Basado en la literatura proporcionada, las ventajas del Random Survival Forest (RSF) y XGBoost (especialmente en su variante AFT o Cox-gradient boosting) frente al modelo tradicional de Cox Proportional Hazards (Cox-PH) al analizar datos con interacciones complejas y alta dimensionalidad se pueden categorizar en cuatro áreas clave:

1. Captura Automática de No-Linealidades e Interacciones

La ventaja más destacada de los modelos basados en árboles (RSF y XGBoost) es su capacidad intrínseca para modelar relaciones no lineales sin necesidad de especificación previa por parte del investigador.

- **Limitación de Cox-PH:** El modelo de Cox asume que el logaritmo del riesgo (*log-risk*) es una función lineal de las covariables 1. Esto implica que las interacciones y efectos no lineales deben ser conocidos *a priori* y modelados explícitamente (por ejemplo, mediante transformaciones de covariables o términos de interacción), lo cual es difícil y propenso a sesgos en entornos complejos 2, 3.
- **Ventaja de ML:** Los algoritmos de aprendizaje automático como RSF y XGBoost permiten que los datos "dicten la forma del modelo", capturando automáticamente todas las posibles interacciones y efectos no lineales entre las variables a través de la división recursiva de nodos 4.
- **Evidencia empírica:** En un estudio sobre cáncer de mama, se demostró mediante valores SHAP que la superioridad de XGBoost (C-index 0.73) frente a Cox (C-index 0.63) se debía a la capacidad de XGBoost para capturar efectos de interacción complejos (ej. entre la edad y el estadio del tumor) que el modelo de Cox, al asumir independencia de características, no pudo detectar 5, 6.

2. Manejo de Alta Dimensionalidad

Cuando el número de predictores es alto (datos de alta dimensionalidad, como datos ómicos o registros médicos electrónicos extensos), los modelos tradicionales encuentran dificultades significativas.

- **Limitación de Cox-PH:** El modelo de Cox tradicional no está diseñado para manejar conjuntos de datos complejos de alta dimensionalidad; se vuelve inestable o inadecuado cuando el número de características excede el número de instancias de datos, o cuando existe multicolinealidad 7, 2.
- **Ventaja de RSF:** El RSF es particularmente eficaz en entornos de alta dimensionalidad porque selecciona un subconjunto aleatorio de predictores candidatos en cada división del nodo del árbol, lo que decorrelaciona los árboles individuales y reduce la varianza 8.
- **Ventaja de XGBoost:** XGBoost implementa regularización (penalización de la complejidad del modelo) y algoritmos eficientes que le permiten manejar grandes cantidades de predictores y evitar el sobreajuste mejor que los modelos lineales estándar 9.

3. Independencia de Supuestos Restrictivos (Riesgos Proporcionales)

Los modelos de *Machine Learning* ofrecen flexibilidad frente a las rígidas asunciones estadísticas de los modelos clásicos.

- **Violación de PH:** El modelo de Cox depende del supuesto de riesgos proporcionales (las curvas de supervivencia de dos grupos no se cruzan). Si este supuesto se viola, las predicciones de Cox pueden ser inexactas 10, 11.
- **Flexibilidad de ML:**
- **RSF:** Es un método que no asume riesgos proporcionales (non-PH), permitiendo que el ranking de riesgo entre individuos cambie con el tiempo (es decir, las curvas de supervivencia pueden cruzarse) 12, 13.
- **XGBoost-AFT:** La implementación de modelos de Tiempo de Falla Acelerada (AFT) dentro de XGBoost permite modelar directamente el tiempo hasta el evento y ofrece un mejor ajuste cuando el supuesto de riesgos proporcionales no se cumple, capturando patrones que Cox podría perder 14, 15.

4. Eficiencia Computacional y Escalabilidad (Específico de XGBoost)

Aunque el RSF es potente, puede ser computacionalmente costoso. XGBoost ofrece ventajas técnicas específicas en el procesamiento de grandes volúmenes de datos.

- **Velocidad y Hardware:** XGBoost utiliza derivadas de segundo orden para acelerar la convergencia y es capaz de utilizar GPUs (Unidades de Procesamiento Gráfico) de NVIDIA, logrando una aceleración sustancial (6-7 veces más rápido que en CPU) en conjuntos de datos grandes 16, 15, 17.
- **Escalabilidad:** A diferencia de los métodos tradicionales que pueden requerir pasos costosos para estimar la función de riesgo base (como el estimador de Breslow) para predecir el tiempo, XGBoost con función de pérdida AFT puede predecir tiempos de supervivencia directamente usando solo los parámetros ajustados, lo cual es más eficiente en contextos de *Big Data* 18, 14.

En resumen, mientras que Cox-PH es valorado por su interpretabilidad en escenarios simples, RSF y XGBoost son superiores cuando los datos presentan **relaciones no lineales desconocidas**, **interacciones complejas** o **alta dimensionalidad**, ya que estos algoritmos optimizan la precisión predictiva sin las restricciones paramétricas tradicionales 19, 20.