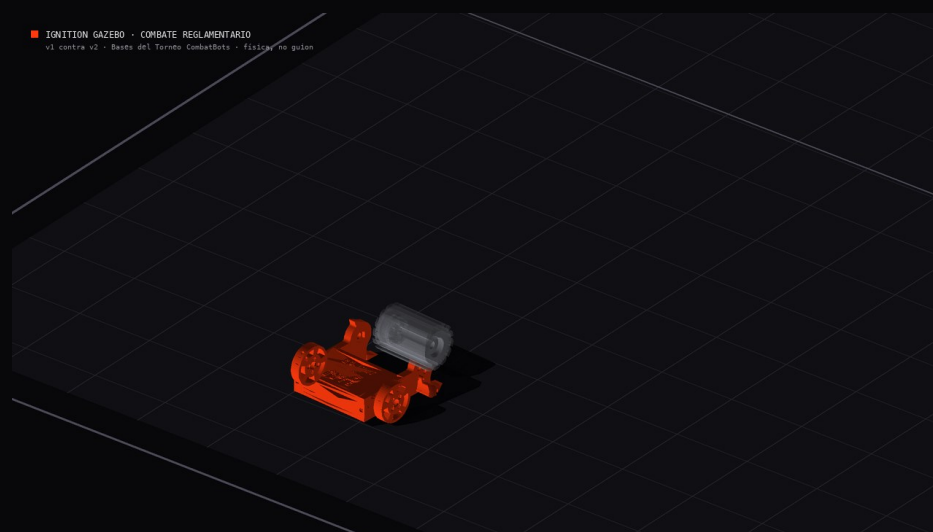


Figura 11. Puntos de cada combate. La ventaja es constante, no depende de la semilla.



El mayor de la tanda es de v2, 55,2 J. v1 no pasa de 35,6 J en ningún combate.

Figura 12. El golpe de mayor energía de cada combate. v1 no pasa de 35,6 J en ninguno; v2 llega a 55,2 J.



Fotograma de un clip. Vídeo en

<https://combatbots.ai/research/banco-de-impacto/media/torneo-completo.mp4>

Figura 13. Un combate entero, de la salida al KO: los 29 s del más largo de los diez, a velocidad real y con la arena completa en cuadro. v1 encaja los 240 J de destrucción y queda fuera de combate. Los rebotes contra los paneles son restitución 0,40: el recinto devuelve el golpe.

El §3.3 traduce energía a daño con un modelo declarado, y ese modelo es una hipótesis. Así que se repitió la tanda escalándolo a la mitad y al doble.



El modelo de daño §3.3, escalado a la mitad, al doble y sin tocar.

El resultado aguanta, pero a la mitad de severo v1 arranca un combate.

Figura 14. El mismo torneo con el modelo de daño a $\times 0,5$, $\times 1$ y $\times 2$.

El resultado aguanta, pero no es indiferente: **a la mitad de severo, v1 arranca un combate**. Y hay un aviso que no se puede ocultar: en las diez semillas **no se corrió el orden inverso**, así que el desvío de rumbo del piloto no se cancela y ese sesgo queda dentro del resultado.

5.2 La liguilla, y el desorden que devuelve

La liguilla —todos contra todos bajo el reglamento— devolvió el resultado que obliga a repensar el montaje.

Puesto	Diseño	Ganados	Perdidos	Qué es
1	v9	5	0	descartado pese a batir a ForceKnight
2	v3	3	2	el campeón del bucle
3	v5	3	2	descartado pese a batir a ForceKnight
4	v2	3	2	el kit publicado
5	fk	1	4	ForceKnight, el rival externo
6	v2ia	0	5	el Kit AI v2, fuera de clase

PUESTO DISEÑO GANADOS PERDIDOS

1	v9	5	0	descartado pese a batir a ForceKnight
2	v3	3	2	el campeón del bucle
3	v5	3	2	descartado pese a batir a ForceKnight
4	v2	3	2	el kit publicado
5	fk	1	4	ForceKnight, el rival externo
6	v2ia	0	5	el Kit AI v2, fuera de clase

v9 gana los cinco enfrentamientos; no todos con margen concluyente.

Figura 15. La clasificación por enfrentamientos ganados.

EL BANCO ORDENA ASÍ		LA ARENA, ASÍ	
v3	3.61 J	1	v9
v2	3.532 J	2	v3
v5	3.529 J	3	v5
v9	3.442 J	4	v2
fk	2.904 J	5	fk
		6	v2ia

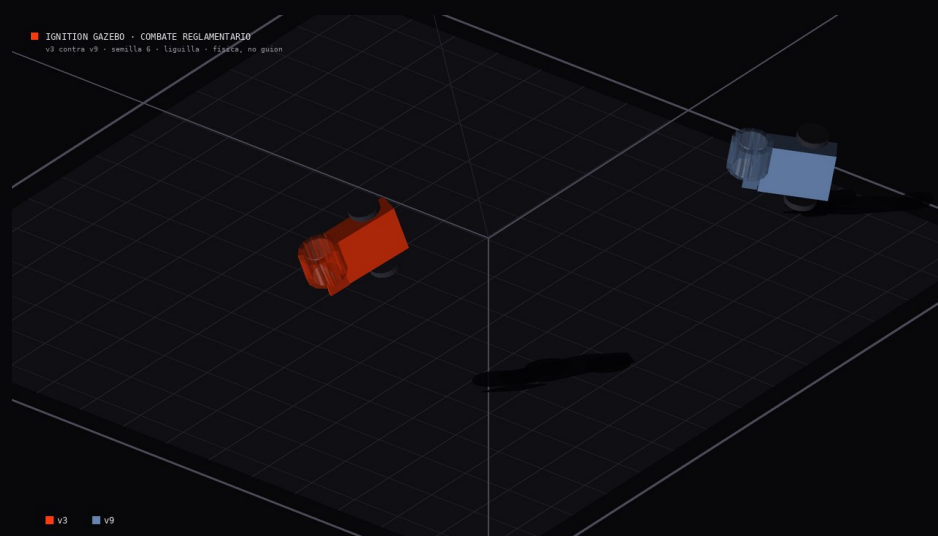
v9 entrega MENOS energía que v3 en el banco y aun así lidera la liguilla.

La métrica sustituta no ordena como el resultado.

Figura 16. Las mismas versiones ordenadas por el banco y por la arena. No coinciden.

v9 gana los cinco enfrentamientos y encabeza con un margen de liga de $+1,60 \pm 0,50$ puntos, y sin embargo entrega **menos** energía a la diana que v3 (3,44 frente a 3,61 J). La métrica sustituta no ordena como el resultado.

Y hay algo peor para cualquier bucle que escale una sola cima: la liguilla contiene un **ciclo**.



Fotograma de un clip. Vídeo en

<https://combatbots.ai/research/banco-de-impacto/media/liguilla-v3-v9.mp4>

Figura 17. v3 contra v9: el que menos energía entrega en el banco gana aquí.

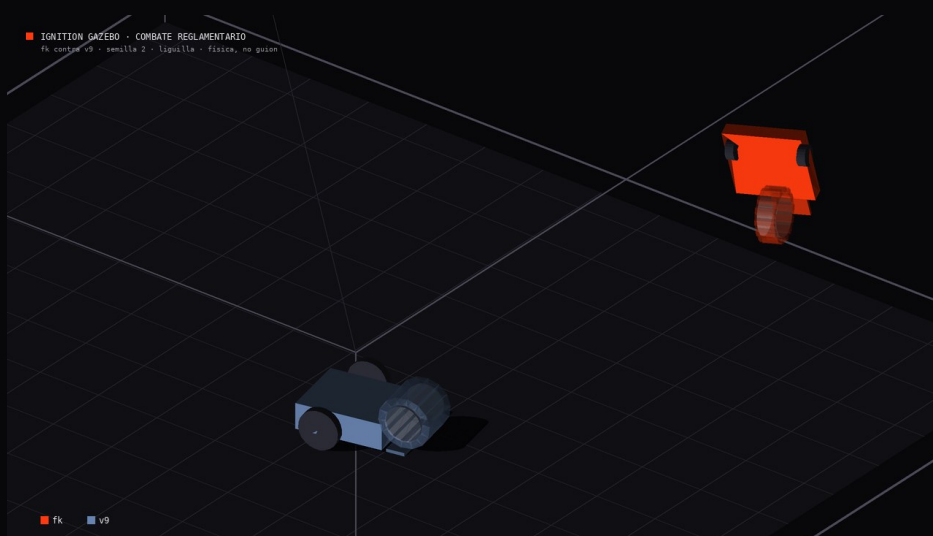


v2 gana a v5, v5 gana a v3, v3 gana a v2

Con ciclos, «mejor que» no ordena.

Figura 19. v2 gana a v5, v5 gana a v3, v3 gana a v2. Sólo uno de los tres márgenes es concluyente.

Con ciclos, “mejor que” no ordena. No es un problema nuestro: la coevolución lleva décadas describiendo estados estables mediocres y ciclos intransitivos, y las soluciones conocidas pasan por mantener un archivo de rivales en vez de un solo campeón [14]. Un ranking escalar tampoco lo arregla, porque un sistema tipo Elo no puede representar relaciones no transitivas [20].



Fotograma de un clip. Vídeo en

<https://combatbots.ai/research/banco-de-impacto/media/liguilla-fk-v9.mp4>

Figura 18. v9 contra el rival externo, éste sí con margen concluyente.

Lo que esto invalida del método

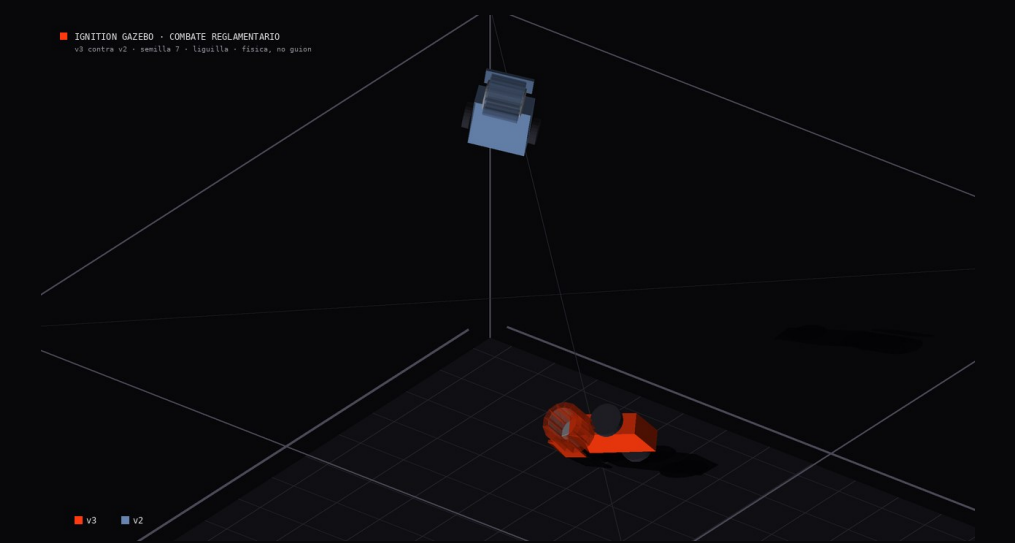
El bucle usaba el banco como criba barata y la arena como juez, y adoptaba sólo lo que batía al campeón vigente. Con el banco desordenando y con ciclos en la arena, esa regla descartó diseños que sí baten al rival externo — v5 y el propio v9 entre ellos. No es un matiz: es una razón para rehacer el criterio de selección, y está sin rehacer.

Conviene además dimensionar la confianza que merece una liguilla de este tamaño. **Seis de los quince emparejamientos no alcanzan el listón** de la t unilateral al 95 %, las semillas no están repartidas por igual —de 5 a 12 según la pareja— y uno de los seis contendientes, el Kit AI, pesa 555 g contra 426-450 y compite con un arma de un tercio de energía: **no compite en igualdad**. Para comparar, en la liga de simulación de RoboCup establecer un orden defendible entre ocho equipos exigió más de mil partidos **por cada pareja** [26]. Aquí hay treinta combates en total.

6 El suelo se mueve

6.1 Ni el motor de física ni el ordenador son neutrales

Queda el nivel más incómodo: el suelo sobre el que corre todo. Se compararon siete configuraciones de física —tres detectores de colisión, dos solucionadores y tres motores distintos— sobre dos emparejamientos de control: el decisivo (v2 contra v3) y el reñido (v3 contra v9).



Fotograma de un clip. Vídeo en <https://combatbots.ai/research/banco-de-impacto/media/liguilla-v2-v3.mp4>

Figura 20. El único tramo concluyente del ciclo: v3 sobre v2, +2,75 puntos.

	v2 vs v3 (el decisivo)	v3 vs v9 (el refido)
ode	= v3	= v9
fcl	= v3	= v9
bullet	= v3	= v9
pgs	≠ sin ganador	= v9
mujoco	≠ v2	= v9
pybullet	= v3	≠ v3
physx	≠ v2	≠ v3

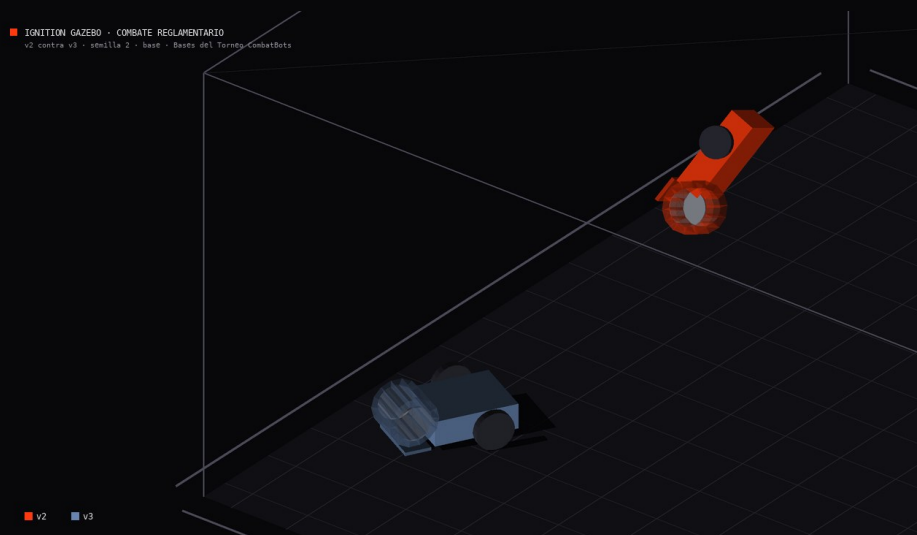
≠ : el motor da otro ganador que la referencia, o ninguno.

Cuatro de siete configuraciones discrepan en al menos un emparejamiento.

Figura 21. Acuerdo de cada configuración con la referencia. En rojo, las que dan otro ganador o ninguno.

Cuatro de siete discrepan en al menos un emparejamiento (tabla A.4). En el decisivo, el solucionador PGS no declara ganador, y MuJoCo y PhysX lo dan la vuelta.

En el reñido, PyBullet y PhysX eligen al otro. Que motores distintos den resultados distintos sobre el mismo modelo está documentado desde hace una década [11]; lo que aquí se añade es que la diferencia basta para cambiar quién gana un torneo.



Fotograma de un clip. Vídeo en

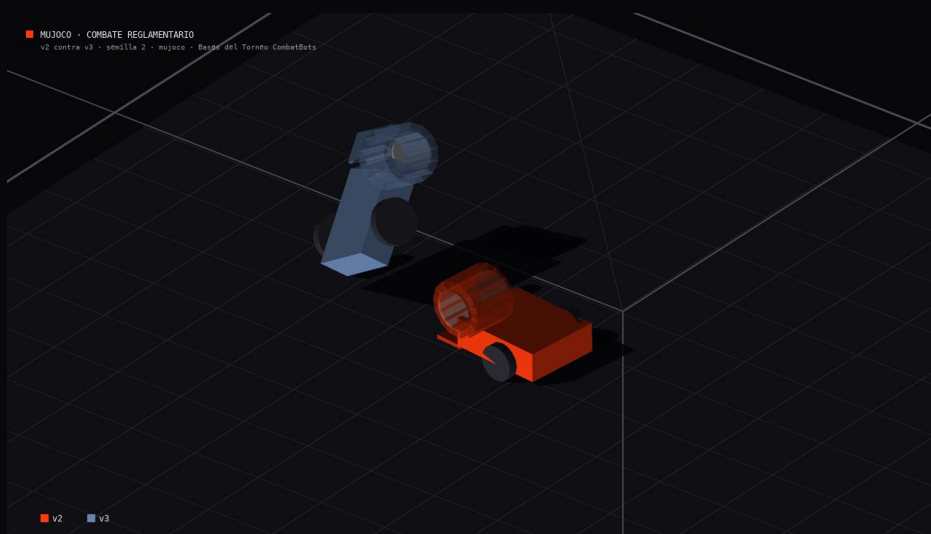
<https://combatbots.ai/research/banco-de-impacto/media/motor-base-v2-v3.mp4>

Figura 22. El emparejamiento decisivo con el motor de referencia: gana v3.

A esto hay que añadir dos hallazgos sobre el instrumento que sólo aparecen mirándolo de cerca: el detector de colisiones por defecto **no era el que la documentación decía** —se comprobó por identidad bit a bit en semillas sin contacto— y un motor alternativo carga sin protestar y deja las juntas a velocidad cero exacta durante mil pasos, es decir, congelado y en silencio [9].

Y la máquina. Se probó una más rápida para acortar las iteraciones. Con software idéntico al bit, las actas salieron **distintas en tres de tres semillas**: el KO se movía hasta 8,3 s y el reparto de puntos cambiaba en las tres. La causa es de precisión: las rutinas trigonométricas de la biblioteca matemática difieren en el último bit según la microarquitectura. La no asociatividad del punto flotante como fuente de irreproducibilidad está estudiada en computación de altas prestaciones, y su efecto en algoritmos iterativos es exactamente éste [29].

La máquina rápida no es el mismo banco corriendo antes: es otro banco. Sirve para explorar, nunca para comparar con lo publicado.



Fotograma de un clip. Vídeo en

<https://combatbots.ai/research/banco-de-impacto/media/motor-mujoco-v2-v3.mp4>

Figura 23. El mismo emparejamiento con MuJoCo, que resuelve el contacto de forma blanda en vez de rígida: gana **v2**. Mismo diseño, mismas reglas, otro ganador.

6.2 Lo que costó, y lo que no compró velocidad

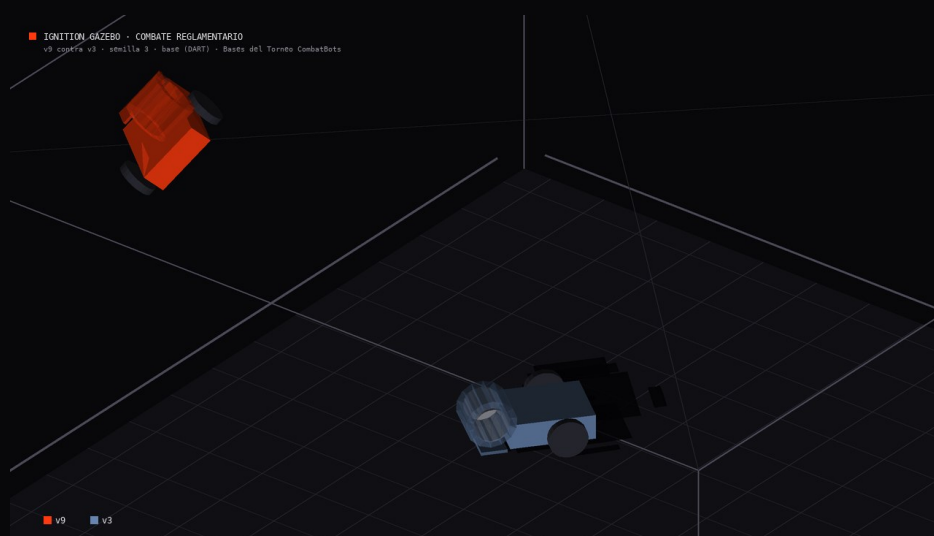
Una iteración completa cuesta entre 33 y 67 minutos según la tabla de arriba, y el grueso es arena, no banco. Se buscaron palancas para abaratarla y **dos pistas resultaron falsas**, que es información igual de útil [8]:

- **Un detector de colisiones distinto parecía un 20 % más barato.** Repetido bajo la misma carga, el ahorro desaparecía: era contención de la máquina, no del algoritmo.
- **El paso adaptativo prometía multiplicar por dos.** Se descartó al medirlo: cambiar el paso voltea el ganador en una de cada cuatro semillas.

Y la palanca que todo el mundo sugiere, la GPU, está medida y descartada: los motores usados integran en CPU, la integración temporal es serie pura y el problema de cada paso es demasiado pequeño para amortizar el envío. Lo que compra velocidad aquí son núcleos, no tarjeta.

7 Discusión

El resultado de diseño, contra la literatura. La mejora de masa efectiva es de 28,6 a 43,8 g, un factor 1,53, y produce un 96 % más de energía transferida. El orden de magnitud es coherente con lo que predice la mecánica del impacto: aumentos de masa efectiva en el rango 1,5–4 tienen un efecto sobre la velocidad de salida comparable al de aumentar la velocidad de golpeo, y el nuestro está en el extremo



Fotograma de un clip. Vídeo en

<https://combatbots.ai/research/banco-de-impacto/media/motor-base-v3-v9.mp4>

Figura 24. El emparejamiento reñido con la referencia: gana v9, por poco.

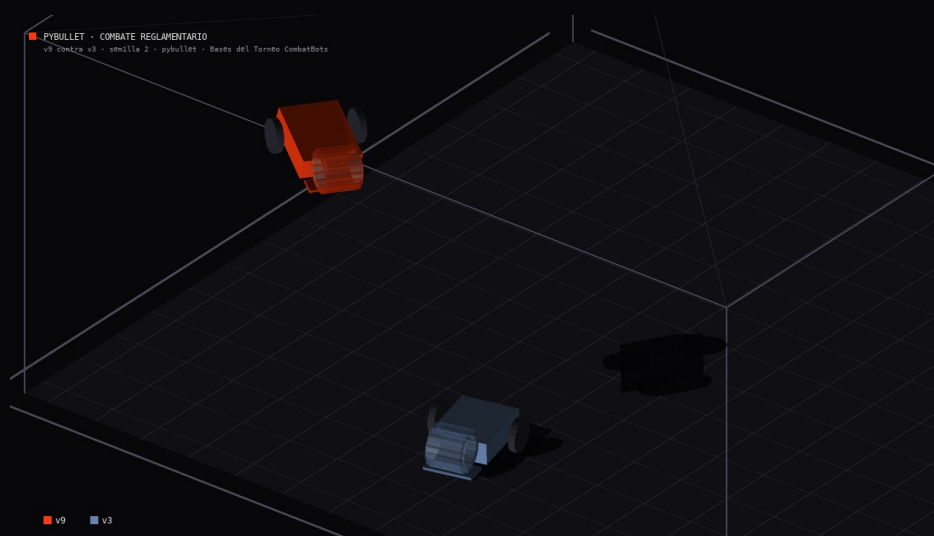
bajo de ese rango. La ganancia adicional viene de que la velocidad de punta también sube, de 40,9 a 54,4 m/s, al alejar el radio de impacto.

Y una divergencia con el análogo del volante de inercia que merece atención.

Si el objetivo fuera la energía almacenada por unidad de masa, la teoría de rotores empuja hacia perfiles de tensión repartida [21]. Aquí el objetivo es otro y el óptimo se desplaza: **gana el anillo de pared fina, no el disco**. El contraejemplo está en nuestros propios datos —el rival externo tiene el mayor diámetro de la serie, Ø60, y una de las masas efectivas más bajas, 28,2 g, porque su pared de 26 mm concentra la masa cerca del eje— y confirma que la intuición del volante no se traslada sin más al arma de combate. La restricción que sí liga es la tensión de aro, que en nuestro diseño queda en 3,67 MPa frente a un techo estimado de 15.

Sobre la magnitud de la mejora, conviene no exagerar lo que significa. Un 96 % suena mucho frente a las mejoras que reporta la optimización de forma de rotores, pero la comparación es injusta en nuestro favor: el punto de partida no era un diseño optimizado, sino **un tambor cuya masa estaba mal colocada** —el 1,4 % de sus vértices pasa de 25 mm de radio— y buena parte de la ganancia es corregir eso. La mejora mide la distancia desde un diseño ingenuo, no desde el estado del arte.

El instrumento, contra la literatura de validación. El trabajo reciente sitúa el error de un simulador de sólidos rígidos frente a impactos reales instrumentados en torno al 3,1 % en el salto de velocidad [34], con evaluaciones sistemáticas de varios simuladores contra medidas reales [1]. Ese es el listón. Nuestro trabajo no lo alcanza porque no tiene validación física, pero sí aporta un dato de la misma familia y que no hemos encontrado reportado: **la dispersión intrínseca inducida por el ángulo de contacto llega a un factor 3,7 dentro del mismo diseño**. Es una magnitud que no es error del simulador sino varianza del fenómeno, y que domina sobre el



Fotograma de un clip. Vídeo en

<https://combatbots.ai/research/banco-de-impacto/media/motor-pybullet-v3-v9.mp4>

Figura 25. El mismo, con PyBullet: gana v3. En el par reñido, cambiar de solucionador basta para invertir el resultado.

efecto que se quiere medir. En un banco de impacto con geometría dentada, cualquier comparación de una sola tirada carece de sentido.

Sobre el protocolo experimental. El patrón de cribar con pocas repeticiones y confirmar con muchas es práctica establecida en experimentación computacional [35], y la asignación óptima del presupuesto entre candidatos y réplicas tiene teoría propia [3]. Aquí se aplicó de forma heurística —ocho y treinta y dos— y el efecto del sesgo de selección resultó ser de 40 puntos porcentuales, lo que sugiere que el cribado a ocho estaba infradimensionado.

Alcance. Este trabajo optimiza morfología con control fijo, no codiseño [30][28]. Los dos desacuerdos que documentamos —banco frente a arena, y ciclos de no transitividad— sugieren que esa separación tiene un coste: el mejor rotor aislado no es necesariamente el que gana combates, porque el resultado depende también de cómo se pilota y de contra quién.

8 Lo que este trabajo no dice

No dice si algo rompe. El simulador es de sólidos rígidos: mide impulso, energía y vuelo, no fractura. La tensión de aro se calcula aparte y con un techo que es una estimación.

No dice cómo se comporta contra un rival que embiste. La diana está quieta y es un ladrillo.

No hay paredes en el banco. La diana vuela más de 1,3 m; en la arena de 150 ×

150 cm habría chocado.

El **modelo de daño es una hipótesis**, y el piloto es el mismo autómatas para los dos robots: no hay nada del piloto humano, que es media competición. Tampoco hay fiabilidad —ni electrónica que falle, ni conectores que se suelten, ni batería que caiga.

Y una honestidad que va contra la tesis de este artículo: **la brecha entre simulación y realidad no es igual en todos los fenómenos**. Para locomoción sobre superficies duras la transferencia funciona razonablemente y hay técnicas maduras para estrecharla [32][24]. El impacto a 40 m/s entre piezas rígidas es de los casos peores, no de los típicos [12]. Que aquí el instrumento pese tanto no autoriza a concluir que la simulación no sirva.

Y el límite que engloba a todos: ninguna cifra de este artículo se ha comprobado golpeando nada. Todo es simulación.

8.1 Trabajo futuro: la validación física

La validación pendiente es concreta y barata. **Construir los dos diseños, enfrentarlos a una misma diana instrumentada y medir**: un acelerómetro en el blanco da la energía transferida, y una cámara rápida sobre el rotor da la caída de revoluciones. Son las dos magnitudes que el banco predice, y bastan para decidir si acierta el orden, el factor, o ninguno de los dos.

Ese ensayo tiene además un valor que va más allá de confirmar una cifra: **la caída de revoluciones es un contraste independiente del solucionador**, porque sale de la conservación del momento angular y no del reparto de impulso. Si el banco acierta en ésta y falla en la energía, el error está localizado en el modelo de contacto y no en el diseño.

El trabajo deja además tres mejoras de protocolo identificadas y sin ejecutar: repetir las tandas en **orden inverso** para cancelar el sesgo de rumbo del piloto, **igualar el número de repeticiones** entre emparejamientos, y sustituir el criterio de cima única por un **archivo de rivales** que tolere los ciclos, siguiendo lo que la literatura de coevolución propone desde hace dos décadas [14].

8.2 Conclusiones

Se ha diseñado y optimizado el arma de un robot de combate de 500 g enteramente en simulación, con una mejora del **96 % en energía transferida** confirmada sobre 32 repeticiones y validada en competición con un 10-0 bajo reglamento. El criterio de diseño que gobierna esa mejora es la **distribución radial de la masa y no su**

cantidad: en esta clase, el límite de peso no es la restricción activa, y el mejor rotor de la serie no es el de mayor diámetro.

La aportación metodológica es la que se generaliza. Una medida simulada de impacto sólo significa algo si el actuador está limitado en par, si el paso de integración está convergido, si el ángulo de contacto se marginaliza en lugar de fijarse y si el modelo paramétrico reproduce la pieza real. Omitir cualquiera de los cuatro produce un artefacto mayor que el efecto buscado, y ninguno de ellos avisa: la simulación corre, escribe números y no se queja.

Los tres desacuerdos documentados —banco frente a arena, ciclos de no transitividad, y motores de contacto que eligen ganadores distintos— no invalidan el diseño, pero sí acotan qué se puede afirmar con este instrumento: **sirve para ordenar diseños dentro de un mismo montaje, y no para predecir un resultado absoluto**. Es, en sentido estricto, un banco de pruebas: un instrumento cuyo valor está en enseñar dónde falla el modelo antes de gastar material [6]. Y mientras no exista la validación física descrita arriba, esa afirmación es todo lo que sostiene el trabajo.

9 Apéndice

Las tablas que el texto cita y que no cabían en el hilo. Todas salen de los ficheros del banco; ninguna cifra está medida en el mundo real.

9.1 Tabla A.1 · Parámetros de los diseños

Cotas y magnitudes derivadas de cada contendiente. *m.* efectiva es la masa que el rival nota en el impacto, I/r^2 , y no la masa del tambor: es la diferencia entre *v1* y *v2* mejor resumida.

Diseño	Masa (g)	r cuerpo (mm)	r punta (mm)	Pared (mm)	Dientes	Ancho (mm)	m. efectiva (g)	Punta (m/s)	Almacena (J)
v1	426.0	17.87	28.58	13.87	2	71.44	28.6	40.9	23.97
v2	426.0	27.0	38.0	6.23	2	71.44	43.8	54.4	64.84
v3	445.0	27.0	38.0	6.23	2	71.44	43.8	54.4	64.84
v5	445.0	27.0	38.0	6.23	2	71.44	43.8	54.4	64.84
v9	445.0	27.0	38.0	6.23	2	71.44	43.8	54.4	64.84
fk	450.0	29.98	37.5	25.98	1	28.0	28.2	53.7	40.67
v2ia	555.4	22.8	25.0	7.99	2	62.5	36.2	35.8	23.21

9.2 Tabla A.2 · Las 32 tiradas de cada finalista

Estadística de la confirmación. El **factor** es el cociente entre la peor tirada y la mejor del mismo diseño, y es la cifra que obliga a marginalizar la fase en vez de fijarla.

Candidato	n	Media (J)	Error estándar	Mínima	Máxima	Factor
v1 (referencia)	32	1.800	±0.115	0.65	2.41	× 3.7
A · ganador del bucle	32	3.532	±0.331	1.25	10.98	× 8.8
B · A + 3S (11,4 V)	32	3.759	±0.281	1.13	8.64	× 7.6
C · A + 454 g	32	3.675	±0.242	0.74	6.18	× 8.3
D · A + 3S + 454 g	32	3.450	±0.252	0.30	5.42	× 17.9

9.3 Tabla A.3 · Los veinte ensayos de la búsqueda manual

Todos los candidatos de las ocho rondas. Δ vs vigente es la mejora sobre el diseño en curso en esa ronda, no sobre v1. **Dispersión** es el recorrido entre las ocho tiradas de la criba, en porcentaje de la media: la columna que explica por qué ocho no bastan para publicar. **Caída del arma** es cuánta velocidad pierde el tambor en el impacto, y es el control de sanidad: si no cae, la energía la está poniendo el controlador y no la inercia.

9.4 Tabla A.4 · Comparación de motores de física

Ganador declarado por cada configuración en los dos emparejamientos de control, con su margen en puntos. La marca *discrepa* señala las que **no coinciden** con la referencia. Márgenes negativos indican que gana el segundo del par.

Referencias

- [1] Brian Acosta, William Yang, and Michael Posa. Validating robotics simulators on real-world impacts. *IEEE Robotics and Automation Letters*, 7(3):6471–6478, 2022. doi: 10.1109/LRA.2022.3174367.
- [2] Gavin C. Cawley and Nicola L. C. Talbot. On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, 11:2079–2107, 2010.
- [3] Chun-Hung Chen, Jianwu Lin, Enver Yücesan, and Stephen E. Chick. Simulation budget allocation for further enhancing the efficiency of ordinal optimization. *Discrete Event Dynamic Systems*, 10(3):251–270, 2000. doi: 10.1023/A:1008349927281.
- [4] COMBATBOTS. El banco de impacto: cómo se decide qué entra en v2. Documentación interna, robots/1.COMBATBOTS/opt/README.md, 2026.
- [5] COMBATBOTS. Bases y normativa del torneo COMBATBOTS. Categoría 1 lb Plastic. Documento de competición, 3D Printer Party Valladolid, 2026.
- [6] COMBATBOTS. Capacidad frente a agresividad como doctrina: Ciber-guerreros de combate. COMBATBOTS · Investigación, ensayo n.º 0. <https://combatbots.ai/research/capacidad-vs-agresividad/>, 2026.

- [7] COMBATBOTS. Contrato del bucle evolutivo: proponer, cribar, remedir, arena, decidir. Documentación interna, `opt/evolucion/CONTRATO.md`, 2026.
- [8] COMBATBOTS. Dónde se va el tiempo: coste de la simulación. Documentación interna, `opt/coste-de-la-simulacion.md`, 2026.
- [9] COMBATBOTS. Reconocimiento de motores de física. Documentación interna, `opt/motores/reconocimiento.md`, 2026.
- [10] Rod Cross. Impact of a ball with a bat or racket. *American Journal of Physics*, 67(8):692–702, 1999. doi: 10.1119/1.19354. Deriva la masa efectiva en el punto de impacto de un cuerpo extendido: $m_{ef} = M/(1 + r^2/k^2)$.
- [11] Tom Erez, Yuval Tassa, and Emanuel Todorov. Simulation tools for model-based robotics: Comparison of Bullet, Havok, MuJoCo, ODE and PhysX. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 4397–4404, 2015. doi: 10.1109/ICRA.2015.7139807.
- [12] Nima Fazeli, Elliott Donlon, Evan Drumwright, and Alberto Rodriguez. Empirical evaluation of common contact models for planar impact. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 3418–3425, 2017. doi: 10.1109/ICRA.2017.7989389.
- [13] Nima Fazeli, Samuel Zapolsky, Evan Drumwright, and Alberto Rodriguez. Fundamental limitations in performance and interpretability of common planar rigid-body contact models. arXiv:1710.04979; ISRR, 2017.
- [14] Sevan G. Ficici and Jordan B. Pollack. A game-theoretic memory mechanism for coevolution. In *Genetic and Evolutionary Computation (GECCO)*, pages 286–297, 2003. doi: 10.1007/3-540-45105-6_35.
- [15] G. Gilardi and I. Sharf. Literature survey of contact dynamics modelling. *Mechanism and Machine Theory*, 37(10):1213–1239, 2002. doi: 10.1016/S0094-114X(02)00045-9.
- [16] Sami Haddadin, Alin Albu-Schäffer, and Gerd Hirzinger. Requirements for safe robots: measurements, analysis and new insights. *The International Journal of Robotics Research*, 28(11–12):1507–1527, 2009. doi: 10.1177/0278364909343970. Establece la masa efectiva y la velocidad como los dos parámetros que gobiernan la severidad de un impacto.
- [17] Illinois-Indiana ASEE. Design and fabrication of beetleweight combat robot with hubless equator blade. In *ASEE Illinois-Indiana Regional Conference*, 2021. Trabajo docente de grado: diseño conceptual y fabricación, sin optimización paramétrica.
- [18] Hiroaki Kitano, Minoru Asada, Yasuo Kuniyoshi, Itsuki Noda, Eiichi Osawa, and Hitoshi Matsubara. RoboCup: A challenge problem for AI. *AI Magazine*, 18(1), 1997. Define la competición como un «standard problem» de investigación, no como espectáculo.

- [19] Sylvain Koos, Jean-Baptiste Mouret, and Stéphane Doncieux. The transferability approach: Crossing the reality gap in evolutionary robotics. *IEEE Transactions on Evolutionary Computation*, 17(1):122–145, 2013.
- [20] Marc Lanctot, Kate Larson, Yoram Bachrach, Luke Marris, Zun Li, Avishkar Bhoopchand, Thomas Anthony, Brian Tanner, and Anna Koop. Evaluating agents using social choice theory. arXiv:2312.03121, 2023.
- [21] Xiaojun Li and Alan Palazzolo. A review of flywheel energy storage systems: state of the art and opportunities. *Journal of Energy Storage*, 2021. doi: 10.1016/j.est.2021.103576.
- [22] Hod Lipson and Jordan B. Pollack. Automatic design and manufacture of robotic lifeforms. *Nature*, 406:974–978, 2000. doi: 10.1038/35023115.
- [23] Marco Antonio Meggiolaro. *RioBotz Combat Robot Tutorial*. CreateSpace, 2 edition, 2009. Manual técnico, no revisado por pares; es el texto de referencia de la comunidad de combate robótico.
- [24] Fabio Muratore, Fabio Ramos, Greg Turk, Wenhao Yu, Michael Gienger, and Jan Peters. Robot learning from randomized simulations: A review. *Frontiers in Robotics and AI*, 9:799893, 2022. doi: 10.3389/frobt.2022.799893.
- [25] Sardar Pashaie, Tobias Jung, and Hamid Golmohammadi. Tracking scientific output in robotic soccer and AI: A bibliometric overview. *AI and Tech in Behavioral and Social Sciences*, 3(3), 2025.
- [26] Mikhail Prokopenko, Peter Wang, Sebastian Marian, Aijun Bai, Xiao Li, and Xiaoping Chen. RoboCup 2D soccer simulation league: Evaluation challenges. In *RoboCup 2017: Robot World Cup XXI, LNCS 11175*, pages 325–337, 2018.
- [27] Ludovic Righetti and Vincent Boulanin. Navigating the dual-use dilemma in open-source robotics. *IEEE Spectrum*, 2025.
- [28] Charles Schaff, David Yunis, Ayan Chakrabarti, and Matthew R. Walter. Jointly learning to construct and control agents using deep reinforcement learning. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 9798–9805, 2019. doi: 10.1109/ICRA.2019.8793537.
- [29] Sanjif Shanmugavelu, Mathieu Tallefumier, Christopher Culver, Oscar Hernandez, Mark Coletti, and Ada Sedova. Impacts of floating-point non-associativity on reproducibility for HPC and deep learning applications. In *SC '24 Workshops*, pages 170–179, 2024. doi: 10.1109/SCW63240.2024.00028.
- [30] Karl Sims. Evolving virtual creatures. In *SIGGRAPH '94*, pages 15–22, 1994. doi: 10.1145/192161.192167.
- [31] James E. Smith and Robert L. Winkler. The optimizer’s curse: Skepticism and postdecision surprise in decision analysis. *Management Science*, 52(3):311–322, 2006. doi: 10.1287/mnsc.1050.0451.

- [32] Josh Tobin, Rachel Fong, Alex Ray, Jonas Schneider, Wojciech Zaremba, and Pieter Abbeel. Domain randomization for transferring deep neural networks from simulation to the real world. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 23–30, 2017. doi: 10.1109/IROS.2017.8202133.
- [33] Diego Torricelli et al. Making bipedal robot experiments reproducible and comparable: The Eurobench software approach. *Frontiers in Robotics and AI*, 2022. doi: 10.3389/frobt.2022.951663.
- [34] Jari van Steen, Daan Stokbroekx, Nathan van de Wouw, and Alessandro Saccon. Impact-aware robotic manipulation: Quantifying the sim-to-real gap for velocity jumps. arXiv:2411.06319, 2024. Error medio del 3,1 % entre la velocidad post-impacto simulada y la medida.
- [35] Felipe A. C. Viana. A tutorial on latin hypercube design of experiments. *Quality and Reliability Engineering International*, 32:1975–1985, 2016. doi: 10.1002/qre.1924.

Ronda	Alteración	Energía (J)	Δ vs vigente	EE	Dispersión m. efectiva	ω régi- men	Caída del arma	
R0	v1 (línea de base)	1.74	+0.0 %	± 17.3 %	106 %	28.6 g	1348 rad/s	61.5 %
R1	cuerpo Ø54	3.59	+107.0 %	± 8.3 %	63 %	49.0 g	1328 rad/s	43.4 %
R1	cuerpo Ø60	3.40	+95.7 %	± 18.0 %	139 %	51.2 g	1272 rad/s	40.7 %
R1	cuerpo Ø48	3.04	+75.0 %	± 13.7 %	94 %	42.6 g	1344 rad/s	57.2 %
R1	cuerpo Ø42	2.51	+44.6 %	± 13.3 %	87 %	34.8 g	1347 rad/s	63.6 %
R2	454 g · pared	3.84	+6.8 %	± 15.9 %	132 %	60.1 g	1287 rad/s	40.4 %
R2	440 g · pared	3.43	-4.6 %	± 19.9 %	137 %	55.1 g	1308 rad/s	42.5 %
R3	3S · 11,4 V	4.23	+17.8 %	± 21.9 %	182 %	49.0 g	1648 rad/s	33.4 %
R3	3S · kv 1400	4.15	+15.5 %	± 20.5 %	154 %	49.0 g	1606 rad/s	36.6 %
R3	kv 2300 · 2S	3.33	-7.2 %	± 11.2 %	109 %	49.0 g	1279 rad/s	44.8 %
R4	3 dientes	2.86	-20.5 %	± 12.2 %	99 %	51.8 g	1320 rad/s	34.2 %
R4	4 dientes	2.22	-38.1 %	± 16.4 %	115 %	54.6 g	1310 rad/s	29.7 %
R4	6 dientes	1.63	-54.6 %	± 18.6 %	166 %	60.2 g	1286 rad/s	21.0 %
R5	punta 38 mm	4.08	+13.5 %	± 13.0 %	129 %	43.8 g	1207 rad/s	40.1 %
R5	punta 32 mm	3.49	-2.9 %	± 17.6 %	125 %	46.9 g	1314 rad/s	46.5 %
R5	punta 35 mm	3.19	-11.3 %	± 19.6 %	145 %	44.9 g	1278 rad/s	36.7 %
R6	eje a 32 mm	4.14	+1.4 %	± 17.8 %	167 %	43.8 g	1208 rad/s	36.9 %
R6	eje a 20 mm	4.02	-1.3 %	± 8.8 %	85 %	43.8 g	1208 rad/s	40.4 %
R6	eje a 26 mm	3.73	-8.5 %	± 11.5 %	85 %	43.8 g	1208 rad/s	38.4 %
R7	500 g · tambor más largo	3.41	-16.4 %	± 13.5 %	98 %	48.4 g	1112 rad/s	42.4 %

Configuración	v2 vs v3	Margen	v3 vs v9	Margen
ode	v3	-3.2	v9	-2.4
fcl	v3	-2.2	v9	-1.5
bullet	v3	-1.8	v9	-1.0
pgs	sin ganador · discrepa	0.0	v9	-2.4
mujoco	v2 · discrepa	1.8	v9	0.6
pybullet	v3	-1.2	v3 · discrepa	1.8
physx	v2 · discrepa	0.0	v3 · discrepa	0.4
<i>referencia</i>	v3	-1,6	v9	—